

Regelungstechnik

von

Prof. Dr.-Ing. Oliver Nelles



Mein beruflicher Werdegang

- 1969 Geboren in Frankfurt/M
- 1988 Abitur
- 1993 Dipl.-Ing. Elektrotechnik (Regelungstechnik) an der TU Darmstadt
- 1999 Dr.-Ing. (Automatisierungstechnik) an der TU Darmstadt
Thema: Identifikation nichtlinearer Systeme mit neuronalen Netzen
- 1999 – 2000 Postdoktorand am Mechanical Engineering Dept. an der UC Berkeley
- 2000 – 2001 Projektleiter bei Siemens Automobiltechnik in Regensburg
- 2001 – 2004 Gruppenleiter bei Siemens VDO Automotive in Regensburg
Bereich: Elektronische Getriebesteuerungen, Funktionsentwicklung
- 2004 – ... Professor für Mess- und Regelungstechnik, Mechatronische Systeme,
Fachbereich Maschinenbau, Universität Siegen

Paul-Bonatz-Str. 9–11 • Raum PB-A 248, 249 • Sekretariat: Tel.: 0271 / 740 - 4630 •
Tel.: 0271 / 740 - 4045 • E-mail: oliver.nelles@uni-siegen.de

Lehrangebote

- **Regelungstechnik** (WS, 5 CP) 5. Sem. B.Sc.
- **Signalverarbeitung** (SS, 5 CP) 2. Sem. M.Sc.
- 2 Versuche **Messtechniklabor** 4. Sem B.Sc.
- 2 Versuche **Maschinenlabor** 5. Sem B.Sc.

- **Digitale Regelung** (SS, 3 CP) B.Sc. und M.Sc.
- **Systemidentifikation** (SS, 3 CP) B.Sc. und M.Sc.
- **Neuronale Netze und Fuzzy-Systeme** (WS, 3 CP) M.Sc.

- **Sensorics** (SS, 5 CP) 2. Sem. M.Sc. Mechatronics

Inhaltsverzeichnis RT

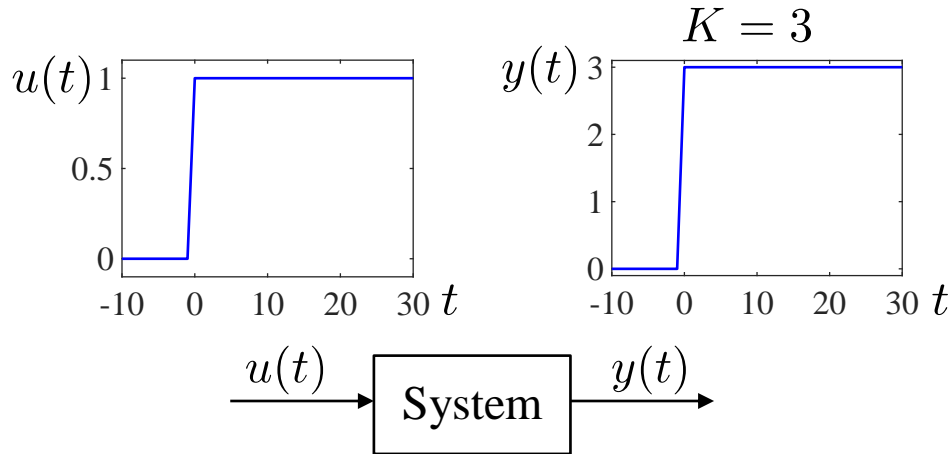
1. Einführung in die Regelungstechnik
 2. Modellierung linearer dynamischer Systeme
 3. Linearisierung nichtlinearer Systeme
 4. Laplace-Transformation
 5. Übertragungsfunktion
 6. Frequenzgang und Ortskurve
 7. Wichtige dynamische Systeme
 8. Stabilität linearer Systeme
 9. Quantitative Stabilitätskriterien
 10. Einfache lineare Regler
 11. Reglerentwurf mittels Optimierung oder Einstellregeln
 12. Reglerentwurf mittels Kompensation
 13. Wurzelortskurve (WOK)
 14. *Reglerentwurf im Frequenzbereich
 15. Vertiefungen und Erweiterungen
 16. Nichtlineare Regelung
 17. *Digitale Regelung
- * = wird in der Vorlesung nicht behandelt, und ist auch kein Klausurstoff.

Inhalt Kapitel 1

$$\text{Änderungsrate von } y(t): \dot{y}(t) = \frac{dy(t)}{dt}$$

Was ist ein dynamisches System?

Statisch:



- Algebraischer Zusammenhang: Funktion

- Im linearen Fall:

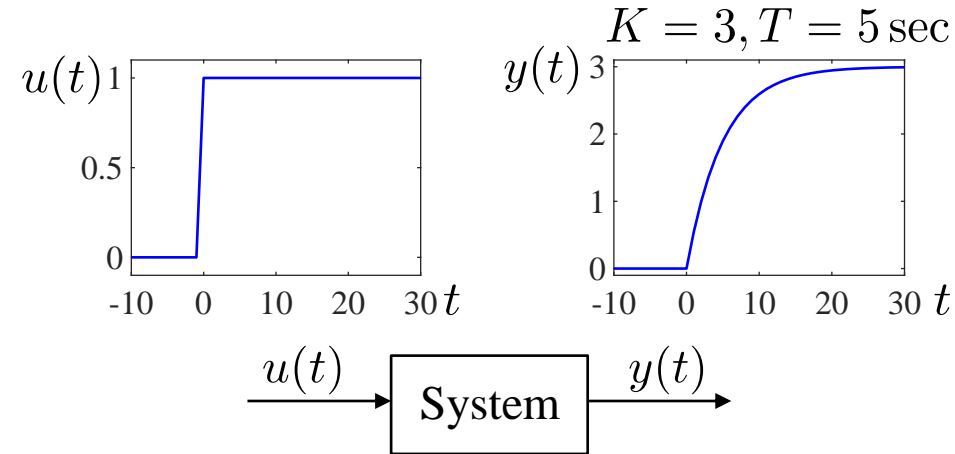
$$y(t) = K \cdot u(t)$$

- Mögliche nichtlineare Fälle:

$$y(t) = K \cdot \sqrt{u(t)}$$

$$y(t) = K \cdot u^2(t)$$

Dynamisch:



- Differentialgleichung (DGL)

- Im linearen Fall, einfachstes Bsp.:

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

$T \rightarrow 0$: statisch

$T \rightarrow \infty$: integrierend $y(t) \approx \frac{K}{T} \int u(t) dt$

- Nichtlineare Fälle \rightarrow Kapitel 3

1. Einführung in die Regelungstechnik

Inhalt Kapitel 1

Einführung in die Regelungstechnik

- 1.1 Faszination Regelungstechnik
- 1.2 Geschichte der Regelungstechnik
- 1.3 Regelkreis
 - a) Sensoren: Die Augen
 - b) Aktoren: Die Muskeln
 - c) Regler: Das Gehirn
- 1.4 Steuern versus Regeln
- 1.5 Prinzip Rückkopplung
- 1.6 Automatisierungstechnik
- 1.7 Kybernetik
- 1.8 Arbeitsmarkt
- 1.9 Wichtige Begriffe in englischer Sprache
- 1.10 Literatur

1.1 Faszination Regelungstechnik

Anwendungsgebiete der Regelungstechnik:

- Chemie-, Verfahrens- und Prozesstechnik
 - Energietechnik
 - Luft- und Raumfahrt
 - Militär
 - Automobil- und Verkehrstechnik
 - Mechatronik
 - Werkzeugmaschinen, Robotik
 - Bio-/Medizintechnik
 - Kraftwerksautomatisierung
 - Fertigungsautomatisierung
 - Gebäudeautomatisierung
 - Volkswirtschaft, Sozialwissenschaften
- Viele **verschiedene** Gebiete, **eine** gemeinsame Theorie!

1.1 Faszination Regelungstechnik

Windkraftwerk, Dänemark



Ölbohrinsel, Golf von Mexiko



Wasserkraftwerk Itaipu

Kernkraftwerk Grohnde



1.1 Faszination Regelungstechnik

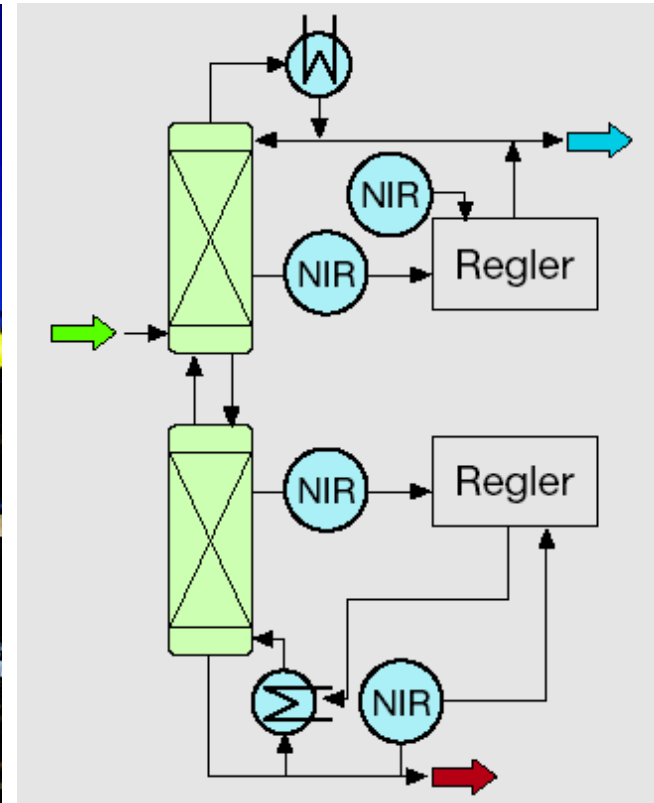
Eine moderne großindustrielle Anlage: Ein Teil der OMV-Ölraffinerie in Österreich



1.1 Faszination Regelungstechnik

Merkmale vieler biotechnologischer Prozesse:

- Materialstrom oft nicht kontinuierlich sondern als (Semi-)Batch-Prozess.
- Interessierende Größen oft nicht (in Echtzeit) messbar: Konzentrationen.
- Modellierung schwierig.



1.1 Faszination Regelungstechnik

Stahlproduktion und -verarbeitung

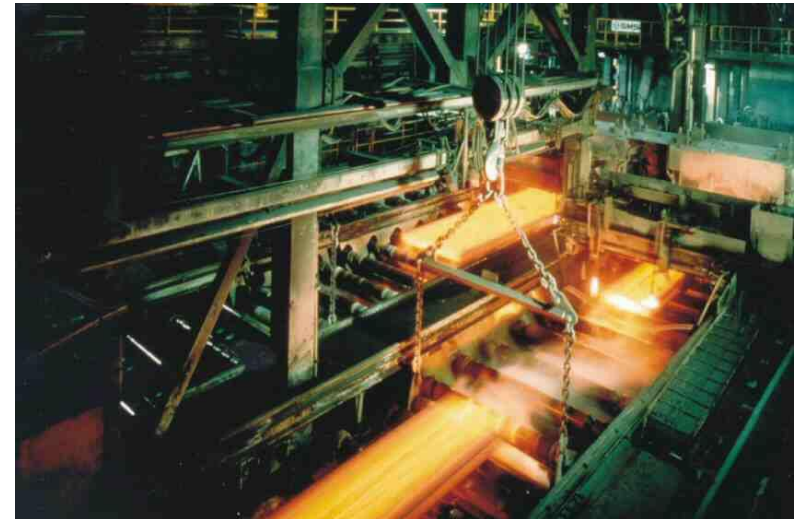
Hochofen 5 Rogesa, Dillingen



Elektroofen der Georgsmarienhütte



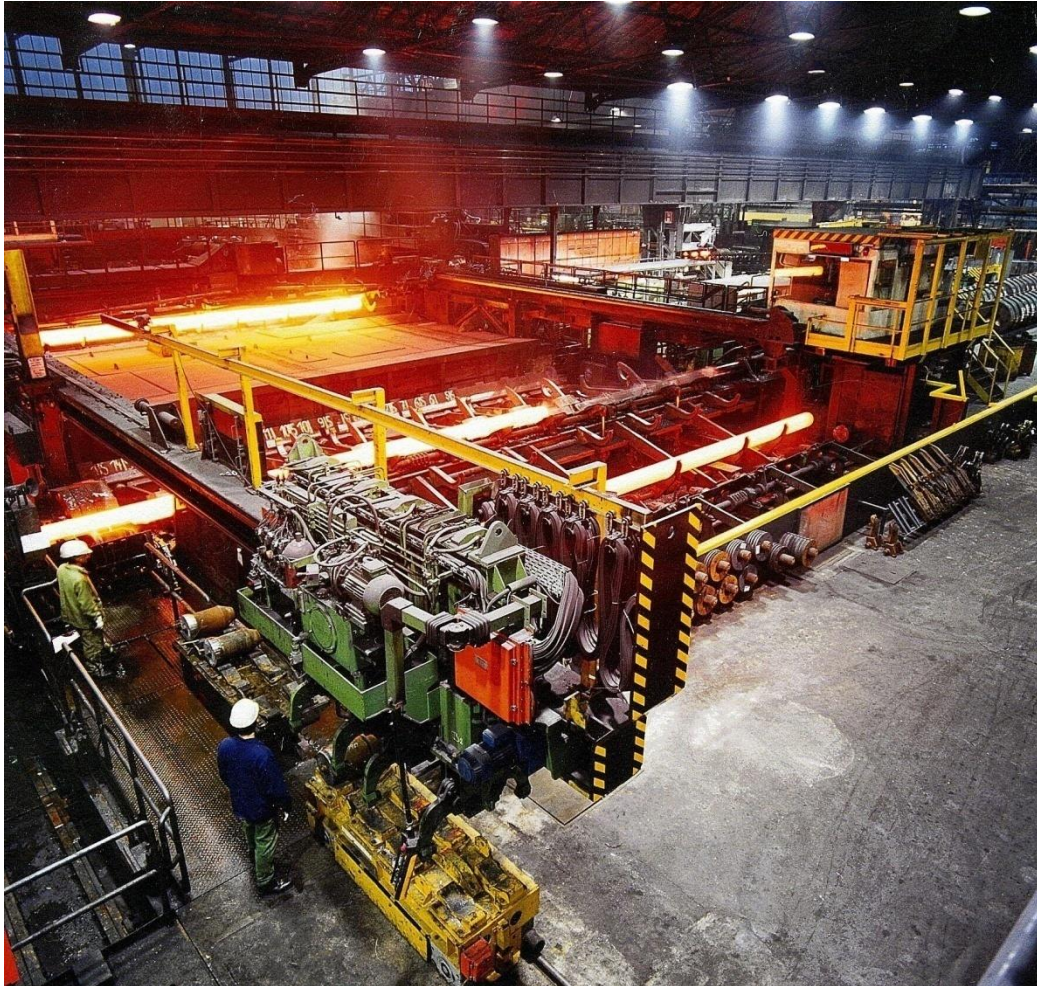
Auslauf einer Stranggießanlage



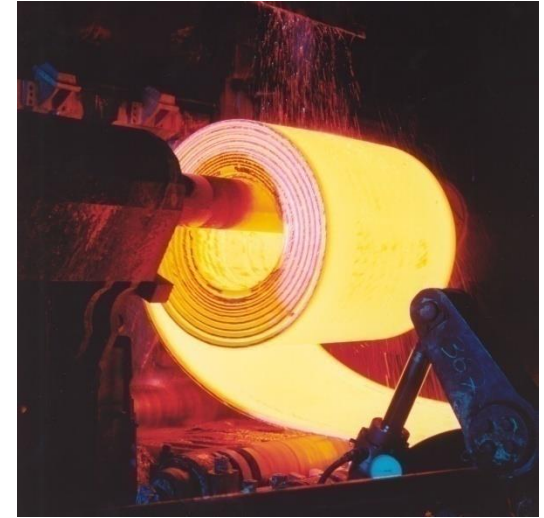
1.1 Faszination Regelungstechnik

Stahlproduktion und -verarbeitung

Rohrwalzwerk



Warmbandcoil



Leitstand zur Überwachung von Strangguß



1.1 Faszination Regelungstechnik

Stahlproduktion und -verarbeitung

Kaltwalzwerk



Bramme



Versandfertige Coils



1.1 Faszination Regelungstechnik



Quelle der nächsten 3 Seiten:

[http://www05.abb.com/global/scot/scot227.nsf/veritydisplay/97910b41591f1f99c125789d003c1ec3/\\$file/techn_regelungen_walzwerke_3bdd017150de_final.pdf](http://www05.abb.com/global/scot/scot227.nsf/veritydisplay/97910b41591f1f99c125789d003c1ec3/$file/techn_regelungen_walzwerke_3bdd017150de_final.pdf)

„Wichtige Qualitätsparameter beim Walzen

Die bedeutendsten Qualitätsparameter beim Walzen sind Oberfläche, Planheit, Materialhomogenität und die Banddickentoleranzen. Zur Kostenoptimierung und zur maximalen Materialausnutzung sind enge Dickentoleranzen von höchster Bedeutung; denn dadurch kann das Band so nah wie möglich auf die zulässige Mindestdicke heruntergewalzt werden.



1.1 Faszination Regelungstechnik



Technologische Regelungen für den Kaltwalzprozess Schlüsselfunktionen für höchste Dicken- und Planheitsqualität

Walzspaltregelung

Für die Walzspaltregelung können zwei Betriebsarten, Positions- oder Walzkraftregelung, angewählt werden. Die Auswahl des Modus hängt von der Art der übergeordneten automatischen Dickenregelung (AGC) und der Einfädelstrategie ab.

Um eine konstante Anregelzeit über den gesamten Stellbereich sicherstellen zu können, sind verschiedene prozessbezogene Adaptionen und Überwachungen enthalten.

AGC für Reduktionsgerüste

In Abhängigkeit des Sensor-Konzepts, basierend auf den Kundenanforderungen, sind verschiedene Dickenregelungsstrategien möglich:

- Monitorregelung
- Dickenvorsteuerung
- Massenflussregelung
- Geschwindigkeitsvorsteuerung
- Walzenexzentrizitätskompensation
- Zugvorsteuerung

Geschwindigkeitsvorsteuerung

Mit der Geschwindigkeitsvorsteuerung werden geschwindigkeitsabhängige Prozessveränderungen kompensiert, insbesondere in den Beschleunigungs- und Abbremsphasen.

Walzenexzentrizitätskompensation

Durch Ungleichförmigkeiten an der Walzengeometrie entstehen periodische Schwankungen im Walzspalt, die zu Dickenabweichungen führen können. Die Funktion kompensiert die periodischen Störungen, die aufgrund der Asymmetrien der Stützwalzen entstehen. Eine automatische Adaption zum Ausgleich von Änderungen während des Walzvorgangs wie z. B. aufgrund von Wärmeveränderungen, Verschleiß oder Last ist implementiert.

Zugvorsteuerung

Die Zugvorsteuerung berechnet einen Momentenkorrekturwert für die Haspel in Abhängigkeit von der einlaufenden Banddickenabweichung. Dadurch werden die Verkopplungen zwischen Walzspalt und Zug berücksichtigt.

1.1 Faszination Regelungstechnik



Monitorregelung

Grundregelungsstrategie ist die Monitorregelung, die auf der gemessenen Dickenabweichung im Auslauf des Walzgerüsts basiert.

Die Verzögerungszeit, die durch die Transportzeit des Bandes vom Walzspalt zum Banddickenmessgerät entsteht, beeinflusst im wesentlichen die Dynamik des Regelkreises, speziell bei geringen Anlagengeschwindigkeiten. Zur Minimierung dieser Verzögerungszeit kann die Regelung mit einem prediktiven Modell ergänzt werden, welches das dynamische Verhalten des Monitorreglers erheblich verbessert.

Dickenvorsteuerung

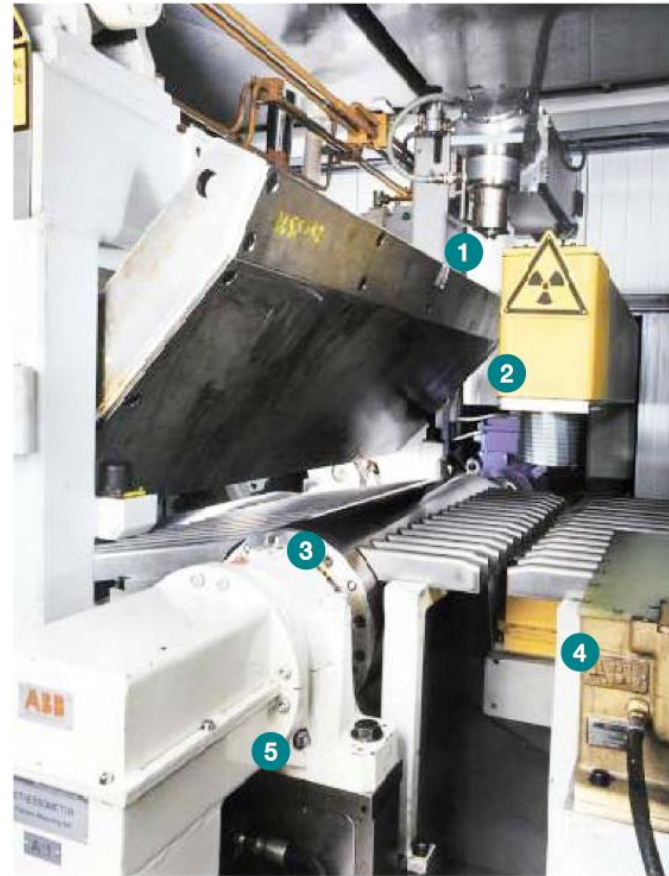
Mit Hilfe eines Banddickenmessgerätes vor dem Walzgerüst kann zusätzlich eine Dickenvorsteuerung implementiert werden. Sie ermöglicht die Erkennung und Ausregelung von einlaufenden Banddickenänderungen. Basierend auf der einlaufenden Dickenabweichung wird ein Korrekturwert berechnet und zeitrichtig aufgeschaltet, sobald die Dickenabweichung im Walzspalt eintrifft.

Massenflussregelung

Durch die Anwendung des Massenflussgesetzes werden für ein Bandsegment im Moment des Walzens die austretende Banddicke aus der einlaufenden Banddicke und der Ein- und Auslaufgeschwindigkeit berechnet und daraus geeignete Korrekturwerte für die Stellglieder ermittelt.

Mit der Massenflussregelung kann das Konzept der Monitorregelung und Dickenvorsteuerung optimal ergänzt werden, zur Erhöhung der Regelgenauigkeit für die Dickentoleranzen.

Wichtige Sensoren in Kaltwalzwerken



- 1 Geschwindigkeitssensor
- 2 Dickenmessgerät
- 3 ABB-Stressometer zur Planheitsmessung
- 4 ABB-Kantenmesser
- 5 ABB-Zugmessung

1.1 Faszination Regelungstechnik



Regelungen beim Kaltwalzen

Vorteile des neuen Regelkonzepts:

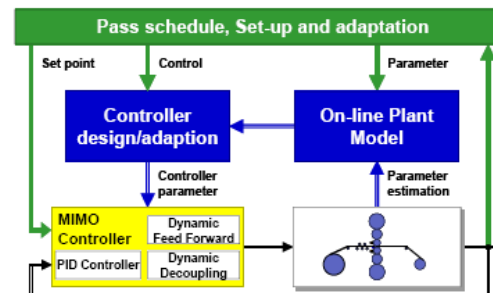
- Schneller im Toleranzfenster, damit Verringerung der Abmaßlängen
- Unterbrechungsfreie Anlagenbeschleunigung und damit Durchsatzerhöhung
- Bessere Störungsunterdrückung und damit bessere Banddickentoleranzen über das gesamte Band
- Ermöglicht höhere Betriebsgeschwindigkeiten und damit Erhöhung des Durchsatzes

Neues ABB Dickenregelungskonzept

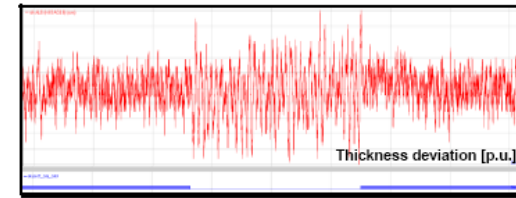
Das neue Dickenregelungskonzept von ABB setzt den Standard für höchste Anforderungen an Materialqualität und Anlagenproduktivität.

Schlüsselkomponenten der neuen Banddickenregelungsstrategie sind:

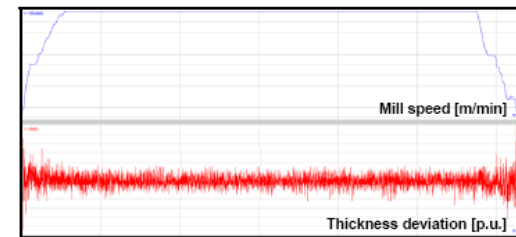
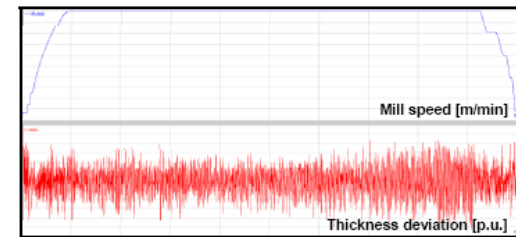
- Online Schätzung der zeitvarianten Prozessparameter
- Online MIMO Regleradaption auf Basis eines Online- Prozessmodells
- Dynamische Vorsteuerungsstrategie zur besseren Störungsunterdrückung
- Dynamisch entkoppelter MIMO Regler
- Überwachungsfunktionen für die Reglerparameter



MIMO Reglerkonzept mit dynamisch entkoppelten PID Reglern, dynamischen Vorsteuerungen und online Parameteradaption.



Einfluss auf die Banddickenqualität beim Zu- und Abschalten der neuen MIMO Regelung.



Resultat nach dem 4ten Stich, gewalzt unter normalen Betriebsbedingungen: Längenbasierter Dickschrieb für klassische Regelung (oben) und neue MIMO Regelung (unten).

3-Sigma Verbesserung von:
1,3% Klassische Regelung (oben) zu
0,7% für die neue MIMO Regelung (unten)

1.1 Faszination Regelungstechnik

Prozessautomatisierung eines Zementwerks

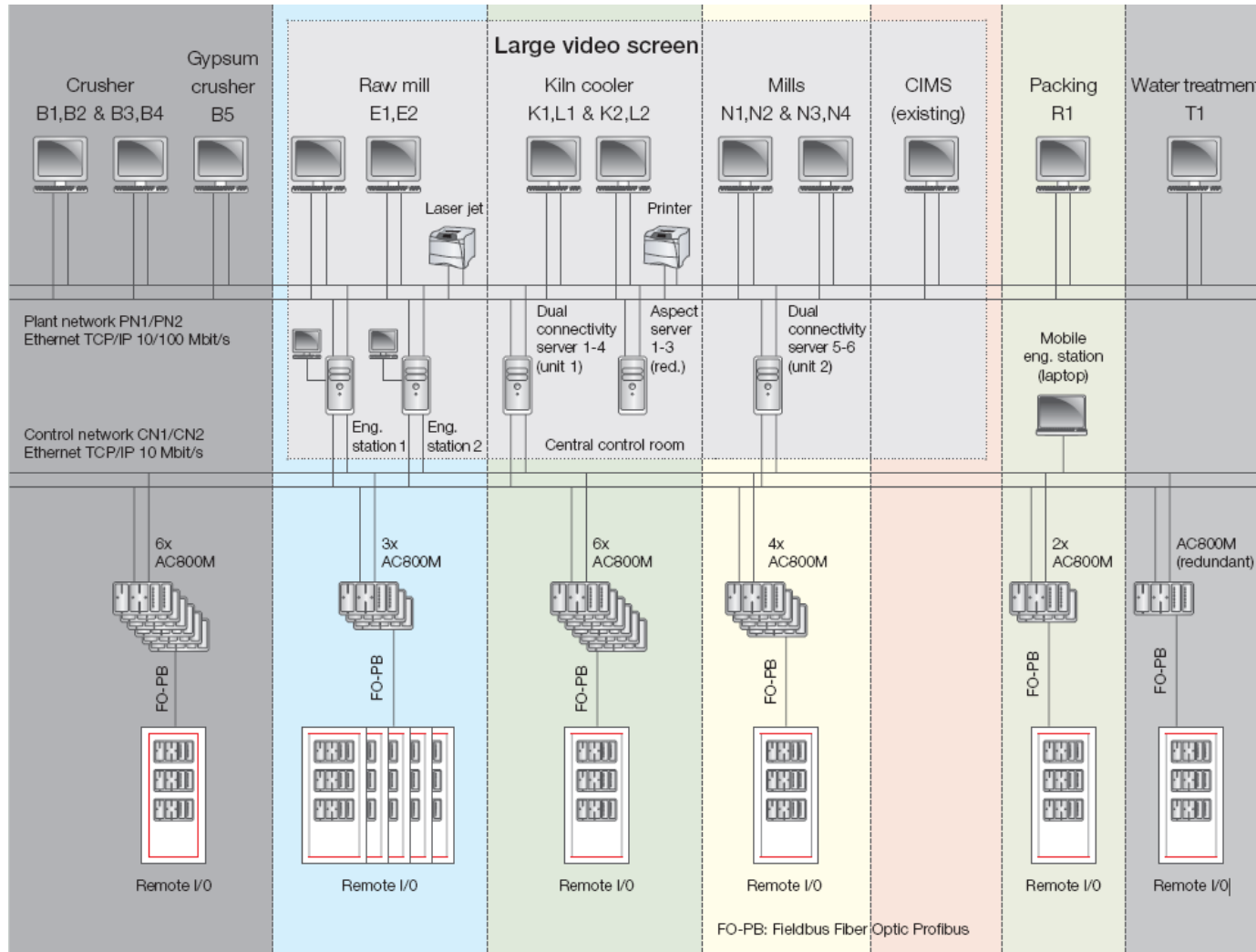


Eastern Province Cement Company,
Saudi-Arabien

1.1 Faszination Regelungstechnik



Prozessautomatisierung eines Zementwerks



1.1 Faszination Regelungstechnik

Prozessautomatisierung eines Zementwerks



1.1 Faszination Regelungstechnik

Automatisierung mit Pneumatik von Festo

Nahrungsmittel- und Verpackungsindustrie:

Hochgeschwindigkeitsanwendungen:



Ventil



Überprüfung



Elektromechanischer Antrieb



Vakuumgenerator



Servopneumatik

1.1 Faszination Regelungstechnik

Eurofighter: Mechanisch instabil konstruiert. Nur mit Regelung zu stabilisieren.



1.1 Faszination Regelungstechnik

Infos zum Eurofighter:

- Im Cockpit der Phantom dominieren analoge Rundinstrumente: Im Eurofighter arbeitet der Pilot mit multifunktionalen Displays. Der Eurofighter ist gezielt instabil ausgelegt und damit wesentlich wendiger. Bei herkömmlichen Flugzeugen liegt der Schwerpunkt immer hinter dem Hauptauftriebspunkt der Tragflächen. Das Flugzeug wird sozusagen gezogen und stabilisiert sich selbst. Beim Eurofighter ist es genau umgekehrt. Er wird praktisch angeschoben. Dadurch will er ständig ausbrechen.
- Dieser Balanceakt ist von Menschenhand nicht mehr zu bewältigen. Und so korrigiert - selbst bei normalem Geradeausflug - der Bordcomputer ständig den Eurofighter. Im Langsamflug zeigt sich, was die Rechner permanent leisten müssen. Die Vorflügelklappen, die den Auftrieb der Tragflächen variieren, sind ständig in Bewegung. 83 Computer interpretieren und realisieren die Steuereingaben des Piloten.



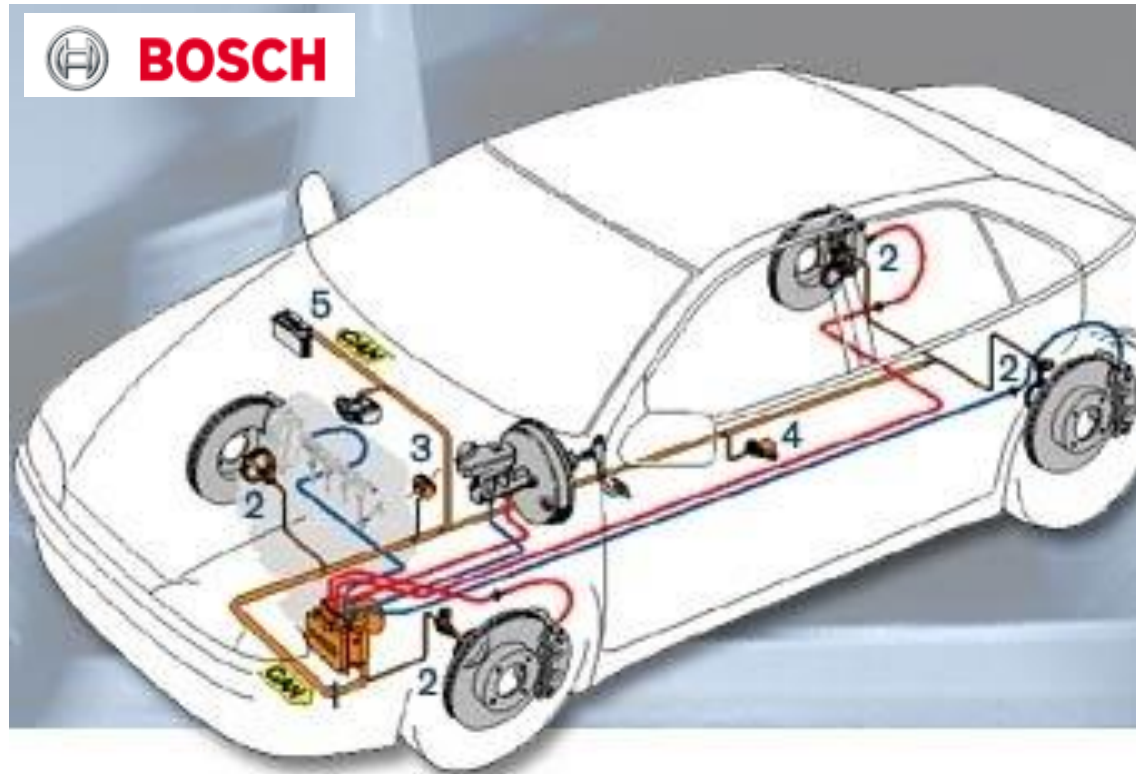
1.1 Faszination Regelungstechnik

ABS, ASR:

- Regelung des Brems- bzw. Antriebsschlupfes der Räder

ESP:

- Verhindert Ausbrechen des Autos
- Zustandsraummethoden erforderlich



Hydroaggregat mit Anbausteuergerät



Raddrehzahlsensoren



Lenkwinkelsensor



Drehraten- und Querbeschleunigungssensor



1.1 Faszination Regelungstechnik

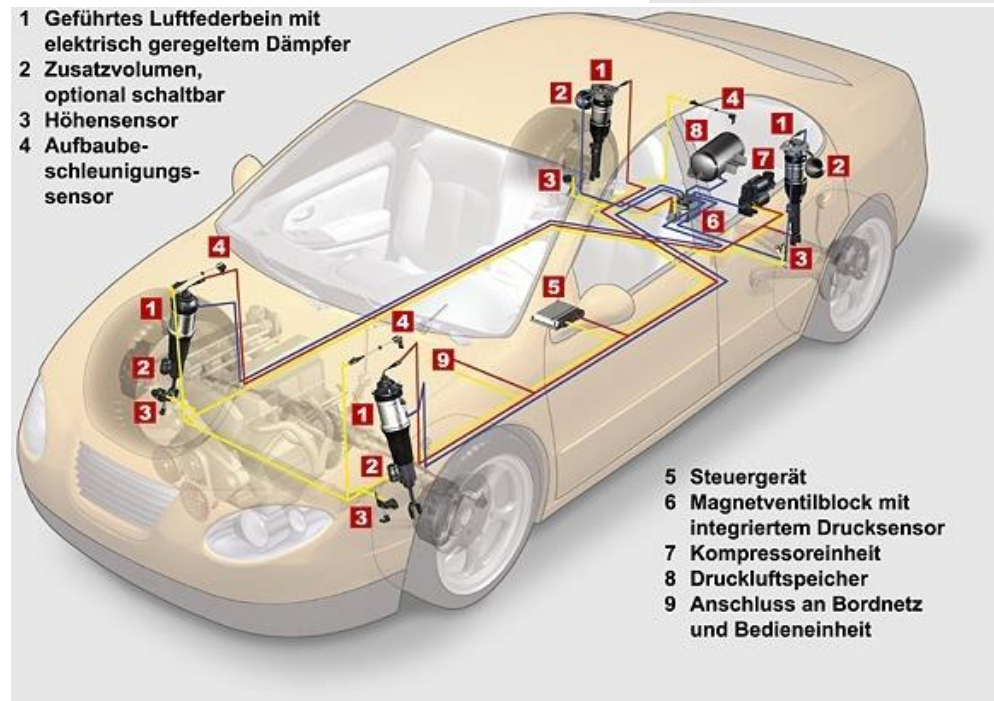


Elektronische Luftfedersysteme

Mit EAS (Electronic Air Suspension System) nutzt man die Möglichkeiten der Elektronik für Fortschritte am Fahrwerk, die höchste Fahrsicherheit, besten Komfort und sportliches Handling bieten. EAS passt Federkennlinie, Dämpfungskennlinie und Karosserieniveau automatisch an wechselnde Fahrzustände und den Beladungszustand des Fahrzeugs an.

Vorteile von EAS

- Reduzierung von Wank- und Nickbewegungen
- Reduzierung sonstiger Aufbau-
bewegungen
- Reduzierung der Radlastschwankungen
- Spürbarer Zugewinn an Fahrdynamik
und -komfort



1.1 Faszination Regelungstechnik

Weiterentwicklung Motorsteuerungen

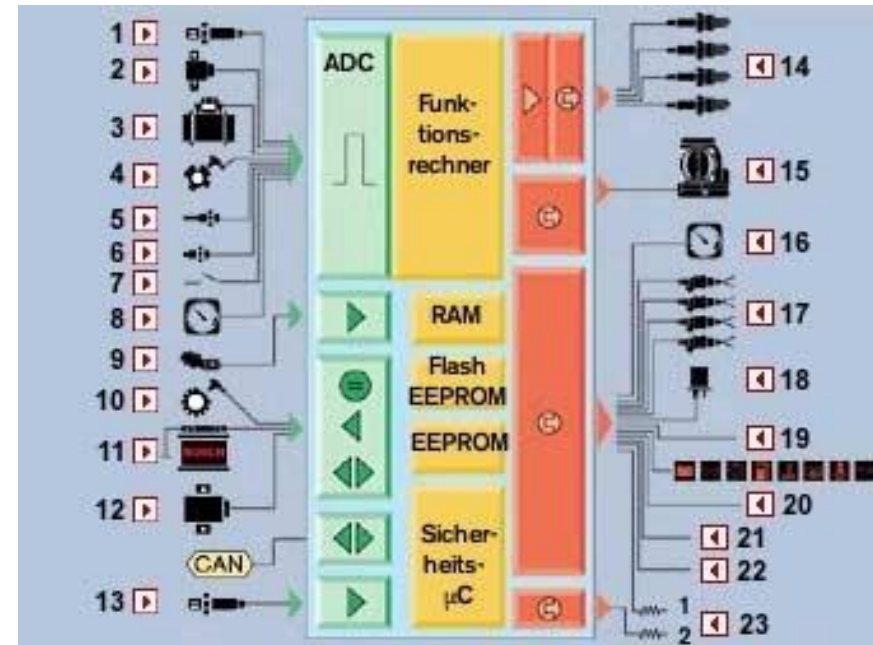
50 und mehr Steuergeräte im Kfz für:

- Motor, Bremse, Lenkung, Batterie, Getriebe, Verteilergetriebe, ...
- ESP (inkl. ABS, ASR), Airbags, Schließsystem, Reifendruck, ...
- Komfortfunktionen, Klimaanlage, Sitze, Türen, Cockpit, ...



Nicht nur Steuerung und Regelung, auch Diagnosefunktionen. Motorsteuerung:

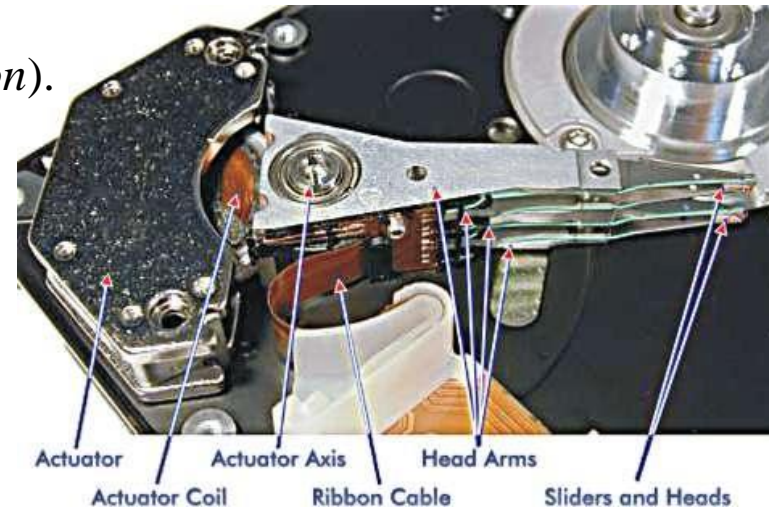
- 1 Lambda-Sonde 2, 2 Drosselklappenstellung, 3 Luftmasse,
4 Nockenwellenstellung, 5 Ansauglufttemperatur,
6 Motortemperatur, 7 Getriebestufe,
8 Fahrzeuggeschwindigkeit, 9 Klopfintensität,
10 Kurbelwellendrehzahl und oberer Totpunkt (OT),
11 Batterie, 12 Hauptrelais, 13 Lambda-Sonde 1,
14 Zündkerzen, 15 EGAS-Steller, 16 Motordrehzahl,
17 Einspritzventile, 18 Kraftstoffpumpenrelais,
19 Nockenwellensteuerung, 20 Tankentlüftung,
21 Saugrohrumschaltung, 22 Sekundärlufteinblasung und
Abgasrückführung, 23 Heizung Lambda-Sonde



1.1 Faszination Regelungstechnik

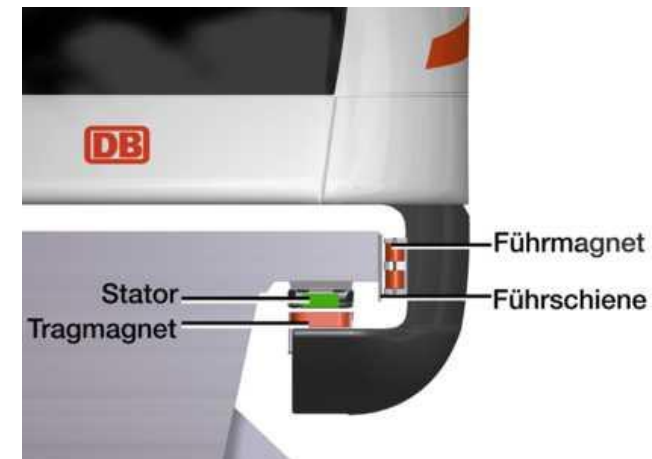
Regelung des Schreib/Lesekopfes einer Harddisk:

- Auf der Spur bleiben: Festwertregelung (*Regulation*).
- In neue Spur fahren: Folgeregelung (*Tracking*).
- Sehr hohe Geschwindigkeitsanforderungen.
- Große Störungen: Periodische durch Exzentrizität und stoßweise.



Magnetschwebbahn Transrapid:

- Regelung in vertikaler Richtung (Sollwert Abstand Stator-Tragmagnet 10 mm).
- Regelung in horizontaler Richtung mittels Führungsmagnet.



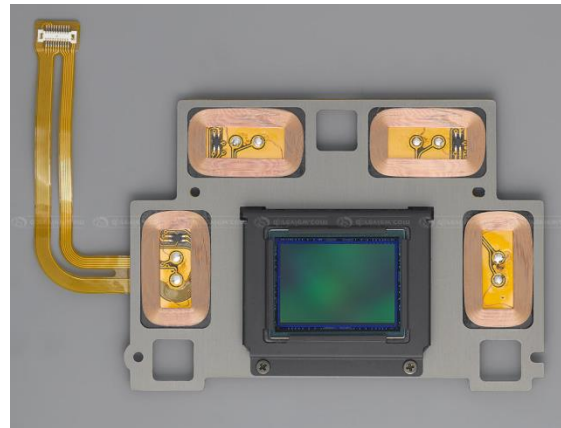
1.1 Faszination Regelungstechnik

Optischer Bildstabilisator bei Digitalkameras:

- Ermittlung der Vibrationen der Kamera durch Messung der Winkelgeschwindigkeiten des Bildsensors um seine beiden Achsen mittels mikromechanisch hergestellten Drehratensensoren nach dem Gyroskop-Prinzip (Ausnutzung der Coriolis-Kraft).
- Kompensation (d.h. Regelung auf Wunschposition) der durch Vibrationen verursachten Verschiebung der Bildpunkte durch horizontale und vertikale Verschiebung des Bildsensors. Alternative: Verstellung der Linsen (Canon: IS, Nikon: VR, Panasonic: OIS)
- Als Aktor für die Verstellung des Bildsensors können piezoelektrische (Minolta/Sony) oder elektromagnetische (Pentax/Samsung) Stellglieder zum Einsatz kommen (bis $\sim 10\text{Hz}$).



Super Steady Shot (Sony)



Shake Reduction (Pentax)



1.1 Faszination Regelungstechnik

Quelle: www.digitalkamera.de



Die im Folgenden beschriebenen Tests veranschaulichen die Effektivität eines optischen Bildstabilisators: Auf einem kleinen Leuchtpult wurde eine schwarze Maske aufgeklebt, die in der Mitte ein winziges Loch hat (0,2 mm Durchmesser, siehe Bild 7).

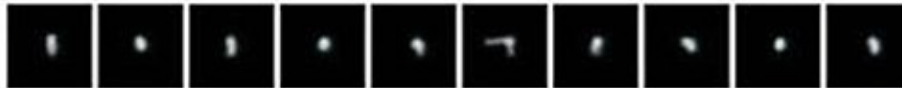
Das Licht hinter dem winzigen Loch in der Mitte ist intensiv genug, um auch bei starkem Verwackeln eine Leuchtspur auf dem Bildsensor zu hinterlassen.

Im ersten Test wurde eine D200 mit Nikkor 18-200 mm VR auf 200 mm gezoomt und frei stehend jeweils zehnmals fotografiert. Auf das KB-Format umgerechnet, ist das ein 300mm-Tele, und nach der alten Faustregel bräuchte man dafür für Verwacklungsfreiheit eine

maximale Verschlusszeit von 1/300 s. Wenn man diese Verschlusszeit um die vier Stufen verlängert, welche Nikon mit der Vibration Reduction VR II verspricht, dann bekommt man rund 1/20 s ($300 > 150 > 75 > 37,5 > 18,75$).



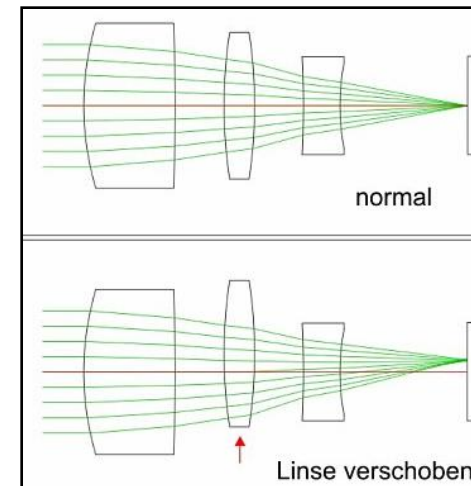
VR OFF – 300mm (umger. auf KB), 1/20s



VR normal – 300mm (umger. auf KB), 1/20s



VR active – 300mm (umger. auf KB), 1/20s



Jedes der schwarzen Quadrate ist nur ein winziger Ausschnitt von 30 x 30 Pixeln (das gesamte Bild ist 3872 x 2592 Pixel groß). Wie man in der ersten Reihe (Bild 8) sieht, kann es ohne VR leicht zu 30 Pixel langen Wischern kommen; das Bild wäre totaler Ausschuss. Bei den nächsten zwei Reihen wird ganz deutlich, wie das Verwackeln dramatisch reduziert wird. Nikons VR II funktioniert also wie versprochen (VR normal, Bild 9). Dabei ist – wie erwartet und von Nikon auch so beschrieben – der "active mode" (Bild 10) ein bisschen weniger effektiv für einen ruhigen Test im Labor. Dafür soll er aber besser sein, wenn man von einem fahrendem Auto, einem Boot oder einem Hubschrauber aus fotografiert. Mit diesem Test lässt sich das aber nicht nachprüfen. Eine weitere Beobachtung von den vielen Testreihen ist, dass praktisch in jeder Folge von zehn Aufnahmen mit VR jeweils irgendwo ein Ausreißer mit dabei ist. Die Lehre daraus ist, dass man selbst mit VR sicherheitshalber ein paar extra Bilder machen, oder zumindest das Ergebnis auf dem LCD kontrollieren sollte.

1.1 Faszination Regelungstechnik

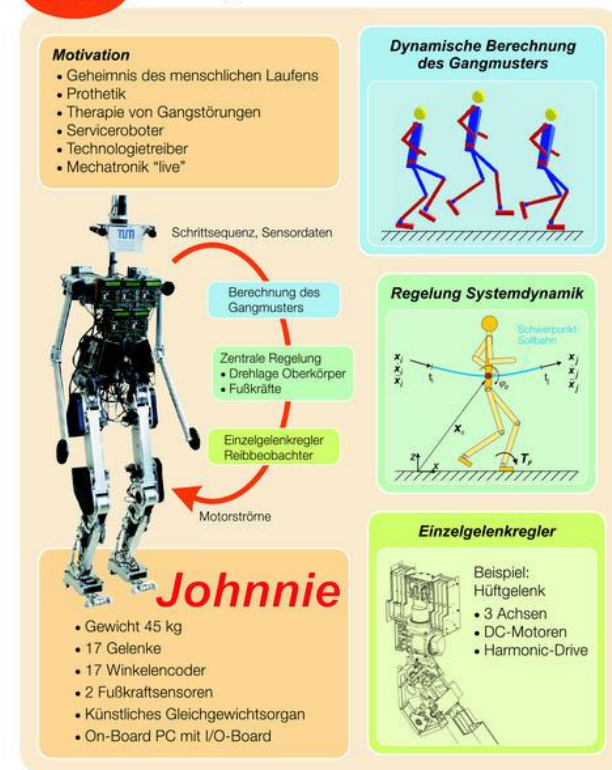
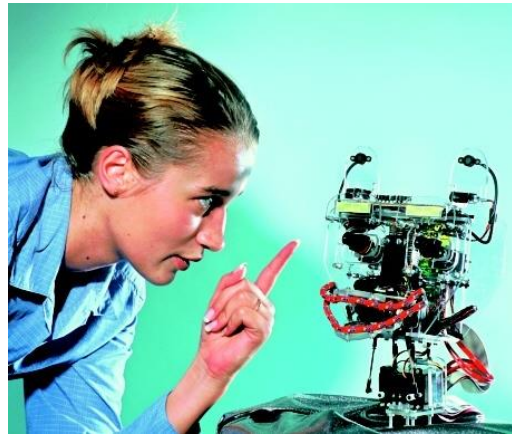
Industrieroboter:

- Einsatz in der Fertigung
- Kraft- und positionsgeregelt
- Gute (nichtlineare) Modelle



Humanoide Roboter:

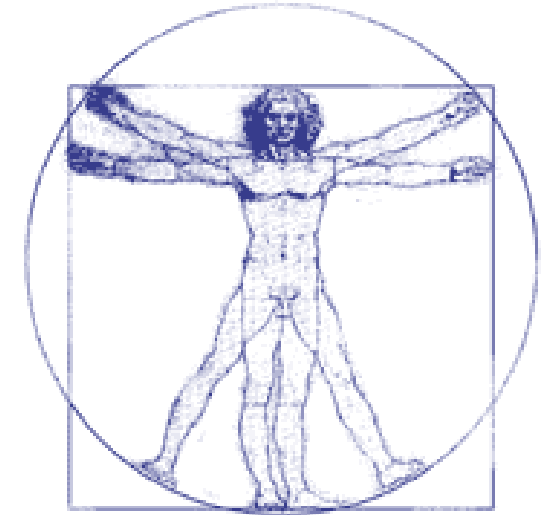
- Nachbildung menschlicher Fähigkeiten: 2-beiniges Gehen, Fühlen, Sehen, ...
- Nachbildung bestimmter tierischer Fähigkeiten, z.B. Krabbeln durch Pipelines zur Hindernis oder Lecksuche.



1.1 Faszination Regelungstechnik

Menschlicher Körper enthält eine Vielzahl an Regelkreisen:

- Körpertemperatur
- Blutzucker
- Blutdruck
- Herzfrequenz
- Alle Hormone ...



Starke Variation zw. verschiedenen menschlichen Körpern und große Änderungen der Umwelt erfordern sehr **robuste Funktionsweise** → **Regelung!**

Medizin (insbesondere innere) ist oft das Reparieren oder Kompensieren defekter bzw. aus dem Gleichgewicht geratener Regelkreise, z.B.:

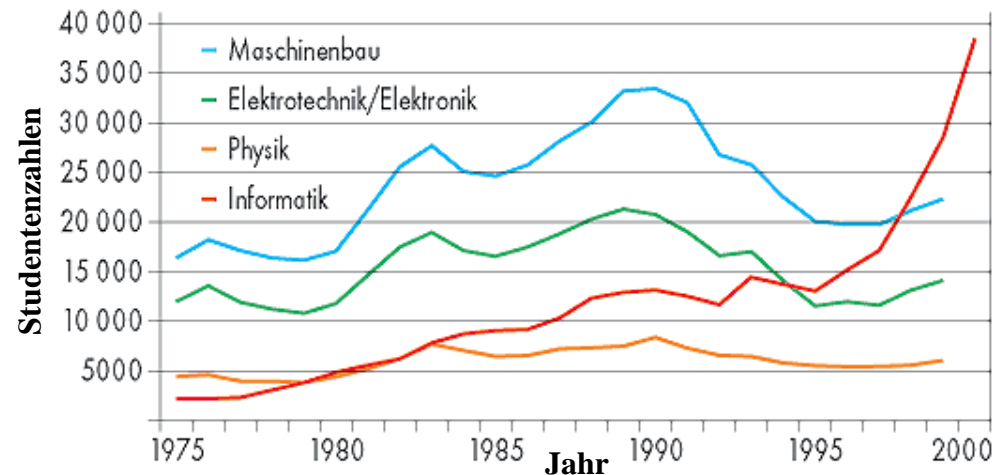
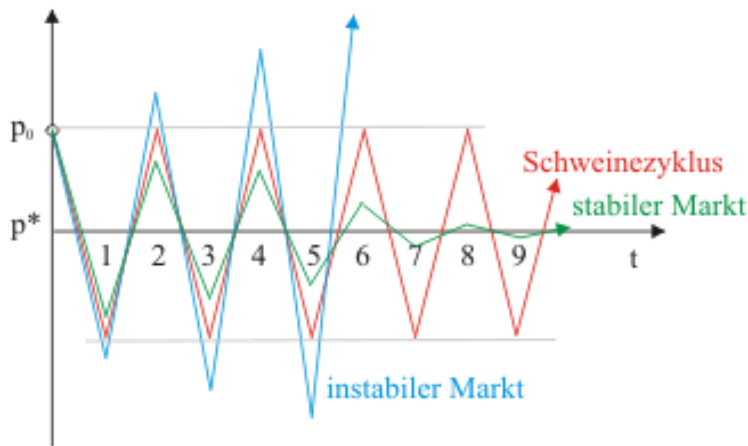
- Diabetes Typ I: Bauchspeicheldrüse (Stellglied) defekt. Insulin (Stellgröße) muss von außen zugeführt werden.
- Diabetes Typ II: Insulin hat reduzierte Wirkung; die Zellen nehmen Zucker schlecht auf → Regelstrecke defekt. Reparatur der Zellen wünschenswert; ansonsten: Erhöhte Dosis Insulin von Außen.

1.1 Faszination Regelungstechnik

Viele wirtschaftliche oder soziale Kennziffern verlaufen in Zyklen (sog. Schweinezyklus, 1927):

- Wachstumsraten, Inflation, Devisenkurse, Aktienkurse, Insolvenzen, ...
 - Hase/Fuchs-Population, Studentenanfänger-/Neueinstellungszahlen, ...
- Diesen Dynamiken liegen ähnliche Differentialgleichungen zugrunde!

1. Viele Füchse fressen viele Hasen → weniger Hasen
2. Füchse sterben an Hunger, weil nicht genug Hasen da sind → weniger Füchse
3. Wenige Füchse können nicht viele Hasen fressen → mehr Hasen
4. Füchse haben viel Nahrung → mehr Füchse

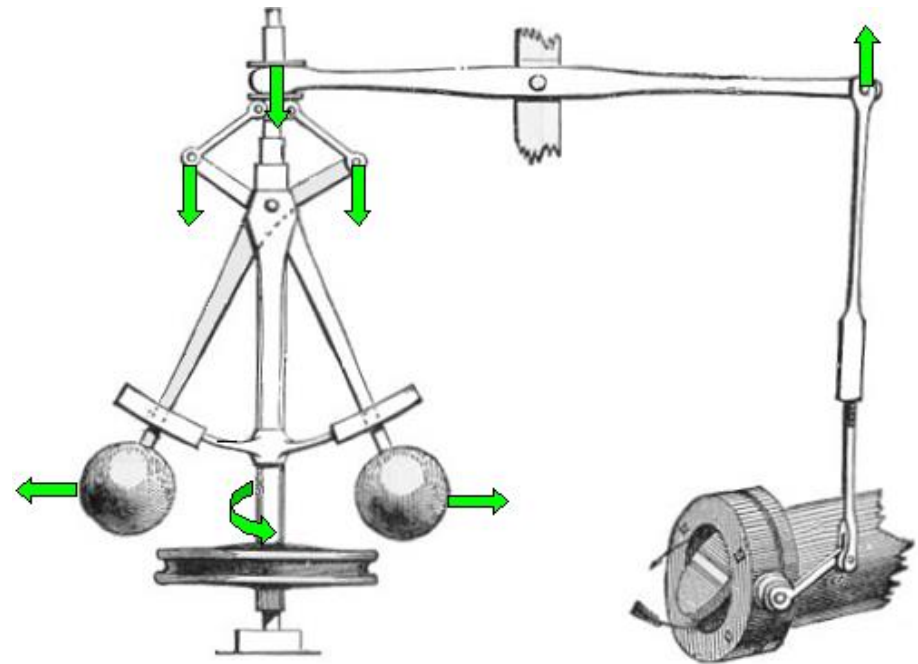


1.2 Geschichte der Regelungstechnik

Technologien / Methoden:

- -300 Griechen → Wasserkanäle
- 1700 Industrielle Revolution → Dampfmaschine
- 1935 Kriegsjahre → Elektrischer Verstärker (Ein-/Ausgangsmethoden)
- 1960 Computerzeitalter → Raumfahrt (Zustandsraummethoden)
- 1975 Computermassenmarkt → Adaptive Regelung
- 1985 Systematisierung beim Reglerentwurf → Robuste Regelung
- 1995 Leistungsfähige Berechnungsmethoden → Konvexe Optimierung

James Watt, 1788:
Fliehkraftregler für
Dampfmaschinen



Quelle: <http://de.wikipedia.org/wiki/Fliehkraftregler>

1.2 Geschichte der Regelungstechnik

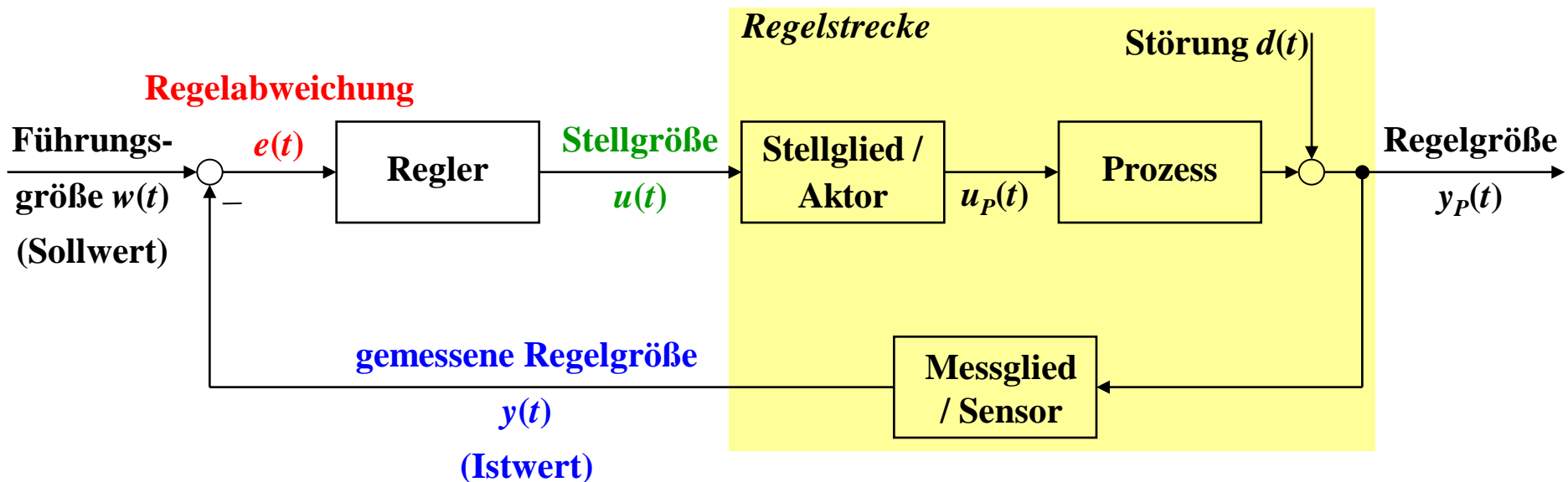
Regelungstechnische Meilensteine:

- 1788 James Watt: Fliehkraftregler
- 1868 Maxwell: DGL-Beschreibung verschiedener Regler
- 1895 Hurwitz: Stabilitätskriterium basierend auf Nennerpolynom
- 1922 Minorsky: PID-Regler
- 1932 Nyquist: Stabilitätskriterium basierend auf Ortskurve
- 1942 Ziegler/Nichols: Einstellregeln für PID-Regler
- 1942 Wiener: Filter und Prädiktion
- 1945 Bode: Frequenzganganalyse
- 1948 Evans: Wurzelortskurve
- 1956 Pontryagin: Maximumprinzip (zeitoptimale nichtlineare Steuerung)
- 1960 Kalman: Kalman-Filter, Dualität zw. Zustandsregler und -beobachter
- 1962 Bellman: Dynamische Programmierung (numerische Lsg. dyn. Opt.)

1.3 Regelkreis

Eigenschaften einer Regelung:

- **Messen:** Regelgröße wird mittels Sensor gemessen oder aus gemessenen Größen berechnet.
- **Vergleichen:** Regelgröße (Istwert) wird mit Führungsgröße (Sollwert) verglichen: Regelabweichung $e(t) = w(t) - y(t)$.
- **Stellen:** Aus Regelabweichung wird Stellgröße berechnet und auf Stellglied/ Aktor gegeben.

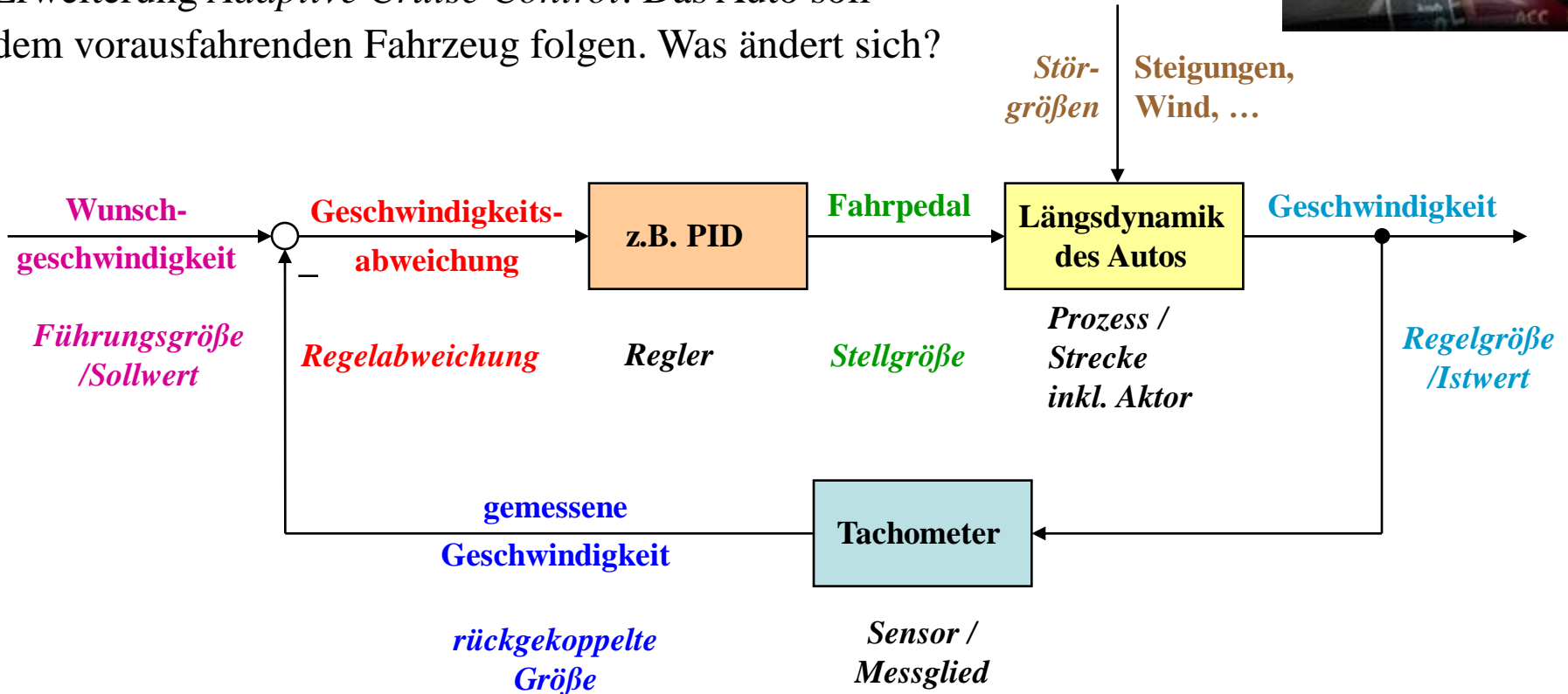
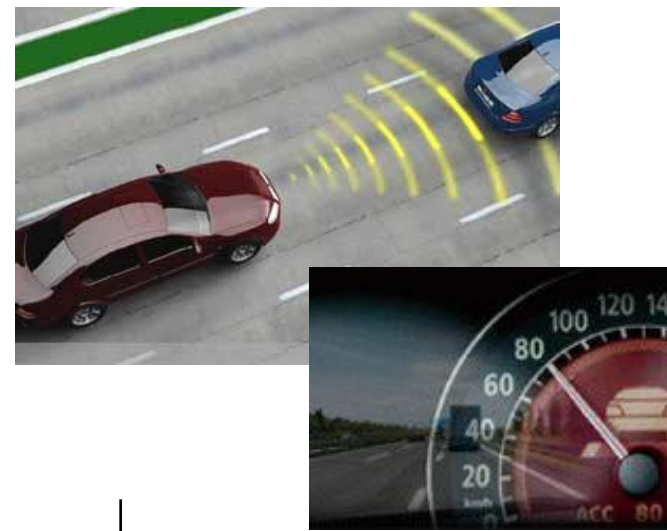


1.3 Regelkreis

Beispiel Geschwindigkeitsregelung (Tempomat)

Ziel des Tempomats: Die Geschwindigkeit des Autos soll möglichst gut der vom Fahrer vorgegebenen Wunschgeschwindigkeit folgen.

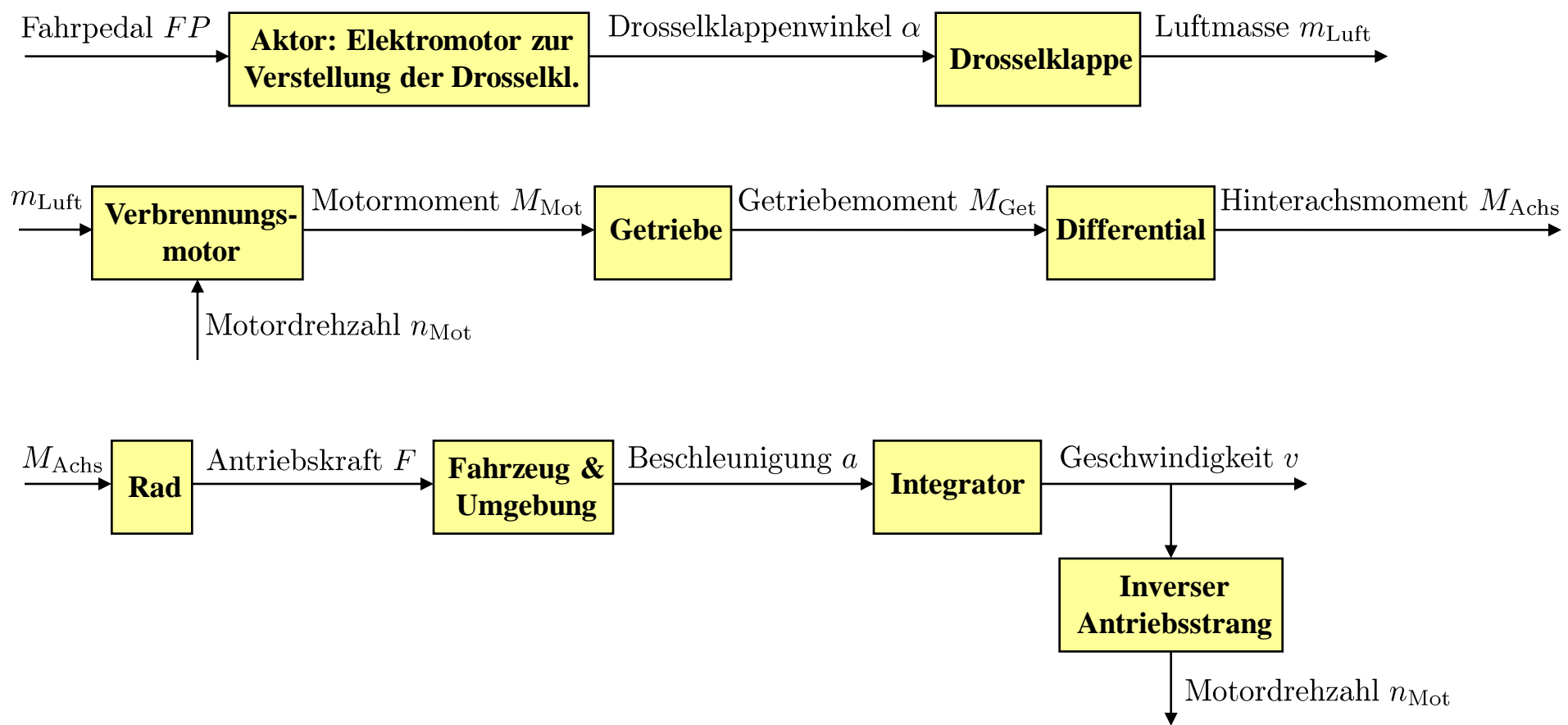
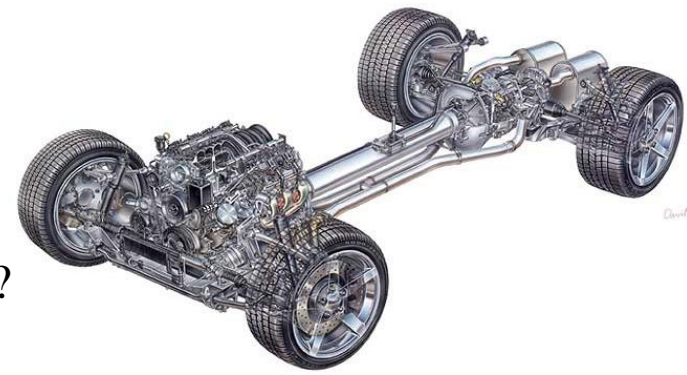
Erweiterung *Adaptive Cruise Control*: Das Auto soll dem vorausfahrenden Fahrzeug folgen. Was ändert sich?



1.3 Regelkreis

Beispiel Geschwindigkeitsregelung (Tempomat)

Was genau verbirgt sich hinter dem Block Prozess / Strecke?



1.3 Regelkreis

Beispiel Geschwindigkeitsregelung (Tempomat)

Vereinfacht ergeben sich folgende formelmäßige Zusammenhänge:

Aktor: Elektromotor zum Bewegen des Fahrpedals

Drosselklappenwinkel $\alpha \sim$ Fahrpedal FP

Aktordynamik vernachlässigt!

Drosselklappe

Luftmasse $m_{\text{Luft}} \sim \cos(\text{Drosselklappenwinkel } \alpha)$

Luftpaddynamik vernachlässigt!

Verbrennungsmotor

Motormoment $M_{\text{Mot}} = f(m_{\text{Luft}}, n_{\text{Mot}})$

Motordynamik vernachlässigt!

Getriebe

Getriebemoment $M_{\text{Get}} = i_{\text{Get}} \cdot \text{Motormoment } M_{\text{Mot}}$

Antriebsstrangdynamik vernachlässigt!

Differential

Hinterachsmoment $M_{\text{Achse}} = i_{\text{Dif}} \cdot \text{Getriebemoment } M_{\text{Get}}$

Rad

Antriebskraft $F = \frac{\text{Radmoment } M_{\text{Rad}}}{\text{Radradius } r_{\text{Rad}}}$ Radmoment $M_{\text{Rad}} = \frac{1}{2} \text{Hinterachsmoment } M_{\text{Achse}}$

Fahrzeug & Umgebung

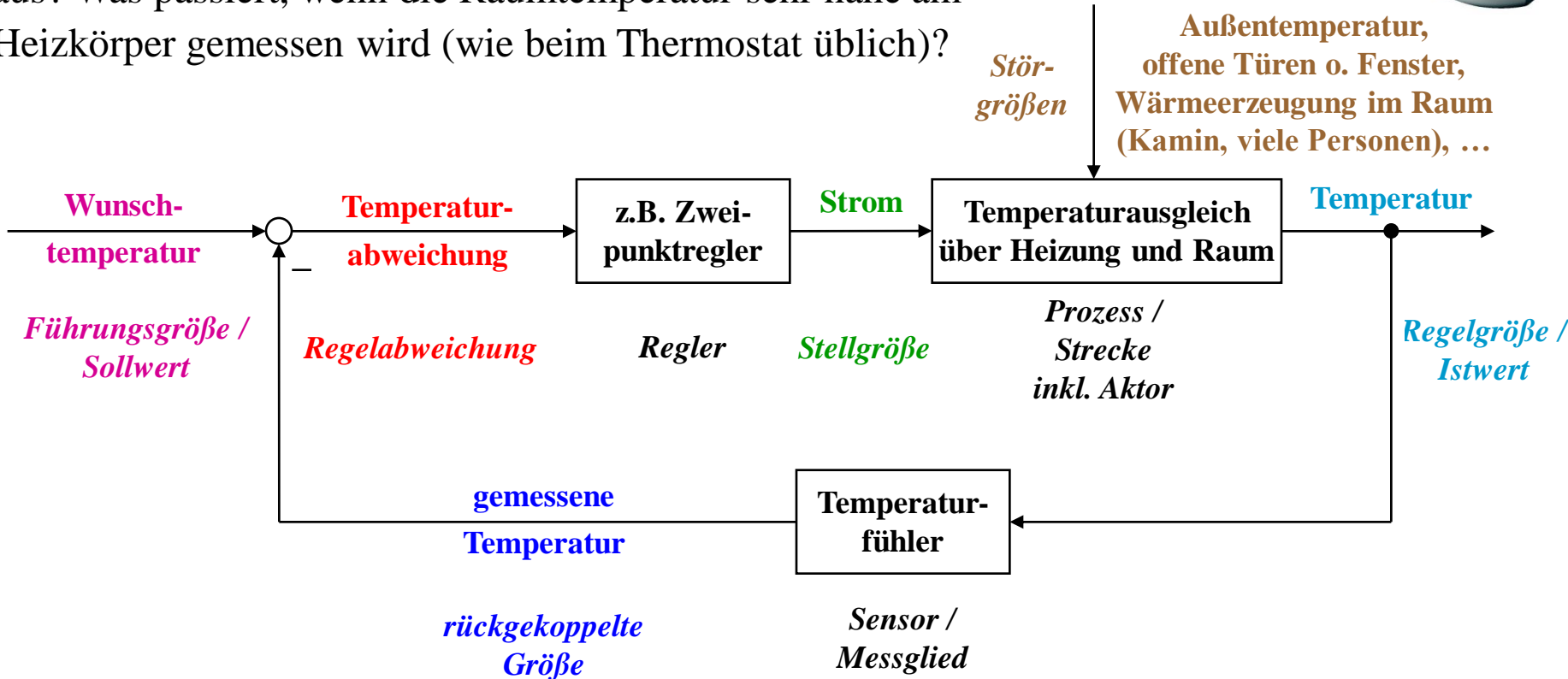
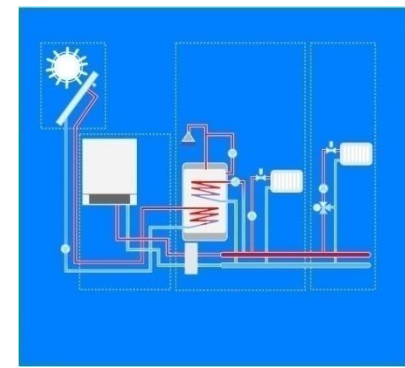
Antriebskraft $F = \text{Fahrzeugmasse } m \cdot \text{Beschleunigung } a + \text{Luftwiderstand} + \text{Reibung}$
Geschwindigkeit $v = \int \text{Beschleunigung } a \, dt$

1.3 Regelkreis

Beispiel Temperaturregelung mit elektrischer Heizung

Ziel: Die Temperatur des Raumes soll möglichst gut der vorgegebenen Wunschtemperatur folgen.

Wie wirkt sich eine bessere Wärmedämmung des Hauses aus? Was passiert, wenn die Raumtemperatur sehr nahe am Heizkörper gemessen wird (wie beim Thermostat üblich)?



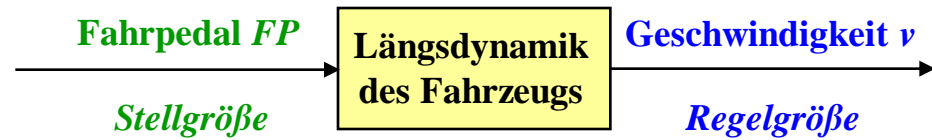
1.3 Regelkreis

Beispiel: Regelung eines Elektrofahrzeugs

Eigenschaften:

- Auto hat ein Getriebe mit nur einen Gang.
- Auto hat ein „Fahrpedal“ von -100% bis $+100\%$. Im positiven Bereich wird die E-Maschine als Motor (Antrieb) genutzt; im negativen Bereich als Generator (Bremse).
- Geschwindigkeiten sind so klein, dass der Luftwiderstand vernachlässigbar ist:
 - People Mover (Flughafen),
 - Kran,
 - Gabelstapler, ...
- Regelgröße: Geschwindigkeit des Fahrzeugs v (Tempomat)
- Stellgröße: Fahrpedal FP
- Führungsgröße: gewünschte Geschwindigkeit v_w
- Regelabweichung: $e = v_w - v$

1.3 Regelkreis



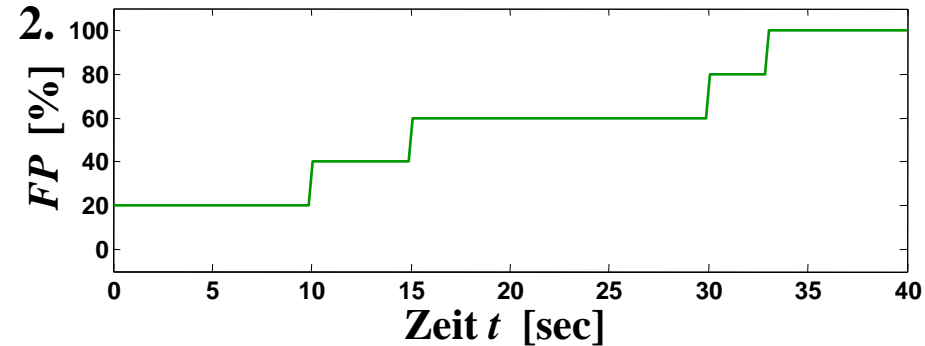
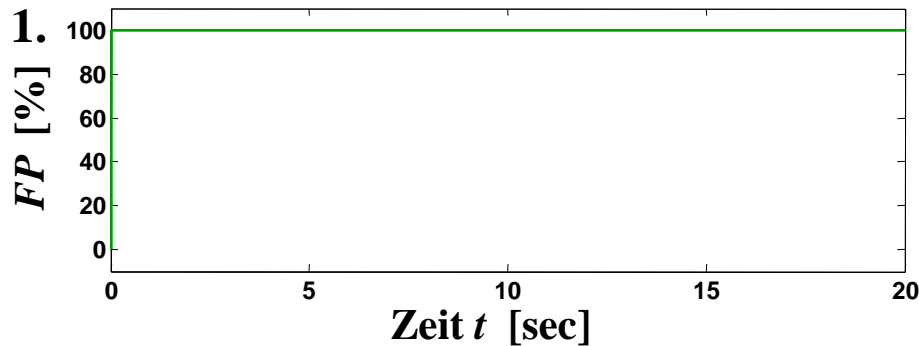
Beispiel: Messungen am Fahrzeug (ohne Regelung)

1. Messung: Vollgas geben zum Zeitpunkt $t = 0$ sec.

→ Fahrzeug beschleunigt in ca. 10 sec auf seine Höchstgeschwindigkeit $v_{\max} = 40$ km/h.

2. Messung: Stufenweises Hochbeschleunigen.

→ Je nach Länge der „Stufe“ erreicht das Fahrzeug seine Endgeschwindigkeit oder nicht.



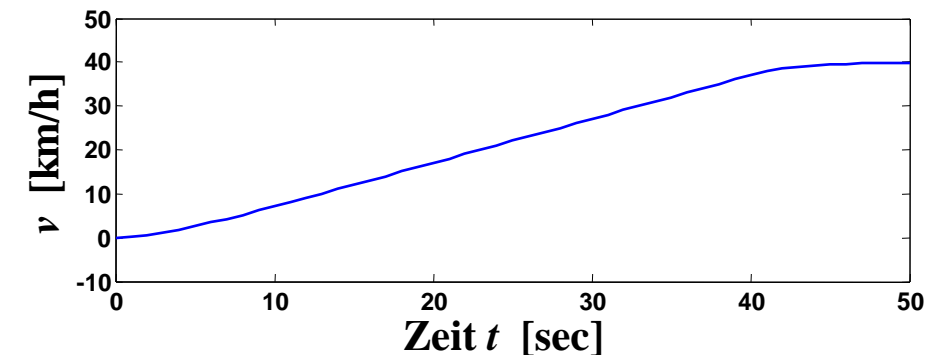
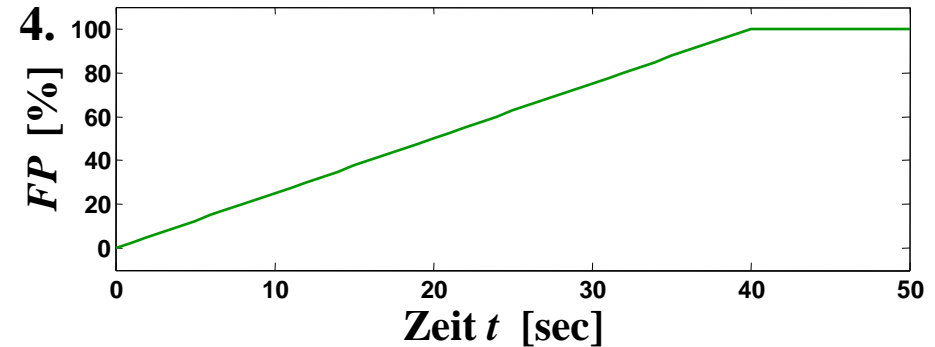
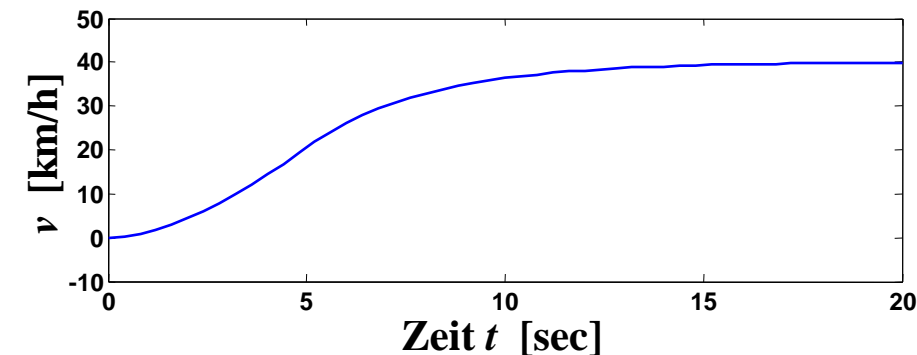
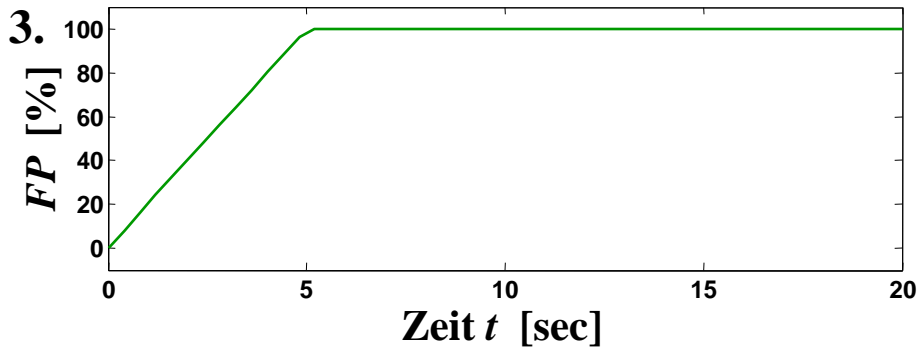
1.3 Regelkreis

3. Messung: Kontinuierliches Gasgeben von 0% bis 100% von $t = 0$ sec bis $t = 5$ sec.

→ Fahrzeug „hinkt“ dem Fahrerwunsch stark hinterher und erreicht erst deutlich nach $t = 5$ sec seine Höchstgeschwindigkeit v_{\max} .

4. Messung: Kontinuierliches Gasgeben von 0% bis 100% von $t = 0$ sec bis $t = 40$ sec.

→ Fahrzeug folgt dem Fahrerwunsch und erreicht kurz nach $t = 40$ sec sein v_{\max} .



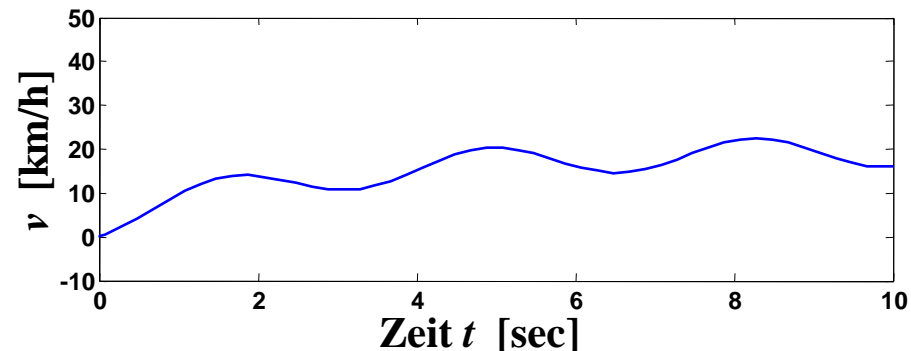
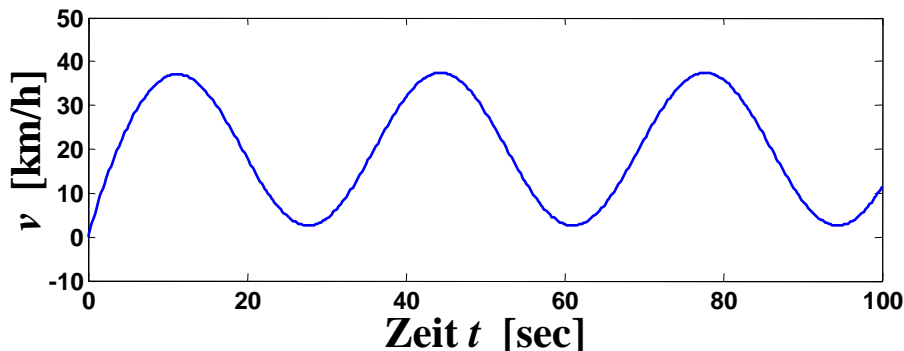
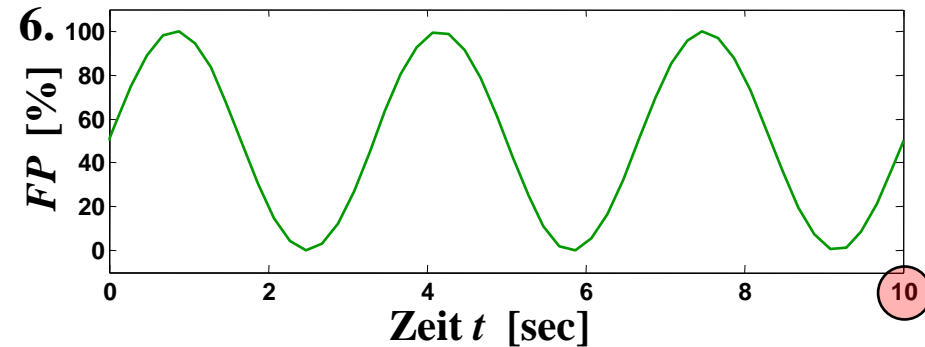
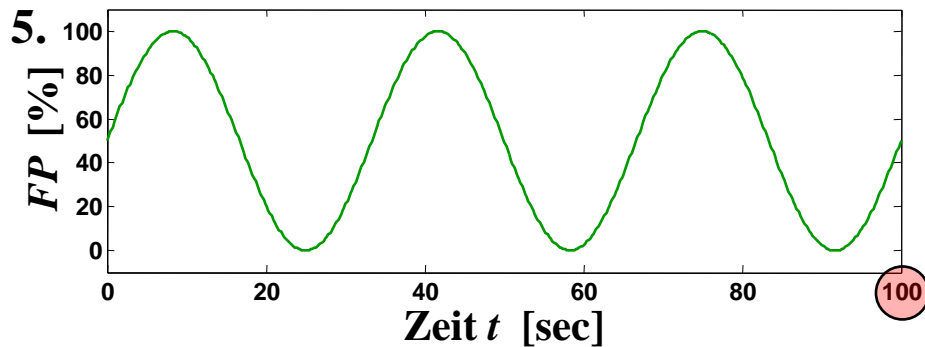
1.3 Regelkreis

5. Messung: Langsames sinusförmiges Gasgeben zwischen 0% bis 100%.

→ Fahrzeug „hinkt“ leicht nach, erreicht aber fast v_{\max} .

6. Messung: Schnelles sinusförmiges Gasgeben zwischen 0% bis 100%.

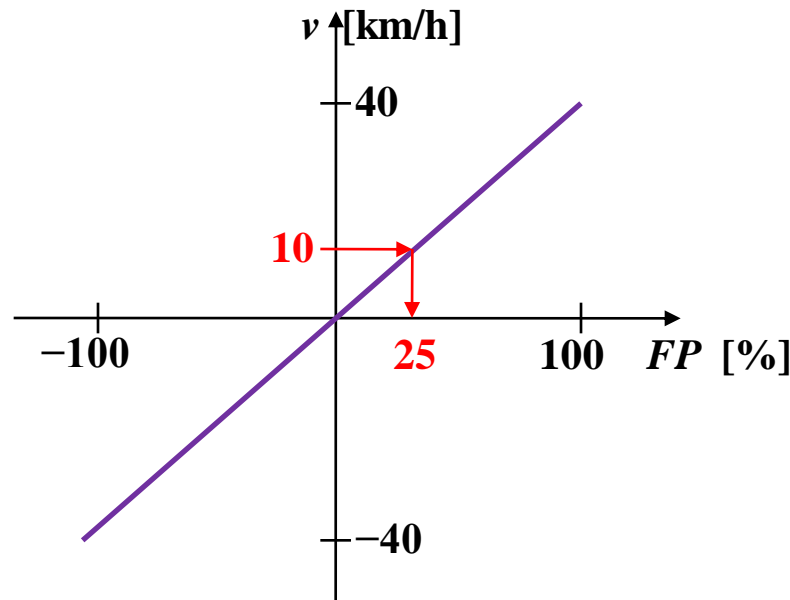
→ Fahrzeug kann Fahrpedal kaum folgen; schwingt um 20 km/h (nach Einschwingphase).



1.3 Regelkreis

Beispiel: Statischer Zusammenhang zwischen Stell- und Regelgröße

Vernachlässigt man den quadratisch wirkenden Luftwiderstand (was bei kleinen Geschwindigkeiten realistisch ist) und alle anderen Nichtlinearitäten, so ergibt sich folgender linearer Zusammenhang zwischen Stellgröße und Regelgröße:

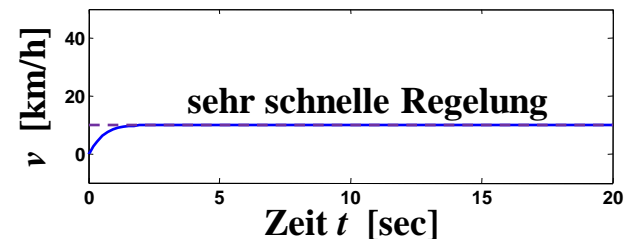
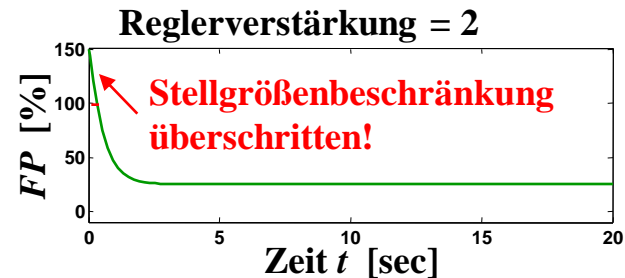
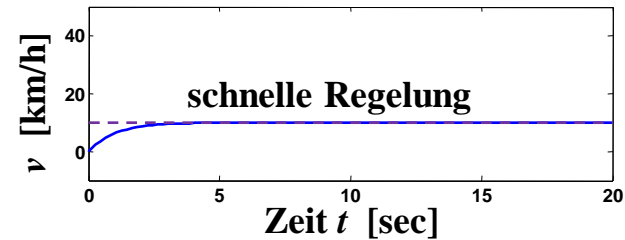
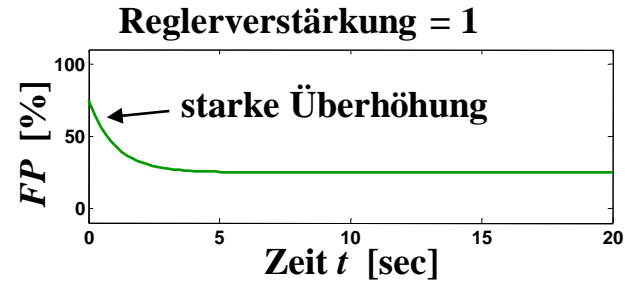
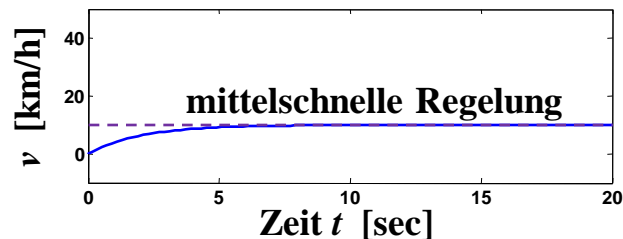
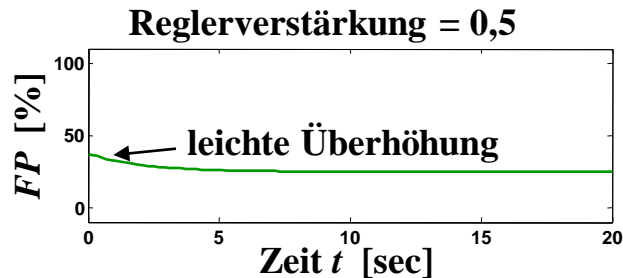
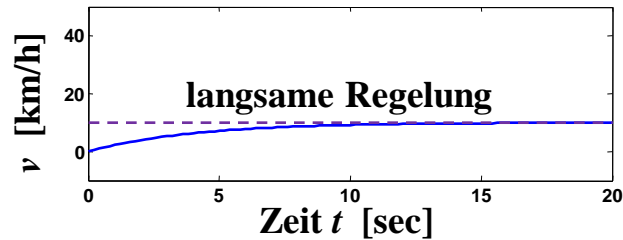
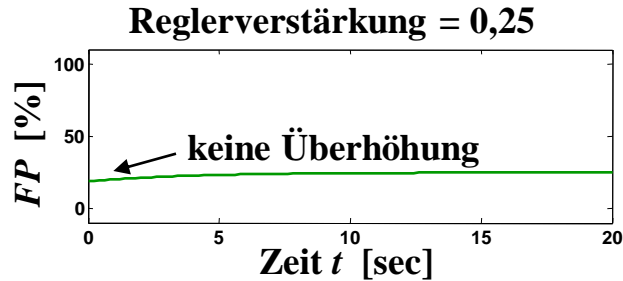


Aus dieser statischen Kennlinie kann man ablesen, wie viel Prozent Gas man geben muss, um asymptotisch (d.h. für $t \rightarrow \infty$) eine gewünschte Geschwindigkeit zu erreichen.

Z.B.: $FP = 100\% \rightarrow v = 40 \text{ km/h}$, $FP = 25\% \rightarrow v = 10 \text{ km/h}$.

1.3 Regelkreis

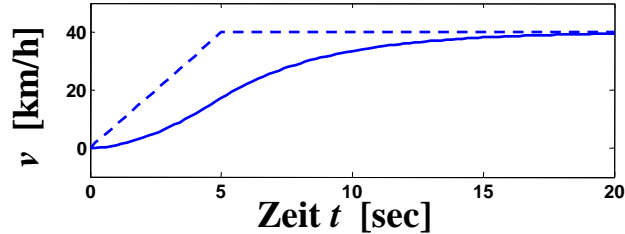
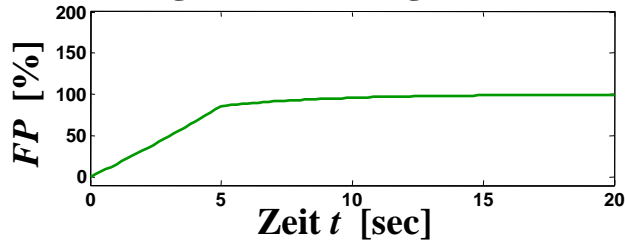
Beispiel: Messungen am geregelten Fahrzeug (mit PI-Regler), $v_w = 10 \text{ km/h}$



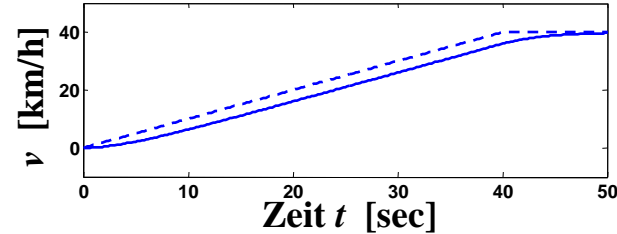
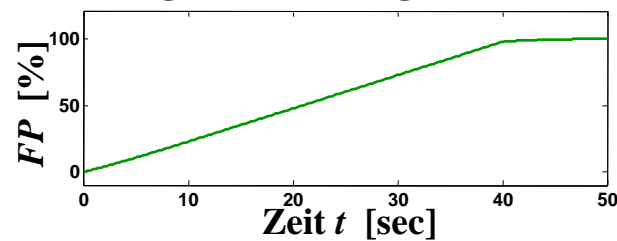
1.3 Regelkreis

Beispiel: Messungen am geregelten Fahrzeug, $v_w = \text{Rampe}$ (schnell/langsam)

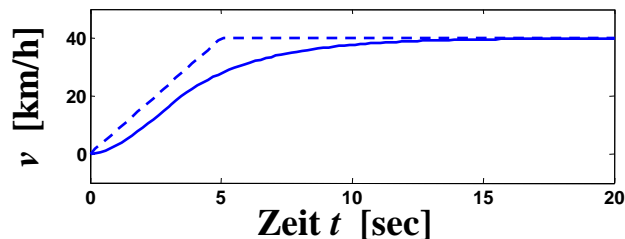
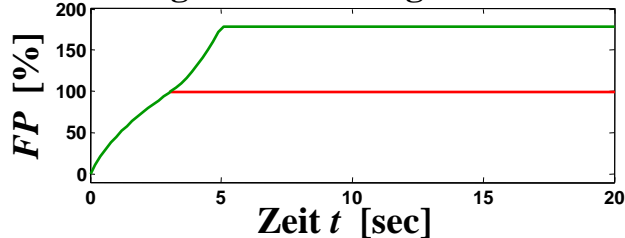
Reglerverstärkung = 0,25



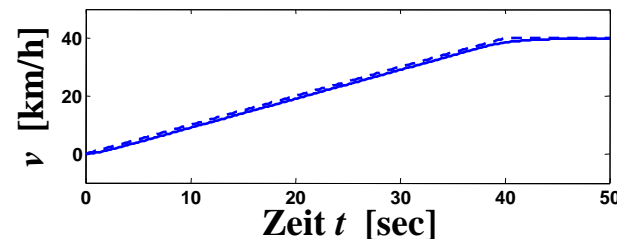
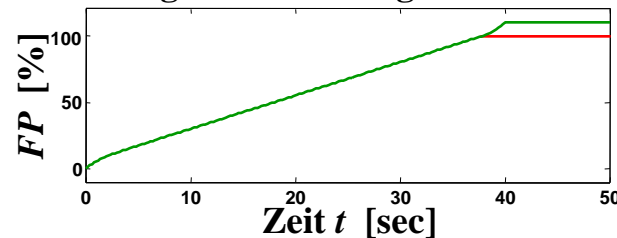
Reglerverstärkung = 0,25



Reglerverstärkung = 1



Reglerverstärkung = 1



1.3 Regelkreis

Beispiel: Erkenntnisse aus Messungen und Regelung am Fahrzeug

Messungen an der Regelstrecke (Fahrzeug):

- Typischer Tiefpass-Charakter, d.h. langsam veränderlichen (niederfrequenten) Eingangssignalen kann gefolgt werden; schnell veränderliche (hochfrequente) Eingangssignale werden stark gedämpft.
- Die Kennlinie gibt den statischen Zusammenhang (im eingeschwungenen Zustand) zwischen Eingangs- und Ausgangsgröße an.
- Nach Abwarten der Einschwingphase antwortet die Strecke auf ein Sinussignal auch mit einem Sinussignal der selben Frequenz, allerdings frequenzabhängig mit anderer Amplitude und Phase.

Geschwindigkeitsregelung des Fahrzeugs:

- Schnelligkeit der Regelung kann über Reglerverstärkung eingestellt werden.
- Schnelle Regelung erfordert Stellgrößenüberhöhung.
- Zu starke Stellgrößenüberhöhung führt zu Überschreitung der Stellgrößenbeschränkung.
- Zu schnellen Führungsgrößenänderungen kann auch die Regelung nicht folgen.

1.3 Regelkreis

Ergänzende Bemerkungen:

- Eigentlich möchte man die wirkliche Regelgröße $y_p(t)$ regeln und nicht $y(t)$. Man würde sich also für den Sensor ein Übertragungsverhalten = 1 wünschen. Es ist allerdings meist möglich, die statische und dynamische Verzerrung des Sensors zu vernachlässigen bzw. (teilweise) zu kompensieren, so dass $y(t) \approx y_p(t)$!
- Das Stellglied ist bei manchen Problemstellungen schon vorgegeben. Mehr Freiheiten hat der Regelungstechniker, wenn das Stellglied noch gewählt werden kann.
ACHTUNG: Am Prozess kommt immer nur die Ausgangsgröße des Stellglieds $u_p(t)$ an. Wenn das Stellglied langsam ist und/oder einen kleinen Stellbereich (von Minimal- bis Maximalanschlag) aufweist, können diese Einschränkungen mit keinem noch so „tollen“ Regler überwunden werden.
- Grundprinzip einer Regelung ist die **Rückkopplung** der Regelgröße. Die Stellgröße wird so verändert, dass sich die rückgekoppelte Regelgröße in der gewünschten Weise (entsprechend der Führungsgröße) entwickelt. Ohne Messung ist keine Rückkopplung möglich. ABER: Messungen müssen nicht immer zur Rückkopplung verwendet werden.
→ Prinzip der Steuerung.

1.3 Regelkreis

Warum Regeln? Vorteile:

- Regelgröße folgt Führungsgröße.
- Störungen werden eliminiert (weggeregelt); auch wenn sie nicht gemessen werden.
- Eine ungefähre Beschreibung (Modell) der Regelstrecke ist ausreichend. Ungenauigkeiten eines Modells der Regelstrecke werden (in gewissen Grenzen) von der Regelung kompensiert. Man sagt auch:
Eine Regelung ist robust
in Bezug auf Modellungenauigkeiten.
- Instabile Regelstrecken können nur durch Rückkopplung stabilisiert werden.

Nachteile:

- Die Rückkopplung macht die Analyse des Systems komplizierter.
- Die Regelung kann **instabil** werden, d.h. die Stell- und Regelgröße können gegen unendlich bzw. ihre physikalischen Grenzen laufen.
- Genaue und dynamisch schnelle Messung kann aufwändig oder teuer sein.

1.3 Regelkreis

Physikalische Größen:

- Messung von: el. Strom, el. Spannung, Weg, Winkel, Füllstand, Geschwindigkeit, Drehzahl, Beschleunigung, Winkelbeschleunigung, Kraft, Drehmoment, Druck, Temperatur, Volumen-/Massenstrom, ...

Messglied / Sensor (die Augen):

- Drehspule, Generator, Potentiometer, kapazitiver/induktiver Sensor, Schwimmer, Lichtschranke, Wirbelstrom, Dehnungsmessstreifen, Widerstandsthermometer, Thermoelement, Stauscheibe, Heissfilm, ...



Stellglied / Aktor (die Muskeln):

- Schrittmotor, Thyristor, Elektromagnet, Stelltransformator, Ventil, Pumpe, Hubkolben, Drehkolben, Membranantrieb, Piezokristall, ...



Regler (das Gehirn):

- Analog: Operationsverstärkerschaltungen, Hydraulik, Pneumatik
- Digital: Software

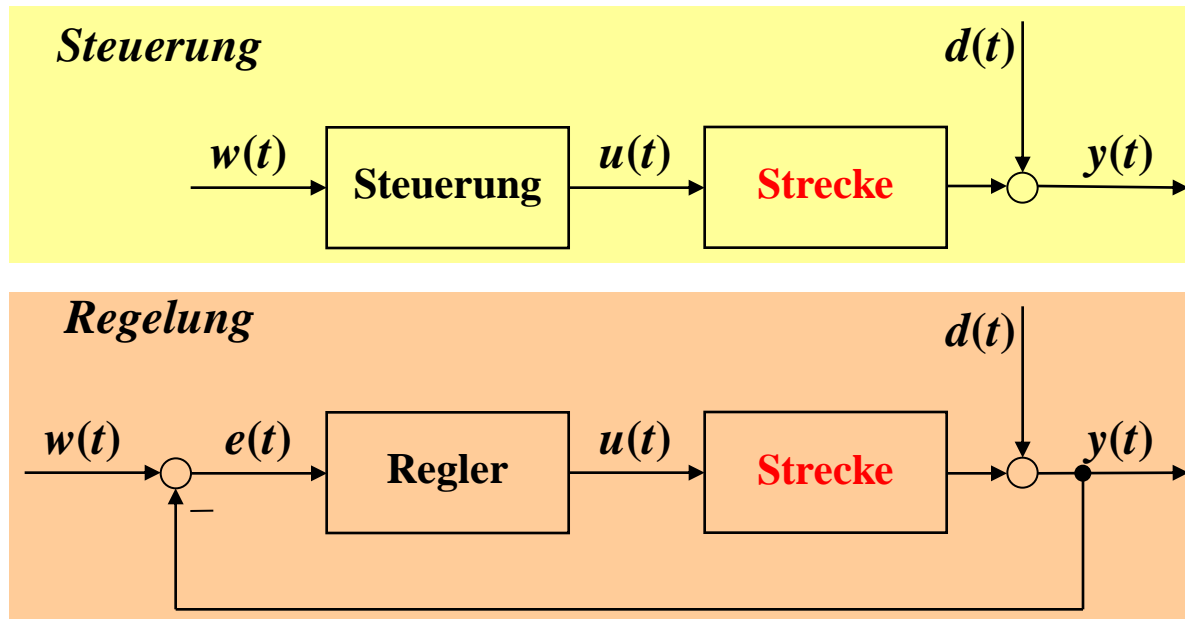


1.4 Steuern versus Regeln

Steuerung:

- Keine Rückkopplung.
 - Keine Messung notwendig (trotzdem kann auch eine Steuerung Messwerte verarbeiten; nur darf *nicht* die Stellgröße $u(t)$ aus der „Regelgröße“ $y(t)$ bestimmt werden).
- Keine Kompensation von Störungen und Modellungenauigkeiten!

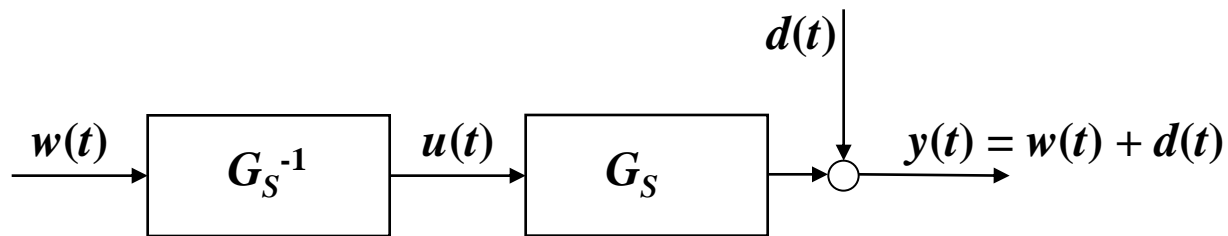
Info: **Strecke** = Stellglied/Aktor + Prozess



1.4 Steuern versus Regeln

Optimale Steuerung:

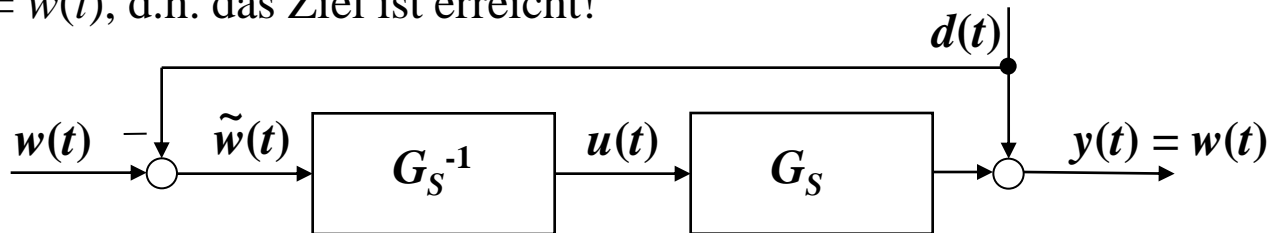
- Forderung: Regelgröße soll Führungsgröße folgen, d.h.: $y(t) = w(t)$.
 - Es sei $y(t) = G_S u(t)$. G_S beschreibt das *Verhalten* der Strecke. Dazu später mehr...
 - Optimale Steuerung: **Inverse** der Strecke, d.h. $u(t) = G_S^{-1} w(t)$.
- $y(t) = w(t) + d(t)$, d.h. das Ziel ist erreicht, nur die Störung $d(t)$ ist nicht beseitigt.



Wenn Störung $d(t)$ messbar (ansonsten ist nichts besseres machbar!):

- Dann kann sie durch Subtraktion von der Führungsgröße kompensiert werden.

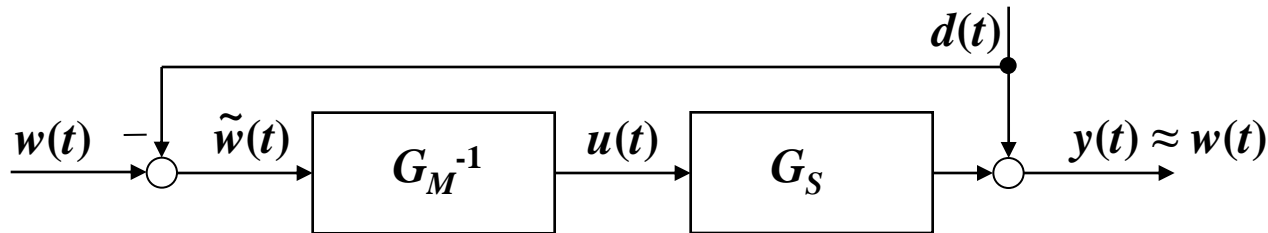
→ $y(t) = w(t)$, d.h. das Ziel ist erreicht!



1.4 Steuern versus Regeln

Einschränkungen zur optimalen Steuerung:

- Das Verhalten der Strecke G_S ist niemals exakt bekannt; bestenfalls kennt man ein gutes Modell G_M der Strecke mit $G_M \approx G_S$.



- Manche Störungen lassen sich nicht messen.
- Teile des Streckenverhaltens dürfen bzw. können gar nicht invertiert werden, weil:
 1. Die Inverse instabil wäre.
 2. Die Inverse in die Zukunft schauen müsste.
 3. Die Inverse rein differenzierendes Verhalten aufweisen würde.

Problem 1 und 2 können auftreten, hängen von dem Streckenverhalten ab. Problem 3 tritt in der Praxis fast immer auf. Dazu später mehr...

1.4 Steuern versus Regeln

- Trotz dieser Einschränkungen ist die Steuerung mit einer näherungsweise Inversen \tilde{G}_M das Beste, was man mit einer Steuerung erreichen kann. Es gibt verschiedene Methoden eine näherungsweise Inverse zu berechnen, die verschiedene Vor- und Nachteile aufweisen. Dazu später mehr...
- Zusammenfassend nochmals die Vor- und Nachteile einer Steuerung im Vergleich zur Regelung:
 - + Keine Messung der Regelgröße erforderlich.
 - + Einfacher Entwurf durch näherungsweise Inversion des Streckenmodells.
 - + Dynamisch schnell, d.h. $y(t)$ folgt $w(t)$ mit nur geringer (im theoretischen Idealfall ganz ohne) Zeitverzögerung.
 - Nicht messbare Störungen können nicht kompensiert werden.
 - Modellungenauigkeiten wirken sich voll aus.

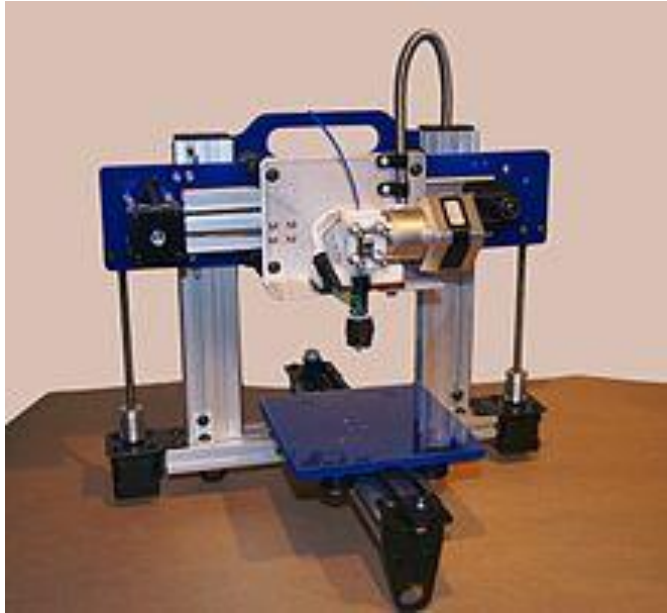
Zu letztgenannten schwerwiegenden Nachteil ein triviales Beispiel mit $d(t) = 0$:

$$G_S = K \quad G_M = K + \varepsilon \quad \rightarrow \quad y(t) = G_M^{-1} G_S w(t) = K / (K + \varepsilon) w(t) = 1 / (1 + \varepsilon/K) w(t)$$

d.h. die Regelgröße folgt der Führungsgröße mit einem systematischen Fehler.

1.4 Steuern versus Regeln

Steuerung: 3D-Drucker

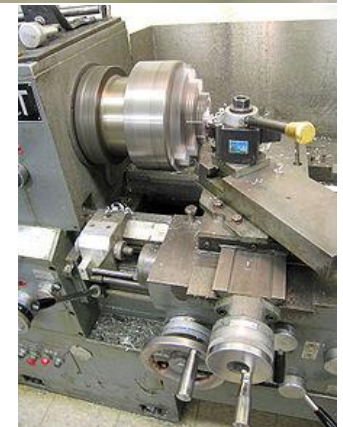


- einfach
- Schrittmotoren
- begrenzte Genauigkeit
- Störungen werden vernachlässigt

Regelung: CNC-Drehmaschine



- komplex
- positionsgeregelt
- hochgenau
- Störungen werden ausgeregelt



1.4 Steuern versus Regeln

Beispiele für Steuerungen

- Gasherd, Mikrowelle, manuelle Klimaanlage (Auto) und Hausheizung, aber NICHT: Kühlschrank, Backofen, moderne Hausheizung: Temperatur soll genau eingehalten werden; daher wird sie gemessen und geregelt!
- Dimmer zur Beleuchtung: Mensch agiert als Regler indem der den Dimmer so lange dreht bis gewünschte Helligkeit erreicht ist.
- Schrittmotoren zur Positionierung: Zähler der Schritte erlaubt Rückschluss auf Position auch ohne Messung, allerdings nur für genügend kleine Störungen (Drehmomente).
- Verbrennungsmotor: Regelung wäre wünschenswert aber lässt sich kaum realisieren, da mögliche Regelgrößen (Motordrehmoment oder Verbrennungsdruck im Zylinder) nicht in Serie sondern nur am Motorprüfstand gemessen werden können (zu teuer, zu wenig robust). AUSNAHMEN: Lambda-Regelung, Leerlaufregelung.
- Straßenbeleuchtung: Evtl. eingeschaltet bei Unterschreiten einer Helligkeitsgrenze, ABER die Helligkeit wird nicht geregelt!
- Ampel: Evtl. verkehrsflussgesteuert durch Fahrzeugmessung mittels Induktionsschleifen.

1.5 Prinzip Rückkopplung

Die beiden Hauptnachteile der Steuerung werden von der Regelung aufgehoben!

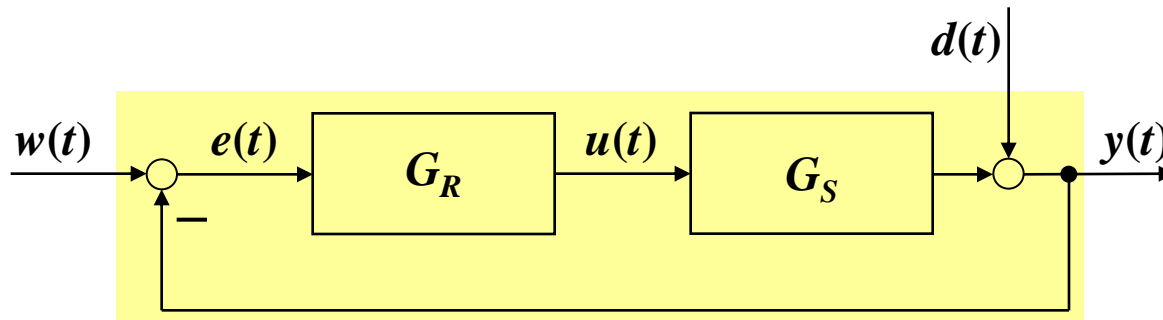
Wenn die Inverse der Strecke der optimalen Steuerung entspricht, welche Rolle spielt sie bei einer Regelung?

Ohne Störung (also $d(t) = 0$) ergibt sich:

$$y(t) = G_R G_S e(t) = G_R G_S [w(t) - y(t)] \rightarrow (1 + G_R G_S) y(t) = G_R G_S w(t)$$

$$\rightarrow y(t) = \frac{G_R G_S}{1 + G_R G_S} w(t) = \frac{1}{1 + \frac{1}{G_R G_S}} w(t)$$

→ Für große Werte von $G_R G_S$ ist $y(t) \approx w(t)$, d.h. der **Regler** zusammen mit der **Rückkopplung erzeugt** implizit eine **näherungsweise Inverse**.



1.5 Prinzip Rückkopplung

Der optimale Regler wäre demnach $G_R \rightarrow \infty$.

Und das Beste dabei ist, dass $y(t) \approx w(t)$ selbst dann gilt, wenn man statt des realen (unbekannten) Streckenverhaltens G_S sein (fehlerbehaftetes) Modell G_M verwendet und eine Störung (also $d(t) \neq 0$) vorliegt:

$$y(t) = G_R G_S [w(t) - y(t)] + d(t) \rightarrow (1 + G_R G_S) y(t) = G_R G_S w(t) + d(t)$$

$$y(t) = \frac{G_R G_S}{1 + G_R G_S} w(t) + \frac{1}{1 + G_R G_S} d(t) = \underbrace{\frac{1}{1 + \frac{1}{G_R G_S}}}_{\mathbf{1}} w(t) + \underbrace{\frac{1}{1 + G_R G_S}}_{\mathbf{0}} d(t)$$

und damit für $G_R \rightarrow \infty$: $y(t) = w(t)$.

Dies bedeutet: Eine **Regelung ist robust** (d.h. unempfindlich) **in Bezug auf Modellfehler** und **Störungen**. Sie werden im Gegensatz zur Steuerung kompensiert, wenn $G_R \rightarrow \infty$. In der Praxis ist es nicht möglich, generell $G_R \rightarrow \infty$ zu realisieren. Aber man kann dies für bestimmte Frequenzen bzw. Frequenzbereiche (exakt bzw. näherungsweise) tun. Dazu später mehr...

1.5 Prinzip Rückkopplung

Beispiele für Regelungen (meist digital in elektronischen Steuergeräten) im Kfz

<i>Name</i>	<i>Stellgröße</i>	<i>Regelgröße</i>
λ -Regelung	Einspritzmenge/Luftmasse	Luft-Kraftstoff-Verhältnis λ
Leerlaufregelung	Einspritzmenge/Luftmasse	Motordrehzahl
Tempomat	Motorleistung und evtl. Gang	Fahrzeuggeschwindigkeit
ABS	Bremsdruck über Hydraulikventile	Bremsschlupf/Raddrehzahl
ASR	Motorleistung	Antriebsschlupf/Raddrehzahl
Aktives Fahrwerk	Federkonstante des Fahrwerks	Wankwinkel des Fahrzeugs
Niveauregulierung	Federkonstante des Fahrwerks	Abstand zw. Karosserie und Achse
Leuchtweite	Vertikaler Winkel der Scheinwerfer	Nickwinkel des Fahrzeugs
Kurvenlicht	Horizontaler Winkel der Scheinwerfer	Lenkradwinkel
Klimaanlage	Kompressorleistung	Temperatur
Abblend. Spiegel	Reflexionsgrad des LCD	Lichteinfall auf den Innenspiegel
Auto. Scheibenw.	Scheibenwischermotor	Regendichte

1.6 Automatisierungstechnik

Zeitraumen	Ebene	Aufgaben	Tools
Tage	Management	Strategie, Planung, Logistik	SAP
Stunden	Leitsystem	Optimierung der Arbeitspunkte und Profile, Data Mining	Prozessleitsystem
Minuten	langsame Regelkreise	überlagerte Regelung, Diagnose	Adaptive und prädiktive Regelung, Mehrgrößenregelung
Sekunden	schnelle Regelkreise	unterlagerte Regelung, Steuerungen	PID-Regelung

1.6 Automatisierungstechnik

Automatisieren ist mehr als Steuern und Regeln kontinuierlicher Prozesse:

- Steuerung ereignisdiskreter Prozesse (SPS)
- Prozessvisualisierung
- Überwachung, Fehlererkennung, Fehlerdiagnose, Rekonfiguration
- Optimierung
- Datenspeicherung, Data Mining
- Kommunikationssysteme / Vernetzung / Bussysteme
- Messtechnik / Sensorik
- Stellglieder / Aktorik

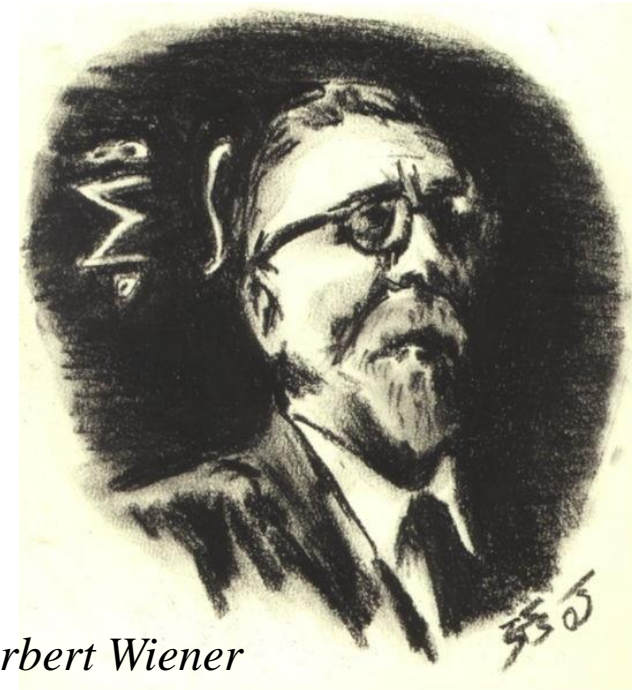
1.7 Kybernetik

Der Begriff Kybernetik (griechisch: Steuermann) wurde von Norbert Wiener 1947 geprägt. Er beschreibt die Wissenschaft *komplexer dynamischer Systeme*, die dazu dienen ein *Gleichgewicht* aufrecht zu erhalten oder ein *Ziel* zu erreichen (*Navigation*).

Das universelle Prinzip der *Rückkopplung* ist damit das zentrale Element der Kybernetik. Sie wird oft die „Kunst des Steuerns“ beschrieben.

Die Kybernetik ist eine Zusammenführung u.a. der folgenden Disziplinen:

- Regelungstechnik und Systemtheorie
- Nachrichtentechnik und Informationstheorie
- Entscheidungs- und Spieltheorie



Norbert Wiener

1.9 Wichtige Begriffe in englischer Sprache

Regelung	<i>feedback control, closed-loop control</i>
Steuerung	<i>feedforward control</i>
offener / geschlossener Regelkreis	<i>open / closed (control) loop</i>
Prozess / Regelstrecke	<i>process / plant</i>
Modell / Modellierung	<i>model / modeling</i>
Regler	<i>controller</i>
Regelgröße	<i>controlled variable, control output</i>
Stellgröße	<i>manipulated variable, control variable</i>
Führungsgröße	<i>reference</i>
Störgröße / Rauschen	<i>disturbance / noise</i>
Regelabweichung	<i>control deviation, control error</i>
Festwertregelung / Folgeregelung	<i>regulation / tracking</i>
Sensor / Stellglied, Aktor	<i>sensor / actuator</i>
Optimierung / Adaption	<i>optimization / adaptation</i>

1.10 Literatur

Grundlage dieser Vorlesung:

- Lunze: „Regelungstechnik 1“, 4. Aufl., Springer, 2004. (635 Seiten).
Web Site: <http://www.esr.ruhr-uni-bochum.de>
- Goodwin, Graebe, Salgado: „Control System Design“, Prentice Hall, 2000. (944 Seiten). Web Site: <http://csd.newcastle.edu.au/control/>
- Dorf, Bishop: „Moderne Regelungssysteme“, Pearson, 2005. (1168 Seiten).
- Interessante Bücher kostenlos als PDF verfügbar. Unter <http://link.springer.com/> nach „Regelungstechnik“ suchen.

Ergänzungen und Erweiterungen:

- Lunze: „Regelungstechnik 2“, 3. Aufl., Springer, 2005. (635 Seiten).
Web Site: <http://www.esr.ruhr-uni-bochum.de>
- Föllinger; „Regelungstechnik“, 8. Aufl., Hüthig, 1994. (646 Seiten).

Nachschlagewerk:

- Lutz, Wendt: „Taschenbuch der Regelungstechnik“, 5. Aufl., H. Deutsch, 2003 (1160 S).

2. Modellierung linearer dynamischer Systeme

Inhalt Kapitel 2

2 Modellierung linearer dynamischer Systeme

2.1 Beispiele

2.2 Linearität und weitere Eigenschaften

2.3 Generelle Vorgehensweise bei der Modellierung

2.4 Allgemeine lineare DGL n -ter Ordnung

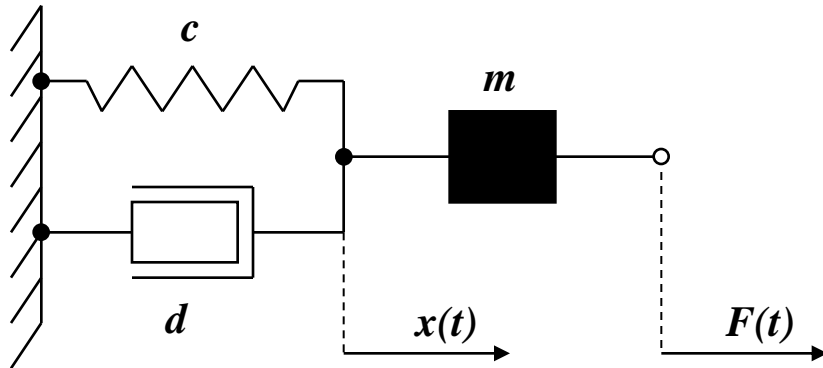
2.5 Homogene und partikuläre Lösung der DGL

2.6 Lösung der DGL 1. Ordnung

2.7 Lösung der DGL 2. Ordnung

2.1 Beispiele

Mechanisches Beispiel



Kräftebilanz in x -Richtung:

$$F(t) - m\ddot{x}(t) - cx(t) - d\dot{x}(t) = 0$$

$$\ddot{x}(t) + \frac{d}{m}\dot{x}(t) + \frac{c}{m}x(t) = \frac{1}{m}F(t)$$

Ausgang

Eingang

Variablen:

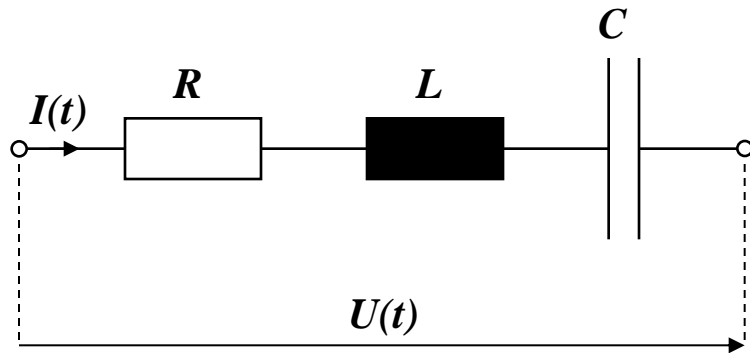
- Eingangsgröße: Kraft F
- Ausgangsgröße: Weg x

Parameter:

- Feder mit Federkonstante c
- Dämpfer mit Dämpfungskonstante d
- Masse m

2.1 Beispiele

Elektrisches Beispiel



Spannungsbilanz:

$$\dot{U}(t) - R\dot{I}(t) - L\ddot{I}(t) - \frac{1}{C}I(t) = 0$$

$$\ddot{I}(t) + \frac{R}{L}\dot{I}(t) + \frac{1}{LC}I(t) = \frac{1}{L}\dot{U}(t)$$

Ausgang

Eingang

Variablen:

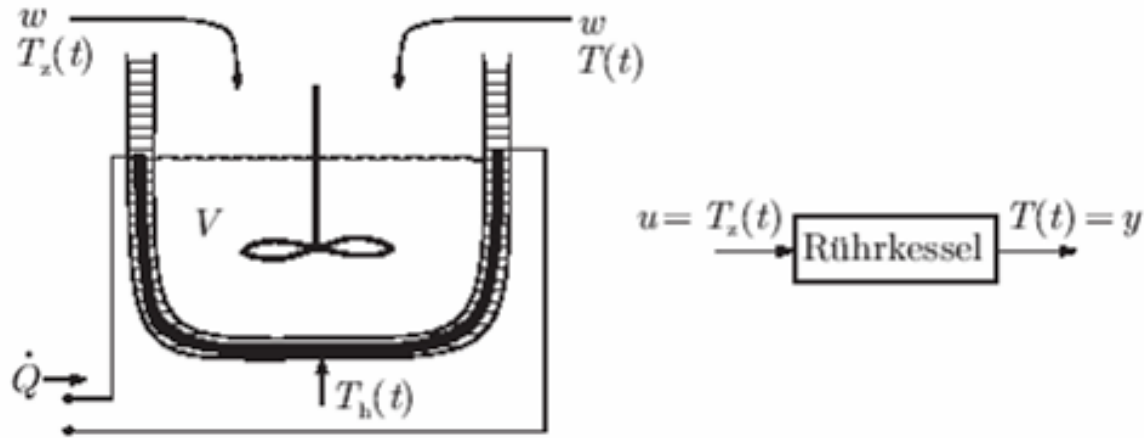
- Eingangsgröße: Spannung U
- Ausgangsgröße: Strom I

Parameter:

- Kondensator mit Kapazität C
- Ohmscher Widerstand R
- Spule mit Induktivität L

2.1 Beispiele

Verfahrenstechnisches Beispiel



→ Modellierung in Beispiel 4.3 aus *Lunze 1*.

Lösung vom selben Typ wie die beiden anderen Beispiele:

$$a_2 \ddot{T}(t) + a_1 \dot{T}(t) + a_0 T(t) = b_1 \dot{T}_z(t) + b_0 T_z(t)$$

2.1 Beispiele

Alle drei Beispiele führen auf eine Gleichung des selben Typs!

Vergleich zwischen mechanischem und elektrischem Beispiel:

- Jeweils ein Energieverbraucher (Energie wird in Wärme umgewandelt):
Dämpfer d \leftrightarrow Widerstand R
- Jeweils zwei Energiespeicher:
 - Feder c (potentielle E.: $\frac{1}{2} \cdot c \cdot x^2$) \leftrightarrow Kondensator C (elektrische E.: $\frac{1}{2} \cdot 1/C \cdot Q^2$)
 - Masse m (kinetische E.: $\frac{1}{2} \cdot m \cdot v^2$) \leftrightarrow Spule L (magnetische E.: $\frac{1}{2} \cdot L \cdot I^2$)

$$v = \dot{x} \leftrightarrow I = \dot{Q}$$

Wenn der Energieverbraucher nicht zu groß ist, wird die Energie von einem Speicher zum anderen hin- und hertransportiert \rightarrow **Schwingkreis**.

Wenn der Energieverbraucher zu groß ist, wird die Energie verbraucht (also in Wärme umgewandelt und damit dem System entzogen) bevor es zum Schwingen kommen kann.

2.2 Linearität und weitere Eigenschaften

Linearität:

Ein Zusammenhang $y = f(x)$ ist **linear**, wenn gilt:

- Additivität (Superpositionsprinzip): $f(x_1) + f(x_2) = f(x_1 + x_2)$
- Homogenität (Verstärkungsprinzip): $K \cdot f(x) = f(K \cdot x)$

Beispiel: 1-dimensionale lineare Funktion: Gerade durch den Ursprung: $y = ax$.

Wir beschäftigen uns zu 95% nur mit linearen DGLs.

Affinität:

Wenn ein Zusammenhang bis auf einen Gleichwert linear ist, dann ist er **affin**.

Beispiel: 1-dimensionale affine Funktion: Gerade: $y = ax + b$.

Durch eine lineare Koordinatentransformation lässt sich aus einem affinen Zusammenhang ein linearer machen. Deshalb können affine und lineare Systeme mit den selben Methoden behandelt werden.

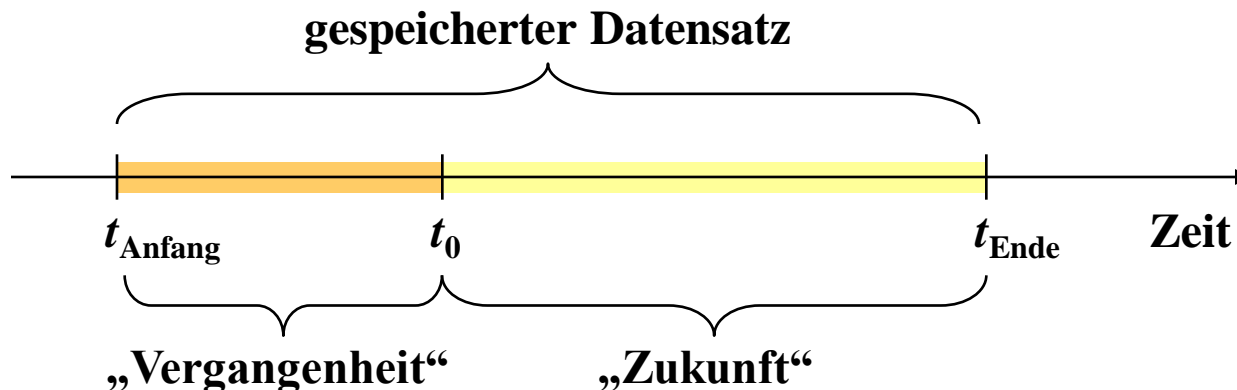
2.2 Linearität und weitere Eigenschaften

Kausalität:

Der Ausgang eines *kausalen* Systems zum Zeitpunkt $t = t_0$ hängt nur von Größen der Gegenwart und der Vergangenheit (also von $t \leq t_0$) ab; insbesondere kann eine Eingangsgröße zum Zeitpunkt $t = t_0$ nur den gegenwärtigen und den zukünftigen (also für $t \geq t_0$) Verlauf der Ausgangsgröße beeinflussen.

Solange Zeitreisen unmöglich sind, ist jedes reale System kausal. Trotzdem gibt es Anwendungen für *akausale* (nicht kausale) Systeme:

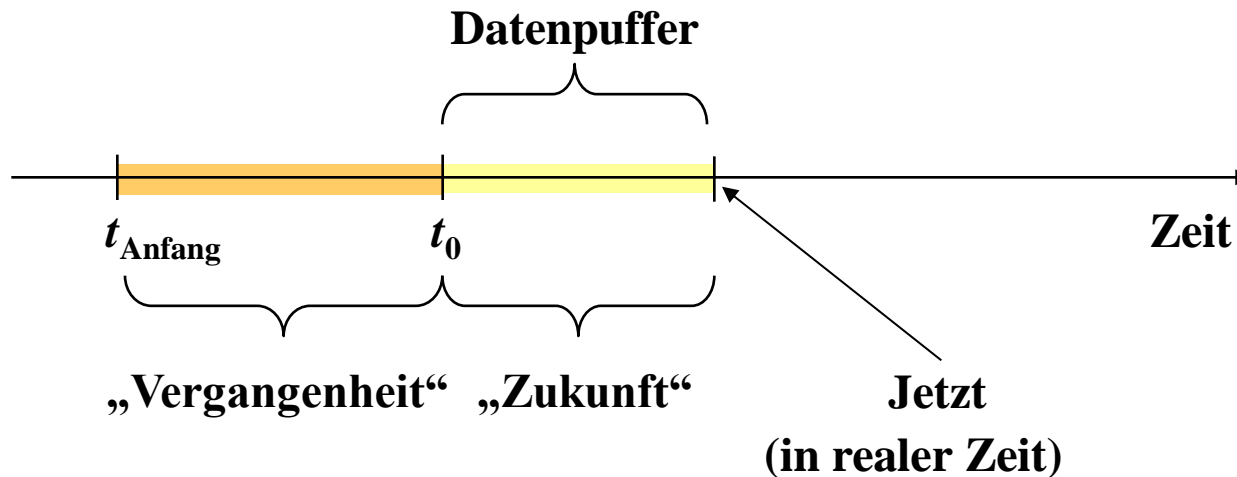
- Wenn im Computer ein zuvor gemessener Datensatz gespeichert wurde, kann man (innerhalb der Datensatzgrenzen) auf „zukünftige Daten“ zugreifen. Somit ist z.B. eine akausale Filterung der Daten möglich.



2.2 Linearität und weitere Eigenschaften

Weiteres Beispiel für Akausalität:

- Messdaten werden nicht sofort bearbeitet sondern in einem Puffer zwischengespeichert. Innerhalb der Grenzen des Puffers kann dann auf die Zukunft zugegriffen werden.



Diese Vorgehensweise ist bei rückgekoppelten Signalen nicht sinnvoll! Warum?

- Pufferung bedeutet Zeitverzögerung. Das verschlechtert die Regelung, weil die Informationen über Regelabweichungen später eintreffen. Dazu später mehr...

2.2 Linearität und weitere Eigenschaften

Zeitinvarianz:

Ein System ist *zeitinvariant*, wenn es sich über die Zeit nicht ändert. D.h. die Antwort $y(t+t_0)$ auf ein identisches Eingangssignal $u(t+t_0)$ ist von t_0 unabhängig.

Beispiel: Typisches *zeitvariantes* Verhalten ist Verschleiß. Wird oft vernachlässigt.

Weiteres Beispiel: Ein typisches *zeitvariantes* System ist eine startende Rakete. Ihre Masse ändert sich durch den verbrannten Treibstoff.

Gleiches gilt für ein Auto; nur hier ist die Masseveränderung im Bereich $< 5\%$ und damit meist vernachlässigbar. Also kann die Längsdynamik eines Autos mit einer zeitinvarianten DGL modelliert werden. Dies gilt nicht mehr, wenn das Auto sehr stark be- oder entladen wird. Im Gegensatz zur Rakete ist hier die Zeitvarianz sprunghaft (zu den Be- oder Entladungszeitpunkten) und nicht kontinuierlich.

Konzentrierte Parameter:

Wir nehmen an, dass die betrachteten Systeme durch gewöhnliche DGLs mit konzentrierten Parametern beschrieben werden können. Partielle DGLs werden hier nicht behandelt.

2.2 Linearität und weitere Eigenschaften

Es scheint eine große Einschränkung zu sein, wenn wir uns auf gewöhnliche, zeitinvariante, kausale, lineare DGLs konzentrieren. Dem ist *nicht* so, weil:

- Jeder reale Prozess ist *kausal*.
- Bei den meisten Prozessen ist eine *Zeitvarianz* so klein in ihrer Auswirkung oder so langsam, dass sie vernachlässigbar ist oder mit der Theorie zeitinvarianter Systeme näherungsweise behandelt werden kann.
- Prozesse mit *verteilten Parametern* können oft näherungsweise mit konzentrierten Parametern beschrieben werden. Andernfalls stellen sie ein Spezialgebiet dar, dessen Behandlung die hier besprochenen Methoden erweitert und damit deren Kenntnis voraussetzt.
- Zwar sind alle Prozesse nichtlinear aber viele lassen sich näherungsweise durch *lineare* Modelle beschreiben. Wo dies nicht möglich ist (bei schnellen, häufigen Arbeitspunktwechseln z.B.) können evtl. mehrere lineare Modelle, zwischen denen gewechselt wird, genügen. Andernfalls müssen sehr viel komplexere nichtlineare Methoden verwendet werden, die aber auch die Kenntnis der linearen Verfahren erfordern!

2.2 Linearität und weitere Eigenschaften

Häufig wurde bisher und wird in der Zukunft der Begriff „*näherungsweise*“ verwendet. Es gibt, zusätzlich zu den oben genannten, weitere Gründe, warum unsere Modelle die betrachtete Regelstrecke nur approximativ beschreiben können:

- Der Prozess ist nicht gut verstanden (insbesondere bei biologischen und chemischen aber auch bei thermo- und fluiddynamischen Prozessen).
- Die Modelle sind zu komplex, als dass sie sinnvoll für einen Reglerentwurf genutzt werden können (Rechenzeit, Überblick, intuitives Verständnis, ...). Es müssen Vereinfachungen durchgeführt werden. → Modellreduktion.
- Die Modelle enthalten unbekannte Parameter, die nicht (oder nur ungenau) aus den vorhandenen Daten bestimmt (geschätzt) werden können.
- Die geschätzten Parameter im Modell weisen statistische Unsicherheiten auf, weil die zur Schätzung verwendeten Daten stark verrauscht waren.

Wir müssen also mit **Modellen** leben, die **nur näherungsweise richtig** sind. Zum Glück haben wir in Kap. 1 gesehen, dass eine **Regelung robust in Bezug auf Modellungenauigkeiten** ist!

2.3 Generelle Vorgehensweise bei der Modellierung

Aufgabe der Modellierung:

- *Systemzerlegung*: System wird in Komponenten zerlegt.
- *Komponentenmodelle*: Gesetze aufschreiben, die jede Komponente beschreiben.
- *Kopplungsbeziehungen*: Beziehungen aufschreiben, die zwischen den Komponenten bestehen.
- *Modellumformung*: Gleichungen zu einer DGL zusammenfassen und nach Eingangs- und Ausgangsgrößen ordnen.


Beim Aufstellen der Gesetzmäßigkeiten können wir unterscheiden:

- *Bilanzgleichungen*: Energiestrombilanz, Impulsbilanz, ...
- *Konstitutive Gleichungen*: u.a. Verbindung von Potential- mit Flussgrößen: Induktionsgesetz, Wärme- / Flüssigkeitsspeicher, Spule / Kondensator, ...
- *Phänomenologische Gleichungen*: irreversible Prozesse: Wärmeleitung, ohmsches Gesetz, ...

2.4 Allgemeine lineare DGL n -ter Ordnung

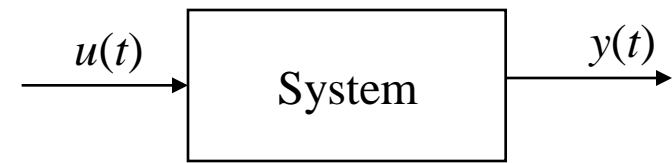
In den Beispielen aus Kap. 2.1 haben wir bereits gesehen, dass sich verschiedenste Prozesse aus eine lineare DGL gleichen Typs zurückführen lassen. Zum größten Teil werden folgende allgemeine gewöhnliche, kausale, zeitinvariante, lineare DGL die Grundlage sein:

$$a_n y^{(n)}(t) + a_{n-1} y^{(n-1)}(t) + \dots + a_1 y^{(1)}(t) + a_0 y(t) = b_m u^{(m)}(t) + b_{m-1} u^{(m-1)}(t) + \dots + b_1 u^{(1)}(t) + b_0 u(t)$$

1. Ableitung 

- In Praxis realisierbar sind nur Systeme mit $n > m$. Für rechnerische Zwischenergebnisse muss diese Bedingung aber nicht erfüllt sein. Dazu später mehr...
- Man kann ohne Einschränkungen $a_n = 1$ setzen (Gleichung durch a_n teilen).
- Zur Lösung der DGL benötigen wir die folgenden Anfangsbedingungen zu einem Zeitpunkt t_0 (wird meist als $t_0 = 0$ gewählt):

$$y(t_0), y^{(1)}(t_0), \dots, y^{(n-1)}(t_0)$$
$$u(t_0), u^{(1)}(t_0), \dots, u^{(m-1)}(t_0)$$



2.5 Homogene und partikuläre Lösung der DGL

Skizze des Lösungswegs:

- Die homogene DGL beschreibt das Verhalten ohne Eingangsgröße (also für $u(t) = 0$), das sog. *Eigenverhalten*, d.h. das System bleibt von den Anfangswerten ausgehend sich selbst überlassen:

$$a_n y^{(n)}(t) + a_{n-1} y^{(n-1)}(t) + \dots + a_1 y^{(1)}(t) + a_0 y(t) = 0$$

- Die Lösung dieser **homogenen** DGL nennen wir $y_{\text{frei}}(t)$.
- Mit Hilfe von $y_{\text{frei}}(t)$ können wir die **partikuläre** Lösung der DGL mittels „Variation der Konstanten“ (siehe Mathematik) berechnen. Sie bezeichnet den Teil der Lösung, der durch ein von Null verschiedenes Eingangssignal ($u(t) \neq 0$) erzeugt wird. Da dieser Lösungsanteil durch das Eingangssignal vom System erzwungen wurde, nennen wir ihn $y_{\text{erzw}}(t)$.
- Die Gesamtlösung ergibt sich aus der Überlagerung der Systemantworten auf die Anfangsbedingungen und auf das Eingangssignal:

$$y(t) = y_{\text{frei}}(t) + y_{\text{erzw}}(t)$$

2.6 Lösung der DGL 1. Ordnung

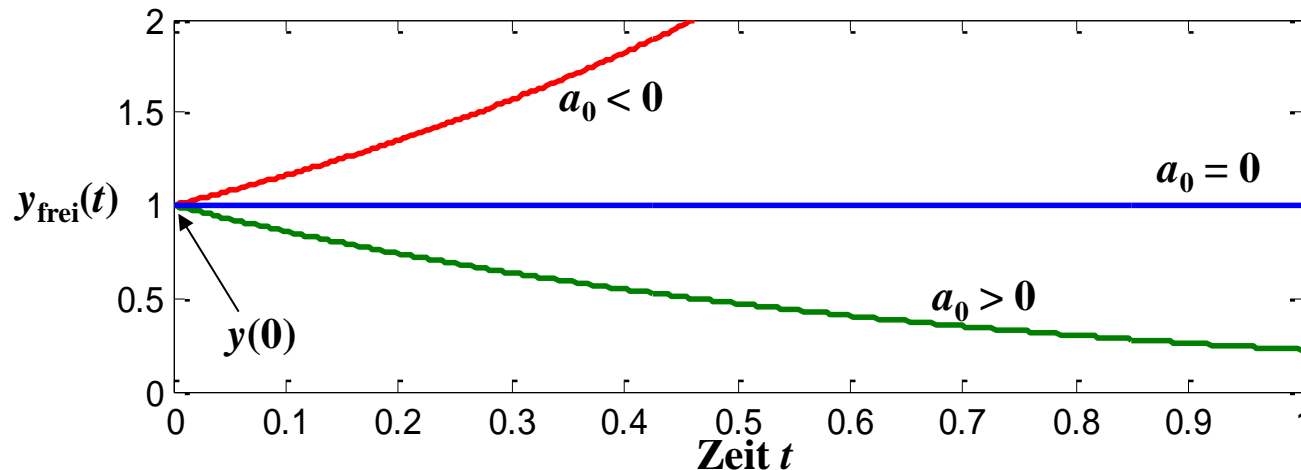
DGL 1. Ordnung: $\dot{y}(t) + a_0 y(t) = b_0 u(t)$

Lösung: $y(t) = \underbrace{e^{-a_0 t} y(0)}_{y_{\text{frei}}(t)} + b_0 \underbrace{\int_0^t e^{-a_0(t-\tau)} u(\tau) d\tau}_{y_{\text{erzw}}(t)}$

← **Faltungsintegral**

3 Fälle für die homogene Lösung:

- $a_0 > 0$: $y_{\text{frei}}(t)$ klingt exponentiell von der Anfangsbedingung $y(0)$ auf 0 ab.
- $a_0 = 0$: $y_{\text{frei}}(t)$ verharrt an der Anfangsbedingung $y(0)$.
- $a_0 < 0$: $y_{\text{frei}}(t)$ wächst exponentiell von der Anfangsbed. $y(0)$ über alle Grenzen.



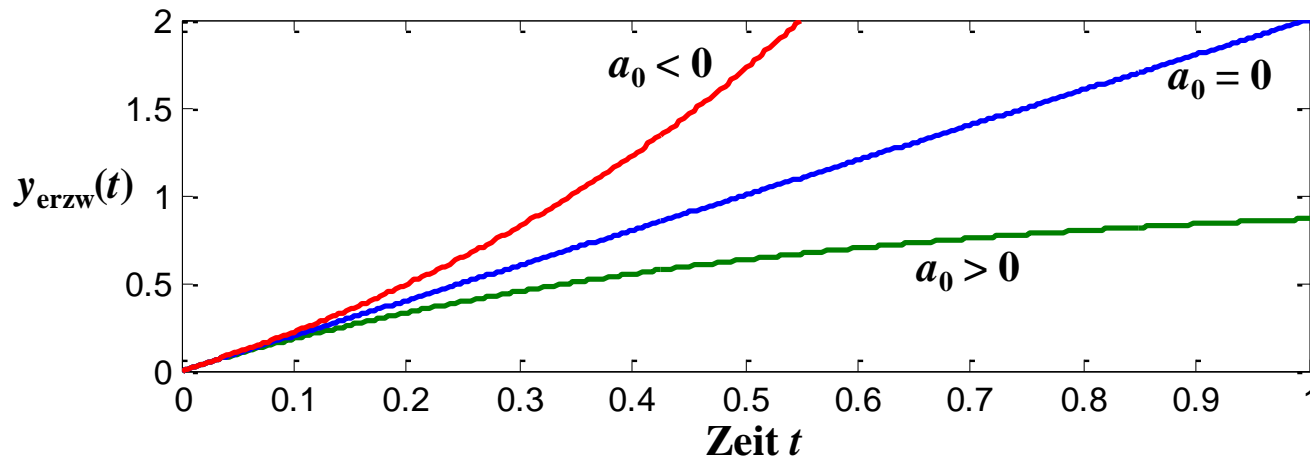
2.6 Lösung der DGL 1. Ordnung

Partikuläre Lösung für Einheitssprung als Eingangssignal $u(t) = \begin{cases} 0 & \text{für } t < 0 \\ 1 & \text{für } t \geq 0 \end{cases}$

$$y(t) = b_0 \int_0^t e^{-a_0(t-\tau)} d\tau = \begin{cases} \frac{b_0}{a_0} (1 - e^{-a_0 t}) & a_0 \neq 0 \\ b_0 t & a_0 = 0 \end{cases}$$

3 Fälle für die partikuläre Lösung:

- $a_0 > 0$: $y_{\text{erzw}}(t)$ nähert sich exponentiell seinem Endwert b_0/a_0 .
- $a_0 = 0$: $y_{\text{erzw}}(t)$ steigt rampenförmig an (verläuft entlang einer Geraden).
- $a_0 < 0$: $y_{\text{erzw}}(t)$ wächst exponentiell über alle Grenzen.



2.7 Lösung der DGL 2. Ordnung

DGL 2. Ordnung: $\ddot{y}(t) + a_1\dot{y}(t) + a_0y(t) = b_1\dot{u}(t) + b_0u(t)$

Wurzeln (Nullstellen) der charakteristischen Gleichung $\lambda^2 + a_1\lambda + a_0 = 0$:

$$\lambda_{1/2} = -\frac{a_1}{2} \pm \sqrt{\frac{a_1^2}{4} - a_0}$$

3 Fälle für die homogene Lösung (Konst. C_1 und C_2 ergeben sich aus den Anfangsbed.):

1. $a_1^2/4 > a_0$: 2 reelle Wurzeln $\lambda_{1/2}$:

**Überlagerung zweier
DGLs 1. Ordnung**

$$y_{\text{frei}}(t) = C_1e^{\lambda_1 t} + C_2e^{\lambda_2 t}$$

2. $a_1^2/4 = a_0$: 2 identische reelle Wurzeln $\lambda_1 = \lambda_2$:

**Exponentiell gewichteter
rampenförmiger Verlauf**

$$y_{\text{frei}}(t) = C_1e^{\lambda_1 t} + C_2te^{\lambda_1 t}$$

3. $a_1^2/4 < a_0$: konjugiert komplexe Wurzeln $\lambda_{1/2} = \alpha \pm i\beta$:

**Exponentiell gewichtete
Schwingung**

$$y_{\text{frei}}(t) = e^{\alpha t} (C_1\cos(\beta t) + C_2\sin(\beta t))$$

2.7 Lösung der DGL 2. Ordnung

Für Fall 1 und 2 (reelle Wurzeln) schreibt man die DGL üblicherweise als:

$$\ddot{y}(t) - (\lambda_1 + \lambda_2)\dot{y}(t) + \lambda_1\lambda_2y(t) = \dots$$

Für Fall 3 (konjugiert komplexe Wurzeln) schreibt man die DGL üblicherweise als:

$$\ddot{y}(t) + 2D\omega_0\dot{y}(t) + \omega_0^2y(t) = \dots$$

D : Dämpfung des Systems; nur $D < 1$ führt auf Fall 3! $D = 1 \rightarrow$ Fall 2. $D > 1 \rightarrow$ Fall 1.

ω_0 : Kreisfrequenz der Schwingung

Mechanisches Beispiel:

$$D = \frac{d}{2\sqrt{mc}}$$

$$\omega_0 = \sqrt{\frac{c}{m}}$$

Elektrisches Beispiel:

$$D = \frac{R}{2} \sqrt{\frac{C}{L}}$$

$$\omega_0 = \frac{1}{\sqrt{LC}}$$

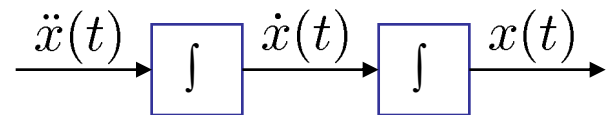
2.7 Lösung der DGL 2. Ordnung

Simulation der partikuläre Lösungen: Beispiel Mechanisches System

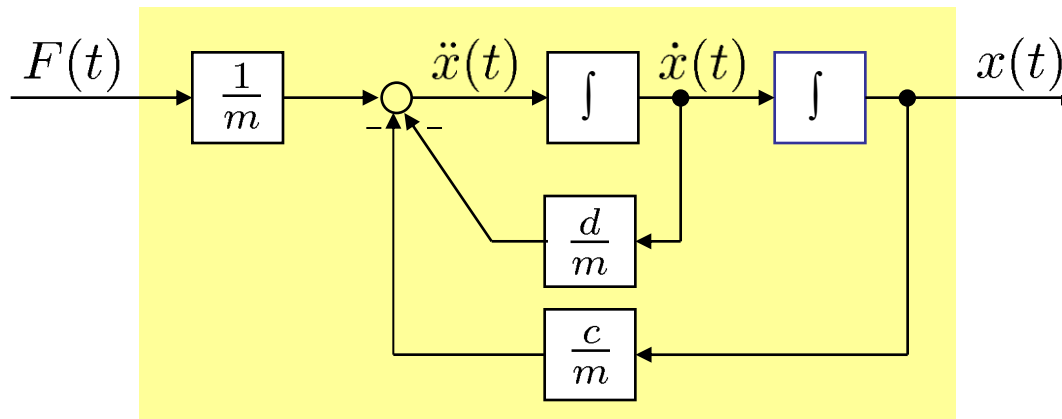
- Zunächst DGL auflösen nach höchster Ableitung des Ausgangs:

$$\ddot{x}(t) = -\frac{d}{m}\dot{x}(t) - \frac{c}{m}x(t) + \frac{1}{m}F(t)$$

- Dann Integratorkette realisieren von höchster Ableitung des Ausgangs bis zu $x(t)$:



- Dann Gleichung für höchste Ausgangsableitung realisieren:



3. Linearisierung nichtlinearer Systeme

Inhalt Kapitel 3

3 Linearisierung nichtlinearer Systeme

3.1 Groß- und Kleinsignalverhalten

3.2 Nichtlineare Dynamik und Kennlinien

3.4 Linearisierung nichtlinearer statischer Systeme

3.5 Linearisierung nichtlinearer dynamischer Systeme

Inhalt Kapitel 3

Ableitungen

Man kann eine Größe y nach jeder Variablen ableiten. In der Schule war das meist die Ableitung von $y(x)$ nach x . Bei uns ist es meist $y(t)$ abgeleitet nach der Zeit t . Dies verdeutlichen wir durch den Punkt über der Größe, also für die 1. und 2. Ableitung nach der Zeit:

$$\dot{y} = \frac{dy}{dt} \quad \ddot{y} = \frac{d^2y}{dt^2}$$

Die Ableitung ist immer die Steigung der Größe in Richtung der Variablen. Deswegen ist die Ableitung am Optimum auch $= 0$.

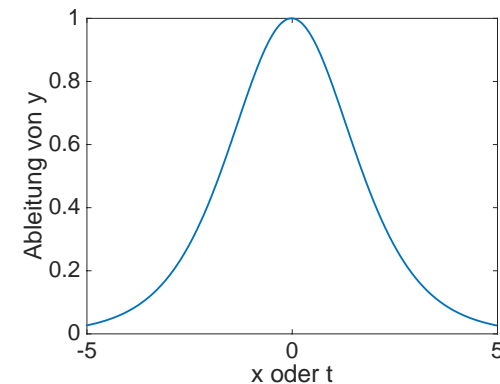
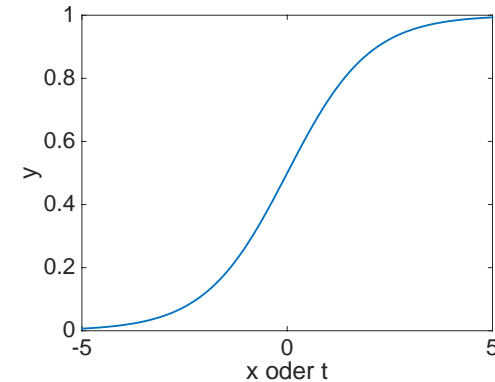
Die Ableitung nach der **Zeit** kann auch als Rate oder Änderungsgeschwindigkeit **interpretiert** werden.

→ Beträgsmäßig große zeitliche Ableitung → Schnell

→ Beträgsmäßig kleine zeitliche Ableitung → Langsam

→ Zeitliche Ableitung $= 0$ → Keine Änderung (stationär, statischer AP)

Bei der **Linearisierung** muss nach den Größen $y, \dot{y}, \ddot{y}, \dots, u, \dot{u}, \ddot{u}, \dots$ selbst abgeleitet werden, weil nichtlineare dynamische Zusammenhänge (nichtlineare DGL) durch eine lineare DGL approximiert werden sollen.



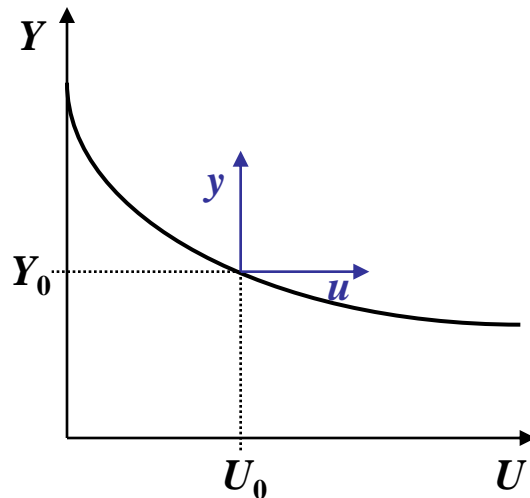
3.1 Groß- und Kleinsignalverhalten

Großsignale

- Wirklich gemessene Signale
- Benannt mit Großbuchstaben: Eingang U , Ausgang Y

Kleinsignale

- Signalabweichungen um einen Arbeitspunkt herum
- Meist wird angenommen, dass diese Abweichungen „klein“ sind
- Benannt mit Kleinbuchstaben Eingang u , Ausgang y



Wie im Bild zu sehen ist gilt:

$$u = U - U_0$$

$$y = Y - Y_0$$

Statischer Zusammenhang:

$$Y = f(U)$$

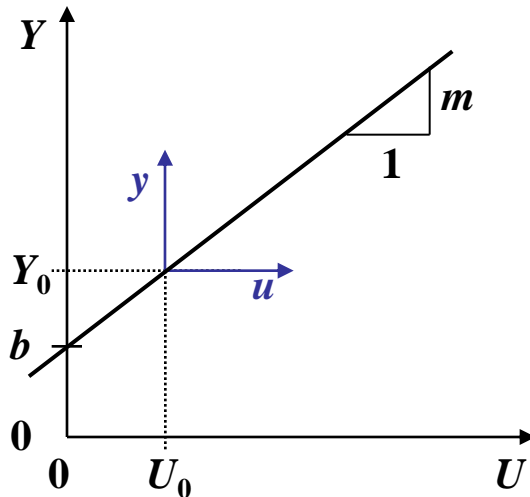
$$Y_0 = f(U_0)$$

3.1 Groß- und Kleinsignalverhalten

Beispiel: Affine Kennlinie

- Gleichung der Kennlinie: $Y = m \cdot U + b$
- Wähle einen beliebigen Arbeitspunkt (U_0, Y_0) auf der Kennlinie: $Y_0 = m \cdot U_0 + b$
- Dies ergibt folgendes Kleinsignalverhalten:
- $Y - Y_0 = m \cdot U + b - (m \cdot U_0 + b) = m \cdot (U - U_0) \rightarrow y = m \cdot u$

→ **Affiner Zusammenhang wird im Kleinsignalverhalten linear.**
(Gleichwert wird unterdrückt)

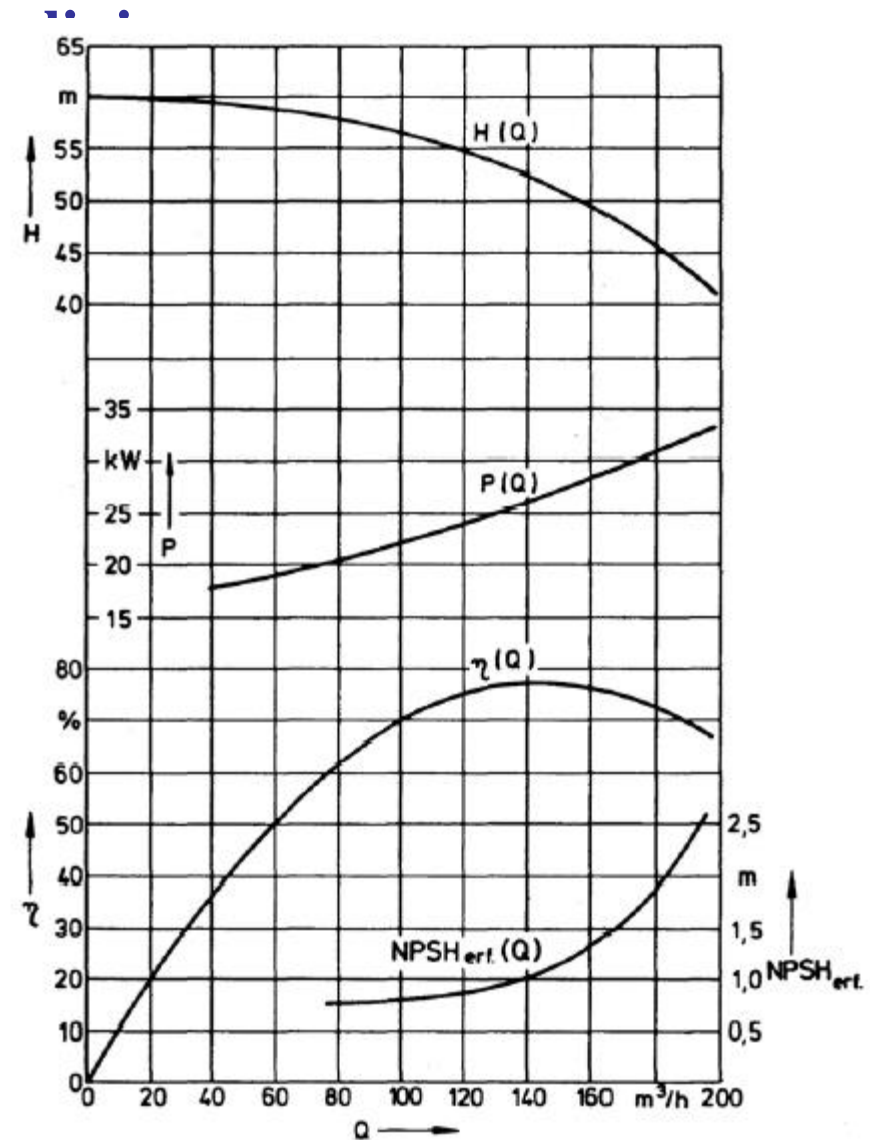
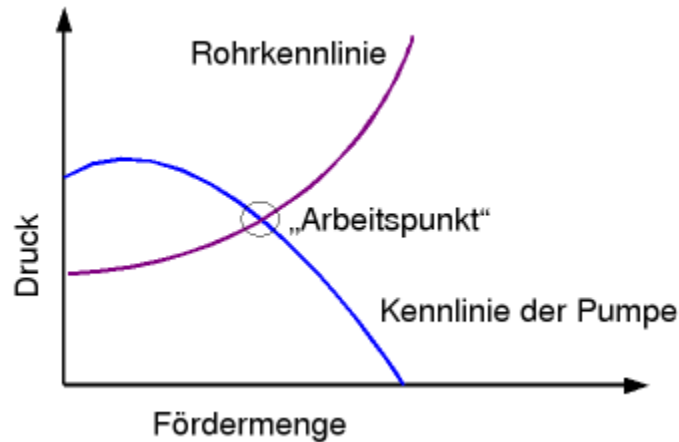


Jeder *nichtlineare* Zusammenhang kann im Kleinsignalverhalten **näherungsweise linear** beschrieben werden.

Die Näherung ist um so besser, je „weniger nichtlinear“ die Kennlinie am Arbeitspunkt ist, und je kleiner u und y sind.

3.2 Nichtlineare Dynamik und Ke

Beispiel: Kreiselpumpe



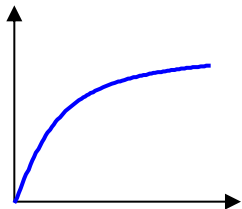
Kennlinien einer einstufigen Kreiselpumpe bei der Drehzahl $n = 1450$ 1/min

3.2 Nichtlineare Dynamik und Kennlinien

Beispiel: Stoßdämpfer

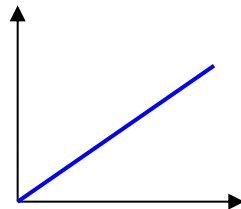
- Nichtlinearität hängt ab von der Geschwindigkeit des Ein- bzw. Ausfederns.
- Nichtlinearität ist unsymmetrisch, d.h. ist unterschiedlich für positive (Einfedern) und negative (Ausfedern) Geschwindigkeiten.
- Wenn sich die Dämpfer der Fahrsituation anpassen, handelt es sich um ein *zeit-variantes* System.
- Man unterscheidet 3 Hauptcharakteristika:

Degressiv

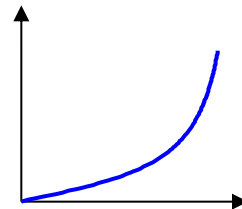


Weich bei großen Unebenheiten

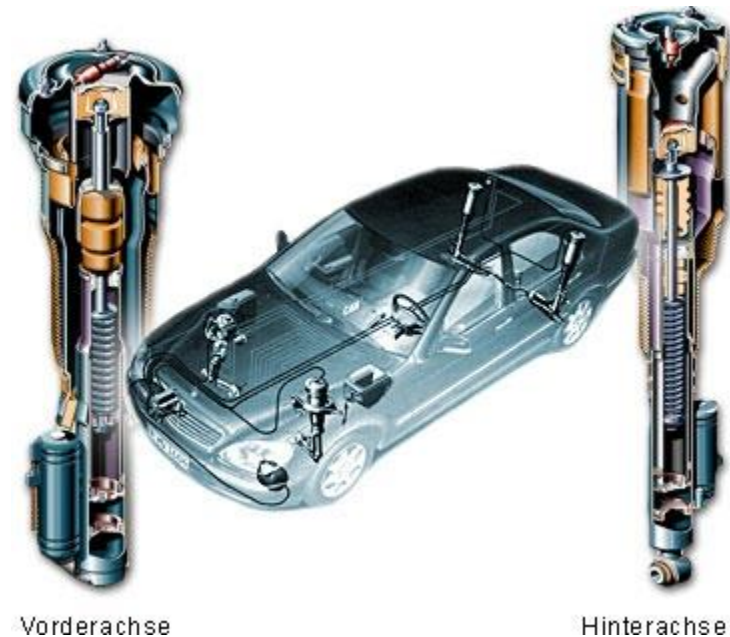
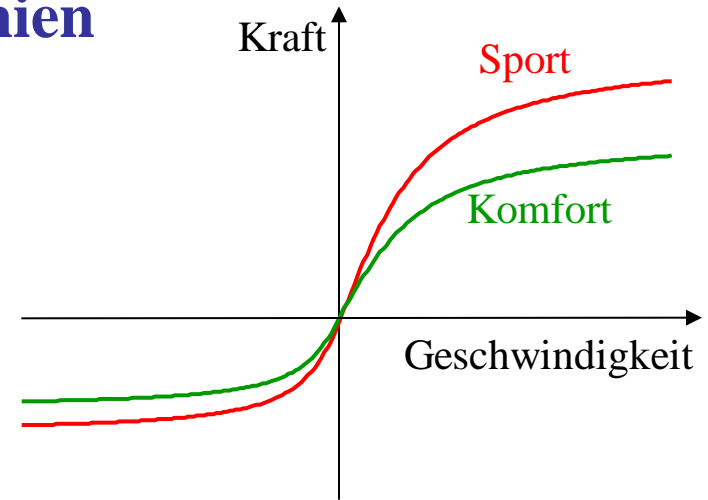
Linear



Progressiv



Weich bei kleinen Unebenheiten



3.2 Nichtlineare Dynamik und Kennlinien

Beispiel: Tank mit Zu- und Abfluss

Druckgleichgewicht:

$$\rho gh(t) = \frac{1}{2} \rho v(t)^2 \quad \rightarrow \quad v(t) = \sqrt{2gh(t)}$$

Geschwindigkeit

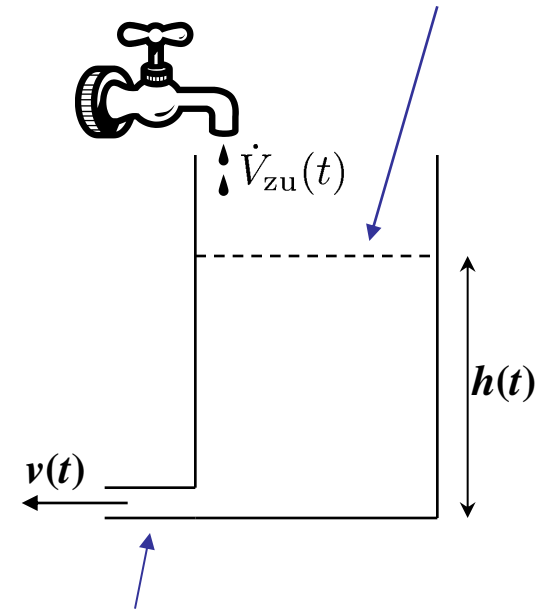
Kontinuität:

$$av(t) + A\dot{h}(t) = \dot{V}_{\text{zu}}(t)$$

$$\rightarrow \dot{h}(t) = \frac{1}{A} \dot{V}_{\text{zu}}(t) - \frac{a}{A} \sqrt{2gh(t)}$$

Höhe $h(t)$ taucht als Ableitung und nichtlinear auf.
→ Nichtlinear *dynamisches* Problem!

Behälterquerschnittsfläche A



Rohrquerschnittsfläche a

3.2 Nichtlineare Dynamik und Kennlinien

Beispiel: Pendel

Energien:

$$E_{\text{kin}}(t) = \frac{1}{2}mv(t)^2 = \frac{1}{2}ml^2\dot{\varphi}(t)^2$$

$$E_{\text{pot}}(t) = mgl(1 - \cos\varphi(t))$$

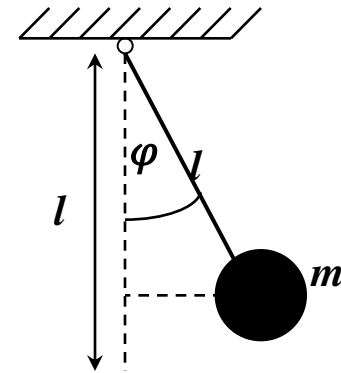
Änderung der Gesamtenergie muss = 0 sein:

$$\frac{d}{dt} (E_{\text{kin}}(t) + E_{\text{pot}}(t)) = ml^2\dot{\varphi}(t)\ddot{\varphi}(t) + mgl\sin\varphi(t)\dot{\varphi}(t) = 0$$

$$\rightarrow \ddot{\varphi}(t) + \frac{g}{l}\sin\varphi(t) = 0$$

Winkel $\varphi(t)$ taucht als 2. Ableitung und nichtlinear auf.

→ Nichtlinear *dynamisches* Problem!



3.2 Nichtlineare Dynamik und Kennlinien

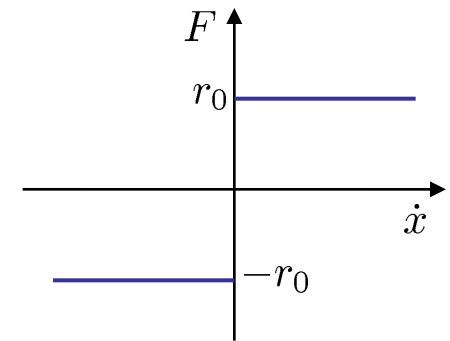
Beispiel: Coulombsche Reibung (trockene Reibung)

Typische Bewegungsgleichung:

$$F(t) = m\ddot{x}(t) + r_1\dot{x}(t) + r_0\text{sign}(\dot{x}(t)) + \dots$$

Viskoser (geschwindigkeits-
proportionaler) Reibterm

Hängt von der
Bewegungsrichtung ab!



Eigenschaften:

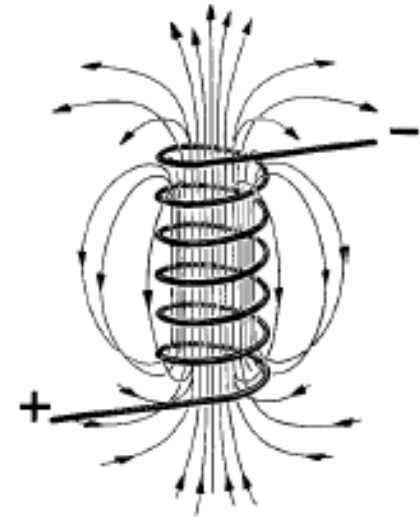
- Coulombsche Reibung wirkt wie eine Konstante (Gleichwert) solange sich die Bewegungsrichtung nicht ändert.
- Im Moment der Änderung der Bewegungsrichtung „springt“ der trockene Reibterm von einem positiven auf einen negativen Wert (bzw. umgekehrt).
- Daher lässt sich dieser Term nicht differenzieren und damit die obige Bewegungsgleichung nicht um alle Arbeitspunkte mit $\dot{x} = 0$ linearisieren!

3.2 Nichtlineare Dynamik und Kennlinien

Beispiel: Magnetische Hysterese

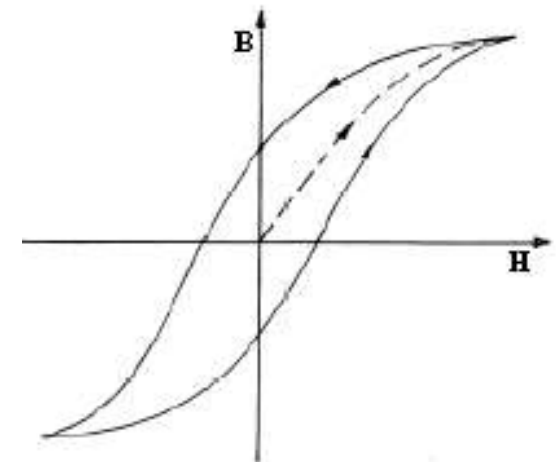
- Magnetische Feldstärke H
(erzeugt vom fließenden Strom und dazu proportional)
- Magnetische Flußdichte B
(Änderung von B induziert eine Spannung)

$$B = \mu \cdot H$$



Eigenschaften der Hysterese:

- μ hängt sowohl von H ab
als auch von der Vorgeschichte von H
- μ ist stark materialabhängig (kann sich um viele
Größenordnungen unterscheiden)
- Keine einfache Linearisierung möglich
- Modellierung von Hysterese generell schwierig



3.2 Nichtlineare Dynamik und Kennlinien

Stabile und instabile Arbeitspunkte

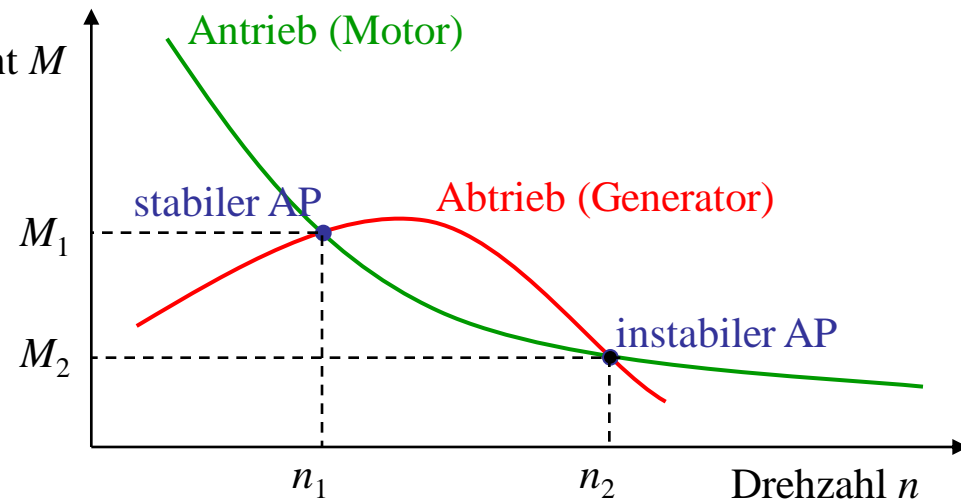
Der Arbeitspunkt (AP) ergibt sich aus dem Schnittpunkt der Kennlinien des Energie abgebenden Teils einer Anlage (Antrieb, Motor, Pumpe, ...) und des Energie aufnehmenden Teils (Abtrieb, Generator, Rohrleitungssystem, ...).

Stabile Arbeitspunkte (APs): Drehmoment M

Nach kleinen Veränderungen der Eingangs- oder Ausgangsgrößen strebt die Anlage wieder in den AP zurück.

Instabile Arbeitspunkte (APs):

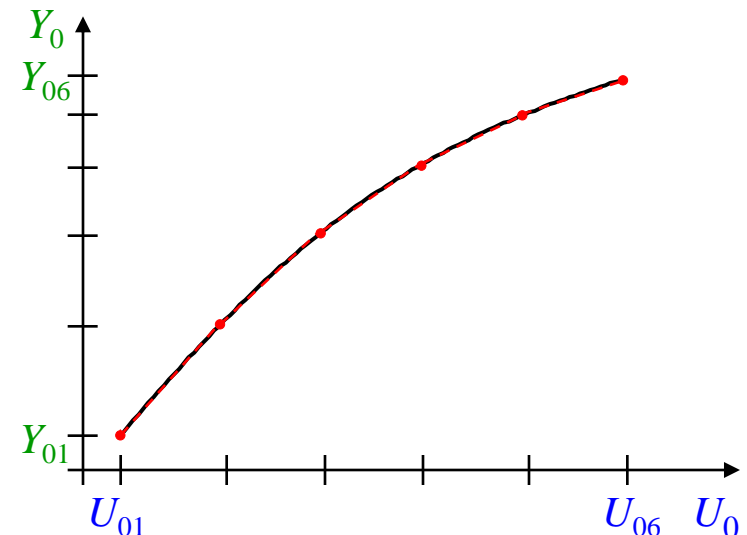
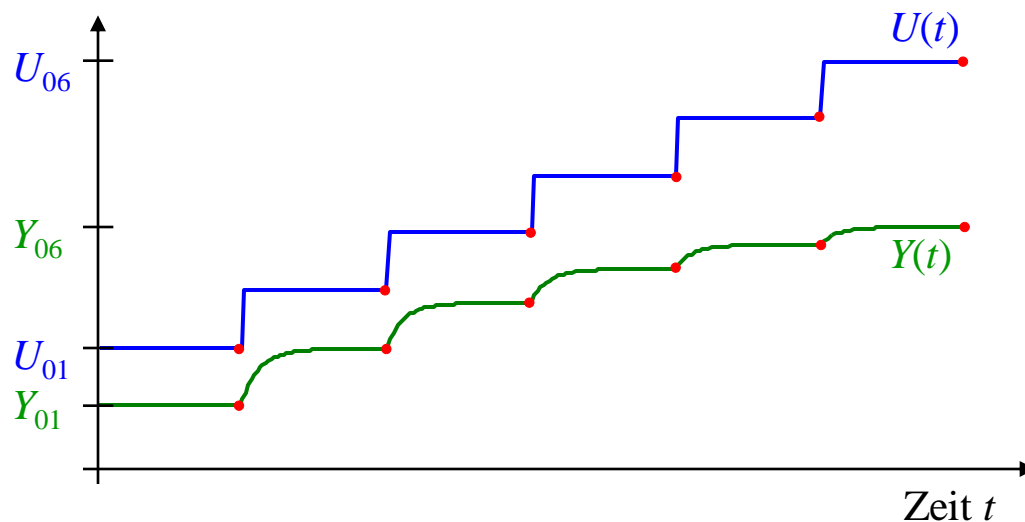
Nach kleinen Veränderungen der Eingangs- oder Ausgangsgrößen strebt die Anlage weiter vom AP weg.



3.2 Nichtlineare Dynamik und Kennlinien

Vermessung von Kennlinien

- Eingangssignal muss solange konstant gehalten werden, bis Ausgangswert eingeschwungen ist. Dann kann ein Punkt der U_0 - Y_0 -Kennlinie abgelesen werden.
→ Zeit zum Erfassen eines Kennlinienpunktes ist hoch!
- Kennlinie wird meist als „Tabelle“ (*lookup table*) mit linearer Interpolation (rot gestrichelter Verlauf) implementiert. Alternative: Polynome, neuronale Netze, ...
- Kennlinien für mehr als 1 Eingang heißen Kennfelder. Sie werden oft **rasterförmig** vermessen, d.h. z.B. 6×6 Eingangsgrößenkombination durchfahren (im 2-dim. Fall).



3.3 Linearisierung nichtlinearer statischer Systeme

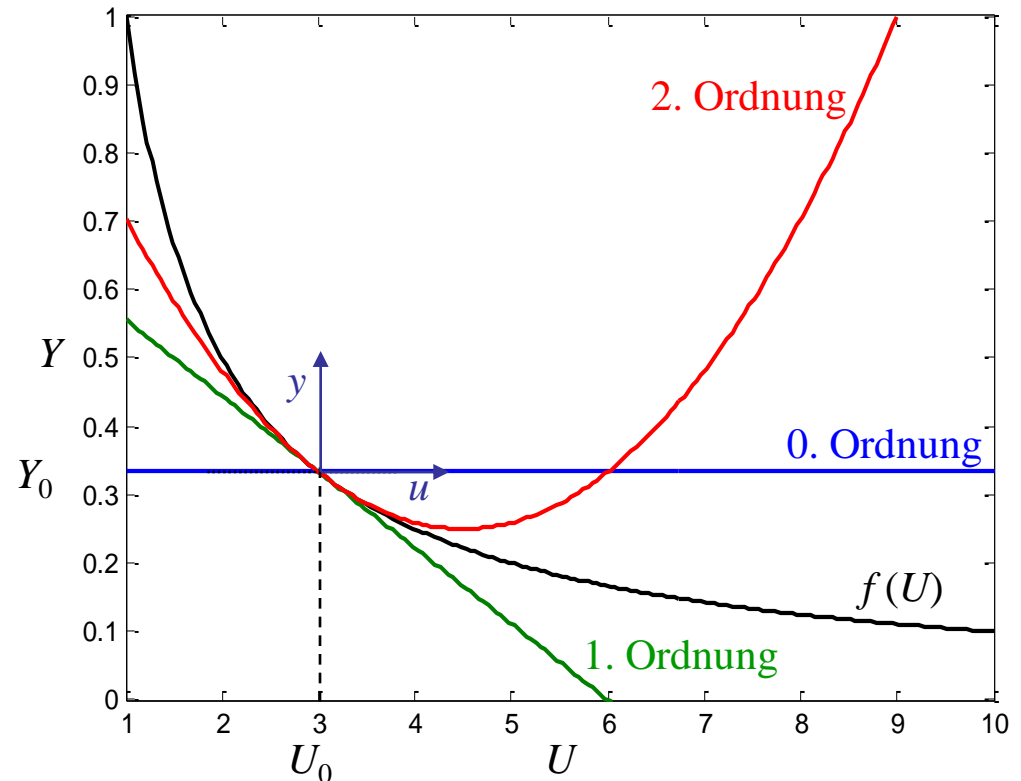
Taylor-Reihenentwicklung um den Punkt (U_0, Y_0) :

$$Y = \underbrace{Y_0}_{0. \text{ Ordn.}} + \underbrace{f'(U_0)(U - U_0)}_{1. \text{ Ordnung}} + \underbrace{\frac{1}{2} f''(U_0)(U - U_0)^2}_{2. \text{ Ordnung}} + \dots$$

In Kleinsignalfom:

$$y = f'(U_0)u + \frac{1}{2} f''(U_0)u^2 + \dots$$

Approximation 1. Ordnung ist eine *Linearisierung* um den Arbeitspunkt (U_0, Y_0) . Ein anderer Arbeitspunkt führt normalerweise zu einer anderen linearisierten Gleichung, weil sich die Steigung von $f(U)$ ändert.



3.3 Linearisierung nichtlinearer statischer Systeme

Erweiterung auf Systeme mit p Eingangsgrößen $\underline{U} = [U_1 \ U_2 \ \dots \ U_p]^T$

Taylor-Reihenentwicklung um den Punkt (\underline{U}_0, Y_0) :

$$Y = \underbrace{Y_0}_{0. \text{ Ordn.}} + \underbrace{\underline{g}^T|_{\text{AP}} (\underline{U} - \underline{U}_0)}_{1. \text{ Ordnung}} + \underbrace{\frac{1}{2} (\underline{U} - \underline{U}_0)^T \underline{H}|_{\text{AP}} (\underline{U} - \underline{U}_0)}_{2. \text{ Ordnung}} + \dots$$

mit der 1. Ableitung (Gradient) ausgewertet am Arbeitspunkt (AP) (\underline{U}_0, Y_0) :

$$\underline{g}^T|_{\text{AP}} = \left[\frac{\partial f}{\partial U_1} \ \frac{\partial f}{\partial U_2} \ \dots \ \frac{\partial f}{\partial U_p} \right] \Big|_{\text{AP}}$$

und der 2. Ableitung ($p \times p$ Hesse-Matrix) ausgewertet am Arbeitspunkt (AP) (\underline{U}_0, Y_0) :

$$\underline{H}|_{\text{AP}} = \left[\begin{array}{cccc} \frac{\partial^2 f}{\partial U_1^2} & \frac{\partial^2 f}{\partial U_1 \partial U_2} & \dots & \frac{\partial^2 f}{\partial U_1 \partial U_p} \\ \frac{\partial^2 f}{\partial U_2 \partial U_1} & \frac{\partial^2 f}{\partial U_2^2} & \dots & \frac{\partial^2 f}{\partial U_2 \partial U_p} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial U_p \partial U_1} & \frac{\partial^2 f}{\partial U_p \partial U_2} & \dots & \frac{\partial^2 f}{\partial U_p^2} \end{array} \right] \Big|_{\text{AP}}$$

3.3 Linearisierung nichtlinearer statischer Systeme

Zur Linearisierung mit p Eingängen:

- Nur die ersten beiden Terme (0. und 1. Ordnung) der Taylorreihe werden benötigt.
- In Kleinsignalform schreibt sich die linearisierte Gleichung:

$$y = g^T \Big|_{AP} \underline{u} = \frac{\partial f}{\partial U_1} \Big|_{AP} \cdot u_1 + \frac{\partial f}{\partial U_2} \Big|_{AP} \cdot u_2 + \dots + \frac{\partial f}{\partial U_p} \Big|_{AP} \cdot u_p$$

- Diese Gleichung beschreibt keine Gerade (wie im 1-dim. Fall) sondern eine (Hyper)-Ebene, also die p -dimensionale Verallgemeinerung einer Geraden.
- Gradient enthält Ableitungen des Ausgangs nach allen Eingängen U_1, U_2, \dots, U_p .

Nebenbemerkung: Reihenentwicklung bis 2. Ordnung:

- Für uns im Moment uninteressant.
- Basis für sehr viele anderen Anwendungen, insbesondere für Näherungs- und Optimierungsverfahren. Approximation 2. Ordnung wird im Newton-Verfahren zur Nullstellen- und Optimumssuche verwendet. Auch in der Least-Squares-Methode.
- Die Gleichungen 2. Ordnung stellen eine (Hyper)-Parabel dar.

3.4 Linearisierung nichtlinearer dynamischer Systeme

Linearisierung um einen stationären Arbeitspunkt (AP):

- „Stationär“ heißt, dass die Ableitungen aller Ein- und Ausgangsgrößen = 0 sind:

$$U|_{AP} = U_0, \quad U^{(1)}|_{AP} = 0, \quad \dots, \quad U^{(m)}|_{AP} = 0$$

$$Y|_{AP} = Y_0, \quad Y^{(1)}|_{AP} = 0, \quad \dots, \quad Y^{(n)}|_{AP} = 0$$

- Standardform der nichtlinearen DGL beschrieben durch:

$$f(Y, Y^{(1)}, \dots, Y^{(n)}, U, U^{(1)}, \dots, U^{(m)}) = 0$$

- Linearisierung um einen stationären AP liefert:

$$\begin{aligned} \frac{\partial f}{\partial Y} \Big|_{AP} (Y - Y_0) + \frac{\partial f}{\partial Y^{(1)}} \Big|_{AP} (Y^{(1)} - Y_0^{(1)}) + \dots + \frac{\partial f}{\partial Y^{(n)}} \Big|_{AP} (Y^{(n)} - Y_0^{(n)}) + \\ \frac{\partial f}{\partial U} \Big|_{AP} (U - U_0) + \frac{\partial f}{\partial U^{(1)}} \Big|_{AP} (U^{(1)} - U_0^{(1)}) + \dots + \frac{\partial f}{\partial U^{(m)}} \Big|_{AP} (U^{(m)} - U_0^{(m)}) = 0 \end{aligned}$$

3.4 Linearisierung nichtlinearer dynamischer Systeme

- Wegen Stationarität des AP: $Y_0^{(i)} = 0$ ($i = 1, \dots, n$) und $U_0^{(i)} = 0$ ($i = 1, \dots, m$)
Daher gilt auch:

$$\frac{\partial f}{\partial Y} \Big|_{AP} (Y - Y_0) + \frac{\partial f}{\partial Y^{(1)}} \Big|_{AP} Y_0^{(1)} + \dots + \frac{\partial f}{\partial Y^{(n)}} \Big|_{AP} Y^{(n)} +$$
$$\frac{\partial f}{\partial U} \Big|_{AP} (U - U_0) + \frac{\partial f}{\partial U^{(1)}} \Big|_{AP} U^{(1)} + \dots + \frac{\partial f}{\partial U^{(m)}} \Big|_{AP} U^{(m)} = 0$$

Oder in Kleinsignalvariablen:

$$\frac{\partial f}{\partial Y} \Big|_{AP} y + \frac{\partial f}{\partial Y^{(1)}} \Big|_{AP} y^{(1)} + \dots + \frac{\partial f}{\partial Y^{(n)}} \Big|_{AP} y^{(n)} +$$
$$\frac{\partial f}{\partial U} \Big|_{AP} u + \frac{\partial f}{\partial U^{(1)}} \Big|_{AP} u^{(1)} + \dots + \frac{\partial f}{\partial U^{(m)}} \Big|_{AP} u^{(m)} = 0$$

3.4 Linearisierung nichtlinearer dynamischer Systeme

Beispiel: Affiner Zusammenhang (wiederholt)

Mit dieser schematischen Vorgehensweise soll zunächst nochmal die affine Kennlinie vom Anfang des Kapitels behandelt werden. Ein statischer Prozess folge dem Gesetz:

$$Y = mU + b \quad \rightarrow \quad f(Y, U) = Y - mU - b = 0$$

Stationärer Arbeitspunkt AP ist:

$$Y_0 = mU_0 + b$$

Da in $f()$ keine *zeitlichen* Ableitungen wie \dot{U} oder \dot{Y} auftauchen, liefert die Linearisierung um diesen AP einfach:

$$\left. \frac{\partial f}{\partial Y} \right|_{AP} (Y - Y_0) + \left. \frac{\partial f}{\partial U} \right|_{AP} (U - U_0) = 0$$

Die Ableitung nach Y und U sind: $\left. \frac{\partial f}{\partial Y} \right|_{AP} = 1 \quad \left. \frac{\partial f}{\partial U} \right|_{AP} = -m$

Im Großsignalverhalten reproduziert sich die Originalgleichung, weil es in diesem Beispiel keine Nichtlinearitäten gibt. Als Kleinsignalverhalten ergibt sich immer lineares Verhalten:

$$\left. \frac{\partial f}{\partial Y} \right|_{AP} y + \left. \frac{\partial f}{\partial U} \right|_{AP} u = 0 \quad \rightarrow \quad 1 \cdot y - m \cdot u = 0 \quad \rightarrow \quad y = mu$$

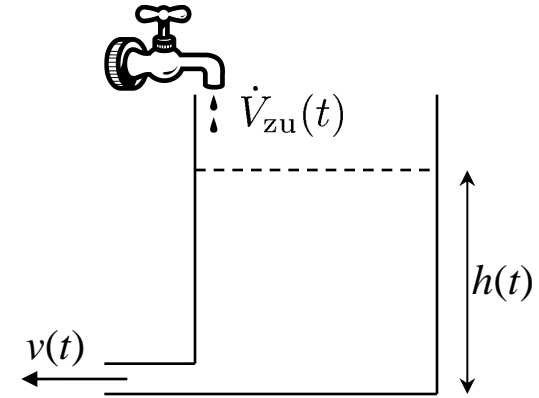
3.4 Linearisierung nichtlinearer dynamischer Systeme

Beispiel: Tank mit Zu- und Abfluss

$$\dot{h}(t) = \frac{1}{A} \dot{V}_{\text{zu}}(t) - \frac{a}{A} \sqrt{2gh(t)}$$

Eingang: Zufluss $U(t) = \dot{V}_{\text{zu}}(t)$

Ausgang: Füllhöhe $Y(t) = h(t)$



Umschreiben in nichtlineare Standardform: $\dot{Y}(t) + \frac{a}{A} \sqrt{2gY(t)} - \frac{1}{A} U(t) = 0$

Für einen stationären Arbeitspunkt gilt: $\frac{a}{A} \sqrt{2gY_0} = \frac{1}{A} U_0 \rightarrow \boxed{Y_0 = \frac{1}{2ga^2} U_0^2}$

Linearisieren:

$$\left. \frac{\partial f}{\partial Y} \right|_{\text{AP}} = \left. \frac{a}{A} \frac{2g}{2\sqrt{2gY(t)}} \right|_{\text{AP}} = \frac{ga}{A\sqrt{2gY_0}}$$

$$\left. \frac{\partial f}{\partial \dot{Y}} \right|_{\text{AP}} = 1 \quad \left. \frac{\partial f}{\partial U} \right|_{\text{AP}} = -\frac{1}{A}$$

Lineare DGL:

$$\boxed{\dot{y}(t) + \frac{ga}{A\sqrt{2gY_0}} y(t) = \frac{1}{A} u(t)}$$

oder

$$\boxed{\dot{y}(t) + \frac{ga^2}{AU_0} y(t) = \frac{1}{A} u(t)}$$

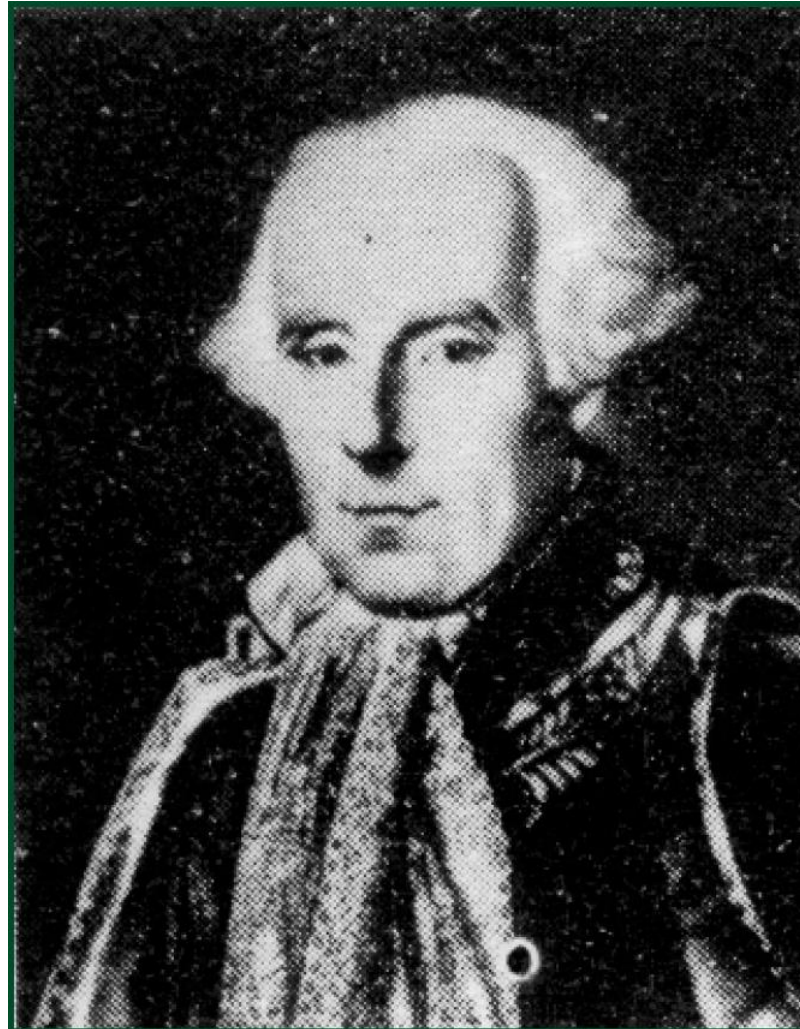
4. Laplace-Transformation

Inhalt Kapitel 4

4 Laplace-Transformation

- 4.1 Motivation
- 4.2 Definition
- 4.3 Eigenschaften
- 4.4 Rücktransformation in den Zeitbereich
- 4.5 Lösen einer DGL im Laplace-Bereich
- 4.6 Rechenregeln
- 4.7 Tabelle Laplace-Transformation

Pierre Simon Laplace (1749-1827)



4. Laplace-Transformation

4.1 Motivation

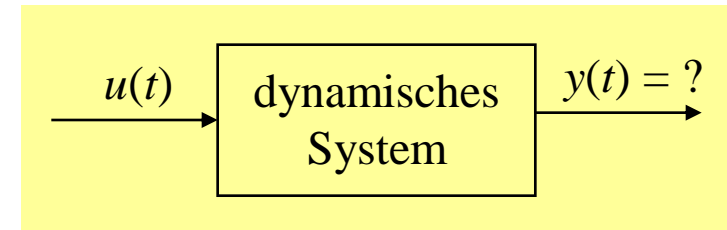
Wichtige Aufgabenstellung beim Rechnen mit dynamischen Systemen:

Gegeben:

- Lineare (linearisierte) DGL, die das dynamische System beschreibt.
- Zeitlicher Verlauf der Eingangsgröße $u(t)$ des Systems.

Gesucht:

- Zeitlicher Verlauf der Ausgangsgröße $y(t)$.



Aus Kapitel 2.6 kennen wir die partikuläre Lösung

für das einfachste dynamische System, beschrieben durch eine DGL 1. Ordnung:

$$y(t) = b_0 \int_0^t e^{-a_0(t-\tau)} u(\tau) d\tau$$

ACHTUNG: Dies ist nur der partikuläre Lösungsanteil $y_{erzw}(t)$, der durch das Eingangssignal $u(t)$ hervorgerufen wird! *Wir nehmen hier und in der Folge an (wenn nicht ausdrücklich anders gesagt), dass die Anfangsbedingungen so gewählt sind, dass die homogenen Lösungsanteile wegfallen. Uns interessiert primär das Ein-/Ausgangsverhalten.*

4.1 Motivation

Analysieren wir die Antwort $y(t)$ auf das Eingangssignal $u(t)$ genauer:

$$y(t) = b_0 \int_0^t e^{-a_0(t-\tau)} u(\tau) d\tau$$

Systemeigenschaften
(signalunabhängig)

- Diese Vorschrift zur Berechnung des Ausgangssignals bei gegebenem Eingangssignal heißt **Faltung**. Man schreibt sie auch symbolisch:

$$y = g * u$$

(Für den Fall einer DGL 1. Ordnung gilt $g(t) = b_0 e^{-a_0 t}$.)

- Die Funktion $g(t)$ nennt man **Gewichtsfunktion** (dazu später mehr...). Sie enthält alle Eigenschaften über das System. Generell gilt für **alle linearen Systeme** (auch mit höherer Ordnung):

$$y(t) = \int_0^t g(t - \tau) u(\tau) d\tau$$

4.1 Motivation

Identische Formulierungen des Faltungsintegrals:

$$\int_0^t g(t-\tau)u(\tau)d\tau \quad \int_0^t g(\tau)u(t-\tau)d\tau \quad \int_{-\infty}^{\infty} g(t-\tau)u(\tau)d\tau \quad \int_{-\infty}^{\infty} g(\tau)u(t-\tau)d\tau$$

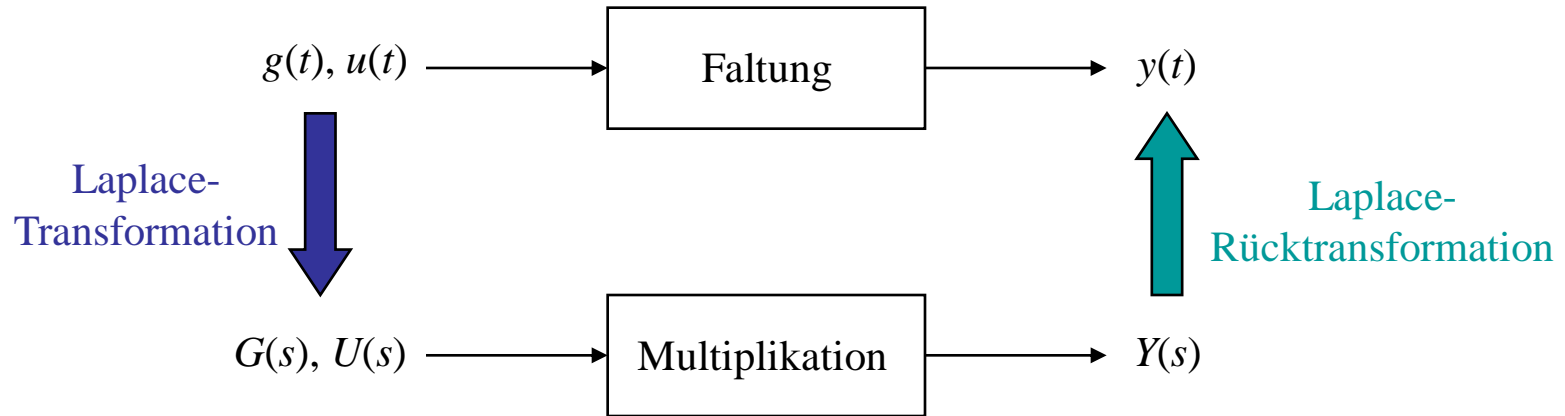
- Die erste Äquivalenz lässt sich leicht mit einer Substitution $\tau' = t - \tau$ zeigen.
- Aus den Annahmen $u(t) = 0$ für $t < 0$ und $g(t) = 0$ für $t < 0$ ergibt sich die Äquivalenz der Integralgrenzen.
 - Die Annahme $u(t) = 0$ für $t < 0$ ist für die Laplace-Transformation nötig. Sie stellt aber keine praktische Einschränkung dar, weil wir die Zeit immer so verschieben können, dass vor $t = 0$ noch nichts passiert ist.
 - Die Annahme $g(t) = 0$ für $t < 0$ entspricht der Annahme, dass das System kausal ist. Davon gehen wir immer aus. Dazu später mehr...

Die Berechnung des Faltungsintegrals ist aufwendig! Die Zeitfunktionen $g(t)$ und $u(t)$ können sehr kompliziert werden und das Faltungsintegral schwierig zu lösen sein.

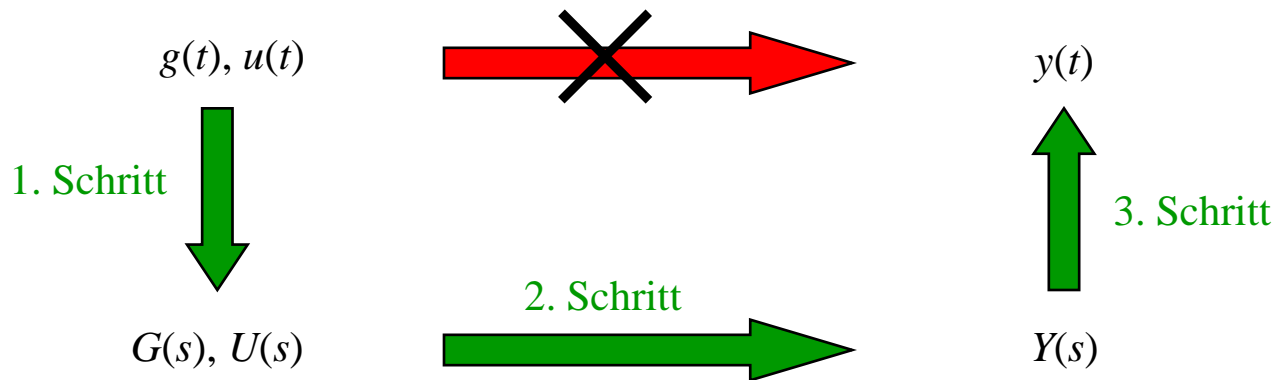
→ Suche nach Ansätzen zur Vermeidung der direkten Berechnung des Faltungsintegrals!

4.1 Motivation

Mit Hilfe der Laplace-Transformation wird aus der Faltung eine Multiplikation:



Deshalb machen wir in Zukunft lieber 3 einfache Schritte als 1 komplizierten:



4.2 Definition

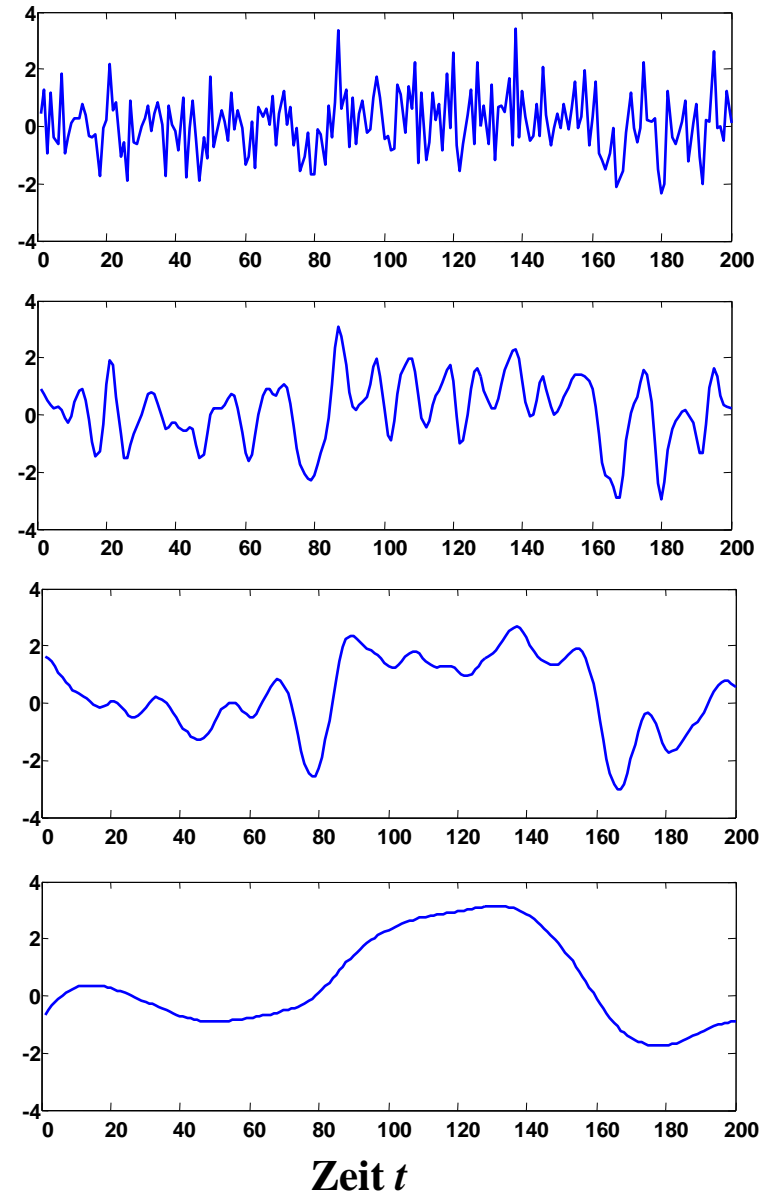
Signale enthalten viele verschiedene Frequenzanteile. Eine Transformation vom Zeitbereich in den Frequenzbereich erlaubt quantitative Aussagen darüber, welche Frequenzen bzw. Frequenzbereiche wie stark im Signal vertreten sind.

Forderungen/Annahmen für eine Transformation in den Frequenzbereich:

- Alle praktisch relevanten Signale sollen transformierbar sein.
- Signale sind Null für $t < 0$.

Dies wird von der Laplace-Transformation erfüllt!

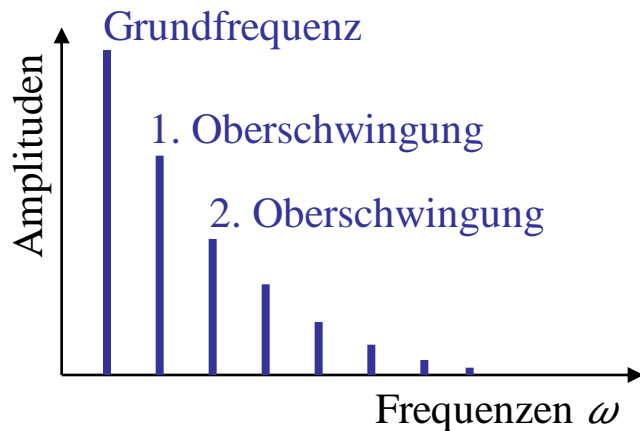
Anteil hoher Frequenzen im Signal



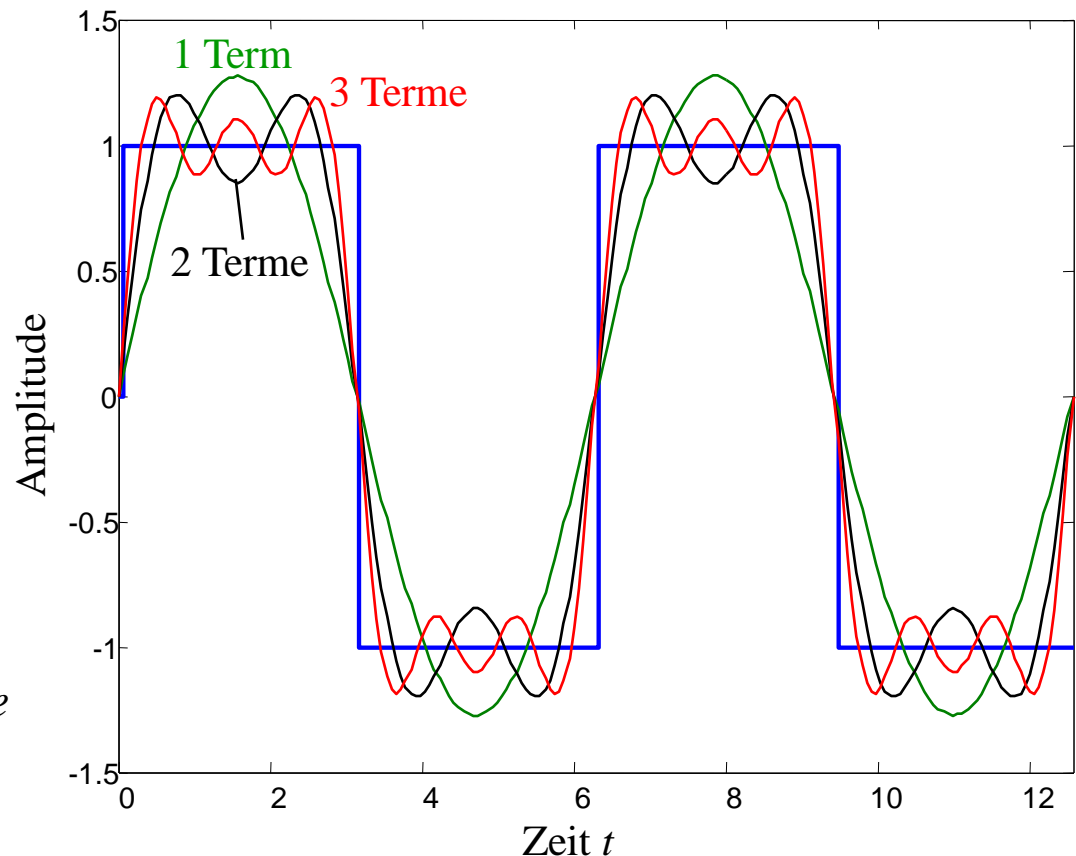
4.2 Definition

Fourier-Reihe

- Zerlegung *periodischer* Signale in ihre Frequenzbestandteile.
- Signal kann in eine unendliche Summe von Sinus- und Cosinus-Termen zerlegt werden.
- Amplitude zu jeder Frequenz gibt an, wie stark diese Frequenz im Signal enthalten ist.



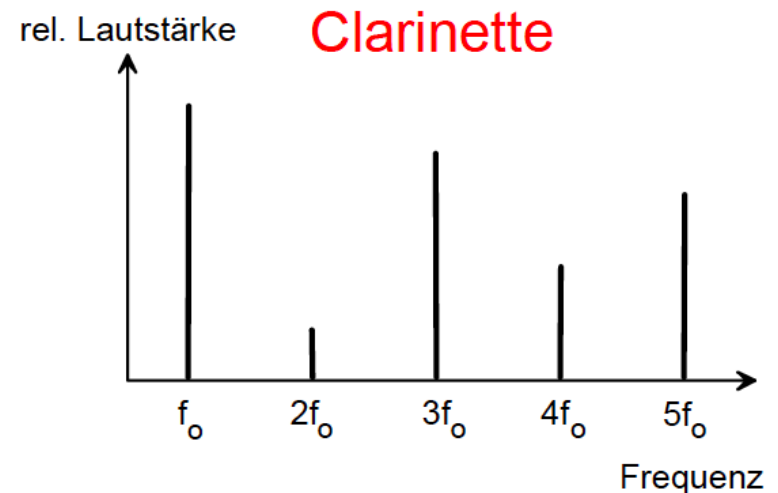
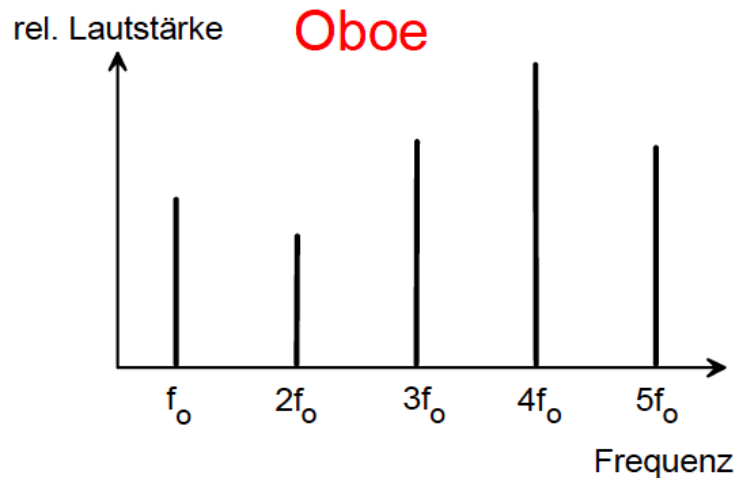
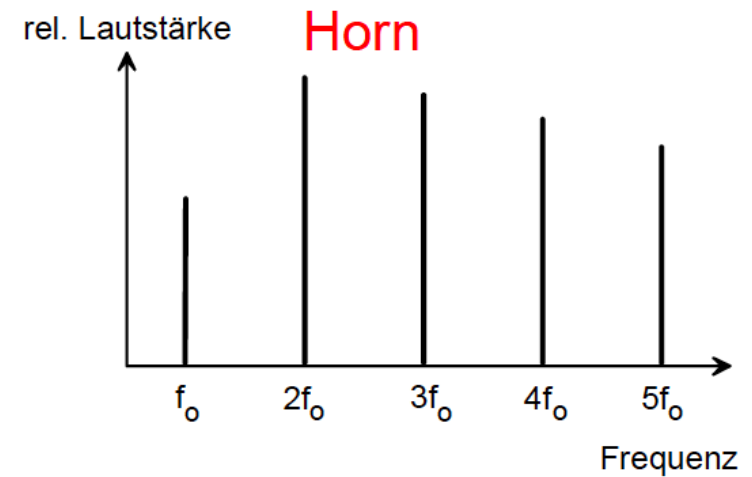
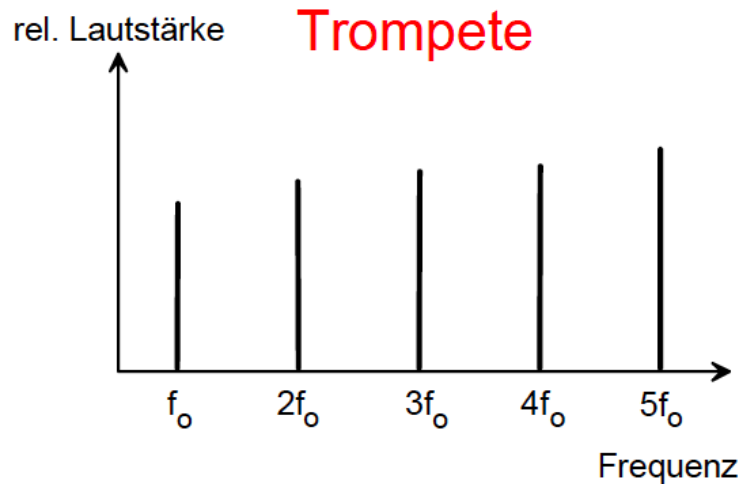
- Wenn man auch *nicht-periodische* Signale behandeln will: Periodenlänge $\rightarrow \infty$, Grundfrequenz $\rightarrow 0$.



4.2 Definition

Quelle: http://eitidaten.fh-pforzheim.de/daten/mitarbeiter/blankenbach/vorlesungen/mathe_2/Fourier_Trafo_kurz_Folien.pdf

Kammerton a : $f_0 = 440$ Hz auf verschiedenen Musikinstrumenten

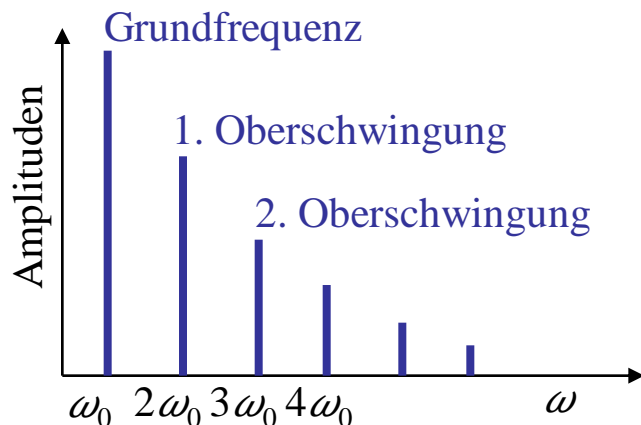


4.2 Definition

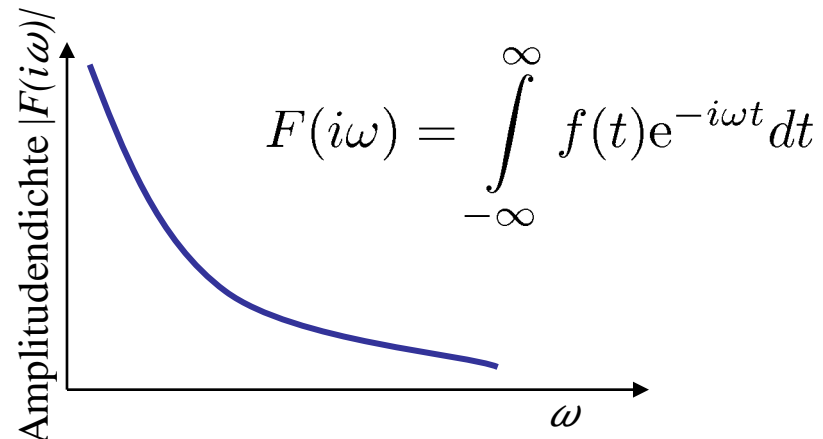
Fourier-Transformation

- Erweiterung der Fourier-Reihe für nicht-periodische Signale.
- Periodenlänge $T \rightarrow \infty$, Grundfrequenz $\omega \rightarrow 0$.
- Das Spektrum besteht nicht mehr aus absoluten *Amplitudenwerten* an *diskreten* Frequenzen $n \cdot \omega_0$ (also Vielfachen der Grundfrequenz), sondern es besteht aus der *Amplitudendichte* über einer *kontinuierlichen* Frequenzachse. (Gleiches gilt für die Phasen.)

Fourier-Reihe

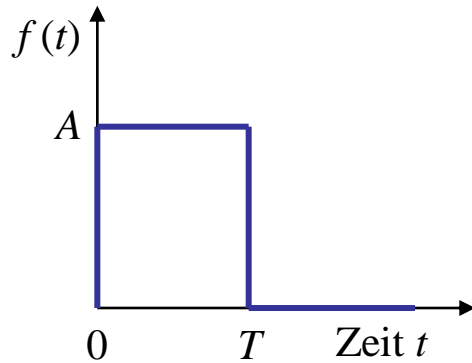


Fourier-Transformation

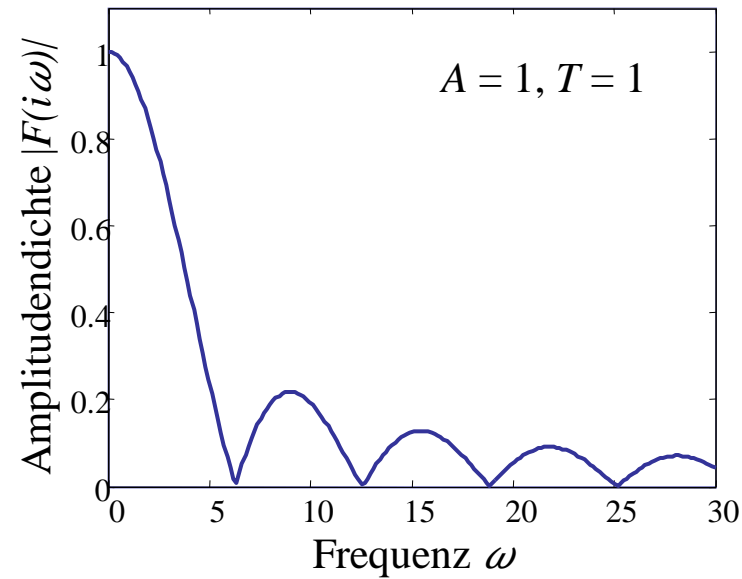


4.2 Definition

Beispiel: Fourier-Transformation



$$f(t) = \begin{cases} 0 & \text{für } t < 0 \\ A & \text{für } 0 \leq t < T \\ 0 & \text{für } t \geq T \end{cases}$$



$$\begin{aligned} F(i\omega) &= \int_{-\infty}^{\infty} f(t)e^{-i\omega t} dt = \int_0^T Ae^{-i\omega t} dt \\ &= -\frac{A}{i\omega} e^{-i\omega t} \Big|_0^T = \frac{A}{i\omega} (1 - e^{-i\omega T}) \end{aligned}$$

$$\begin{aligned} |F(i\omega)| &= \frac{A}{\omega} \sqrt{(1 - \cos\omega T)^2 + \sin^2\omega T} \\ &= \frac{A\sqrt{2}}{\omega} \sqrt{1 - \cos\omega T} \end{aligned}$$

4.2 Definition

Problematik der Fourier-Transformation:

Viele Signale können nicht Fourier transformiert werden, weil das Integral nicht konvergiert.

Beispiel:

Info: Sonderfall $a = 0$ entspricht Einheitssprung:

$$f(t) = \begin{cases} 0 & \text{für } t < 0 \\ e^{at} & \text{für } t \geq 0 \end{cases}$$

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

$$F(i\omega) = \int_{-\infty}^{\infty} f(t)e^{-i\omega t} dt = \int_0^{\infty} e^{at} e^{-i\omega t} dt = \int_0^{\infty} e^{(a-i\omega)t} dt$$

→ Für positive a (selbst für $a = 0$) geht das Integral gegen unendlich!

Ausweg: Wir dämpfen den Integranden mit einem zusätzlich eingeführten Faktor $e^{-\delta t}$.

Dann konvergiert obiges Integral falls wir δ groß genug wählen, also für $\delta > a$:

$$F(i\omega) = \int_0^{\infty} e^{(a-\delta-i\omega)t} dt = \frac{1}{a-\delta-i\omega} e^{(a-\delta-i\omega)t} \Big|_0^{\infty} = \frac{1}{\delta+i\omega-a}$$

Diese neue komplexe Variable nennen wir die **Laplace-Variable** $s = \delta + i\omega$.

4.2 Definition

Da der Term $e^{-\delta t}$ für $t < 0$ aufklingt (nicht dämpft) und damit kontraproduktiv wirkt, starten wir bei der Laplace-Transformation immer erst bei $t = 0$. Dies ist hier aber keine Einschränkung, da die Signale sowieso erst ab $t = 0$ von Null verschieden sind.

Laplace-Transformation einer Zeitfunktion $f(t)$:

$$F(s) = \int_{\underbrace{0}_{-0}}^{\infty} f(t)e^{-st} dt$$

d.h. Zeitpunkt $t = 0$ ist im Integral enthalten!

Laplace-Rücktransformation einer Funktion im komplexen Frequenzbereich $F(s)$:

$$f(t) = \frac{1}{2\pi i} \int_{\delta-i\infty}^{\delta+i\infty} F(s)e^{st} ds$$

$F(s)$ ist die Laplace-Transformierte von $f(t)$. Man schreibt symbolisch:

$$F(s) \bullet \dashrightarrow f(t) \quad \text{oder} \quad F(s) = \mathcal{L}\{f(t)\}$$

$$f(t) \circ \dashrightarrow F(s) \quad \text{oder} \quad f(t) = \mathcal{L}^{-1}\{F(s)\}$$

4.2 Definition

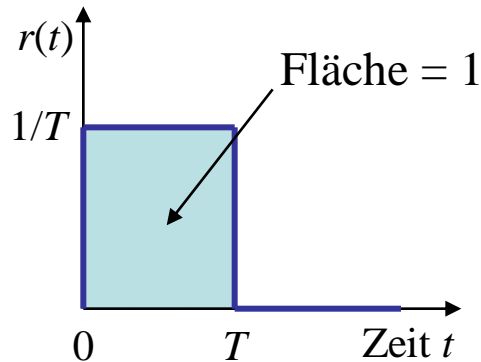
Damit kennen wir schon unsere erste Laplace-Transformierte:

$$f(t) = \begin{cases} 0 & \text{für } t < 0 \\ e^{at} & \text{für } t \geq 0 \end{cases} \quad \circ \bullet \quad F(s) = \frac{1}{s - a}$$

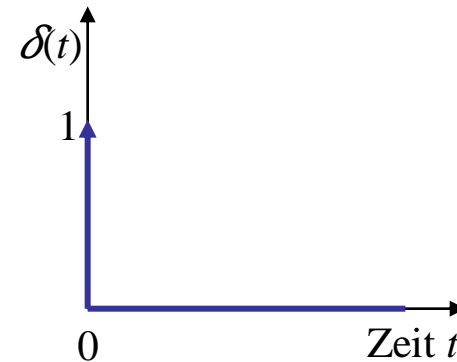
Und als Sonderfall für $a = 0$ den Einheitssprung:

$$f(t) = \sigma(t) = \begin{cases} 0 & \text{für } t < 0 \\ 1 & \text{für } t \geq 0 \end{cases} \quad \circ \bullet \quad F(s) = \frac{1}{s}$$

Eine wichtige weitere Funktion ist der *Dirac-Impuls*. Genau genommen ist das gar keine Funktion, denn der Dirac-Impuls ist unendlich hoch und unendlich kurz. Er ist so definiert:



$$\delta(t) = \lim_{T \rightarrow 0} r(t)$$



ACHTUNG: Das $\delta(t)$ als Bezeichnung des Dirac-Impulses steht in keinem Zusammenhang mit dem δ als Bezeichnung des Realteils der Laplace-Variablen s .

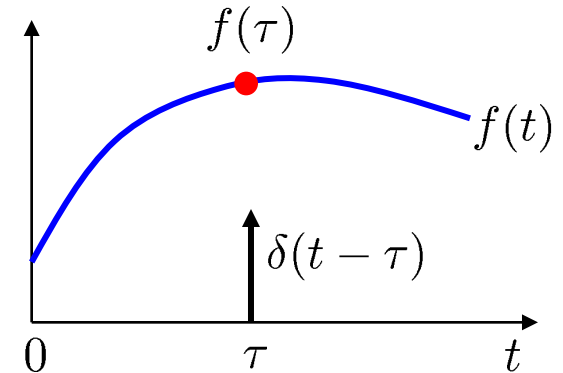
4.2 Definition

Auch wenn der Dirac-Impuls nicht realisierbar ist, stellt er ein wichtiges mathematisches Hilfsmittel dar. Er hat folgende *Ausblendeigenschaft*:

$$\int_{-0}^{\infty} f(t)\delta(t - \tau)dt = f(\tau) \cdot 1 = f(\tau)$$

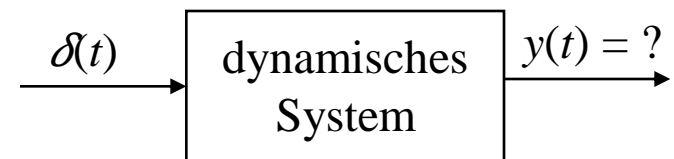
Daraus folgt die Laplace-Transformierte des Dirac-Impulses:

$$\mathcal{L}\{\delta(t)\} = \int_{-0}^{\infty} \delta(t)e^{-st}dt = e^{-s \cdot 0} = 1 \quad \leftarrow \text{gilt für alle Frequenzen } \omega \text{ in } s = \delta + i\omega$$



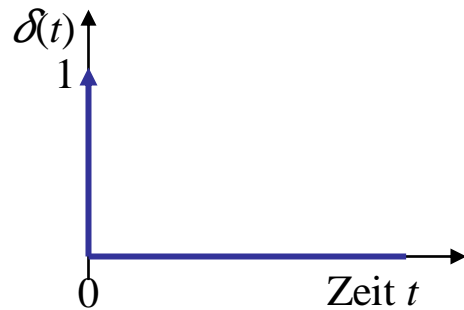
Das bedeutet, alle Frequenzen sind im Spektrum gleich stark vertreten. Damit eignet sich der Dirac-Impuls (theoretisch) gut, um die Eigenschaften eines dynamischen Systems zu ermitteln, denn alle Frequenzen werden gleich stark angeregt. Jede Über- oder Untergewichtung bestimmter Frequenzanteile im Ausgangssignal sind dann auf das Systemverhalten zurückzuführen.

Dazu später mehr...

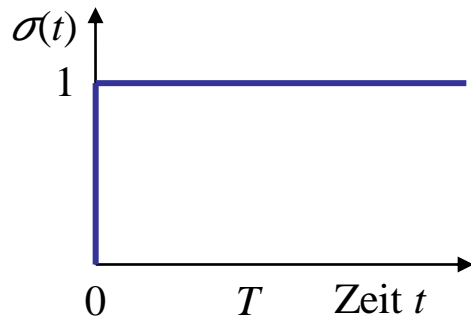
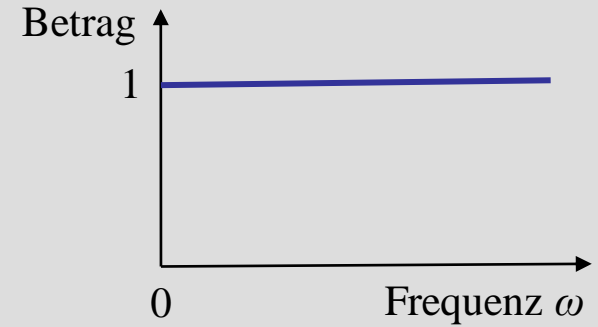


4.2 Definition

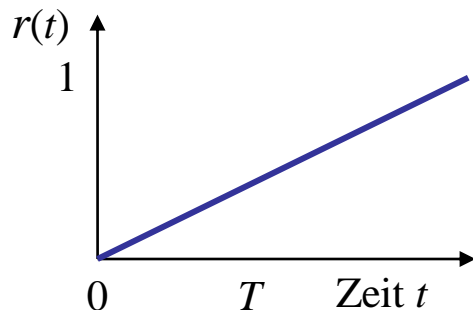
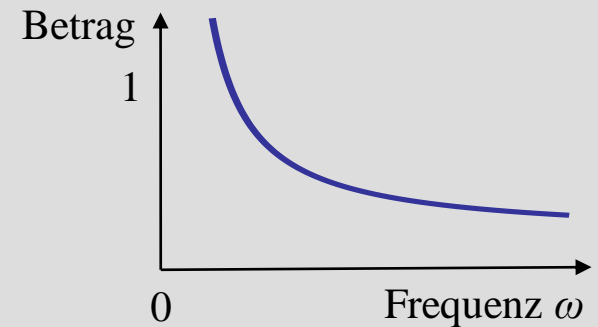
Vergleich der Beträge der Fourier/Laplace-Transformationen der Signale:



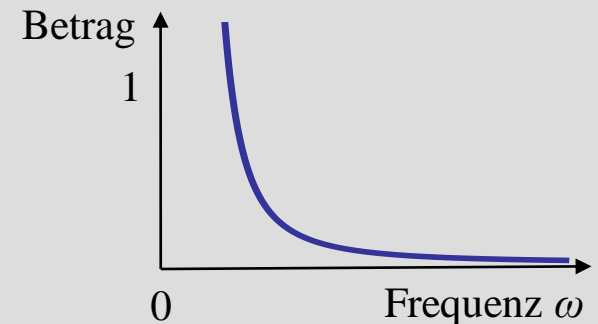
$$\mathcal{L}\{\delta(t)\} = 1$$



$$\mathcal{L}\{\sigma(t)\} = \frac{1}{s}$$

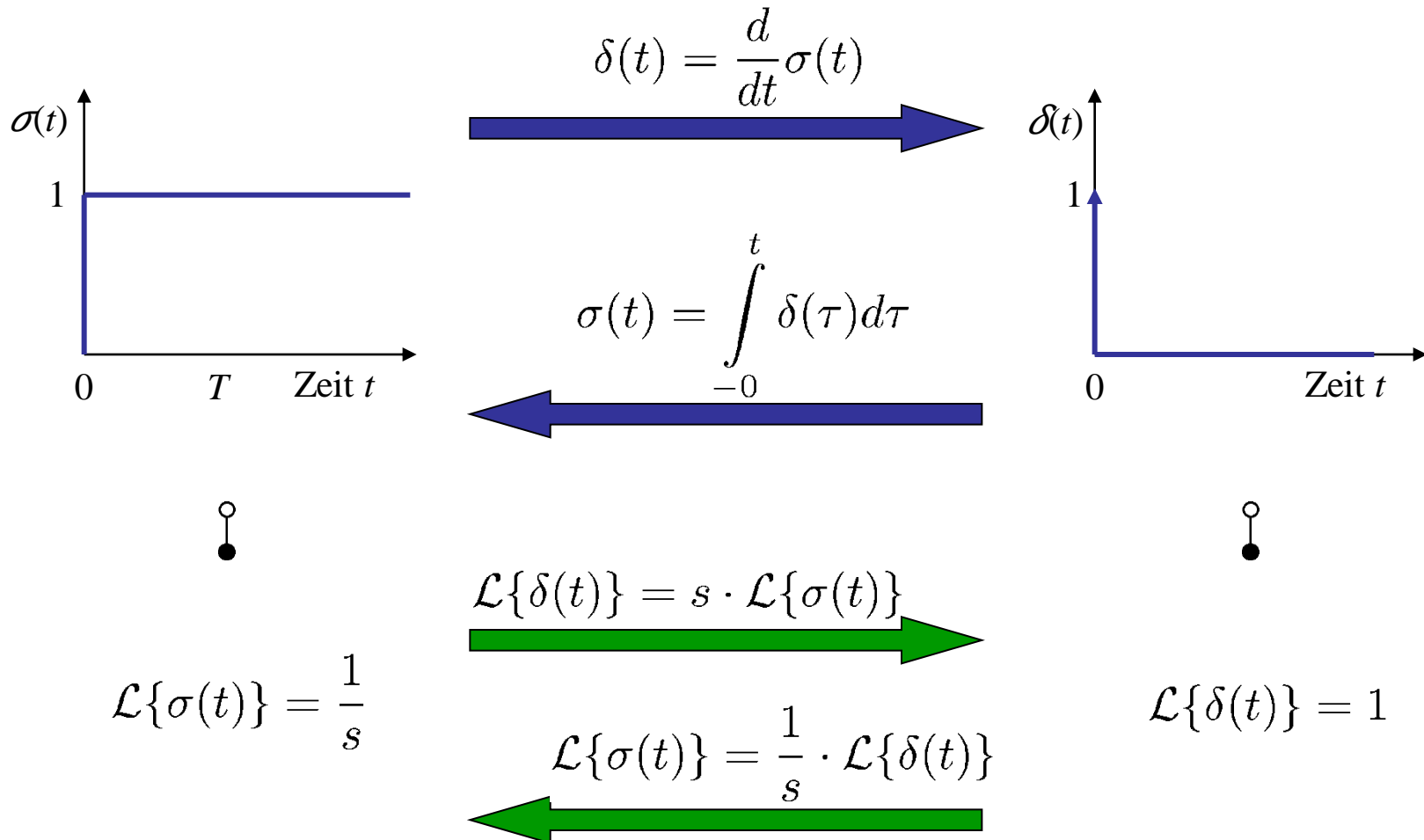


$$\mathcal{L}\{r(t)\} = \frac{1}{s^2}$$



4.2 Definition

Neben dem Einheitssprung $\sigma(t)$ ist der Dirac-Impuls $\delta(t)$ die am häufigsten betrachtete Eingangsgröße für dynamische Systeme. Beide stehen in einer engen Beziehung:



4.3 Eigenschaften

Faltung

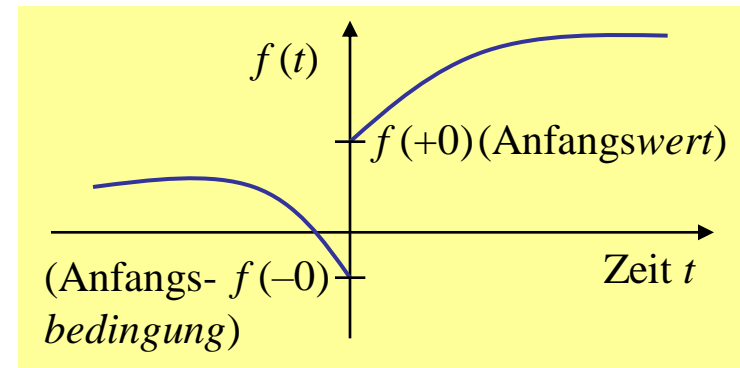
Der Hauptgrund für das Rechnen mit der Laplace-Transformation ist folgende Beziehung:

$$f_1 * f_2 = \int_0^t f_1(t - \tau) f_2(\tau) d\tau \quad \circ \bullet \quad F_1(s) \cdot F_2(s)$$

Anfangswert

Falls die Grenzwerte existieren, lässt sich schnell der Anfangswert des Zeitverlaufs bestimmen:

$$f(+0) = \lim_{t \rightarrow +0} f(t) = \lim_{s \rightarrow \infty} sF(s)$$



Endwert

Falls die Grenzwerte existieren, lässt sich schnell der Endwert des Zeitverlaufs bestimmen:

$$f(\infty) = \lim_{t \rightarrow \infty} f(t) = \lim_{s \rightarrow 0} sF(s)$$

Die Zeit t und die komplexe Frequenz s verhalten sich invers zueinander!

4.3 Eigenschaften

Überlagerung

Die Laplace-Transformation ist eine lineare Transformation:

$$k_1 f_1(t) + k_2 f_2(t) \quad \circ \bullet \quad k_1 F_1(s) + k_2 F_2(s)$$

Skalierung der Zeit- bzw. Frequenzachse (Ähnlichkeitssatz)

Die Zeit t und die komplexe Frequenz s verhalten sich invers zueinander:

$$f(kt) \quad \circ \bullet \quad \frac{1}{k} F\left(\frac{s}{k}\right)$$

Verschiebung der Frequenzachse (Dämpfungssatz)

Eine Frequenzverschiebung führt zu einer exponentiellen Dämpfung der Zeitfunktion:

$$e^{-at} f(t) \quad \circ \bullet \quad F(s + a)$$

Zeitliche Verschiebung

Eine einfache Zeitverzögerung um T wirkt sich im Laplace-Bereich relativ kompliziert aus:

$$f(t - T) \quad \circ \bullet \quad F(s)e^{-sT}$$

4.3 Eigenschaften

Beim Rechnen mit dynamischen Systemen gehört die Differentiation und die Integration neben der Faltung zu den wichtigsten Operationen! Auch sie vereinfachen sich drastisch in eine Multiplikation mit bzw. Division durch die komplexe Frequenzvariable s .

Differentiation

ACHTUNG: Hier geht die Anfangsbedingung $f(-0)$, nicht der Anfangswert $f(+0)$ ein!

$$\frac{df(t)}{dt} \quad \circ \bullet \quad sF(s) - f(-0)$$

Integration

Die Integration ist die zur Differentiation inverse Operation. Deshalb eine Division durch s :

$$\int_0^t f(\tau) d\tau \quad \circ \bullet \quad \frac{1}{s} F(s)$$

4.4 Rücktransformation in den Zeitbereich

Möglichkeiten zur Rücktransformation:

1. Verwendung der Tabelle, d.h. Rückgriff auf die Arbeit Anderer für Standardfälle.
2. Vereinfachung mittels Partialbruchzerlegung; anschließend Verwendung der Tabelle.
3. Explizite Berechnung des Integrals für die Rücktransformation (Residuensatz).

Lösung 1 ist trivial. Lösung 3 behandeln wir hier nicht. Lösung 2 wird jetzt erklärt...

Summen-Standardform eines Signals im Laplace-Bereich:

$$Y(s) = \frac{b_m s^m + \dots + b_1 s + b_0}{a_n s^n + \dots + a_1 s + a_0} e^{-T_t s}$$

Totzeit
(Zeitverzug um T_t)

($n > m$) wobei wir ohne Einschränkungen $a_n = 1$ setzen können (redundanter Parameter).

Produkt-Standardform eines Signals im Laplace-Bereich:

$$Y(s) = k \frac{(s - n_1)(s - n_2) \cdot \dots \cdot (s - n_m)}{(s - p_1)(s - p_2) \cdot \dots \cdot (s - p_n)} e^{-T_t s}$$

mit den m Nullstellen n_1, \dots, n_m und den n Polstellen p_1, \dots, p_n ($n > m$).

4.4 Rücktransformation in den Zeitbereich

Asymmetrie

- **Einfaches Ausmultiplizieren:** Produkt-Standardform \rightarrow Summen-Standardform

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

$$(s + 1)(s + 2)(s + 3) = s^3 + 6s^2 + 11s + 6$$

- **Schwieriges Faktorisieren:** Summen-Standardform \rightarrow Produkt-Standardform

$$s^2 + 3s + 2 = ?$$

$$s_{1,2} = -\frac{3}{2} \pm \sqrt{\frac{9}{4} - 2} = -\frac{3}{2} \pm \sqrt{\frac{1}{4}} = -\frac{3}{2} \pm \frac{1}{2}$$

$$s_1 = -1$$

$$s_2 = -2$$

$$\rightarrow (s + 1)(s + 2)$$

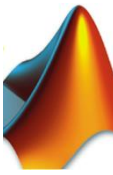
$$s^3 + 6s^2 + 11s + 6 = ? \quad s_{1,2,3} = ?$$

Für Polynome höheren Grades:

- nur mit Nullstellen raten \rightarrow in Praxis unmöglich, da reell
- oder numerisch \rightarrow nur (leicht) mit Computer möglich

4.4 Rücktransformation in den Zeitbereich

MATLAB



Summen-Standardform:

```
System = tf(Zaehler, Nenner); % tf = transfer function
```

$$G(s) = \frac{5s^2 + 3s - 1}{2s^3 - 0.5s + 4} e^{-2s}$$

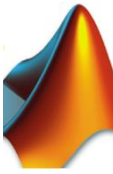
```
Zaehler = [5 3 -1]; % Poly-Koeff. in absteigender Reihenfolge
Nenner = [2 0 -0.5 4]; % 0 für s^2 Koeffizienten
Totzeit = 2;
System = tf(Zaehler, Nenner); % Noch ohne Totzeit!
System.ioDelay = Totzeit; % "Eigenschaft" Totzeit hinzufügen
```

Alternative:

```
s = tf('s'); % Definiere 's' als Laplace-Variable
System = (5*s^2+3*s-1)/(5*s^3-0.5*s+4);
System.ioDelay = 2;

K = dcgain(System); % Berechnet die Verstärkung
```

4.4 Rücktransformation in den Zeitbereich



Produkt-Standardform:

```
System = zpk(Nullstellen, Pole, k);  
% zpk = zeros-poles-k
```

$$G(s) = 5 \frac{(s + 1)(s + 2)}{(s + 3)(s + 1 - i2)(s + 1 + i2)}$$

```
Nullstellen = [-1 -2];
```

```
Pole = [-3 -1+i*2 -1-i*2];
```

```
k = 5;
```

```
% Verstärkung K = G(0) =
```

```
% 5*(1*2)/(3*(1-i2)*(1+i2))
```

```
% = 5*2/15 = 2/3
```

4.4 Rücktransformation in den Zeitbereich

Zerlegung der Produkt-Standardform in Partialbrüche:

- Ziel: Jeder Summand soll so einfach sein, dass er sich mit der Tabelle rücktransformieren lässt.
- Evtl. Totzeiten T werden separat mit dem Satz zur zeitlichen Verschiebung behandelt:

$$\boxed{f(t - T) \quad \longleftrightarrow \quad F(s)e^{-sT}}$$

Bei der Partialbruchzerlegung müssen 3 Fälle unterschieden werden:

1. Alle Pole sind reell und verschieden.
2. Alle Pole sind reell aber es gibt Mehrfachpole.
3. Es gibt auch konjugiert komplexe Pole.

Wir wollen nun also folgendes Signal transformieren (Struktur des Zählers $Z(s)$ ist hier uninteressant, Totzeit wird separat behandelt):

$$Y(s) = \frac{Z(s)}{(s - p_1)(s - p_2) \cdot \dots \cdot (s - p_n)}$$

4.4 Rücktransformation in den Zeitbereich

1. Partialbruchzerlegung (alle Pole reell)

$$Y(s) = \frac{B_1}{s - p_1} + \frac{B_2}{s - p_2} + \dots + \frac{B_n}{s - p_n} \quad \bullet \circ \quad y(t) = B_1 e^{p_1 t} + B_2 e^{p_2 t} + \dots + B_n e^{p_n t}$$

Die unbekannt Parameter B_1, \dots, B_n lassen sich mittels Koeffizientenvergleich ermitteln:

$$Z(s) = B_1(s - p_2) \cdot \dots \cdot (s - p_n) + B_2(s - p_1)(s - p_3) \cdot \dots \cdot (s - p_n) + \dots \\ + B_n(s - p_1) \cdot \dots \cdot (s - p_{n-1})$$

Rechts steht ein Polynom $n-1$. Grades, d.h. mit n Koeffizienten. Dies liefert n Gleichungen für n Unbekannte.

Oder mit einem Trick: Einsetzen einer Polstelle p_i in obige Gleichung liefert direkt einen Parameter, da alle anderen Terme = 0 sind:

$$Z(s = p_i) = B_i(s - p_1) \cdot \dots \cdot (s - p_{i-1}) \cdot (s - p_{i+1}) \cdot \dots \cdot (s - p_n) \\ \rightarrow B_i = [Y(s)(s - p_i)]|_{s=p_i}$$

Dies kann für alle $i = 1, \dots, n$ durchgeführt werden.

4.4 Rücktransformation in den Zeitbereich

Beispiel: 1. Partialbruchzerlegung (alle Pole reell)

Gegeben ist folgendes Signal im Laplace-Bereich:

$$Y(s) = \frac{8}{s(s+2)(s+4)}$$

Ansatz für die Rücktransformation ist:

$$Y(s) = \frac{B_1}{s} + \frac{B_2}{s+2} + \frac{B_3}{s+4} \quad \bullet \circ \quad y(t) = B_1 + B_2 e^{-2t} + B_3 e^{-4t}$$

a) Ermittlung der Parameter mittels Koeffizientenvergleichs:

$$8 = B_1(s+2)(s+4) + B_2s(s+4) + B_3s(s+2)$$

$$8 = (B_1 + B_2 + B_3)s^2 + (6B_1 + 4B_2 + 2B_3)s + 8B_1$$

$$\rightarrow \boxed{B_1 = 1}$$

$$B_2 + B_3 + 1 = 0 \quad \rightarrow \quad B_2 = -B_3 - 1 \quad \rightarrow \quad \boxed{B_2 = -2}$$

$$4B_2 + 2B_3 + 6 = 0 \quad \rightarrow \quad -4B_3 - 4 + 2B_3 + 6 = -2B_3 + 2 = 0 \quad \rightarrow \quad \boxed{B_3 = 1}$$

4.4 Rücktransformation in den Zeitbereich

Beispiel: 1. Partialbruchzerlegung (alle Pole reell)

b) Ermittlung der Parameter mittels Trick:

$$B_1 = \left. \frac{8}{(s+2)(s+4)} \right|_{s=0} = 1$$

$$B_2 = \left. \frac{8}{s(s+4)} \right|_{s=-2} = -2$$

$$B_3 = \left. \frac{8}{s(s+2)} \right|_{s=-4} = 1$$

Ergebnis der Partialbruchzerlegung:

$$Y(s) = \frac{1}{s} - \frac{2}{s+2} + \frac{1}{s+4} \quad \bullet \circ \quad y(t) = 1 - 2e^{-2t} + e^{-4t}$$

4.4 Rücktransformation in den Zeitbereich

2. Partialbruchzerlegung (alle Pole reell, aber Mehrfachpole)

Sortieren wir die Terme so, dass der erste Pol die Vielfachheit p hat:

$$Y(s) = \frac{B_{11}}{s - p_1} + \frac{B_{12}}{(s - p_1)^2} + \dots + \frac{B_{1p}}{(s - p_1)^p} + \frac{B_2}{s - p_2} + \dots + \frac{B_n}{s - p_n}$$

Koeffizientenvergleich liefert auch hier die unbekannt Parameter. Das Polynom hat nun den Grad $n+p-1$, so dass es wieder genauso viele Gleichungen wie Unbekannte gibt.

Der Trick von Fall 1 funktioniert hier nur für die Parameter der einfachen Pole B_2, \dots, B_n . Für die Parameter der Mehrfachpole (Ordnung p) gilt folgende Erweiterung:

$$B_{1i} = \frac{1}{(p - i)!} \frac{d^{p-i}}{ds^{p-i}} [Y(s)(s - p_1)^p] \Big|_{s=p_1}$$

Dies kann für $i = 1, \dots, p$ berechnet werden.

4.4 Rücktransformation in den Zeitbereich

Beispiel: 2. Partialbruchzerlegung (alle Pole reell, aber Mehrfachpole)

Gegeben ist folgendes Signal im Laplace-Bereich:

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

Ansatz für die Rücktransformation ist:

$$Y(s) = \frac{B_1}{s} + \frac{B_{21}}{s+1} + \frac{B_{22}}{(s+1)^2} + \frac{B_3}{s+4} \quad \bullet \circ \quad y(t) = B_1 + B_{21}e^{-t} + B_{22}te^{-t} + B_3e^{-4t}$$

Mit dem Trick berechnen sich die Parameter zu:

$$B_1 = \left. \frac{4}{(s+1)^2(s+4)} \right|_{s=0} = 1 \qquad B_3 = \left. \frac{4}{s(s+1)^2} \right|_{s=-4} = -\frac{1}{9}$$

$$B_{22} = \left. \frac{4}{s(s+4)} \right|_{s=-1} = -\frac{4}{3}$$

$$B_{21} = \left. \frac{d}{ds} \frac{4}{s(s+4)} \right|_{s=-1} = \left. \frac{-4(2s+4)}{s^2(s+4)^2} \right|_{s=-1} = -\frac{8}{9}$$

4.4 Rücktransformation in den Zeitbereich

3. Partialbruchzerlegung (auch konjugiert komplexe Pole)

Sortieren wir die Terme so, dass die ersten beiden Pole ein konjugiert komplexes Polpaar mit $p_{1/2} = p \pm i\omega$ bilden:

$$Y(s) = \frac{B_1 s + B_2}{(s - p_1)(s - p_2)} + \frac{B_3}{s - p_3} + \dots + \frac{B_n}{s - p_n}$$

Dann können wir die Parameter wieder über Koeffizientenvergleich ermitteln. Zu beachten ist, dass $(s - p_1)(s - p_2) = (s - p)^2 + \omega^2$ natürlich reell ist!

Eine Erweiterung des Tricks liefert:

$$B_1 = \frac{1}{\omega} \operatorname{Im} \left\{ [Y(s)(s - p_1)(s - p_2)] \Big|_{s=p_1} \right\}$$

$$B_2 = -B_1 p + \operatorname{Re} \left\{ [Y(s)(s - p_1)(s - p_2)] \Big|_{s=p_1} \right\}$$

4.4 Rücktransformation in den Zeitbereich

Beispiel: 3. Partialbruchzerlegung (auch konjugiert komplexe Pole)

Gegeben ist folgendes Signal im Laplace-Bereich:

$$Y(s) = \frac{5(s+1)}{s(s^2+2s+5)}$$

Ansatz für die Rücktransformation ist:

$$Y(s) = \frac{B_1}{s} + \frac{B_2s + B_3}{(s+1-i2)(s+1+i2)} \quad \bullet \circ \quad y(t) = B_1 + Ke^{pt} \sin(\omega t + \varphi)$$

Kompliziert zu berechnen, siehe Tabelle.

$p = -1$ $\omega = 2$

Mit dem Trick berechnen sich die Parameter zu:

$$B_1 = \left. \frac{5(s+1)}{s^2+2s+5} \right|_{s=0} = 1 \quad B_2 = \frac{1}{2} \operatorname{Im} \left\{ \left[\frac{5(s+1)}{s} \right] \right|_{s=-1+i2} \right\} = -1$$

$$B_3 = -(-1) \cdot (-1) + \operatorname{Re} \left\{ \left[\frac{5(s+1)}{s} \right] \right|_{s=-1+i2} \right\} = 3$$

4.5 Lösen einer DGL im Laplace-Bereich

Lösen einer DGL 1. Ordnung

DGL im Zeitbereich:

$$a_1 \dot{y}(t) + a_0 y(t) = b_0 u(t)$$

DGL transformiert in den Laplace-Bereich:

$$a_1 [sY(s) - y(-0)] + a_0 Y(s) = b_0 U(s)$$

$$Y(s)(a_1 s + a_0) = b_0 U(s) + a_1 y(-0)$$

$$\rightarrow Y(s) = \underbrace{\frac{b_0}{a_1 s + a_0} U(s)}_{Y_{\text{erzw}}(s)} + \underbrace{\frac{a_1}{a_1 s + a_0} y(-0)}_{Y_{\text{frei}}(s)}$$

Uns interessiert jetzt nur das Ein-/Ausgangsverhalten. Wir setzen also $y(-0) = 0$:

$$Y(s) = \frac{b_0}{a_1 s + a_0} U(s)$$

Wir bekommen $U(s)$ durch Laplace-Transformation von $u(t)$. Dann lässt sich $Y(s)$ berechnen.

4.5 Lösen einer DGL im Laplace-Bereich

Wie in Kapitel 2.6 sei $u(t)$ ein Einheitssprung:

$$u(t) = \sigma(t) \quad \circ \bullet \quad U(s) = \frac{1}{s}$$

Transformationstabelle

$$\rightarrow Y(s) = \frac{b_0}{(a_1 s + a_0)s} = \frac{\frac{b_0}{a_0}}{\left(\frac{a_1}{a_0}s + 1\right)s} \quad \bullet \circ \quad y(t) = \frac{b_0}{a_0} \left(1 - e^{-\frac{a_0}{a_1}t}\right)$$

Ergebnis ist identisch mit dem in Kapitel 2.6 für $a_1 = 1$.

Für ein anderes Eingangssignal, z.B. eine Einheitsrampe, bekämen wir:

$$Y(s) = \frac{\frac{b_0}{a_0}}{\left(\frac{a_1}{a_0}s + 1\right)s^2}$$

Die Rücktransformation zu $y(t)$ erfolgt über Partialbruchzerlegung (doppelter Pol bei 0, einfacher Pol bei $-a_0/a_1$) und Tabelle.

→ DGL recht einfach für verschiedenste Eingangssignale lösbar!

4.5 Lösen einer DGL im Laplace-Bereich

Lösen einer DGL n . Ordnung

DGL im Zeitbereich:

$$a_n y^{(n)}(t) + a_{n-1} y^{(n-1)}(t) + \dots + a_1 y^{(1)}(t) + a_0 y(t) = \\ b_m u^{(m)}(t) + b_{m-1} u^{(m-1)}(t) + \dots + b_1 u^{(1)}(t) + b_0 u(t)$$

DGL transformiert in den Laplace-Bereich:

$$a_n [s^n Y(s) - s^{n-1} y(-0) - \dots - s y^{(n-2)}(-0) - y^{(n-1)}(-0)] + \dots \\ + a_1 [s Y(s) - y(-0)] + a_0 Y(s) = \\ b_m [s^m U(s) - s^{m-1} u(-0) - \dots - s u^{(m-2)}(-0) - u^{(n-1)}(-0)] + \dots \\ + b_1 [s U(s) - u(-0)] + b_0 U(s)$$

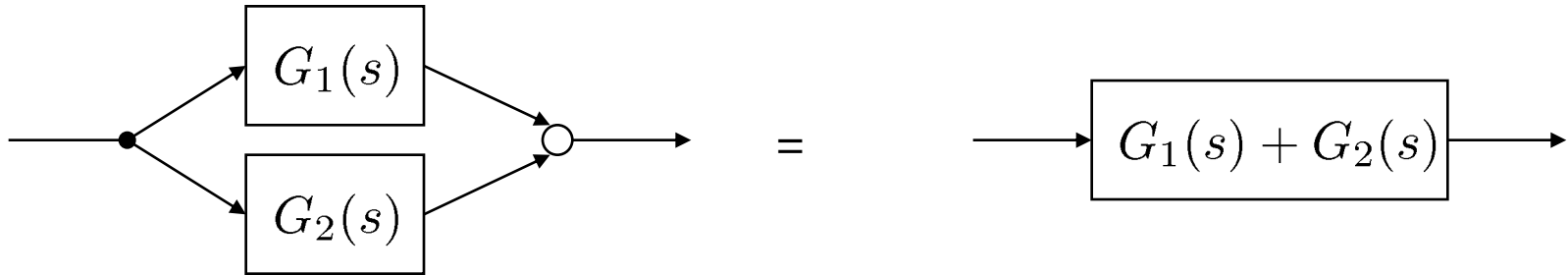
Uns interessiert jetzt nur das Ein-/Ausgangsverhalten. Wir setzen also alle Anfangsbed. = 0:

$$Y(s) = \frac{b_m s^m + \dots + b_1 s + b_0}{a_n s^n + \dots + a_1 s + a_0} U(s)$$

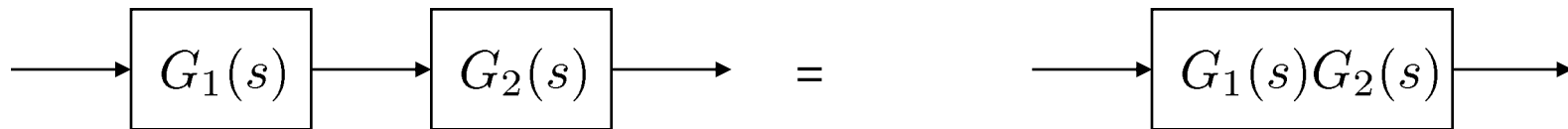
→ *Zeitlösung*: Pole ausrechnen,
Partialbruchzerlegung, Tabelle.

4.6 Rechenregeln

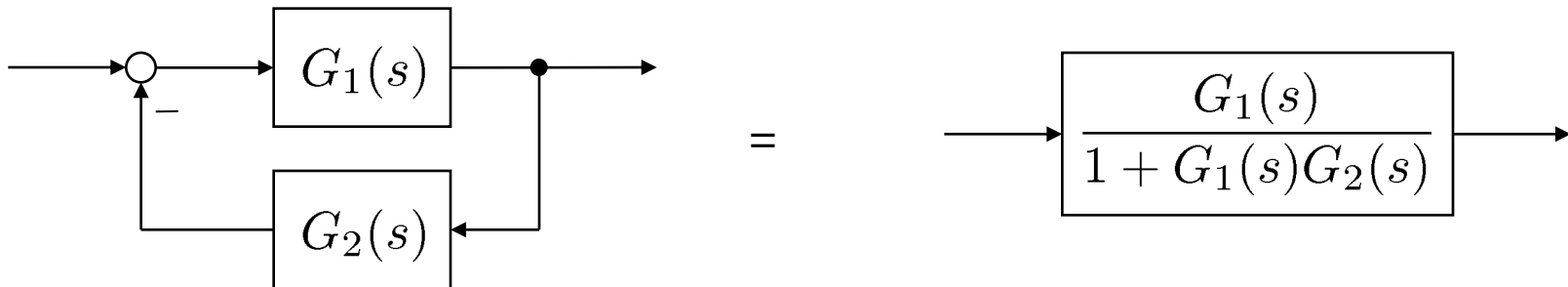
Parallelschaltung



Reihenschaltung



Negative Rückkopplungsschaltung (Gegenkopplung)



4.7 Tabelle Laplace-Transformation

Funktion $f(t)$ mit $f(t) = 0$ für $t < 0$ ← **wichtig!**

$$F(s) = \mathcal{L}\{f(t)\}$$

Dirac-Impuls $\delta(t)$

$$1$$

Einheitssprung $\sigma(t) = 1$

$$1/s$$

Einheitsrampe t

$$1/s^2$$

$$t^2$$

$$2/s^3$$

$$t^n$$

$$n!/s^{n+1}$$

$$e^{-at}$$

$$\frac{1}{s+a}$$

$$te^{-at}$$

$$\frac{1}{(s+a)^2}$$

$$\sin\omega_0 t$$

$$\frac{\omega_0}{s^2 + \omega_0^2}$$

$$\cos\omega_0 t$$

$$\frac{s}{s^2 + \omega_0^2}$$

4.7 Tabelle Laplace-Transformation

Funktion $f(t)$ mit $f(t) = 0$ für $t < 0$

$$F(s) = \mathcal{L}\{f(t)\}$$

$$e^{-at} \sin \omega_0 t$$

$$\frac{\omega_0}{(s+a)^2 + \omega_0^2}$$

$$e^{-at} \cos \omega_0 t$$

$$\frac{s+a}{(s+a)^2 + \omega_0^2}$$

$$\sqrt{\left(\frac{b-a}{\omega_0}\right)^2 + 1} \cdot e^{-at} \sin\left(\omega_0 t + a \tan^{-1} \frac{\omega_0}{b-a}\right)$$

$$\frac{s+b}{(s+a)^2 + \omega_0^2}$$

$$\frac{\omega_0}{\sqrt{1-D^2}} e^{-D\omega_0 t} \sin(\omega_0 \sqrt{1-D^2} t)$$

$$\frac{\omega_0^2}{s^2 + 2D\omega_0 s + \omega_0^2}$$

$$\frac{-1}{\sqrt{1-D^2}} e^{-D\omega_0 t} \sin\left(\omega_0 \sqrt{1-D^2} t - a \tan^{-1} \frac{\sqrt{1-D^2}}{D}\right)$$

$$\frac{s}{s^2 + 2D\omega_0 s + \omega_0^2}$$

4.7 Tabelle Laplace-Transformation

Funktion $f(t)$ mit $f(t) = 0$ für $t < 0$

$$F(s) = \mathcal{L}\{f(t)\}$$

$$\frac{1}{a} (1 - e^{-at})$$

$$\frac{1}{s(s+a)}$$

$$\frac{1}{b-a} (e^{-at} - e^{-bt})$$

$$\frac{1}{(s+a)(s+b)}$$

$$\frac{1}{b-a} (be^{-bt} - ae^{-at})$$

$$\frac{s}{(s+a)(s+b)}$$

5. Übertragungsfunktion

Inhalt Kapitel 5

5 Übertragungsfunktion

- 5.1 Definition der Gewichtsfunktion
- 5.2 Definition der Übertragungsfunktion
- 5.3 Pole der Übertragungsfunktion
- 5.4 Nullstellen der Übertragungsfunktion
- 5.5 Pol-/Nullstellen-Diagramm
- 5.6 Statische Verstärkung der Übertragungsfunktion
- 5.7 Weitere wichtige Eigenschaften der Übertragungsfunktion
- 5.8 Wichtige Übertragungsfunktionen im Regelkreis

5.1 Definition der Gewichtsfunktion

Wie wir bereits aus Kap. 4.2 wissen, gibt es für jedes lineare System eine Funktion $g(t)$, die sog.

Gewichtsfunktion, die *alle* Systemeigenschaften charakterisiert. Für jedes beliebige Eingangssignal

$u(t)$ lässt sich der Ausgang $y(t)$ eines dynamischen Systems mittels des Faltungsintegrals berechnen (Annahme: die Anfangsbedingungen sind Null):

$$y(t) = \int_0^t g(t - \tau)u(\tau)d\tau$$

Wählen wir als Eingangssignal einen Dirac-Impuls $u(t) = \delta(t)$, dann bezeichnet man das zugehörige Ausgangssignal als **Impulsantwort**:

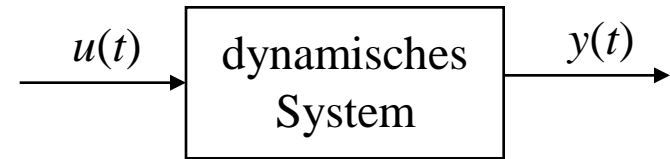
$$y(t) = \int_0^t g(t - \tau)\delta(\tau)d\tau = g(t)$$

Ausblendeigenschaft

Das heißt:

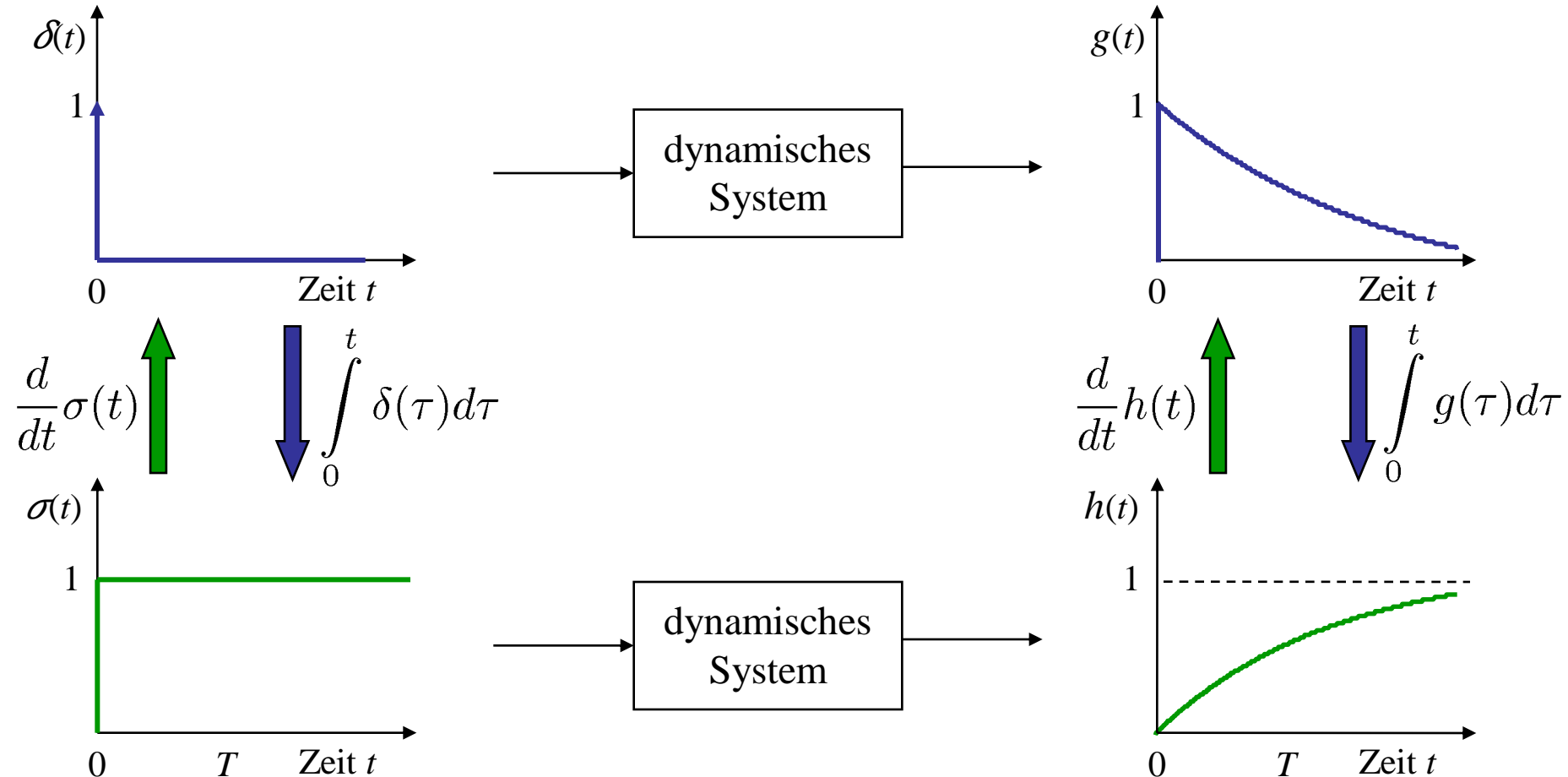
Die Impulsantwort eines linearen dynamischen Systems ist seine Gewichtsfunktion!

In der Mathematik/Physik ist für $g(t)$ die Bezeichnung Greensche Funktion üblich!



5.1 Definition der Gewichtsfunktion

Damit ergeben sich folgende wichtige Zusammenhänge zwischen **Dirac-Impuls** bzw. **Einheitssprung** als Eingangsgrößen und **Impulsantwort** bzw. **Sprungantwort** als Ausgangsgrößen:



5.2 Definition der Übertragungsfunktion

Die Gewichtsfunktion charakterisiert ein lineares dynamisches System vollständig!

Was ist die Entsprechung der Gewichtsfunktion im Frequenzbereich?

Die Faltung im Zeitbereich

$$y(t) = \int_0^t g(t - \tau)u(\tau)d\tau$$

entspricht der Multiplikation im Frequenzbereich

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

Für einen Dirac-Impuls $U(s) = 1$ am Eingang ergibt sich: $Y(s) = G(s) \cdot 1 = G(s)$

Damit wissen wir:

$$\boxed{G(s) \longleftrightarrow g(t)}$$

Die Laplace-Transformierte der Gewichtsfunktion nennen wir *Übertragungsfunktion!*

Sie enthält ebenso alle Informationen über das lineare dynamische System.

5.2 Definition der Übertragungsfunktion

Die Übertragungsfunktion eines linearen dynamischen Systems ist immer eine gebrochen rationale Funktion evtl. multipliziert mit einer e-Funktion, die eine Totzeit beschreibt:

$$G(s) = \frac{Y(s)}{U(s)} = \frac{b_m s^m + \dots + b_1 s + b_0}{a_n s^n + \dots + a_1 s + a_0} e^{-T_t s}$$

Diese Summen-Standardform kann in die Produkt-Standardform umgewandelt werden:

$$G(s) = \frac{Y(s)}{U(s)} = k \frac{(s - n_1)(s - n_2) \cdot \dots \cdot (s - n_m)}{(s - p_1)(s - p_2) \cdot \dots \cdot (s - p_n)} e^{-T_t s}$$

Diese Produktform erlaubt die Zerlegung in Partialbrüche (siehe Kap. 4.4). Daraus lässt sich dann leicht in den Zeitbereich rücktransformieren. Damit ist klar: Die Pole einer Übertragungsfunktion spielen für das dynamische Verhalten eine herausragende Rolle!

Bemerkung: Obige Summen-Standardform enthält einen redundanten Parameter. Man kann ohne Einschränkungen z.B. $a_n = 1$ setzen.

5.2 Definition der Übertragungsfunktion

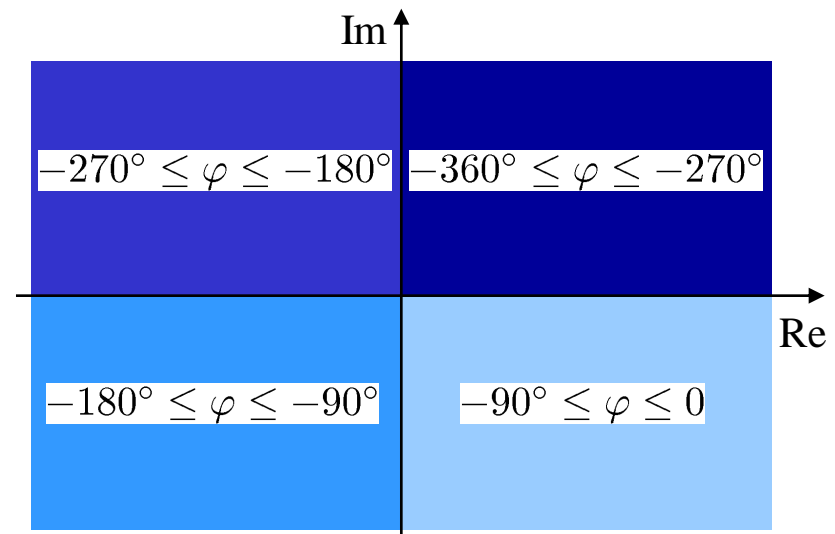
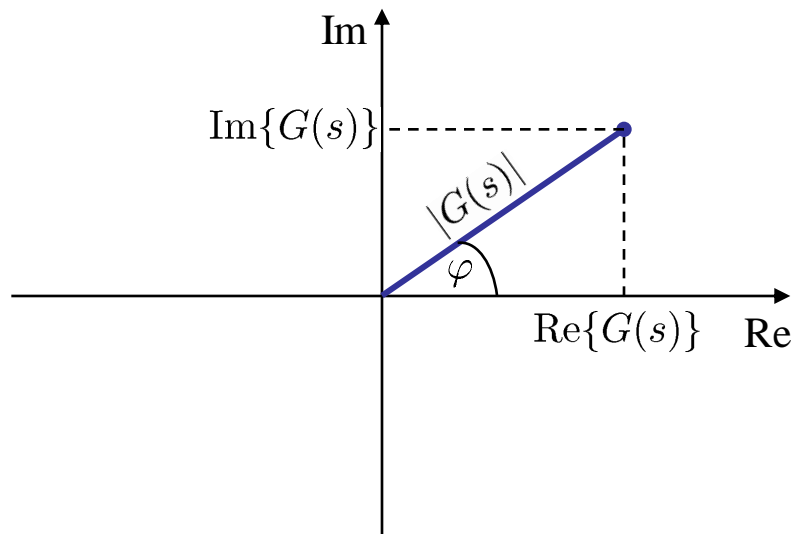
Die Übertragungsfunktion ist eine Funktion, die eine komplexe Variable $s = \sigma + i\omega$ auf die komplexe Zahl $G(s)$ abbildet. Sie kann als Real- und Imaginärteil oder als Betrag und Phase ausgedrückt werden:

$$G(s) = \operatorname{Re}\{G(s)\} + i\operatorname{Im}\{G(s)\} = |G(s)|e^{i\varphi}$$

Diese stehen wie folgt in Beziehung zueinander:

$$|G(s)| = \sqrt{\operatorname{Re}\{G(s)\}^2 + \operatorname{Im}\{G(s)\}^2}$$

$$\varphi = \arctan \frac{\operatorname{Im}\{G(s)\}}{\operatorname{Re}\{G(s)\}}$$



5.3 Pole der Übertragungsfunktion

1. Im einfachsten Fall hat eine Übertragungsfunktion n verschiedene reelle Pole:

$$G(s) = \frac{B_1}{s - p_1} + \frac{B_2}{s - p_2} + \dots + \frac{B_n}{s - p_n}$$

Damit ergibt sich für die Gewichtsfunktion (= Impulsantwort) des Systems:

$$g(t) = B_1 e^{p_1 t} + B_2 e^{p_2 t} + \dots + B_n e^{p_n t}$$

3 Fälle:

- Alle Pole sind negativ, d.h. $p_i < 0$ für alle $i = 1, \dots, n$.
- Ein Pol $p_j = 0$, alle anderen sind negativ.
- Mindestens ein Pol ist positiv, d.h. $p_j > 0$, alle anderen sind irrelevant.

Daraus ergibt sich folgendes charakteristisches Verhalten für die Gewichtsfunktion:

- $g(t)$ strebt exponentiell schnell gegen 0. → **Stabiles** Verhalten.
- $g(t)$ strebt exponentiell schnell gegen den festen Endwert B_j . → **Grenzstabiles** Verhalten.
- $g(t)$ strebt exponentiell schnell gegen unendlich. → **Instabiles** Verhalten.

5.3 Pole der Übertragungsfunktion

2. Übertragungsfunktion mit einem p -fachen reellen Pol, sonst nur einfache reelle Pole:

$$G(s) = \frac{B_{11}}{s - p_1} + \frac{B_{12}}{(s - p_1)^2} + \dots + \frac{B_{1p}}{(s - p_1)^p} + \frac{B_2}{s - p_2} + \dots + \frac{B_n}{s - p_n}$$

Damit ergibt sich für die Gewichtsfunktion (= Impulsantwort) des Systems:

$$g(t) = B_{11}e^{p_1 t} + B_{12}te^{p_1 t} + \dots + \frac{B_{1p}}{(p-1)!}t^{p-1}e^{p_1 t} + B_2e^{p_2 t} + \dots + B_n e^{p_n t}$$

Die selben 3 Fälle wie zuvor aber mit folgender Besonderheit:

b) $p_1 = 0$:

Daraus ergibt sich folgendes charakteristisches Verhalten für die Gewichtsfunktion:

b) $g(t)$ strebt exponentiell schnell gegen eine Funktion t^{p-1} mit dem Vorfaktor $B_{1p}/(p-1)!$, wenn $p_1 = 0$ ein Mehrfachpol ist. → **Instabiles** Verhalten.

5.3 Pole der Übertragungsfunktion

3. Übertragungsfunktion mit einem konjugiert komplexen Polpaar $p_{1/2} = p \pm i\omega$:

$$G(s) = \frac{B_1 s + B_2}{(s - p_1)(s - p_2)} + \frac{B_3}{s - p_3} + \dots + \frac{B_n}{s - p_n}$$

Damit ergibt sich für die Gewichtsfunktion (= Impulsantwort) des Systems:

$$g(t) = B_1 e^{pt} \cos \omega t + \frac{B_2 + B_1 p}{\omega} e^{pt} \sin \omega t + B_3 e^{p_3 t} + \dots + B_n e^{p_n t}$$

Die selben 3 Fälle wie zuvor aber mit folgenden Besonderheiten:

- Alle Pole haben einen negativen *Realteil*...
- Konjugiert komplexes Polpaar hat einen Realteil gleich Null: $g(t)$ führt eine Dauerschwingung mit der Kreisfrequenz ω aus. → **Oszillatorisch grenzstabiles** Verhalten.
- Mindestens ein Pol hat einen positiven Realteil... : Falls dies für das konjugiert komplexe Polpaar der Fall ist: $g(t)$ für eine exponentiell aufklingende Dauerschwingung mit der Kreisfrequenz ω aus. → **Oszillatorisch instabiles** Verhalten.

5.4 Nullstellen der Übertragungsfunktion

Eigenschaften von Nullstellen:

- Entstehen typischerweise durch Parallelschaltungen (parallele Ursache-Wirkungs-Ketten), z.B.:
$$G(s) = G_1(s) + G_2(s) = \frac{1}{s+1} + \frac{1}{s+2} = \frac{2s+3}{(s+1)(s+2)}$$
- Bestimmen (zusammen mit anderen Größen) die Vorfaktoren vor den e-, sin-, cos-Funktionen der Zeitlösungen und beeinflussen damit wie stark welche Dynamikanteile gewichtet werden.
- Nullstellen „schlucken“ bestimmte Anregungssignale weg, weil sie Faktoren im Nenner der Übertragungsfunktion wegekürzen.

Beispiel: Konjugiert komplexe Nullstellen bei $s_{1/2} = \pm i\omega_0$: Zähler $(s+i\omega_0)(s-i\omega_0) = s^2 + \omega_0^2$

$$Y(s) = \frac{Z(s)}{N(s)} U(s) = \frac{\cancel{s^2 + \omega_0^2} \cdot \omega_0}{N(s) \cdot \cancel{s^2 + \omega_0^2}} = \frac{\omega_0}{N(s)} \quad (\text{Schwingung mit } \omega_0 \text{ ist weg!})$$

Anregung mit $\sin \omega_0 t$

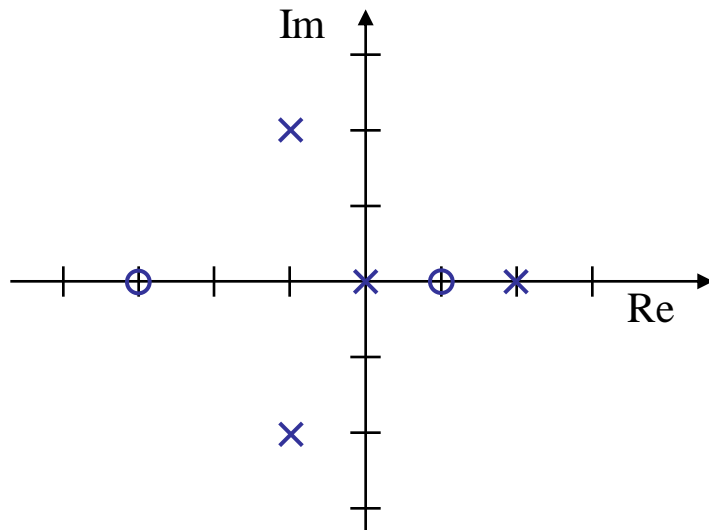
- Pole mit positivem Realteil führen auf Instabilität. Nullstellen mit positivem Realteil führen auch zu einem schwer regelbaren Verhalten. → Nicht phasenminimales Verhalten...

5.5 Pol-/Nullstellen-Diagramm

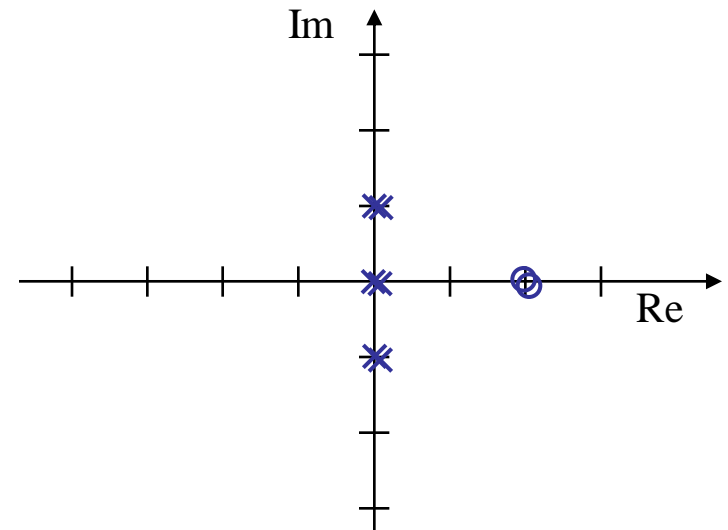
Da die Lage der Pole und Nullstellen für das dynamische Verhalten eines Systems entscheidend sind, werden sie in der komplexen s -Ebene übersichtlich wie folgt dargestellt:

- Pole als Kreuze: ×
- Nullstellen als Kreise: ○

$$G(s) = \frac{(s + 3)(s - 1)}{s(s - 2)(s + 1 - i2)(s + 1 + i2)}$$

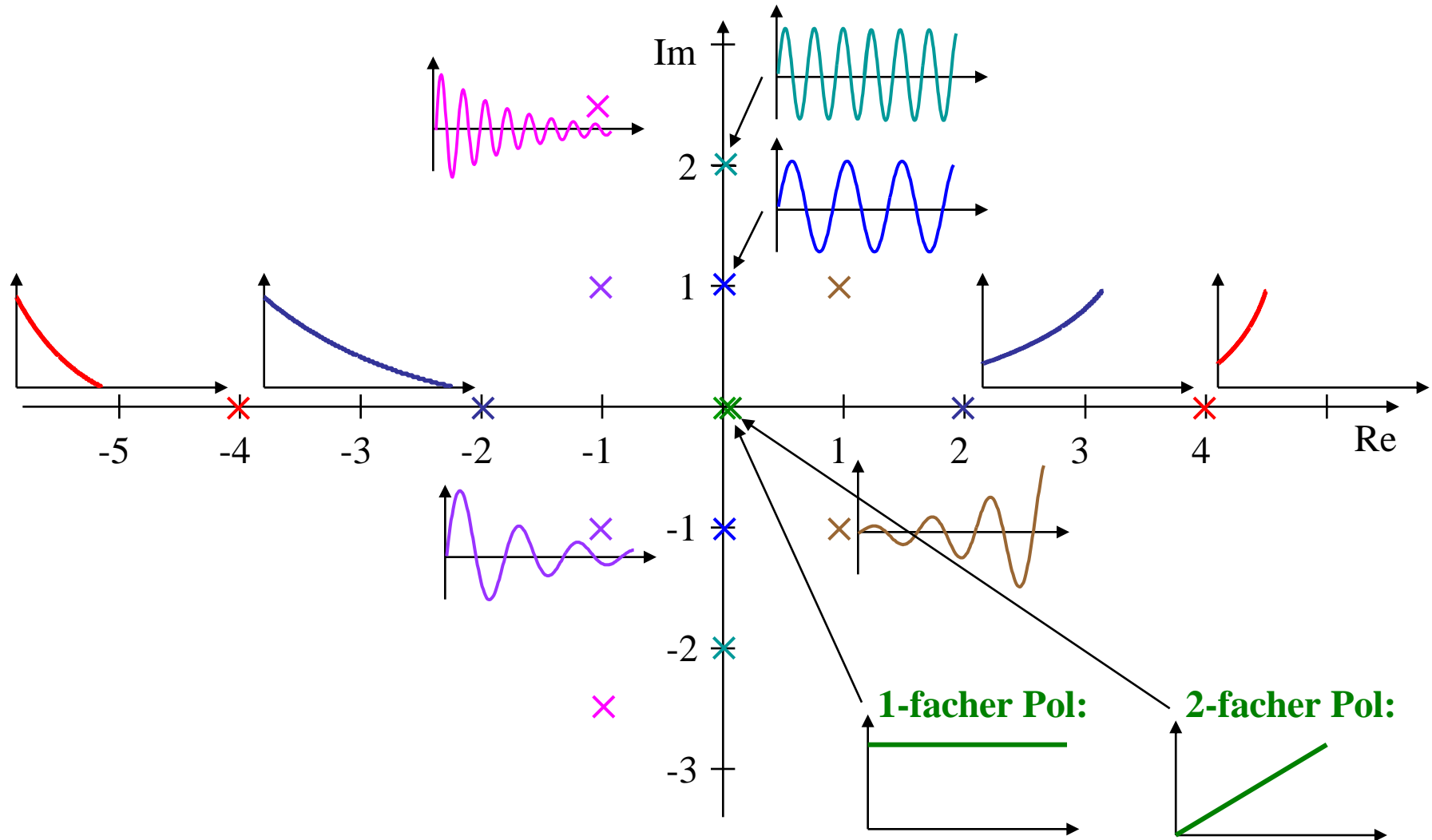


$$G(s) = \frac{(s - 2)^2}{s^2(s - i)^2(s + i)^2}$$

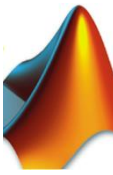


5.5 Pol-/Nullstellen-Diagramm

Zusammenhang: Polage \rightarrow Zeitlösung der zugehörigen DGL (Gewichtsfunktion):



5.5 Pol-/Nullstellen-Diagramm



Berechnung von Nullstellen und Polen:

Suche die Wurzeln eines Polynoms der Ordnung p :

$$c(1)*s^p + c(2)*s^{(p-1)} + \dots + c(p)*s + c(p+1)$$

```
Koeff = [c(1) c(2) ... c(p) c(p+1)]; % Vektor mit Koeffizienten  
Wurzeln = roots(Koeff); % Liefert die p Wurzeln des Poly.
```

ACHTUNG: Nummerierung ist umgedreht zu unserer (bei uns: a_i gehört zur Potenz i : $a_i s^i$).

Multiplikation zweier Polynome:

```
C_Koeff = [c(1) c(2) ... c(p) c(p+1)]; % Vektor mit Koeffizienten  
D_Koeff = [d(1) d(2) ... d(q) d(q+1)]; % Vektor mit Koeffizienten  
conv(C_koeff, D_Koeff); % Polynom der Ordnung p+q
```

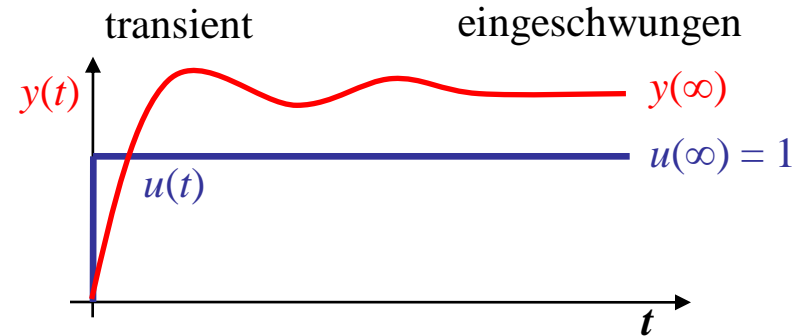
$$C(s)D(s) = (3s^2 + 2s + 1)(4s + 5) = 12s^3 + 23s^2 + 14s + 5$$

```
conv([3 2 1], [4 5]); % Liefert: [12 23 14 5]
```

5.6 Statische Verstärkung der Übertragungsfunktion

Eingangssignal sei ein *Einheitssprung*.

Welcher Wert stellt sich am Ausgang des Systems nach langer Zeit ein?



Endwertsatz der Laplace-Transformation:

$$y(\infty) = \lim_{s \rightarrow 0} sY(s) = \lim_{s \rightarrow 0} sG(s)U(s) = \lim_{s \rightarrow 0} sG(s) \frac{1}{s} = \lim_{s \rightarrow 0} G(s) = G(0) = \frac{b_0}{a_0}$$

Falls sich das System wirklich auf einen Endwert einpendelt, dann nennen wir diesen die *statische Verstärkung* $K = y(\infty) / u(\infty)$ des Systems. Für rein statische Betrachtungen können wir dann ein lineares dynamisches System durch seine statische Verstärkung ersetzen.

ACHTUNG: Es kann vorkommen, dass wir nach obiger Formel einen Endwert berechnen, in Wirklichkeit das System aber gegen unendlich strebt oder eine Dauerschwingung ausführt. Deshalb müssen wir zuvor zusätzlich prüfen, ob das System stabil ist, also alle Komponenten der Zeitlösung einschwingen. Über Stabilität später mehr...

5.7 Weitere wichtige Eigenschaften der Übertragungsfunktion

Übertragungsfunktion:

$$G(s) = \frac{b_m s^m + \dots + b_1 s + b_0}{a_n s^n + \dots + a_1 s + a_0}$$

Bislang haben wir angenommen: Nennergrad > Zählergrad, also $n > m$. Damit ergibt sich für einen Einheitssprung am Eingang stets ein Anfangswert von $G(s \rightarrow \infty) = 0$ (Anfangswertsatz).

Was passiert für $n = m$?

$$G(s) = \frac{b_n s^n + \dots + b_1 s + b_0}{a_n s^n + \dots + a_1 s + a_0} = \frac{b_n}{a_n} + \underbrace{\frac{\tilde{b}_{n-1} s^{n-1} + \dots + \tilde{b}_1 s + \tilde{b}_0}{a_n s^n + \dots + a_1 s + a_0}}_{\tilde{m}=n-1, \text{ also } \tilde{m} < n}$$

Was passiert für $n < m$?

$$G(s) = \frac{b_m}{a_n} s^{m-n} + \dots + \frac{b_{n+1}}{a_n} s + \frac{b_n}{a_n} + \underbrace{\frac{\tilde{b}_{n-1} s^{n-1} + \dots + \tilde{b}_1 s + \tilde{b}_0}{a_n s^n + \dots + a_1 s + a_0}}_{\tilde{m}=n-1, \text{ also } \tilde{m} < n}$$

Was bedeutet das im Zeitbereich?

5.7 Weitere wichtige Eigenschaften der Übertragungsfunktion

Fall $n = m$:

- Impulsantwort enthält einen Dirac-Impuls: $g(t) = \frac{b_n}{a_n} \delta(t) + \dots$
- Sprungantwort springt sofort (bei $t = 0$) auf den Wert b_n/a_n : $h(t) = \frac{b_n}{a_n} + \dots$
- Ein solches System bezeichnet man als *sprungfähig* oder als System mit *Durchgriff*.
- Sprungfähige Systeme gibt es in der Realität nicht, evtl. können sie aber als vereinfachte Idealisierung verwendet werden. Außerdem können solche Systeme natürlich in mathematischen Berechnungen vorkommen.

Fall $n < m$:

- Sprungantwort enthält einen Dirac-Impuls.
- Solche Systeme haben *differenzierenden Charakter*, d.h. sie erzeugen die (erste bis zu $(m-n)$ -te) Ableitung des Eingangssignals (Dirac-Impuls ist Ableitung des Einheitssprungs).
- Differenzierende Systeme gibt es in der Realität nicht. Ein solches System müsste in der Lage sein, auf einen Sprung mit einem unendlich hohen Dirac-Impuls zu antworten.

5.7 Weitere wichtige Eigenschaften der Übertragungsfunktion

Kausalität

Die Totzeit T_t im Faktor $e^{-T_t s}$ darf nicht negativ sein. Ansonsten würden Eingangswerte aus der Zukunft $u(t - T_t)$ den aktuellen Systemausgang $y(t)$ beeinflussen.

Sprungfähigkeit

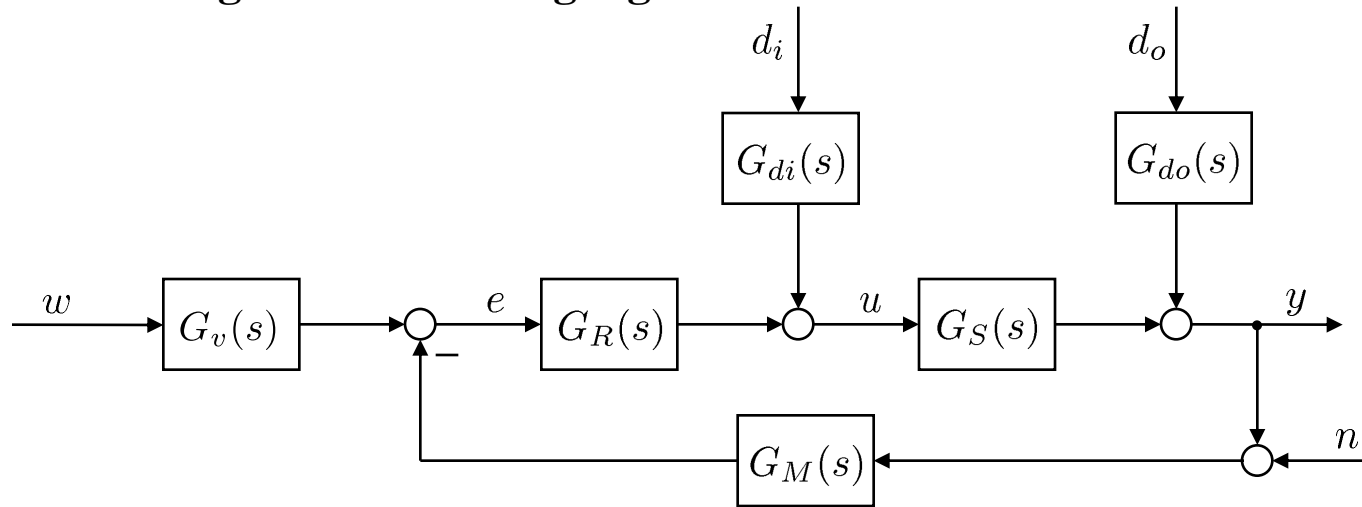
Ein lineares System ist sprungfähig bzw. hat Durchgriff (*proper*), wenn die Zählerordnung gleich groß wie die Nennerordnung ist, also $n = m$. Für nicht sprungfähige (*strictly proper*) muss der Zählergrad niedriger als der Nennergrad sein, also $n < m$.

Realisierbarkeit

Ein realisierbares System muss kausal sein, und die Zählerordnung darf nicht größer als die Nennerordnung sein, es muss also gelten $n \leq m$. Für die allermeisten realisierbaren System gilt sogar $n < m$.

5.8 Wichtige Übertragungsfunktionen im Regelkreis

Regelkreis mit Strecke G_S , Messglied G_M , Regler G_R , Führungsgrößenfilter G_v , Störgrößendynamik G_{di} für die Störung am Prozesseingang, Störgrößendynamik G_{do} für die Störung am Prozessausgang:



$$\frac{Y(s)}{W(s)} = \frac{G_R(s)G_S(s)G_v(s)}{1 + G_R(s)G_S(s)G_M(s)}$$

$$\frac{Y(s)}{D_I(s)} = \frac{G_S(s)G_{di}(s)}{1 + G_R(s)G_S(s)G_M(s)}$$

$$\frac{Y(s)}{N(s)} = \frac{-G_R(s)G_S(s)G_M(s)}{1 + G_R(s)G_S(s)G_M(s)}$$

$$\frac{Y(s)}{D_O(s)} = \frac{G_{do}(s)}{1 + G_R(s)G_S(s)G_M(s)}$$

5.8 Wichtige Übertragungsfunktionen im Regelkreis

Wir stellen fest:

Alle Übertragungsfunktionen enthalten den Faktor $\frac{1}{1 + G_0(s)}$
 mit $G_0(s) = G_R(s)G_S(s)G_M(s)$

Abhängig vom Eingang:

Vorwärtszweig =

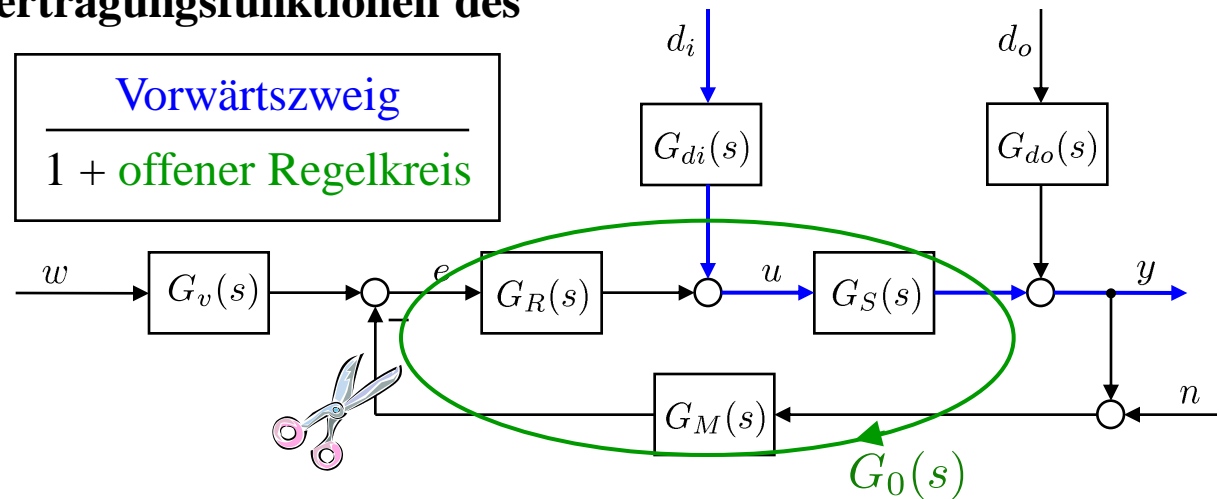
$$\begin{cases} G_v(s)G_R(s)G_S(s) & \text{für } w \\ G_{di}(s)G_S(s) & \text{für } d_i \\ G_{do}(s) & \text{für } d_o \end{cases}$$

Regel zum Erstellen der Übertragungsfunktionen des geschlossenen Regelkreises:

Gilt immer:
offener Regelkreis = $G_0(s)$

Der Faktor

$$S(s) = \frac{1}{1 + G_0(s)}$$



beschreibt den Einfluss der Rückkopplung im Regelkreis oder den Unterschied zwischen offenem und geschlossenem Regelkreis. Wird die Rückkopplung des Regelkreises aufgetrennt (wie mit ✂ eingezeichnet), so fällt dieser Faktor weg und nur der Vorwärtszweig bleibt bestehen. **S** wird Empfindlichkeitsfunktion (*sensitivity function*) genannt!

5.8 Wichtige Übertragungsfunktionen im Regelkreis

Wir stellen auch noch fest:

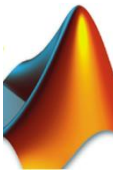
- Die Störung am Prozesseingang lässt sich in eine äquivalente Störung am Prozessausgang umrechnen und umgekehrt:

- Störung nur am Eingang:
$$D_I(s) = \frac{G_{do}(s)}{G_{di}(s)G_S(s)} D_O(s)$$

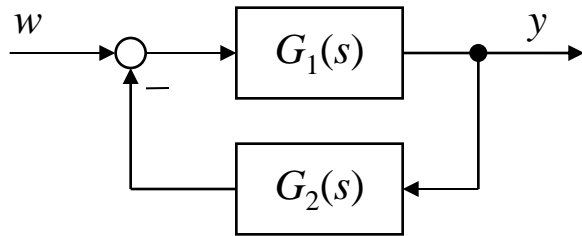
- Störung nur am Ausgang:
$$D_O(s) = \frac{G_{di}(s)G_S(s)}{G_{do}(s)} D_I(s)$$

- Der Regler $G_R(s)$ beeinflusst die Empfindlichkeitsfunktion S .
 - Der Führungsgrößenfilter $G_v(s)$ beeinflusst die Empfindlichkeitsfunktion S *nicht*, ebenso wie die Störfilter $G_{di}(s)$ und $G_{do}(s)$.
- Um das Regelverhalten in Bezug auf die Störgrößen d_i , d_o und das Rauschen n zu beeinflussen, müssen wir den Regler $G_R(s)$ (seine Struktur und/oder Parameter) verändern.
- Um das Regelverhalten in Bezug auf die Führungsgröße w zu beeinflussen, können wir sowohl den Regler $G_R(s)$ als auch das Führungsgrößenfilter $G_v(s)$ verändern.

5.8 Wichtige Übertragungsfunktionen im Regelkreis **MATLAB**



Berechnung der Übertragungsfunktion rückgekoppelter Systeme:



$$G_w(s) = \frac{Y(s)}{W(s)} = \frac{G_1(s)}{1 + G_1(s)G_2(s)}$$

```
s = tf('s');  
G1 = 1/(s^2+2*s+1);  
G2 = 5/(s+5);  
Gw = feedback(G1,G2);
```

*% Berechnet die Übertragungsfunktion
% von w nach y*

$$G_w(s) = \frac{\frac{1}{s^2+2s+1}}{1 + \frac{1}{s^2+2s+1} \frac{5}{s+5}} = \frac{s+5}{(s^2+2s+1)(s+5)+5} = \frac{s+5}{s^3+7s^2+11s+10}$$

6. Frequenzgang und Ortskurve

Inhalt Kapitel 6

6 Frequenzgang und Ortskurve

6.1 Definition des Frequenzgangs

6.2 Beziehung zwischen Frequenzgang und Ortskurve

6.3 Bode-Diagramm

6.4 Konstruktionsregeln für phasenminimale Systeme

6.5 Konstruktionsregeln für nicht phasenminimale Systeme

6.1 Definition des Frequenzgangs

Die Übertragungsfunktion $G(s)$ enthält außer den Anfangsbedingungen alle Informationen über ein lineares dynamisches System. $G(s)$ ist eine Funktion, die die komplexe Variable $s = \delta + i\omega$ auf einen komplexen Wert abbildet. $G(s)$ kann in Real- und Imaginärteil oder in Betrag und Phase zerlegt werden:

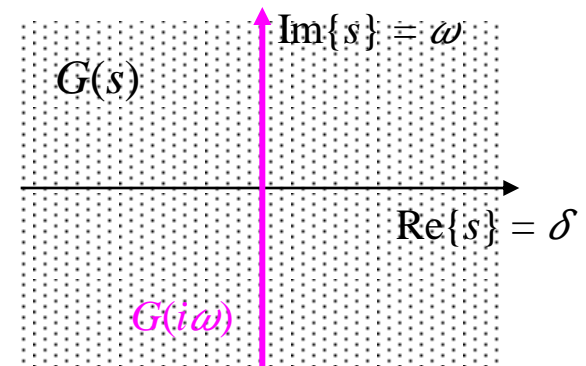
$$G(s) = \operatorname{Re}\{G(s)\} + i\operatorname{Im}\{G(s)\} = |G(s)|e^{i\varphi}$$

Zur Visualisierung von $G(s)$ müssen wir also zwei 3-dimensionale Grafiken zeichnen:

- Betrag $|G(s)|$ in Abhängigkeit von δ und ω .
- Phase (Winkel) $\angle G(s)$ in Abhängigkeit von δ und ω .

Oder alternativ: Real- und Imaginärteil. Das ist aber i.A. weniger anschaulich.

Es zeigt sich, dass alle Informationen von $G(s)$ auch in $G(i\omega)$ enthalten sind, d.h. entlang der imaginären Achse ($\delta = 0$). Es ist viel einfacher mit $G(i\omega)$ zu arbeiten, weil es nur von einer Größe abhängt, nämlich ω ; δ kommt nicht mehr vor.



6.1 Definition des Frequenzgangs

Der Betrag von

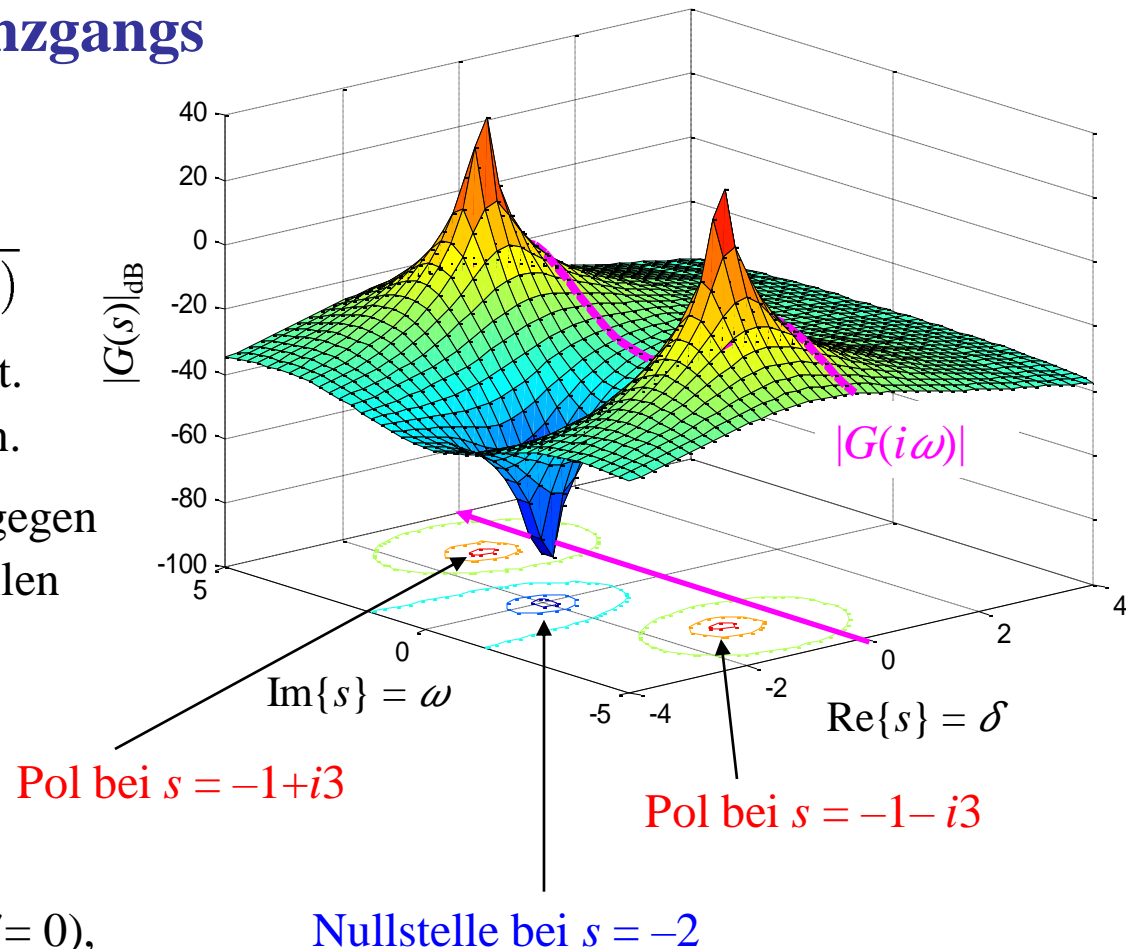
$$G(s) = \frac{s + 2}{(s + 1 - i3)(s + 1 + i3)}$$

ist zur Veranschaulichung dargestellt. Gleiches kann man für die Phase tun.

An den Polstellen strebt der Betrag gegen unendlich ($= \infty$ dB). An den Nullstellen gegen Null ($= -\infty$ dB).

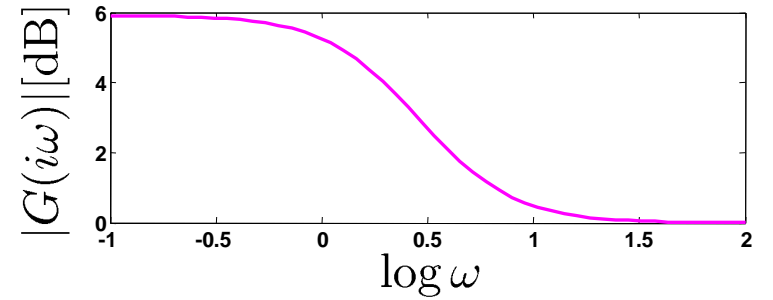
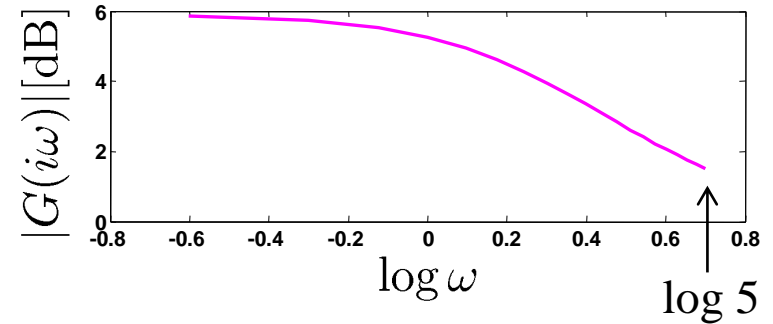
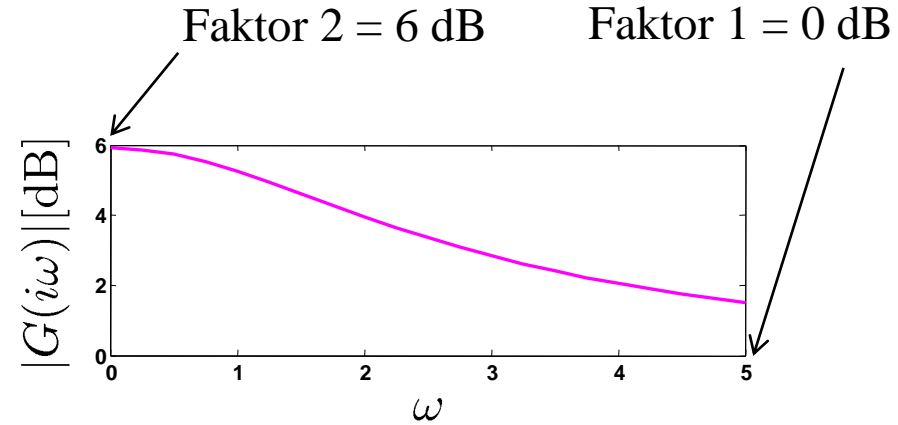
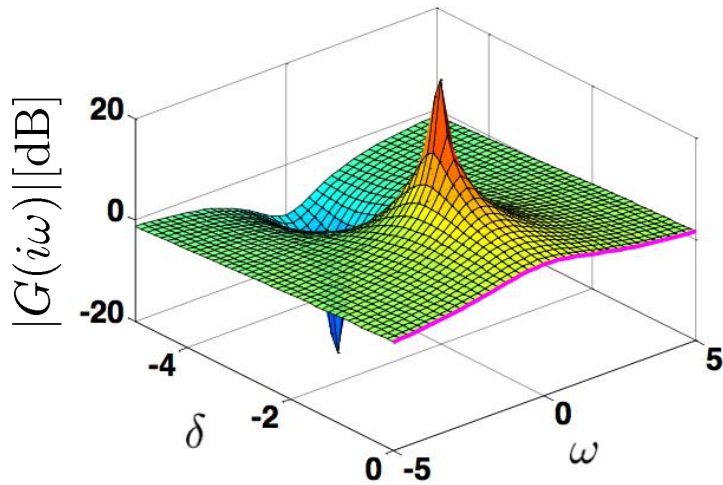
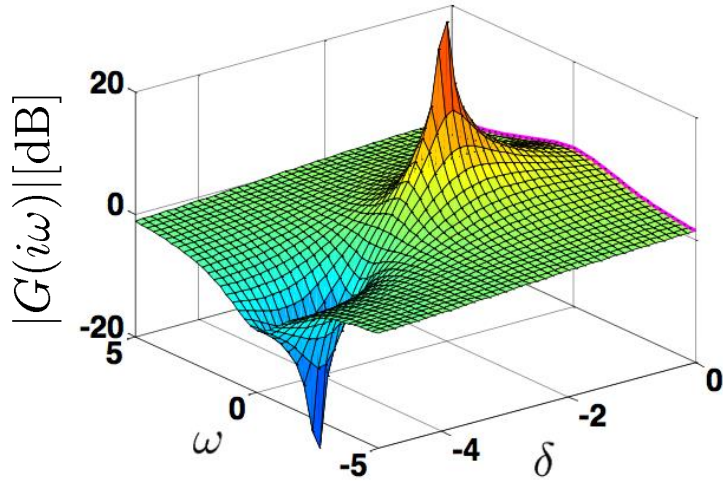
Erstaunlicherweise enthält die (lila) Kurve entlang der Im-Achse $|G(i\omega)|$, also für $s = 0 - i\infty \dots 0 + i\infty$ ($\delta = 0$), die selbe Information wie $|G(s)|$. Gleiches gilt für die Phasen.

Deshalb können wir in Zukunft mit $G(i\omega)$ statt mit $G(s)$ arbeiten, wenn uns dies die Arbeit erleichtert oder anschaulicher macht. Wir benötigen dann nur noch 2-dimensionale Grafiken!



6.1 Definition des Frequenzgangs

Beispiel: Übertragungsfunktion $G(s) = \frac{s + 4}{s + 2}$



6.1 Definition des Frequenzgangs

Wie kann es sein, dass in $G(i\omega)$ die gleiche Information steckt wie in $G(s)$?

- Ähnliche "Merkwürdigkeiten" sind aus der Funktionentheorie reichlich bekannt. Z.B. kommt es bei der Berechnung von Integralen entlang einer geschlossenen Kurve nicht auf den genauen Weg an, sondern nur auf die eingeschlossenen Singularitäten.
- Aus der Funktionentheorie (Analysis für komplexe Zahlen) ist bekannt (Prinzip der analytischen Fortsetzung), dass eine Funktion, definiert auf einem Gebiet (z.B. die imaginäre Achse), nur *eine* Fortsetzung in einem anderen Gebiet (z.B. die ganze s -Ebene) hat, das sich mit dem ersten Gebiet überschneidet.

Dieser Satz erklärt auch, warum es für alle reellen Funktionen eine eindeutige komplexe Erweiterung gibt. Somit ist durch $G(s = i\omega)$ auch $G(s \neq i\omega)$ eindeutig festgelegt und bietet damit keine zusätzliche Information.

- Die Laplace-Rücktransformation benötigt auch nur Informationen entlang einer Linie (parallel zur Im-Achse):

$$g(t) = \frac{1}{2\pi i} \int_{\delta - i\infty}^{\delta + i\infty} G(s) e^{st} ds$$

Offensichtlich reicht dies zur kompletten Berechnung des Zeitsignals aus!

6.1 Definition des Frequenzgangs

Die Abbildung $G(i\omega)$ nennt man **Frequenzgang** eines linearen dynamischen Systems.

Grafische Darstellung des Frequenzgangs:

1. Explizite Funktion der Kreisfrequenz ω :
 - *Amplitudengang*: Betrag des Frequenzgangs in Abhängigkeit von ω .
 - *Phasengang*: Phase des Frequenzgangs in Abhängigkeit von ω .→ Bode-Diagramm
2. Real- und Imaginärteil des Frequenzgangs in der komplexen Ebene mit der Kreisfrequenz ω als Parameter.
→ Ortskurve

Beide Darstellungsweisen sind äquivalent. Die zweite eignet sich nicht so gut für quantitative Analysen, da die genaue Skalierung von ω verloren geht.

6.1 Definition des Frequenzgangs

Interpretation des Frequenzgangs $G(i\omega)$

Anregung mit einem sinusförmigen Eingangssignal:

$$u(t) = \sin(\omega t)$$

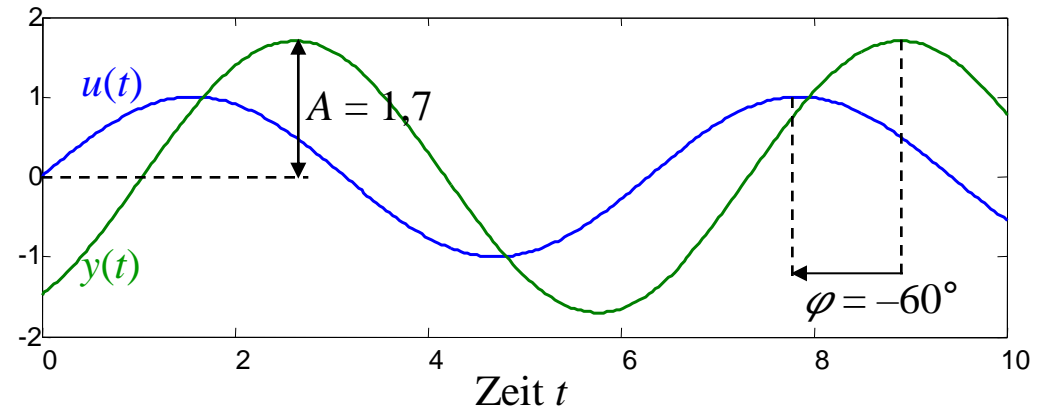
Nach Einschwingen (Abklingen der homogenen Lösung der DGL) ergibt sich ein sinusförmiges Ausgangssignal mit der selben Frequenz:

$$y(t) = A(\omega) \cdot \sin(\omega t + \varphi(\omega))$$

Die Amplitudenverstärkung und die Phasenverschiebung sind frequenzabhängig!

Frequenzgang:

$$G(i\omega) = A(\omega)e^{i\varphi(\omega)}$$

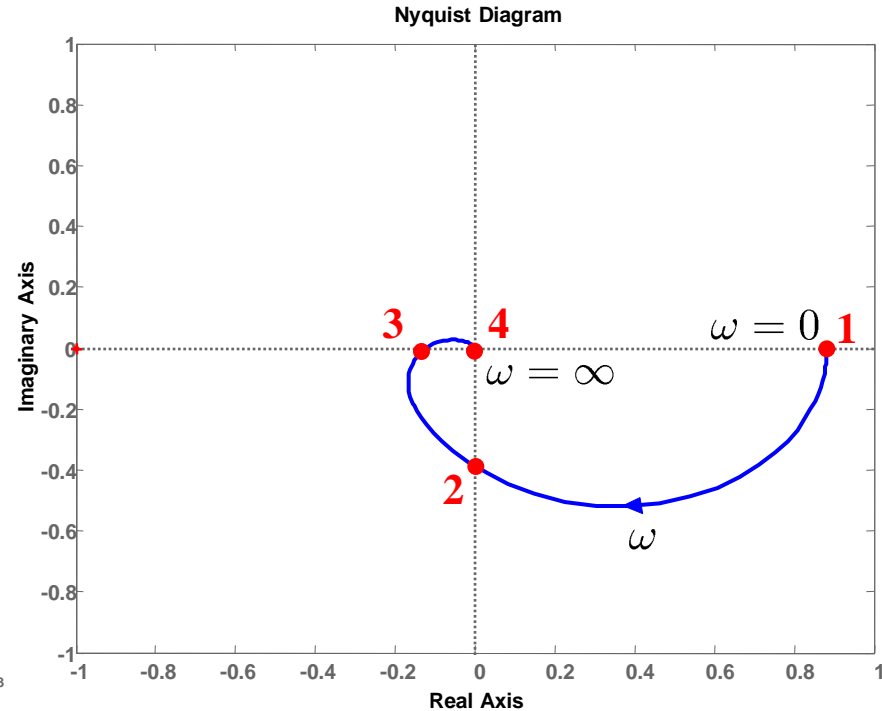
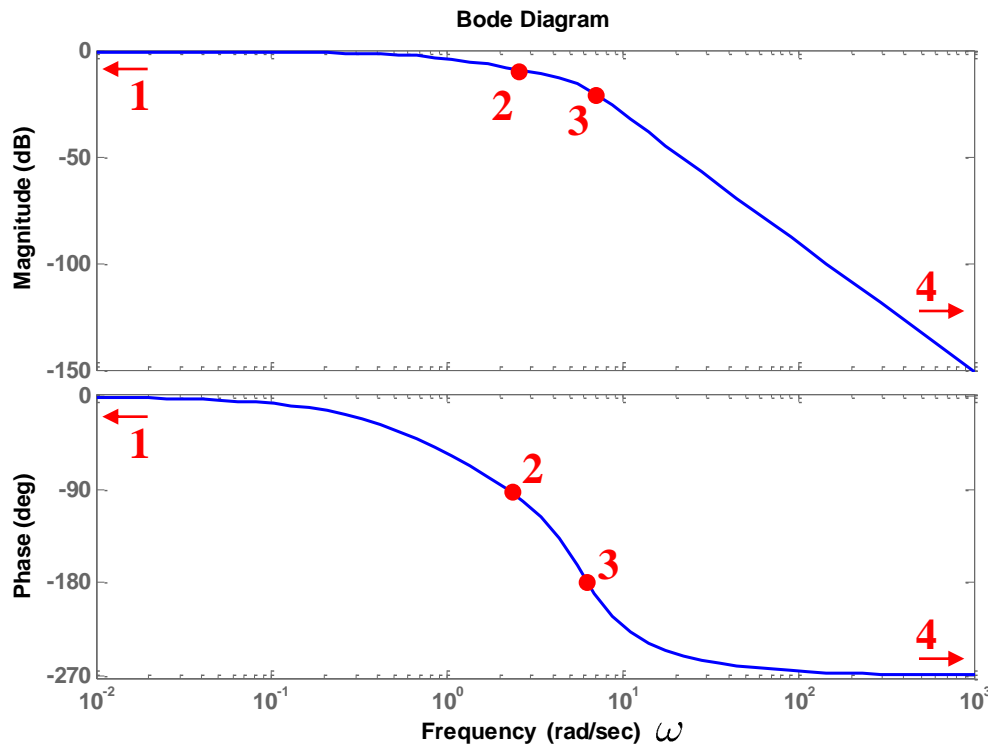


Berechnen oder messen wir $A(\omega)$ für alle Werte von ω und tragen sie in ein Diagramm ein, so ergibt sich der *Amplitudengang*. Tun wir dies für $\varphi(\omega)$, so ergibt sich der *Phasengang*.

6.2 Beziehung zwischen Frequenzgang und Ortskurve

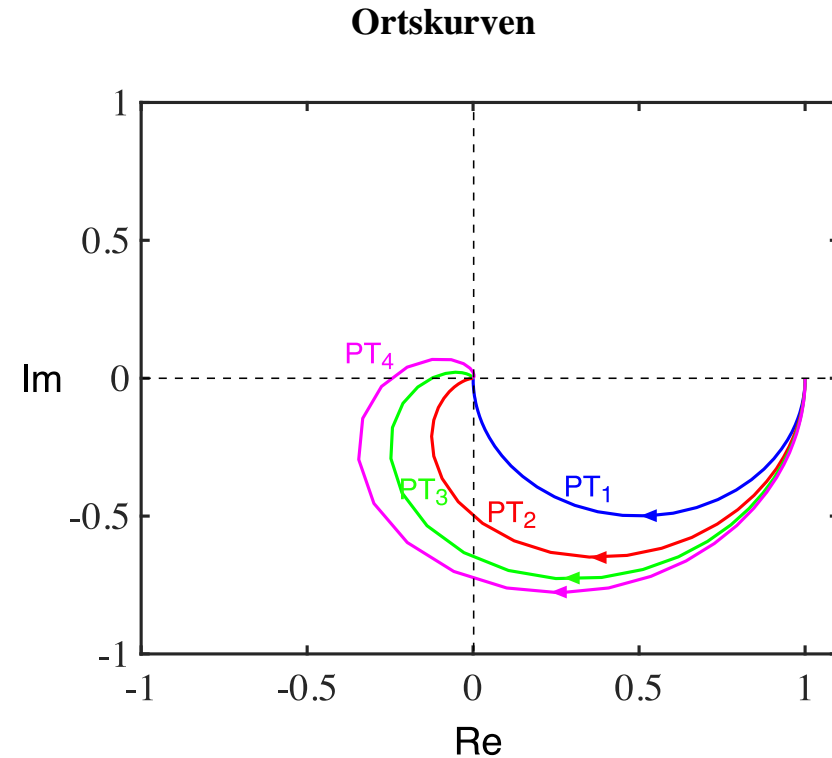
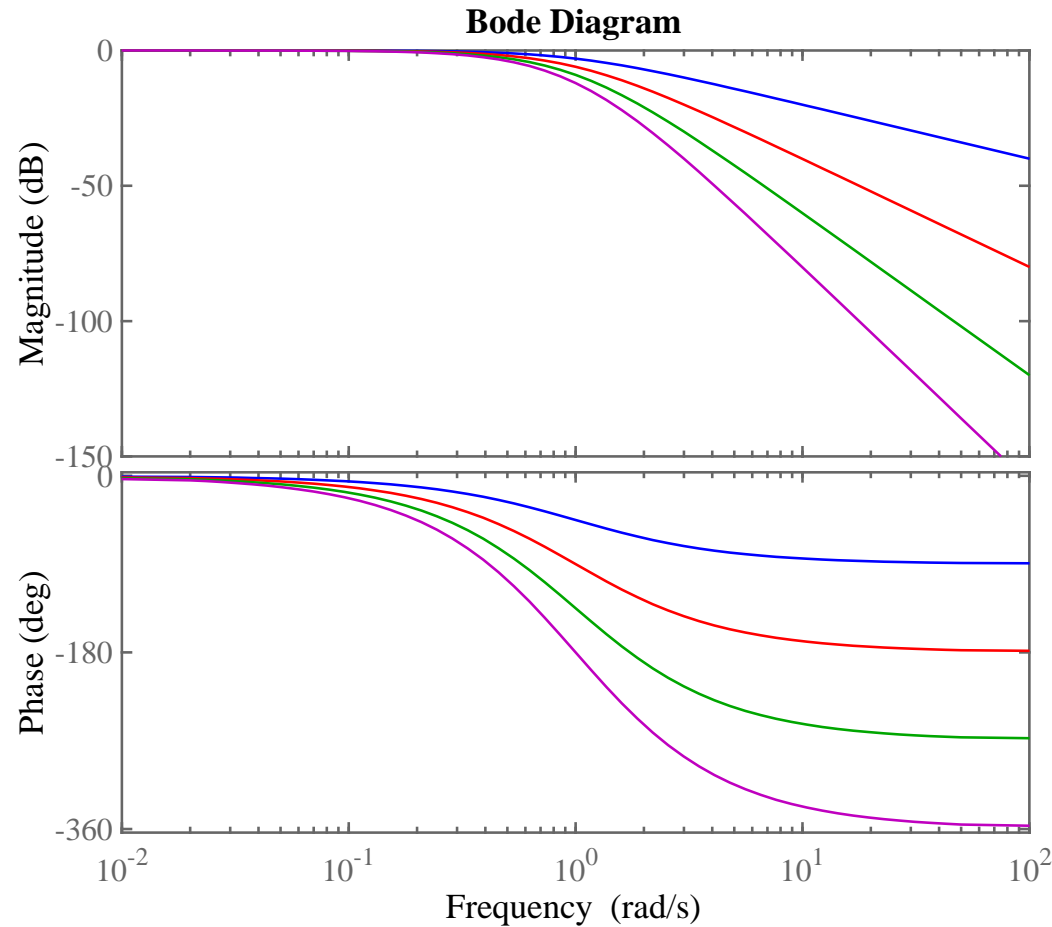
Beispiel: Bode-Diagramm und Ortskurve

$$G(s) = \frac{30}{s^3 + 7s^2 + 40s + 34} = \frac{30}{(s + 1)(s + 3 - i5)(s + 3 - i5)}$$



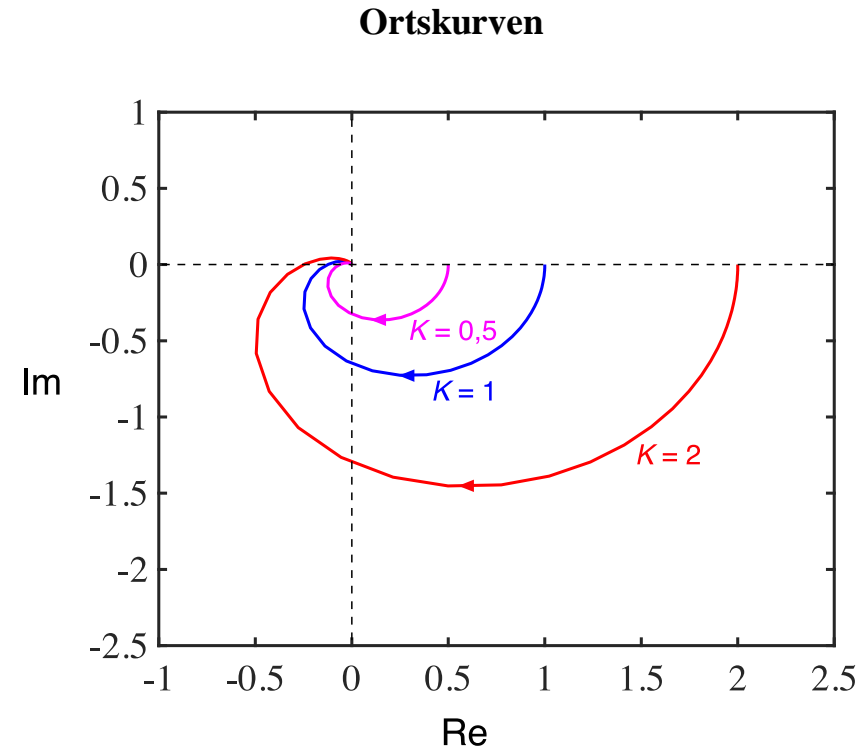
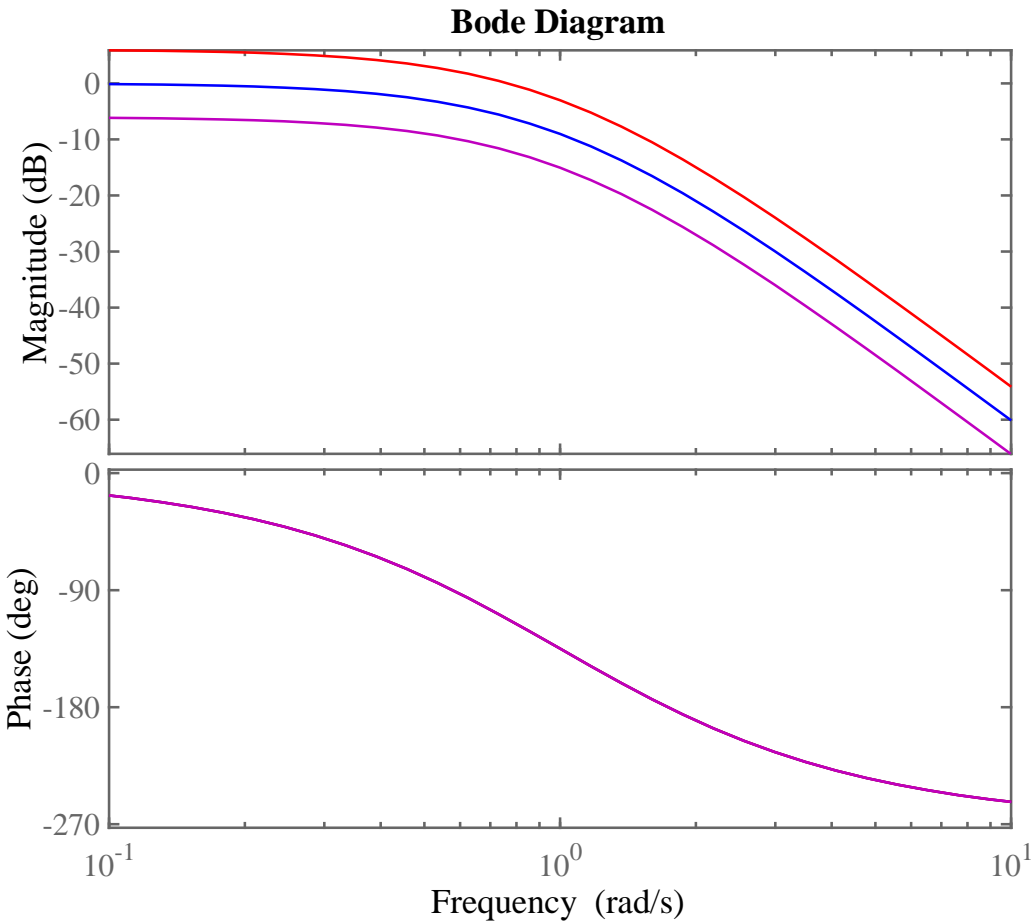
6.2 Beziehung zwischen Frequenzgang und Ortskurve

Verschiedene Ordnungen



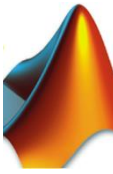
6.2 Beziehung zwischen Frequenzgang und Ortskurve

Verschiedene Verstärkungen



6.2 Beziehung zwischen Frequenzgang und Ortskurve

MATLAB



Zeichnen eines Bode-Diagramms:

```
s=tf('s');  
G = (s+3)/(5*s^3+3*s^2+2*s+1);  
bode(G); % Zeichnet Bode-Diagramm von G  
oder  
[Betrag , Phase, Omega] = bode(G); % Liefert Zahlenwerte; Skalierung  
% der Frequenzachse automatisch
```

Zeichnen einer Ortskurve:

```
nyquist(G); % Zeichnet Ortskurve von G  
% Plot für  $\omega = -\infty \dots \infty$   
oder  
[Realteil, Imaginarteil, Omega] = nyquist(G); % s.o.
```

6.3 Bode-Diagramm

Eigenschaften des Bode-Diagramms:

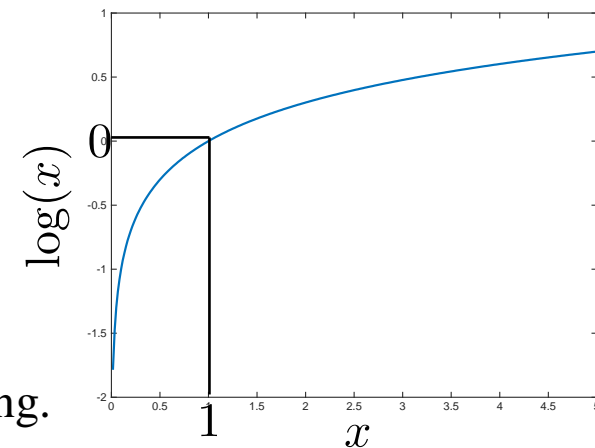
- Amplitude und Phase des Frequenzgangs werden in zwei Diagrammen über der Kreisfrequenz $\omega = 2\pi f$ abgetragen. Die Diagramme nennt man Amplitudengang bzw. Phasengang.
- Kreisfrequenz wird logarithmisch skaliert (... 0,01 – 0,1 – 1 – 10 – 100 ...), um einen großen Frequenzbereich übersichtlich darstellen zu können. D.h. $\omega = 0$ rad/s liegt unendlich weit "links" im Diagramm und $\omega = \infty$ rad/s liegt unendlich weit "rechts".
- Phase wird linear skaliert und typischerweise in der Einheit "Grad" angegeben.
- Amplitude wird logarithmisch skaliert, um einen großen Amplitudenbereich übersichtlich darstellen zu können. D.h. Amplitude 0 liegt unendlich weit "unten" und Amplitude ∞ liegt unendlich weit "oben". Die Amplitude wird in Dezibel angegeben:

$$|G(i\omega)|_{\text{dB}} = 20\log|G(i\omega)|$$

Faktor 10: 20 dB

Faktor 2 : 6 dB

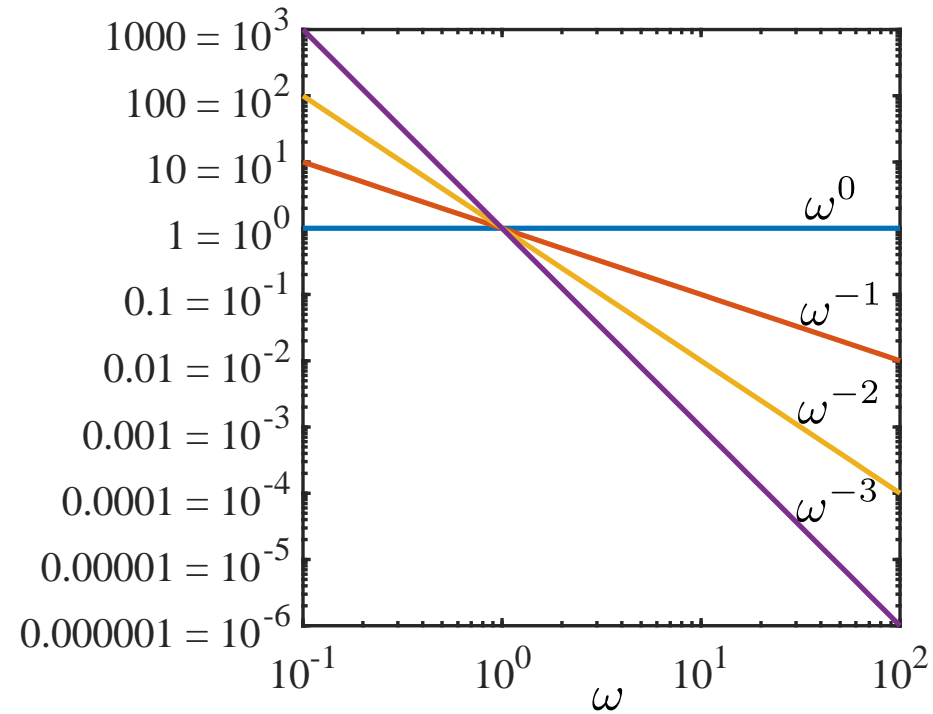
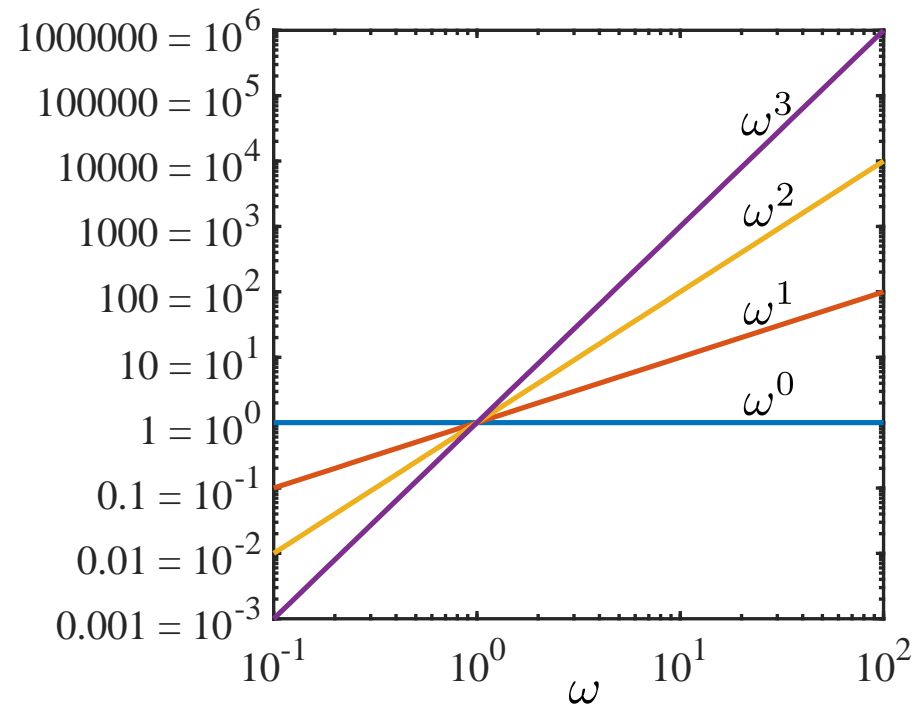
Faktor $\sqrt{2}$: 3 dB



6.3 Bode-Diagramm

Doppelt logarithmische Darstellung

- Alle Potenzgesetze ω^n lassen sich als Geraden darstellen.
- Die Höhe der Potenz n entspricht der Steigung der Geraden.
- Zähler der Übertragungsfunktion = positive Potenzen = Anstieg
- Nenner der Übertragungsfunktion = negative Potenzen = Abfall



6.3 Bode-Diagramm

Logarithmische Skalierung der Amplituden:

- Multiplikation von Übertragungsfunktionen entspricht *Addition* der Amplitudengänge

$$G(i\omega) = G_1(i\omega)G_2(i\omega) \rightarrow |G(i\omega)| = |G_1(i\omega)||G_2(i\omega)|$$

$$\rightarrow 20\log|G(i\omega)| = 20\log(|G_1(i\omega)||G_2(i\omega)|) = 20\log|G_1(i\omega)| + 20\log|G_2(i\omega)|$$

$$\rightarrow \boxed{|G(i\omega)|_{\text{dB}} = |G_1(i\omega)|_{\text{dB}} + |G_2(i\omega)|_{\text{dB}}}$$

Lineare Skalierung der Phasen:

- Multiplikation von Übertragungsfunktionen entspricht *Addition* der Phasengänge

$$G(i\omega) = G_1(i\omega)G_2(i\omega) \rightarrow e^{i\varphi(\omega)} = e^{i\varphi_1(\omega)}e^{i\varphi_2(\omega)} = e^{i\varphi_1(\omega)+i\varphi_2(\omega)}$$

$$\rightarrow \boxed{\varphi(\omega) = \varphi_1(\omega) + \varphi_2(\omega)}$$

6.3 Bode-Diagramm

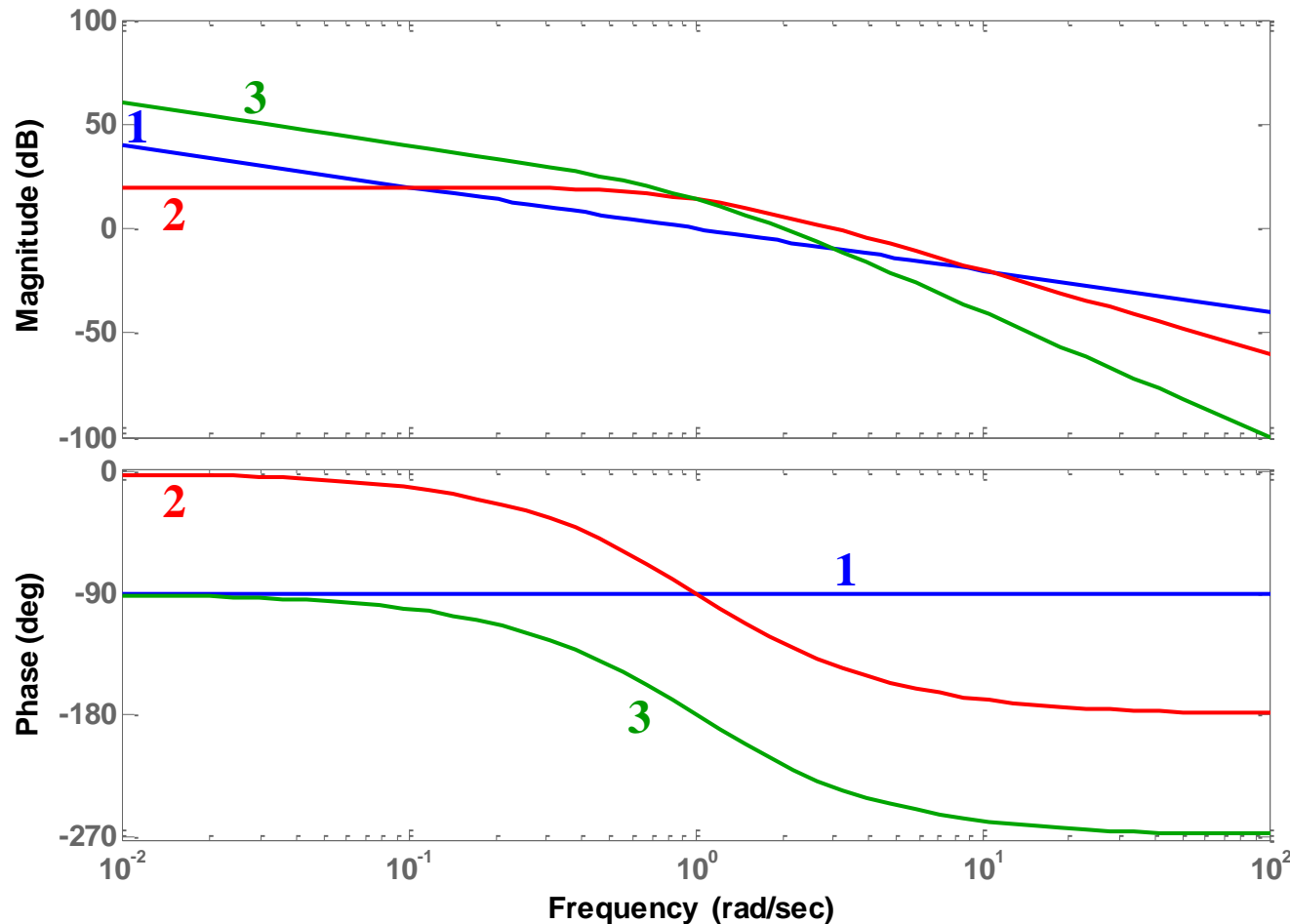
Beispiel:

$$G_1(s) = \frac{1}{s}$$

$$G_2(s) = \frac{10}{s^2 + 2s + 1}$$

$$G(s) = G_1(s)G_2(s)$$

Bode Diagram



**Kurve 3 =
Kurve 1
+ Kurve 2**

6.3 Bode-Diagramm

Beispiel: $G(s) = \frac{K}{1 + Ts}$

Zerlegung in Real- und Imaginärteil (für Ortskurve):

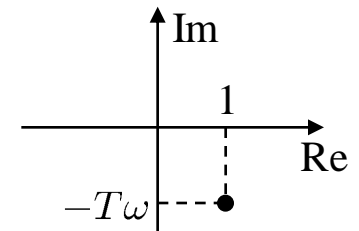
$$G(i\omega) = \frac{K}{1 + iT\omega} = \frac{K(1 - iT\omega)}{(1 + iT\omega)(1 - iT\omega)} = \frac{K(1 - iT\omega)}{1 + T^2\omega^2}$$

$$\operatorname{Re}\{G(i\omega)\} = \frac{K}{1 + T^2\omega^2} \quad \operatorname{Im}\{G(i\omega)\} = \frac{-KT\omega}{1 + T^2\omega^2}$$

Zerlegung in Betrag und Phase (für Bode-Diagramm):

$$|G(i\omega)| = \frac{K}{\sqrt{1 + T^2\omega^2}} \quad \varphi(\omega) = \operatorname{arctan} \frac{-T\omega}{1} = -\operatorname{arctan}(T\omega)$$

Trick: Vorzeichen für Imaginär- und Realteil getrennt verwalten! Daraus ergibt sich der richtige Quadrant.



6.3 Bode-Diagramm

Asymptoten für Betrag

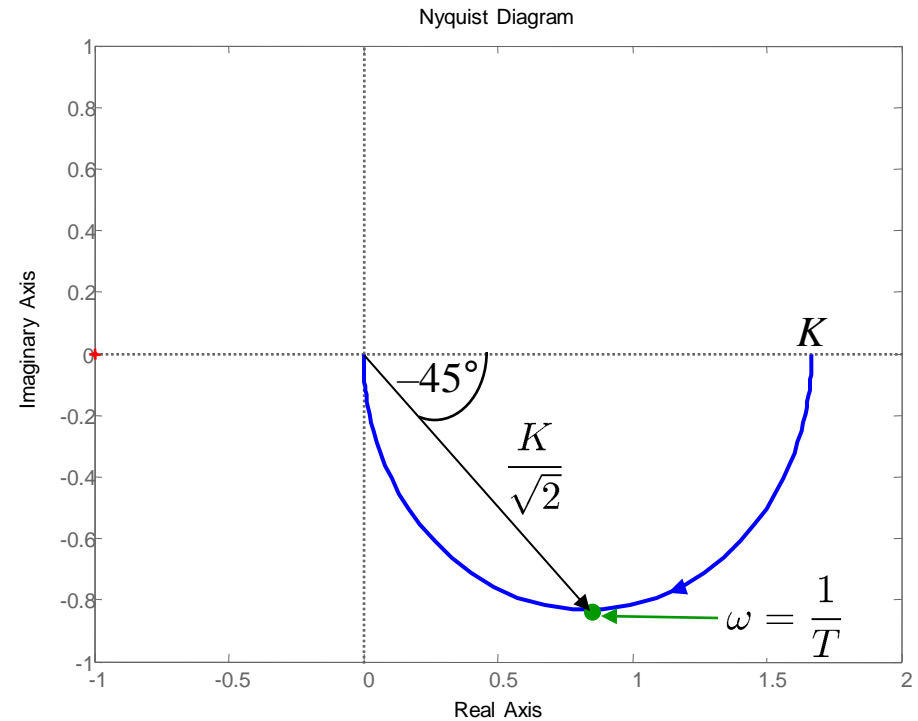
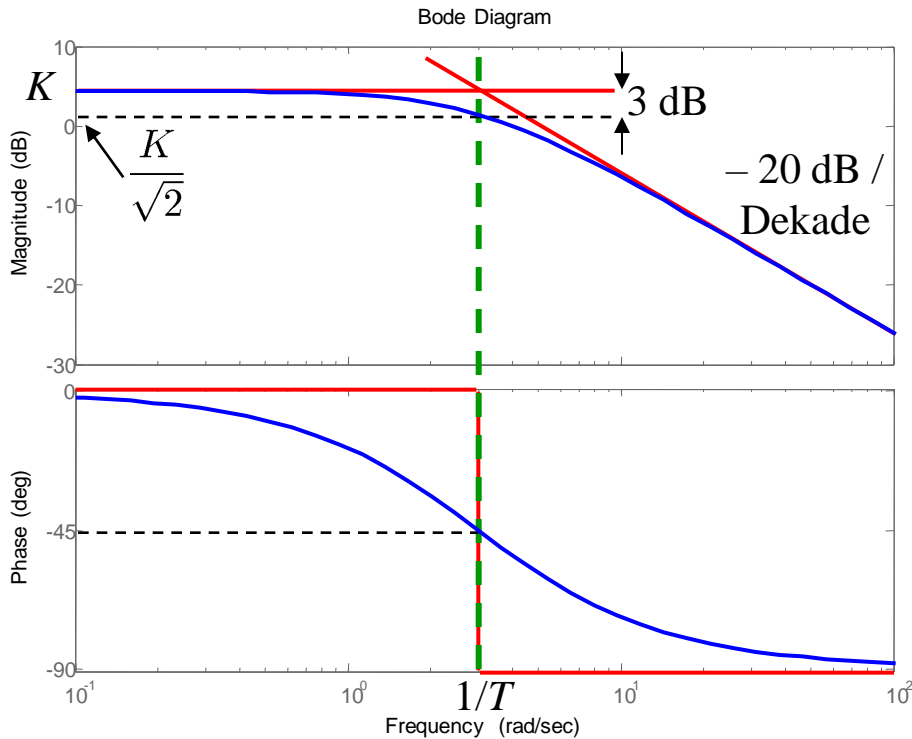
$$\omega \ll \frac{1}{T} : |G(i\omega)| = K$$

$$\omega \gg \frac{1}{T} : |G(i\omega)| = \frac{K}{T} \frac{1}{\omega}$$

Asymptoten für Phase

$$\varphi(\omega) = -\arctan(0) = 0^\circ$$

$$\varphi(\omega) = -\arctan(\infty) = -90^\circ$$



6.3 Bode-Diagramm

Produkt-Standardform einer Übertragungsfunktion:

$$G(s) = k \frac{(s - n_1)(s - n_2) \cdot \dots \cdot (s - n_m)}{(s - p_1)(s - p_2) \cdot \dots \cdot (s - p_n)} e^{-T_t s}$$

Hier tauchen die Pole und Nullstellen explizit auf. Alternativ kann man auch folgende Schreibweise mit *Zeitkonstanten* wählen:

$$G(s) = K \frac{(T_{n1}s + 1)(T_{n2}s + 1) \cdot \dots \cdot (T_{nm}s + 1)}{(T_{p1}s + 1)(T_{p2}s + 1) \cdot \dots \cdot (T_{pn}s + 1)} e^{-T_t s}$$

Die Pole und Nullstellen lassen sich leicht in die Zeitkonstanten umrechnen:

$$T_{n1} = -\frac{1}{n_1}, \quad \dots \quad T_{nm} = -\frac{1}{n_m}; \quad T_{p1} = -\frac{1}{p_1}, \quad \dots \quad T_{pn} = -\frac{1}{p_n}$$

Der Vorteil der Zeitkonstanten-Darstellung ist, dass die Verstärkung K des Systems unabhängig von den anderen Systemparametern wird, während sich k ändern muss, um die Systemverstärkung konstant zu halten, wenn sich ein Pol oder eine Nullstelle verändert:

$$K = G(0) = k \frac{(-n_1)(-n_2) \cdot \dots \cdot (-n_m)}{(-p_1)(-p_2) \cdot \dots \cdot (-p_n)}$$

6.3 Bode-Diagramm

Produkt-Standardform einer Übertragungsfunktion in Zeitkonstanten-Darstellung:

$$G(s) = K \frac{(T_{n1}s + 1)(T_{n2}s + 1) \cdot \dots \cdot (T_{nm}s + 1)}{(T_{p1}s + 1)(T_{p2}s + 1) \cdot \dots \cdot (T_{pn}s + 1)} e^{-T_t s}$$

Logarithmus des Betrags der Übertragungsfunktion:

$$\begin{aligned} 20 \log |G(i\omega)| &= 20 \log \left| K \frac{(iT_{n1}\omega + 1)(iT_{n2}\omega + 1) \cdot \dots \cdot (iT_{nm}\omega + 1)}{(iT_{p1}\omega + 1)(iT_{p2}\omega + 1) \cdot \dots \cdot (iT_{pn}\omega + 1)} e^{-i\omega T_t} \right| \\ &= 20 \log(K) + 20 \sum_{i=1}^m \log |iT_{ni}\omega + 1| - 20 \sum_{i=1}^n \log |iT_{pi}\omega + 1| \end{aligned}$$

Phase der Übertragungsfunktion:

$$\begin{aligned} \angle(G(i\omega)) &= \angle \left(K \frac{(iT_{n1}\omega + 1)(iT_{n2}\omega + 1) \cdot \dots \cdot (iT_{nm}\omega + 1)}{(iT_{p1}\omega + 1)(iT_{p2}\omega + 1) \cdot \dots \cdot (iT_{pn}\omega + 1)} e^{-i\omega T_t} \right) \\ &= \angle(K) + \sum_{i=1}^m \angle(iT_{ni}\omega + 1) - \sum_{i=1}^n \angle(iT_{pi}\omega + 1) - \omega T_t \end{aligned}$$

6.3 Bode-Diagramm

Weit "links" von einer Nullstelle n gilt ($\omega \ll |n|$ oder $|T_n|\omega \ll 1$):

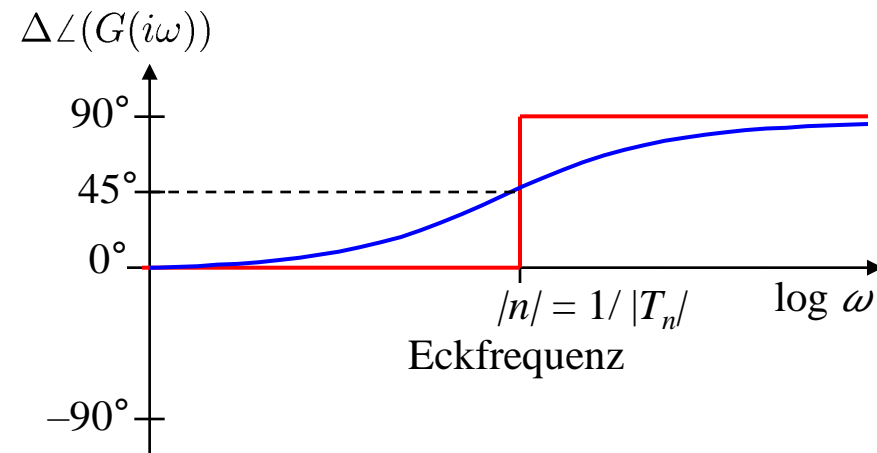
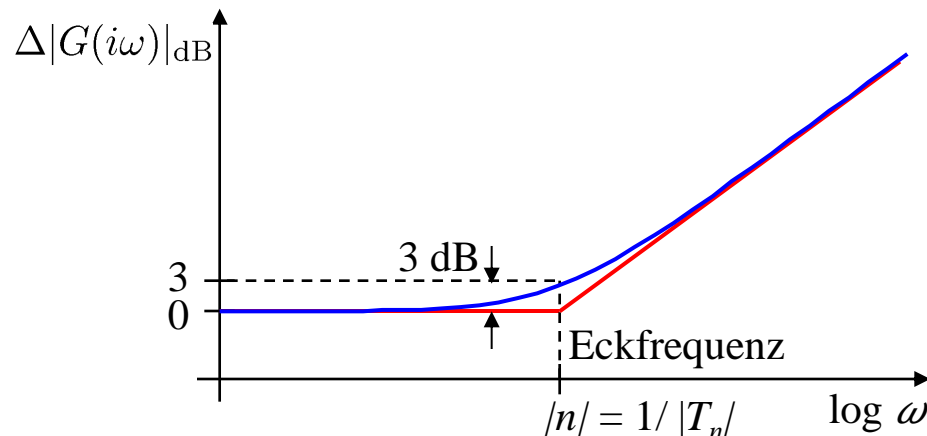
$$\Delta|G(i\omega)|_{\text{dB}} = 20\log|iT_n\omega + 1| \rightarrow +20\log|1| = 0 \text{ dB}$$

$$\Delta\angle(G(i\omega)) = \angle(iT_n\omega + 1) \rightarrow 0^\circ$$

Weit "rechts" von einer Nullstelle n gilt ($\omega \gg |n|$ oder $|T_n|\omega \gg 1$):

$$\Delta|G(i\omega)|_{\text{dB}} = 20\log|iT_n\omega + 1| \rightarrow +20\log|T_n\omega| = 20\log|T_n| + 20\log|\omega|$$

$$\Delta\angle(G(i\omega)) = \angle(iT_n\omega + 1) \rightarrow +90^\circ$$



6.3 Bode-Diagramm

Einfluss eines Pols ist invers zum Einfluss einer Nullstelle!

Weit "links" von einem *Pol* p gilt ($\omega \ll |p|$ oder $|T_p|\omega \ll 1$):

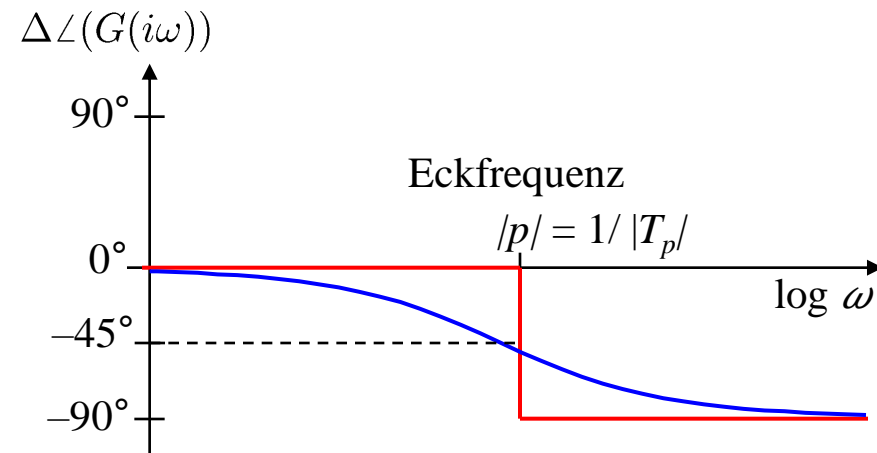
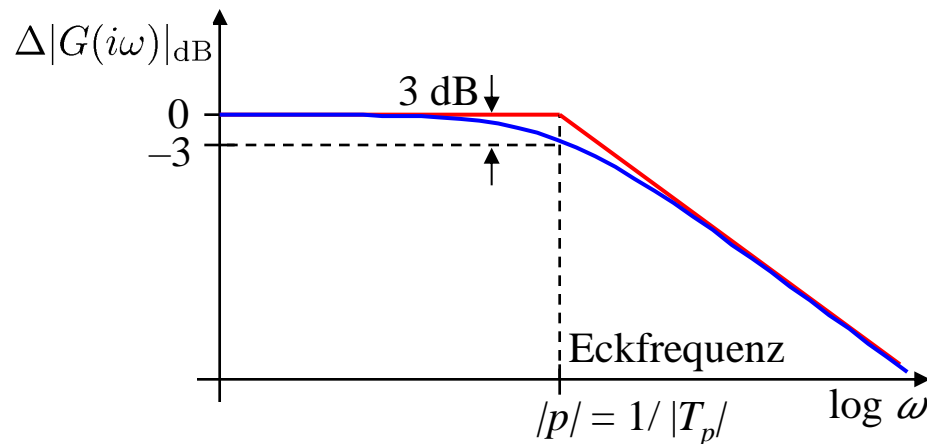
$$\Delta|G(i\omega)|_{\text{dB}} = -20\log|iT_p\omega + 1| \rightarrow -20\log|1| = 0 \text{ dB}$$

$$\Delta\angle(G(i\omega)) = -\angle(iT_p\omega + 1) \rightarrow 0^\circ$$

Weit "rechts" von einem *Pol* p gilt ($\omega \gg |p|$ oder $|T_p|\omega \gg 1$):

$$\Delta|G(i\omega)|_{\text{dB}} = -20\log|iT_p\omega + 1| \rightarrow -20\log|T_p\omega| = -20\log|T_p| - 20\log|\omega|$$

$$\Delta\angle(G(i\omega)) = -\angle(iT_p\omega + 1) \rightarrow -90^\circ$$



6.3 Bode-Diagramm

Behandlung von konjugiert komplexen Nullstellen und Polen:

1. Konjugiert komplexer Pol bzw. Nullstelle werden in folgende Form gebracht:

$$(s - p - i\omega)(s - p + i\omega) = (s^2 + 2D\omega_0s + \omega_0^2) = \omega_0^2 \left(\frac{1}{\omega_0^2}s^2 + \frac{2D}{\omega_0}s + 1 \right)$$

2. Für Zeichnung der Asymptoten: Behandlung wie bei einem reellen *zweifachen* Pol bzw. Nullstelle bei $p = -\omega_0$. D.h. die Asymptoten sind die selben wie für den Ausdruck:

$$(s + \omega_0)^2 = (s^2 + 2\omega_0s + \omega_0^2) = \omega_0^2 \left(\frac{1}{\omega_0^2}s^2 + \frac{2}{\omega_0}s + 1 \right)$$

3. Für den exakten Verlauf des Amplitudengangs ist aber der Wert der Dämpfung D von entscheidender Bedeutung. Für Dämpfungen $D < 1/\sqrt{2}$ weist der Amplitudengang eine Resonanzstelle (Überhöhung) auf. Für größere Dämpfungen nicht. Für Werte $D > 1$ ergeben sich gar keine konjugiert komplexen Pole bzw. Nullstellen sondern reelle.

Weitere Details zu solchen Systemen 2. Ordnung werden in Kap. 7 behandelt...

6.4 Konstruktionsregeln für phasenminimale Systeme

Annahmen (zur Vereinfachung):

- Alle Pole liegen in der linken komplexen Halbebene (haben negativen Realteil).
- Alle Nullstellen liegen in der linken komplexen Halbebene (haben negativen Realteil).
- Keine Totzeit, also $T_t = 0$ s.

Bemerkung: Die beiden ersten Annahmen implizieren, dass keine Pole und Nullstellen bei $s = 0$ liegen. Dadurch wird das Problem vereinfacht. Das ist aber *nicht* notwendig, um Phasenminimalität sicher zu stellen. Wir betrachten diese Fälle etwas später...

Anmerkungen zur Notation:

- Wir wissen bereits: Pole in der linken s -Halbebene führen auf stabiles Verhalten (abklingende e-Funktionen in der Zeitlösung), Pole in der rechten s -Halbebene führen auf instabiles Verhalten (aufklingende e-Funktionen in der Zeitlösung). Deshalb spricht man verkürzt oft auch von *stabilen* und *instabilen Polen*. Obwohl Nullstellen *keine* Auswirkungen auf die Stabilität haben spricht man analog auch von stabilen und instabilen Nullstellen. Im Englischen ist zusätzlich auch die Bezeichnung *minimum-phase* und *nonminimum-phase zeros* geläufig.

6.4 Konstruktionsregeln für phasenminimale Systeme

Konstruktion der Asymptoten des Amplitudengangs im Bode-Diagramm:

1. Schreibe Übertragungsfunktion in Produkt-Standardform.
2. Ordne Pole und Nullstellen in betragsmäßig aufsteigender Reihenfolge.
3. Zeichne eine Gerade der Steigung 0 mit Amplitude K für $\omega = 0$ ($-\infty$ auf logarithmisch skaliertes Achse) bis $\omega =$ Kreisfrequenz des betragsmäßig kleinsten Pols bzw. Nullstelle.
4. Zeichne ausgehend vom rechten Ende der vorherigen Geraden eine neue Gerade bis $\omega =$ Kreisfrequenz des betragsmäßig nächsten Pols bzw. Nullstelle. Die Steigung dieser neuen Geraden ist gleich der Steigung der vorherigen Geraden -20 dB / Dekade für einfache Pole, -40 dB / Dekade für zweifache Pole, ... und $+20$ dB / Dekade für einfache Nullstellen, $+40$ dB / Dekade für zweifache Nullstellen, ...
5. Gehe zu Schritt 4 bis der betragsmäßig größte Pol bzw. Nullstelle erreicht ist. Die sich daran anschließende letzte Gerade wird ebenfalls nach Schritt 4 konstruiert, nur dass sie bis $\omega = +\infty$ fortgeführt wird.

Bemerkung: Konjugiert komplexe Pol- bzw. Nullstellenpaare werden als zweifache Pol- bzw. Nullstelle bei $-\omega_0$ gezählt.

6.4 Konstruktionsregeln für phasenminimale Systeme

Konstruktion der Asymptoten des Phasengangs im Bode-Diagramm:

- Die Einteilung entlang der Frequenzachse ist identisch mit dem Amplitudengang.
- Die Asymptoten sind Geraden mit Steigung 0.
- Das "Runterklappen" der Asymptoten um 20 dB / Dekade des Amplitudengangs bei Polen entspricht einer Verschiebung der Phasengang-Asymptote um -90° , das "Hochklappen" der Asymptoten um 20 dB / Dekade des Amplitudengangs bei Nullstellen entspricht einer Verschiebung der Phasengang-Asymptote um $+90^\circ$.
- D.h. zur Asymptote des Amplitudengangs mit $n \cdot 20$ dB / Dekade Steigung gehört eine Asymptote des Phasengangs mit $n \cdot 90^\circ$ Phase ($n = \dots, -2, -1, 0, 1, 2, \dots$)! Die Phase lässt sich also direkt aus der Amplitude berechnen (oder umgekehrt).

Zur Bezeichnung "phasenminimal":

Phasenminimale Systeme haben von allen Systemen mit gleichem Amplitudengang die minimale Phase. Ein lineares System ist phasenminimal, wenn seine Pole und Nullstellen in der linken s -Halbebene liegen und es keine Totzeit aufweist. Weil sie den kleinstmöglichen Phasenverzug aufweisen, lassen sich phasenminimale Systeme leichter und besser regeln als nicht phasenminimale gleicher Komplexität.

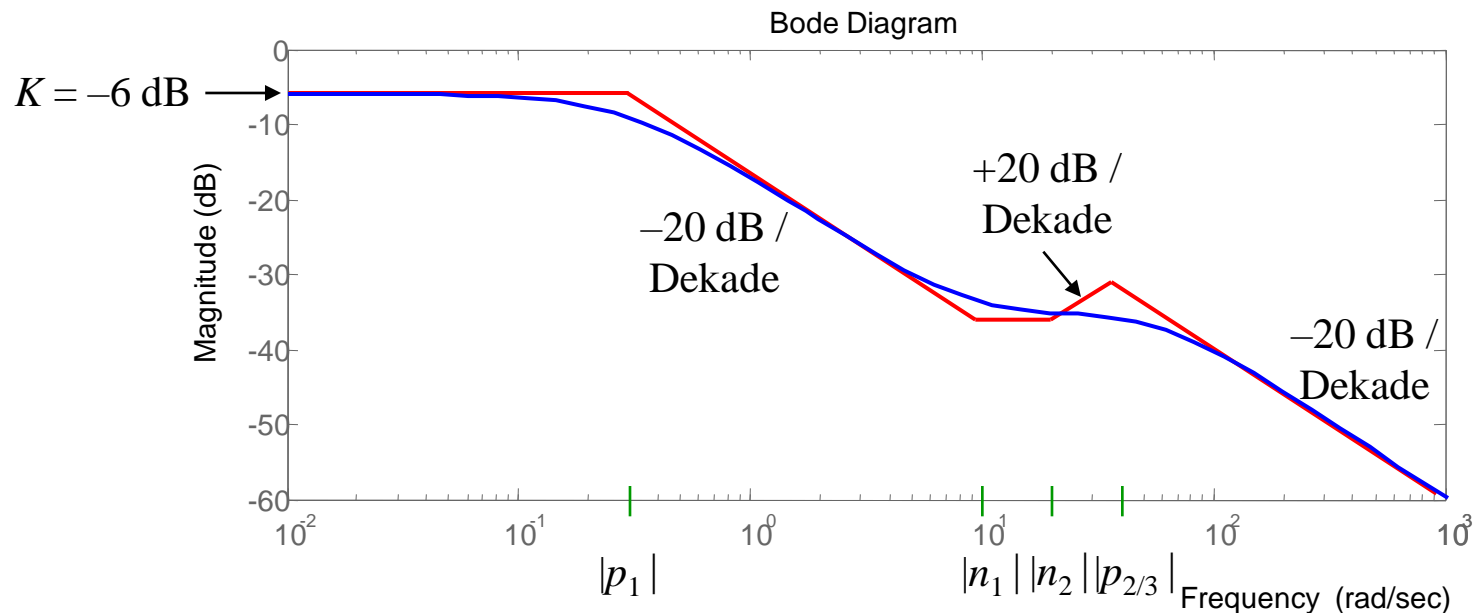
6.4 Konstruktionsregeln für phasenminimale Systeme

Beispiel:
$$G(s) = \frac{6}{5} \frac{(s + 10)(s + 20)}{(s + 0.3)(s + 40)^2}$$

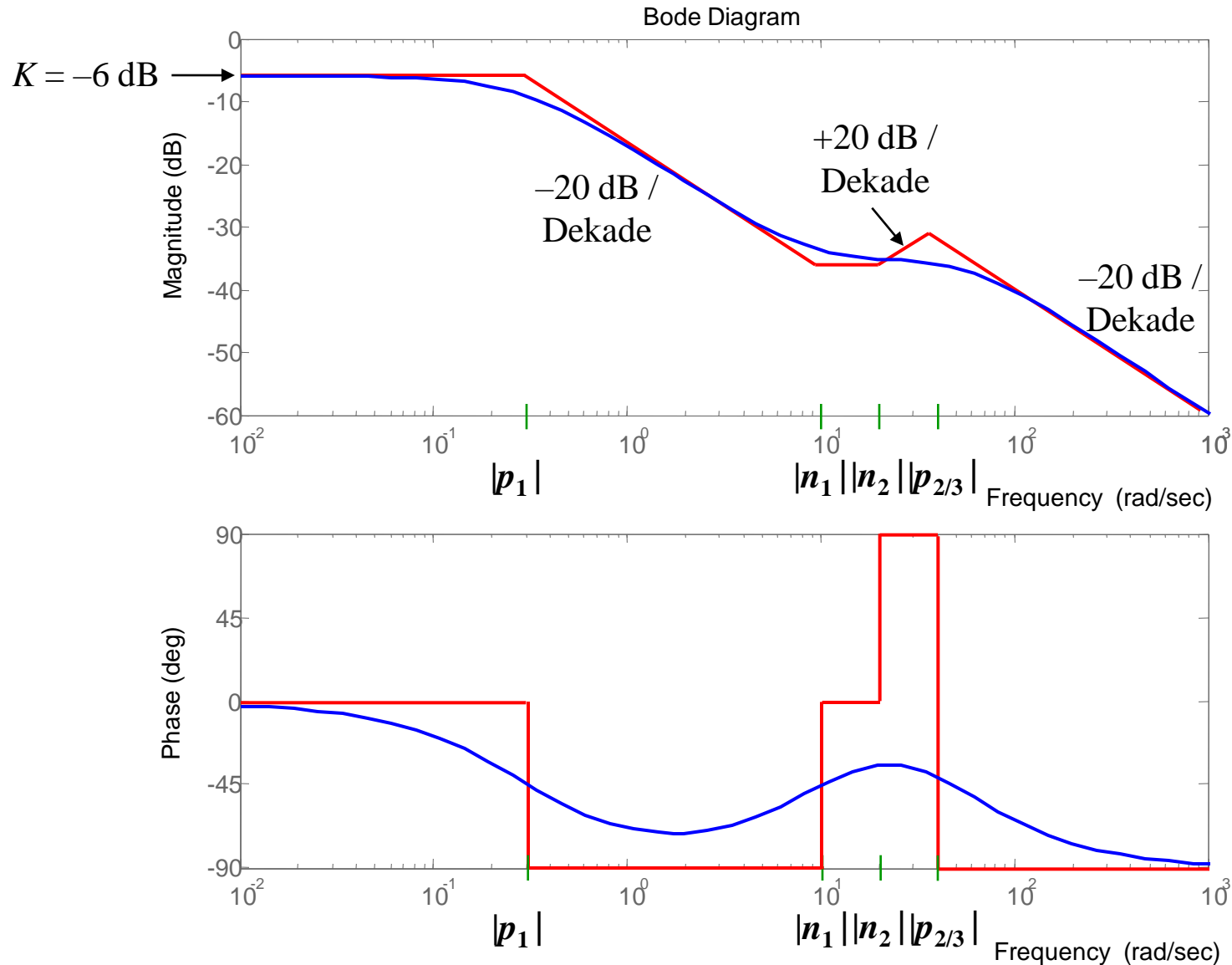
Verstärkung:
$$K = \frac{6}{5} \cdot \frac{10 \cdot 20}{0.3 \cdot 1600} = \frac{1}{2} = -6 \text{ dB}$$

Betragsmäßiges Sortieren der Pole und Nullstellen von unten nach oben:

Pol $p_1 = -0.3$, Nullstelle $n_1 = -10$, Nullstelle $n_2 = -20$, Doppelpol $p_{2/3} = -40$.



6.4 Konstruktionsregeln für phasenminimale Systeme



6.4 Konstruktionsregeln für phasenminimale Systeme

Behandlung von Nullstellen und Polen bei 0 (Faktoren s im Zähler oder Nenner):

- Es gilt immer der Fall weit "rechts" von der Nullstelle ($n = 0$) bzw. dem Pol ($p = 0$).

- Pro Nullstelle bei $n = 0$ gilt daher:

$$\Delta|G(i\omega)|_{\text{dB}} = +20\log|\omega|$$

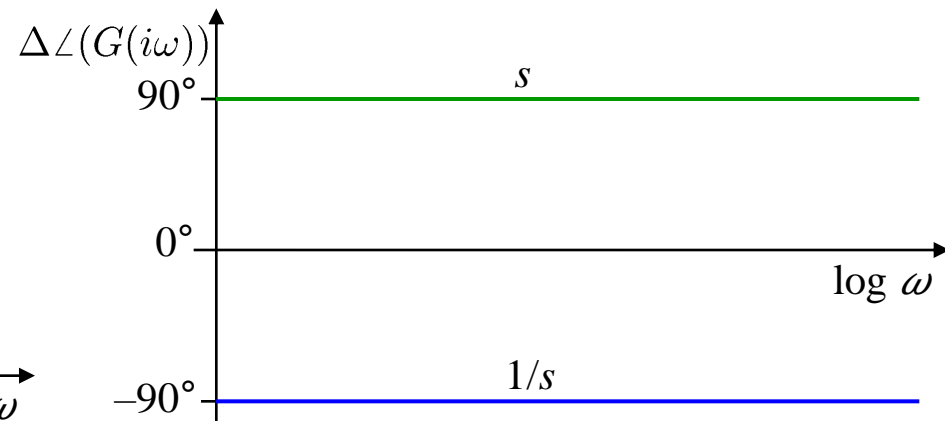
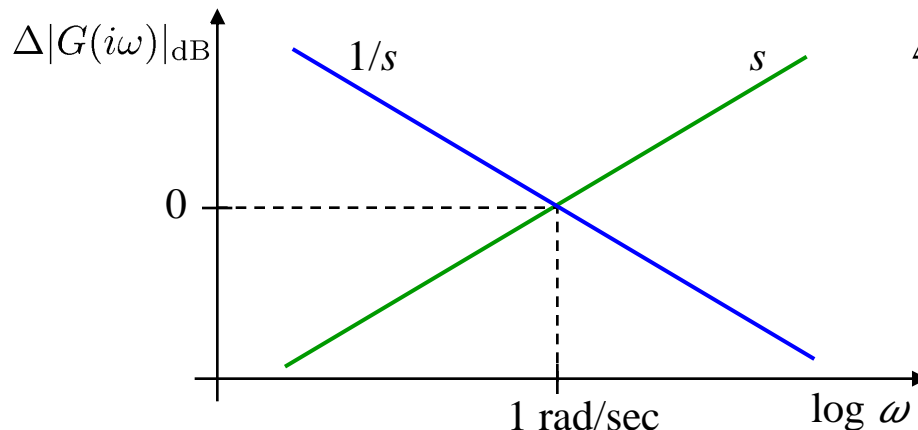
$$\Delta\angle(G(i\omega)) = +90^\circ$$

- Pro Pol bei $p = 0$ gilt daher:

$$\Delta|G(i\omega)|_{\text{dB}} = -20\log|\omega|$$

$$\Delta\angle(G(i\omega)) = -90^\circ$$

- **ACHTUNG:** Diese Werte gelten exakt, nicht nur asymptotisch!



6.4 Konstruktionsregeln für phasenminimale Systeme

Anmerkungen zu Nullstellen und Polen bei 0:

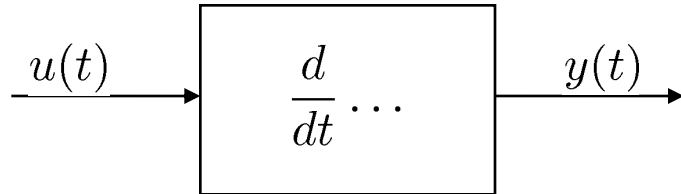
- Nullstelle bei 0 ist ein Differenzierglied: $sX(s) \bullet \circ \frac{d}{dt}x(t)$
 1. Hebt Phase um 90° an.
 2. Hebt Amplitude proportional zur Frequenz an.
- Pol bei 0 ist ein Integrierglied: $\frac{1}{s}X(s) \bullet \circ \int_0^t x(\tau)d\tau$
 1. Senkt Phase um 90° ab.
 2. Senkt Amplitude proportional zur Frequenz ab.

Vertikale Ausrichtung des Amplitudengangs:

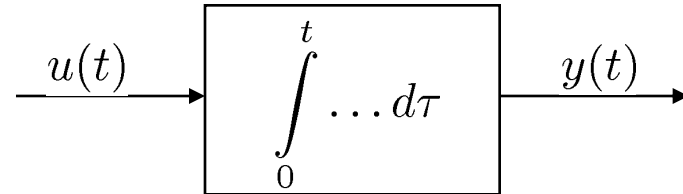
- Nullstelle bei 0: An der Stelle $\omega = 1$ rad/sec hat die Amplitude den Wert $1 = 0$ dB.
- Pol bei 0: An der Stelle $\omega = 1$ rad/sec hat die Amplitude den Wert $1/\omega = 1 = 0$ dB.
- Für das Gesamtsystem hat die Amplitude an der Stelle $\omega = 1$ rad/sec somit den Wert K .
- Liegen links von $\omega = 1$ rad/sec weitere Nullstellen oder Pole, so muss die Gerade in Gedanken verlängert werden und diese Verlängerung bei $\omega = 1$ rad/sec durch den Wert K verlaufen.

6.4 Konstruktionsregeln für phasenminimale Systeme

Differentiation



Integration



Sinus-Anregung: $u(t) = \sin(\omega t)$

Eingeschwungene Antwort: $y(t) = A \sin(\omega t + \varphi)$

$$y(t) = \frac{d}{dt} \sin(\omega t) = \omega \cos(\omega t)$$

$$y(t) = \omega \sin(\omega t + 90^\circ)$$

- Verstärkung hoher Frequenzen proportional zu ω
- Phase eilt um 90° vor

$$y(t) = \int_0^t \sin(\omega \tau) d\tau = -\frac{1}{\omega} \cos(\omega t)$$

$$y(t) = \frac{1}{\omega} \sin(\omega t - 90^\circ)$$

- Verstärkung niedriger Frequenzen anti-proportional zu ω
- Phase eilt um 90° nach

6.4 Konstruktionsregeln für phasenminimale Systeme

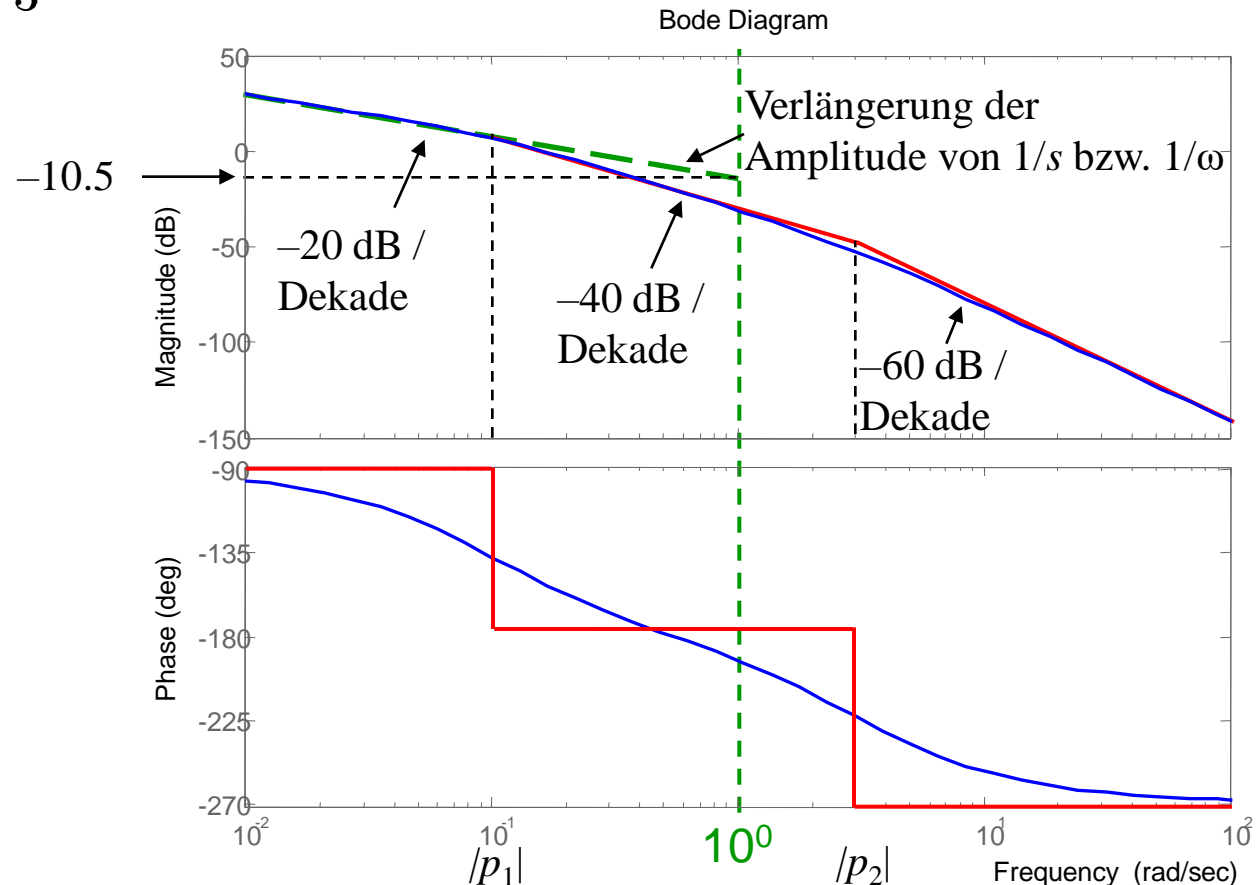
Beispiel: $G(s) = \frac{0.09}{s(s + 0.1)(s + 3)}$

Verstärkung: $K = \frac{0.09}{0.1 \cdot 3} = 0.3 = -10.5 \text{ dB}$

Pol $p_1 = -0.1$

Pol $p_2 = -3$

Zur Berechnung der Verstärkung K werden Nullstellen und Pole bei $s = 0$ ignoriert und dann $s \rightarrow 0$ eingesetzt!



6.5 Konstruktionsregeln für nicht phasenminimale Systeme

Erweiterung auf stabile, nicht phasenminimale Systeme:

- Nullstellen können in der rechten s -Halbebene liegen.
- Es können Totzeiten vorkommen.
- Pole dürfen weiterhin nicht in der rechten s -Halbebene liegen, weil wir für instabile Systeme keinen Frequenzgang messen können. (Mathematisch können wir ihn natürlich dennoch ausrechnen, damit wollen wir uns hier aber nicht beschäftigen.)

Änderungen für Amplitudengang:

- Die selbe Konstruktion wie für phasenminimale Systeme.
- Totzeit spielt auch keine Rolle, da $|e^{-i\omega T_t}| = 1$

Änderungen für Phasengang:

- Konstruktion des Phasenanteils für Nullstellen und Pole in der linken s -Halbebene bleibt.
- Phase für Nullstellen in der rechten s -Halbebene: $\Delta\angle(G(i\omega)) = \angle(iT_n\omega + 1) \rightarrow -90^\circ$, für $\omega \gg |n|$ oder $|T_n|\omega \gg 1$, da nun $T_n < 0$ bzw. $n > 0$.
- Eine Totzeit ruft folgende Phasenänderung hervor: $\Delta\angle(G(i\omega)) = -\omega T_t$.

6.5 Konstruktionsregeln für nicht phasenminimale Systeme

Konsequenzen für die Regelung:

- Eine Nullstelle in der rechten s -Halbebene erzeugt eine Phasenverzögerung um bis zu 90° , im Gegensatz dazu *eilt* die Phase bei einer Nullstelle in der linken s -Halbebene um bis zu 90° *vor*. Daher wirkt eine stabile Nullstelle im allgemeinen stabilisierend auf einen Regelkreis; eine instabile Nullstelle wird destabilisierend und erschwert die Regelung.
- Eine Totzeit bewirkt eine Phasenverzögerung, die linear mit der Frequenz anwächst. (Im Phasengang sieht dies aber durch die logarithmische Darstellung der Frequenz wie ein exponentieller Zusammenhang aus.) Daher wirkt eine Totzeit destabilisierend; und zwar umso stärker, je größer die Totzeit ist. Da die Phasenverschiebung mit zunehmender Frequenz ansteigt, ist die maximal mögliche Schnelligkeit der Regelung beschränkt.

System ohne Phasenverzerrung (Null-Phase):

Systeme, die (in Bezug auf die Im-Achse) spiegelbildliche Nullstellen aufweisen, erzeugen *keine Phasenverschiebung*. Dies kann man bei der Steuerung ausnutzen, wenn man einem Referenzsignal mit möglichst geringem Phasenfehler folgen möchte (*zero phase error tracking controller*). *Beispiel*: Nullstellen bei $s = -a$ und $s = a$:

$$(a + s)(a - s)|_{s=i\omega} = a^2 - s^2|_{s=i\omega} = a^2 + \omega^2 \rightarrow \text{rein reell, Phase} = 0.$$

7. Wichtige dynamische Systeme

Inhalt Kapitel 7

7 Wichtige dynamische Systeme

- 7.1 Bezeichnung linearer dynamischer Systeme
- 7.2 Proportionales Verhalten
- 7.3 Differenzierendes Verhalten
- 7.4 Integrierendes Verhalten
- 7.5 Allpässe und nicht phasenminimale Systeme
- 7.6 Totzeiten

7.1 Bezeichnung linearer dynamischer Systeme

Vorgehensweise:

1. Schreibe Übertragungsfunktion in Produkt-Standardform.
2. Kürze evtl. vorhandene Nullstellen-Pol-Kombinationen.
3. Multipliziere Zähler aus.
4. Dividiere Zählerterme durch evtl. im Nenner vorhandenen Faktor s^k .
5. Gibt es im Zähler einen konstanten Term? Wenn ja \rightarrow P (Proportionales Verhalten).
6. Suche Zählerterm $\frac{c}{s^k}$ mit der höchsten Potenz k . $\rightarrow I_k$ (k -fach integrierendes Verhalten).
7. Suche Zählerterm $c s^l$ mit der höchsten Potenz l . $\rightarrow D_l$ (l -fach differenzierendes Verhalten).
8. Bezeichne die Nennerordnung (ohne s^k) mit n . $\rightarrow T_n$ (n -fach verzögerndes Verhalten).
9. Gibt es eine Totzeit? Wenn ja $\rightarrow T_t$.
10. Das System wird so bezeichnet (wenn $k, l, n = 0$, fällt der entsprechende Teil weg):

$$\boxed{PI_k D_l T_n - T_t}$$

Beispiele: PID, PT_2 , PDT_1 , IT_1 , PI, DT_2 , $PT_3 - T_t$, ...

7.1 Bezeichnung linearer dynamischer Systeme

Prinzipielle Benennung:

- P: $G(s) = a$
- PT₁: $G(s) = \frac{a}{c + ds}$
- D: $G(s) = bs$
- I: $G(s) = \frac{c}{s}$
- PD: $G(s) = a + bs$
- PDT₁: $G(s) = \frac{a + bs}{c + ds} = \frac{\tilde{a} + \tilde{b}s}{1 + \tilde{d}s}$
- PI: $G(s) = a + \frac{c}{s}$
- PID: $G(s) = a + bs + \frac{c}{s}$

Übliche Schreibweise:

- $G(s) = K$
- $G(s) = \frac{K}{1 + Ts}$
- $G(s) = T_D s$
- $G(s) = \frac{1}{T_I s}$
- $G(s) = K(1 + T_D s)$
- $G(s) = K \frac{1 + T_D s}{1 + Ts}$
- $G(s) = K \left(1 + \frac{1}{T_I s} \right) = K \frac{1 + T_I s}{T_I s}$
- $G(s) = K \left(1 + \frac{1}{T_I s} + T_D s \right)$

7.1 Bezeichnung linearer dynamischer Systeme

Beispiel 1: $G(s) = \frac{s^3 + 3s^2 + 2s}{s^2(s + 1)(s + 3)(s + 5)}$

1. Produkt-Standardform: $G(s) = \frac{s(s + 1)(s + 2)}{s^2(s + 1)(s + 3)(s + 5)}$

2. Kürzen: $G(s) = \frac{s + 2}{s(s + 3)(s + 5)}$

4. Dividiere Zähler durch s : $G(s) = \frac{1 + 2/s}{(s + 3)(s + 5)} \rightarrow \text{PIT}_2$

T_2

Beispiel 2: $G(s) = K \frac{1 + T_I s + T_I T_D s^2}{T_I s}$

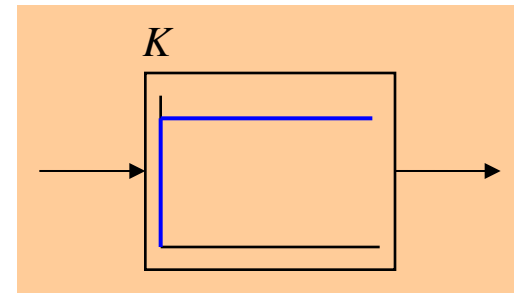
4. Dividiere Zähler durch s : $G(s) = \frac{K}{T_I s} + K + K T_D s \rightarrow \text{PID}$

$\text{I} \quad \text{P} \quad \text{D}$

7.2 Proportionales Verhalten

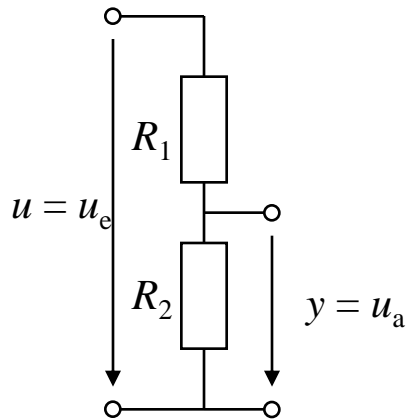
Reines P-Verhalten:

- Übertragungsfunktion: $G(s) = K$
- Einfache Regler haben P-Verhalten.
- Nur ein Parameter: Verstärkung K .
- Linearisierung einer statischen Kennlinie führt auf P-Verhalten ($K = \text{Steigung im AP}$).



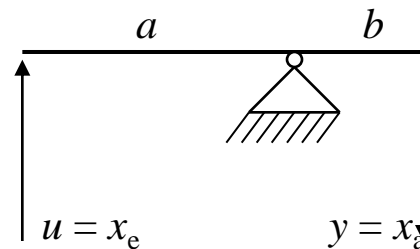
Physikalische Realisierungen:

Spannungsteiler



$$\frac{Y(s)}{U(s)} = K = \frac{R_2}{R_1 + R_2}$$

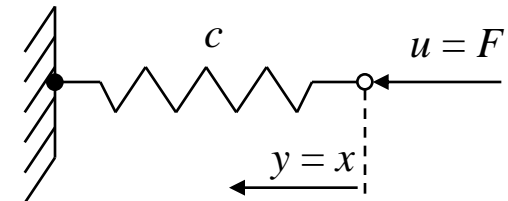
Hebel



Für kleine Auslenkungen:

$$\frac{Y(s)}{U(s)} = K = \frac{b}{a}$$

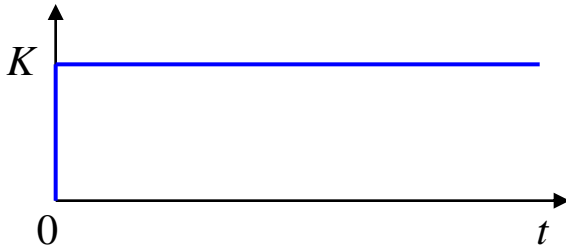
Feder



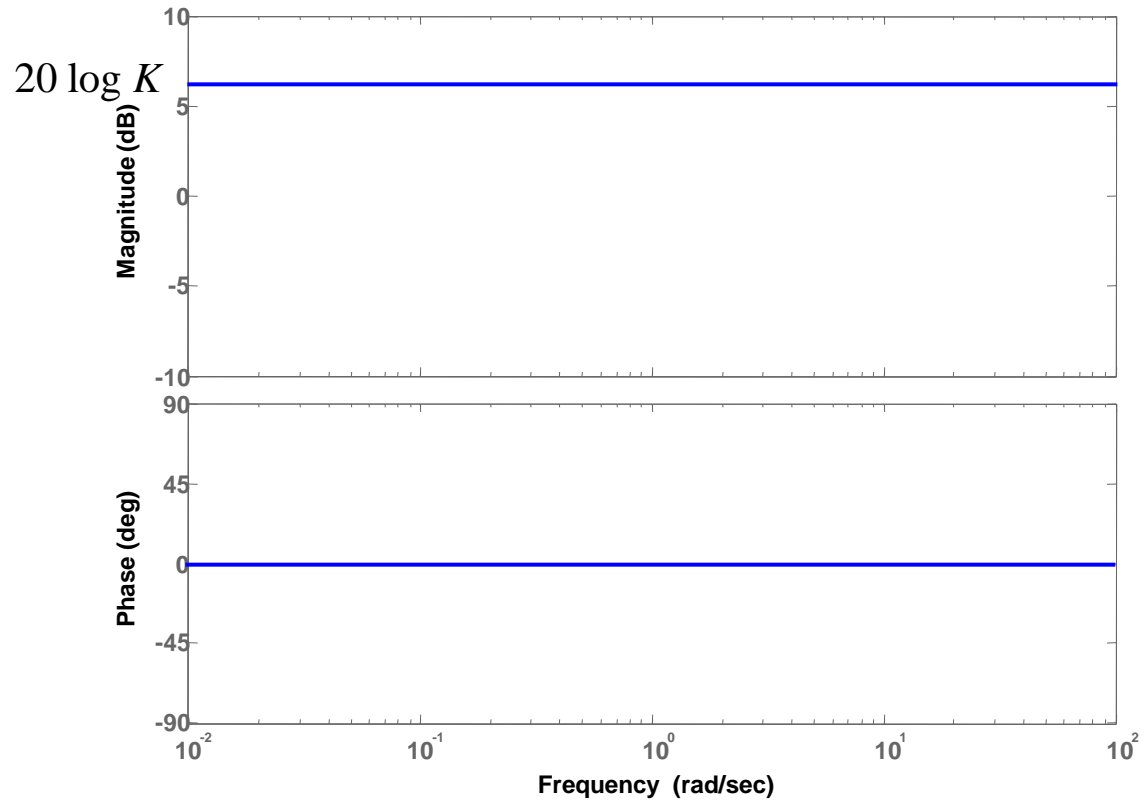
$$\frac{Y(s)}{U(s)} = K = \frac{1}{c}$$

7.2 Proportionales Verhalten

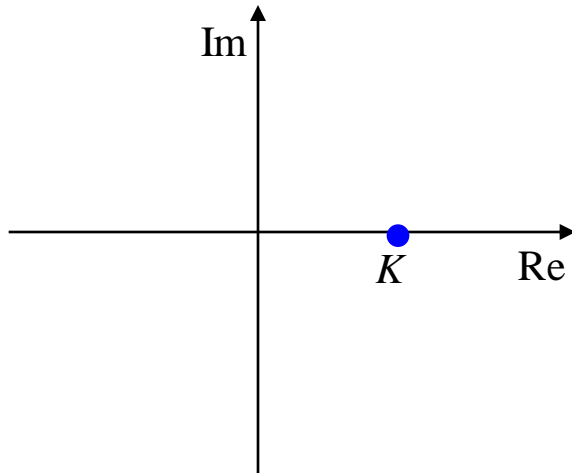
Sprungantwort eines P-Glieds



Bode-Diagramm eines P-Glieds



Ortskurve eines P-Glieds

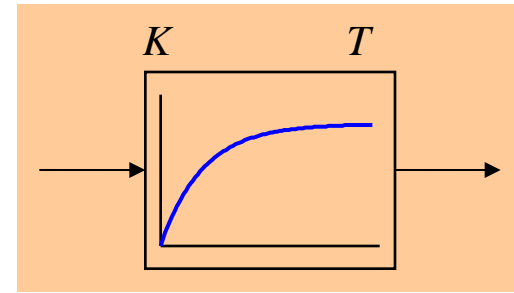


- Kein Abfall des Betrags.
- Keine Phasenverschiebung.

7.2 Proportionales Verhalten

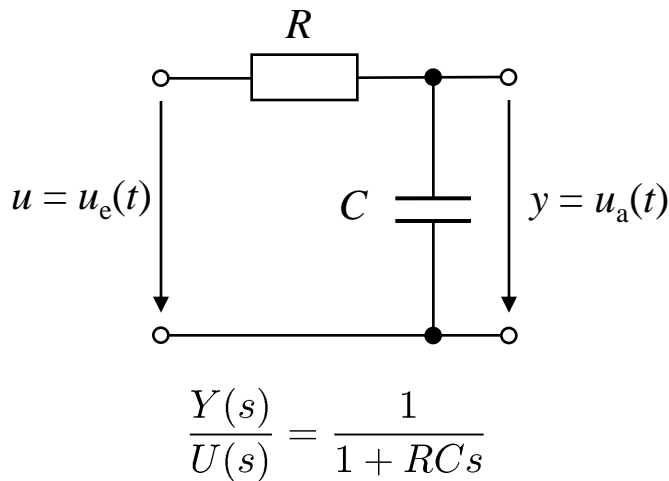
Proportional verzögerndes Verhalten 1. Ordnung (PT₁):

- Übertragungsfunktion: $G(s) = \frac{K}{1 + Ts}$
- Viele einfache Regelstrecken haben ein solches Verhalten bzw. lassen sich näherungsweise damit beschreiben.
- Parameter: Verstärkung K , Zeitkonstante T (~ Langsamkeit des Systems).

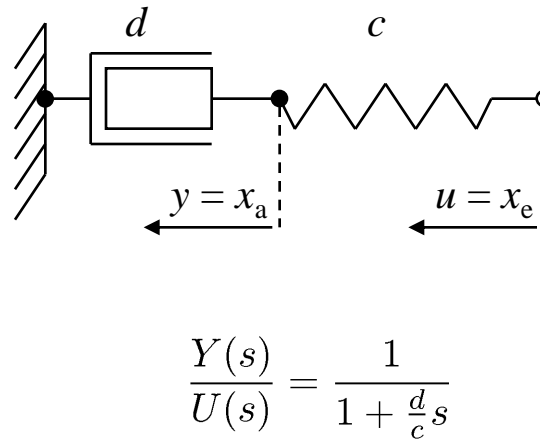


Physikalische Realisierungen:

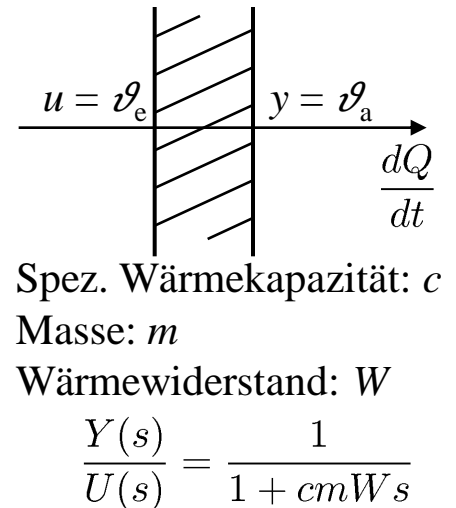
Widerstand-Kondensator-Schaltung



Feder-Dämpfer-Element

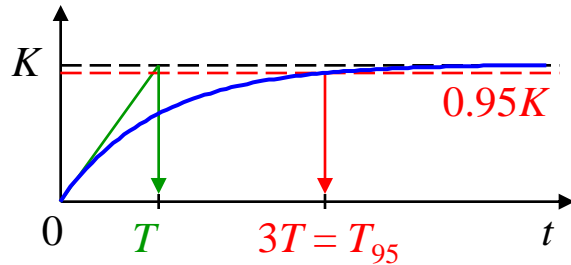


Thermisches System

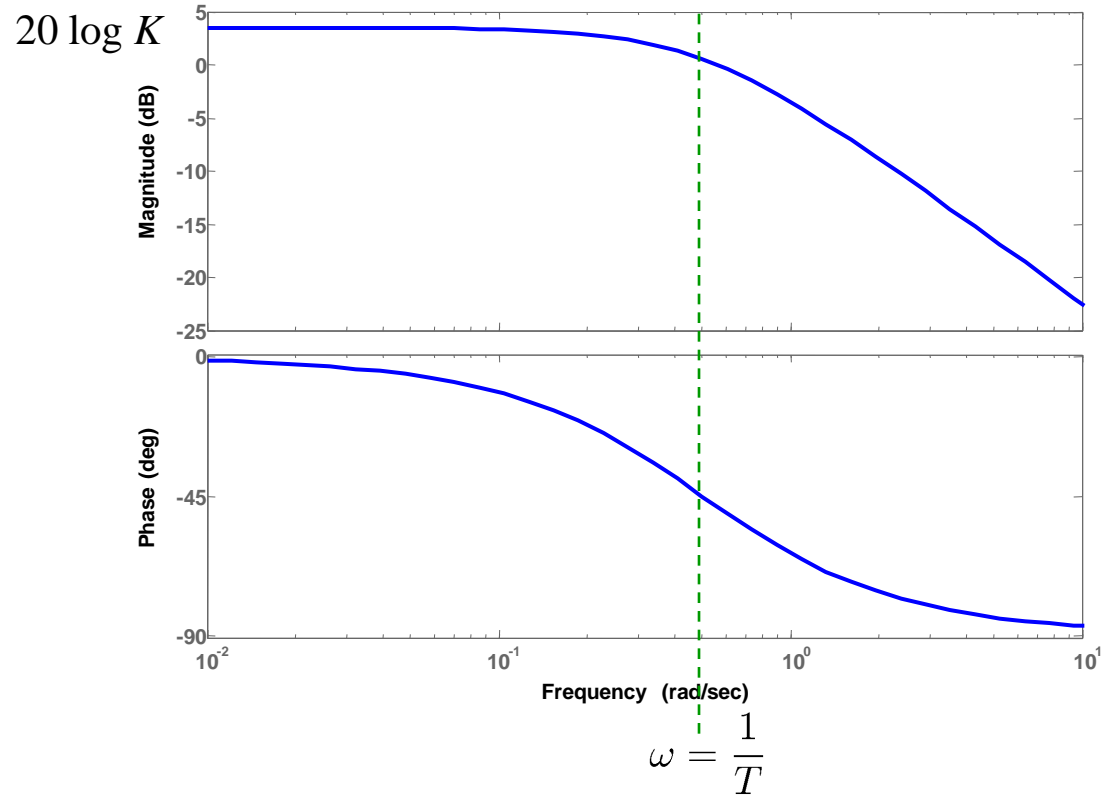


7.2 Proportionales Verhalten

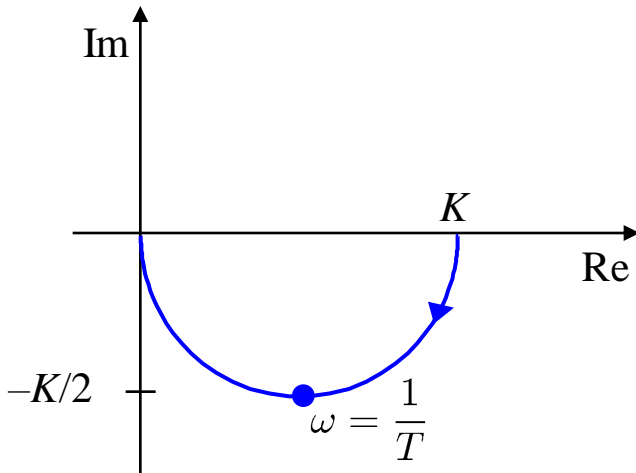
Sprungantwort eines PT₁-Glieds



Bode-Diagramm eines PT₁-Glieds



Ortskurve eines PT₁-Glieds

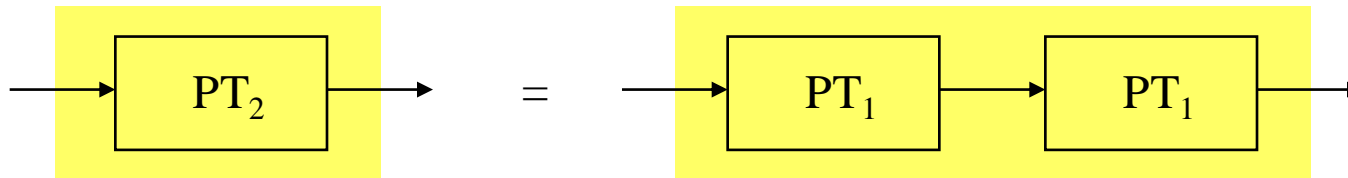


- Abfall des Betrags: 0 bis -20 dB/Dekade.
- Phasenverschiebung: 0° bis -90°.

7.2 Proportionales Verhalten

Proportional verzögerndes Verhalten 2. Ordnung (PT₂):

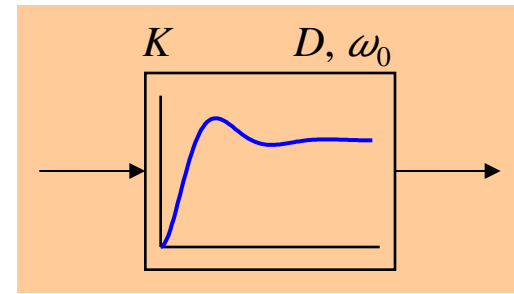
- Übertragungsfunktion:
$$G(s) = \frac{K}{1 + \frac{2D}{\omega_0}s + \frac{1}{\omega_0^2}s^2} = \frac{K\omega_0^2}{\omega_0^2 + 2D\omega_0s + s^2}$$
$$s_{1,2} = -D\omega_0 \pm \sqrt{D^2\omega_0^2 - \omega_0^2} = -D\omega_0 \pm \omega_0\sqrt{D^2 - 1}$$
- Für Dämpfungen $D \geq 1$ ergeben sich reelle Pole (bei $D = 1$ ein reeller Doppelpol) und die Übertragungsfunktion lässt sich auch schreiben als:
$$G(s) = \frac{K}{(1 + T_1s)(1 + T_2s)}$$
- 2 Parameter: Zeitkonstanten T_1 und T_2 bzw. Dämpfung D und Eckfrequenz ω_0 .
- Für $D \geq 1$ *kann* (muss nicht) das PT₂ aus der Hintereinanderschaltung zweier PT₁-Systeme entstanden sein:



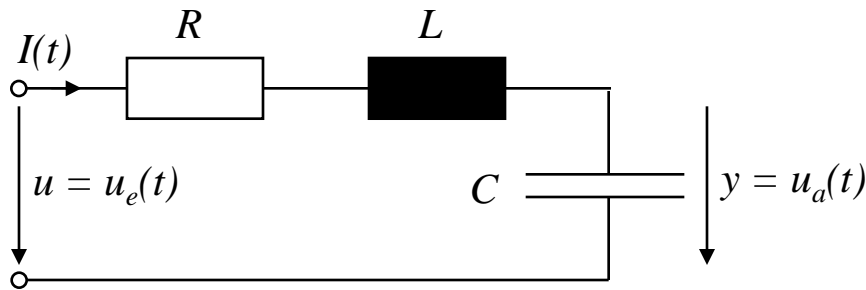
- Für $D < 1$ hat das PT₂ ein konjugiert komplexes Polpaar und ist damit schwingungsfähig. Ein solches Verhalten gibt es bei Systemen 1. Ordnung nicht!

7.2 Proportionales Verhalten

Physikalische Realisierungen eines PT₂-Verhaltens:



Elektrischer Schwingkreis

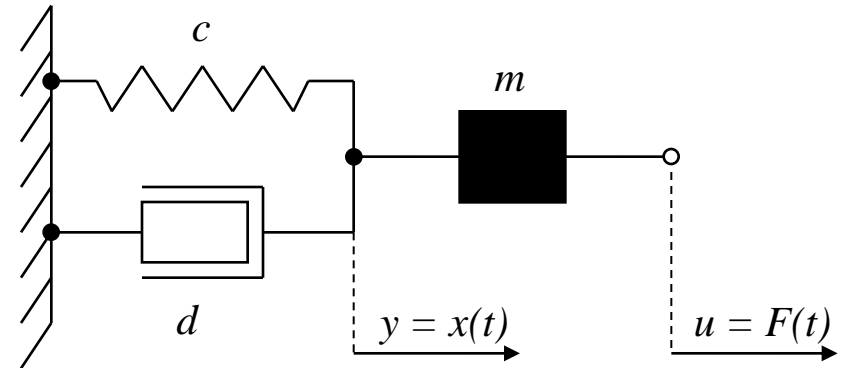


$$\frac{Y(s)}{U(s)} = \frac{1}{1 + RCs + LCs^2} = \frac{\frac{1}{LC}}{\frac{1}{LC} + \frac{R}{L}s + s^2}$$

$$D = \frac{R}{2} \sqrt{\frac{C}{L}}$$

$$\omega_0 = \frac{1}{\sqrt{LC}}$$

Mechanischer Schwinger



$$\frac{Y(s)}{U(s)} = \frac{\frac{1}{c}}{1 + \frac{d}{c}s + \frac{m}{c}s^2} = \frac{\frac{1}{m}}{\frac{c}{m} + \frac{d}{m}s + s^2}$$

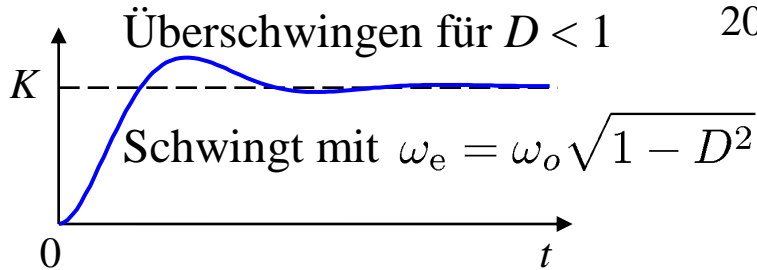
$$D = \frac{d}{2\sqrt{mc}}$$

$$\omega_0 = \sqrt{\frac{c}{m}}$$

Wenn R oder d groß sind, ist das System nicht mehr schwingungsfähig!

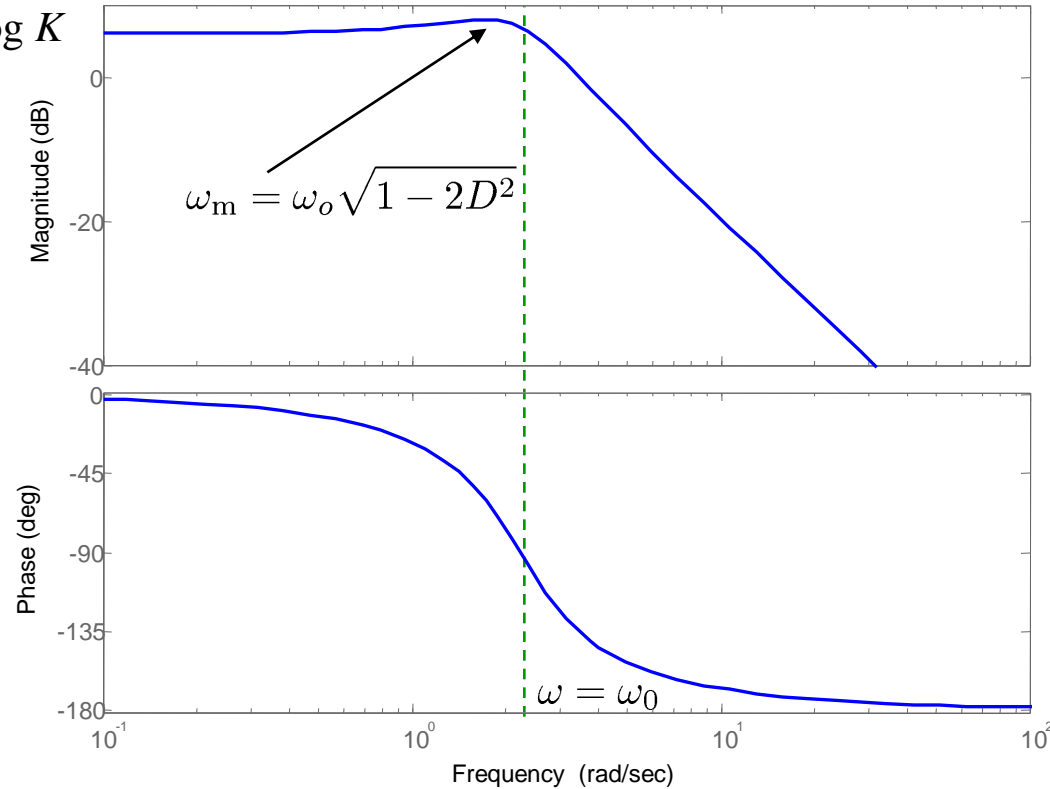
7.2 Proportionales Verhalten

Sprungantwort eines PT₂-Glieds

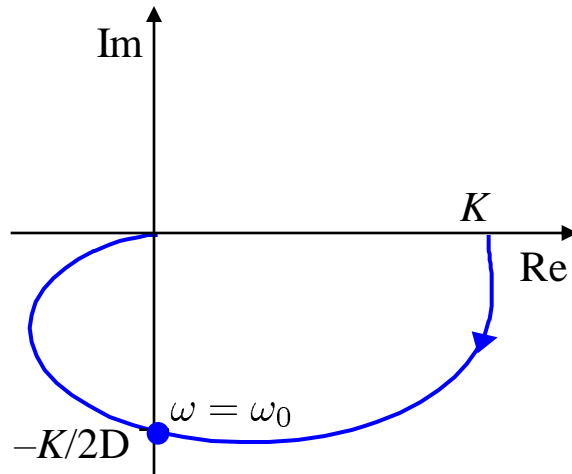


Bode-Diagramm eines PT₂-Glieds

Überhöhung für $D < 1/\sqrt{2}$



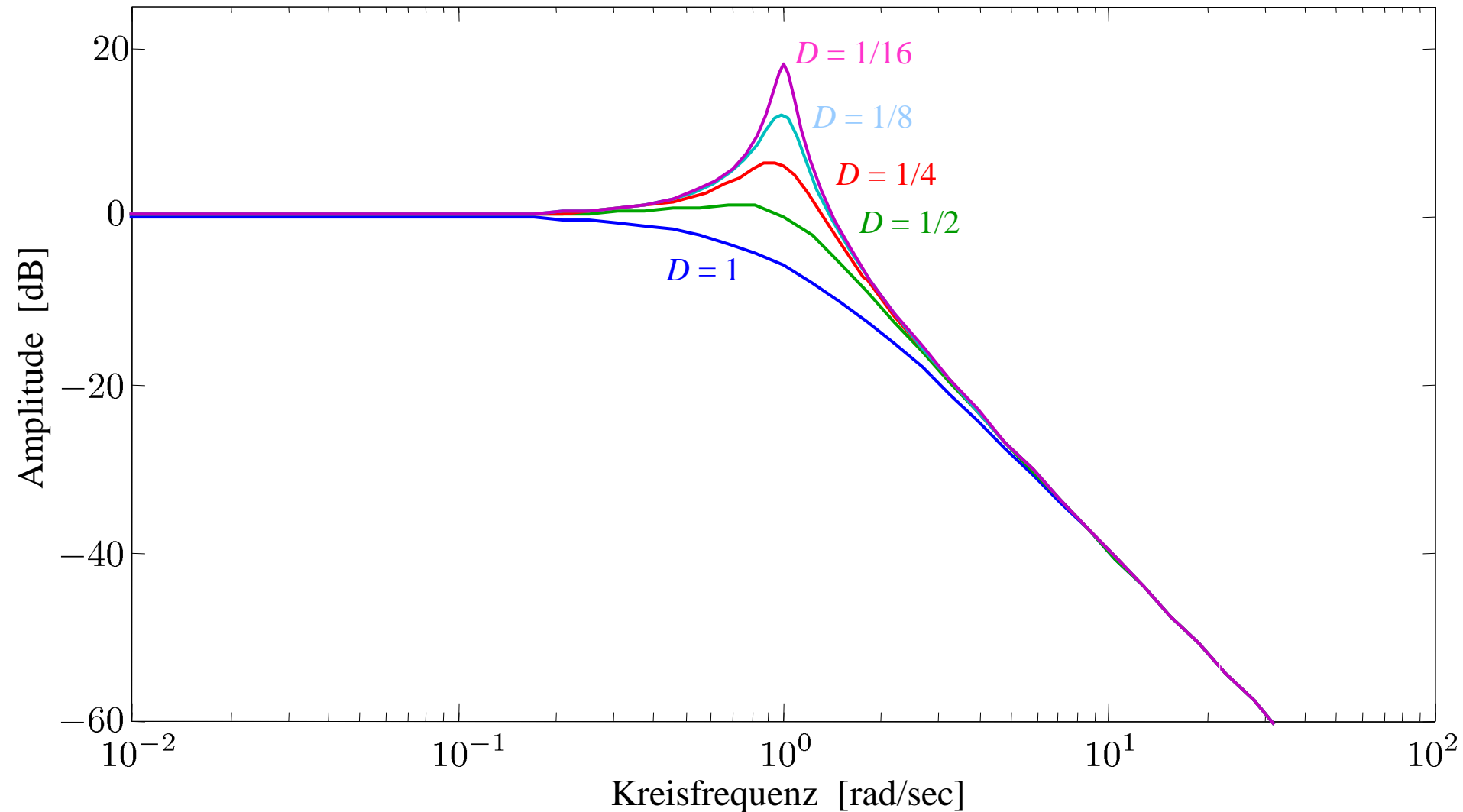
Ortskurve eines PT₂-Glieds



- Abfall des Betrags: 0 bis -40 dB/Dekade.
- Phasenverschiebung: 0° bis -180°.

7.2 Proportionales Verhalten

Verschiedene Dämpfungen



7.2 Proportionales Verhalten

Resonanzkatastrophe: Beispiel *Tacoma Bridge*, 1940



7.2 Proportionales Verhalten

Bei welcher Kreisfrequenz und Dämpfung ist der Amplitudengang überhöht?

$$G(s) = \frac{K\omega_0^2}{\omega_0^2 + 2D\omega_0 s + s^2} \rightarrow G(i\omega) = \frac{K\omega_0^2}{\omega_0^2 - \omega^2 + i2D\omega_0\omega}$$

unabhängig von ω

$$|G(i\omega)| = \frac{K\omega_0^2}{\sqrt{(\omega_0^2 - \omega^2)^2 + (2D\omega_0\omega)^2}} \rightarrow |G(i\omega)|^2 = \frac{(K\omega_0^2)^2}{(\omega_0^2 - \omega^2)^2 + (2D\omega_0\omega)^2}$$

Suche Maximum (identisch für $|G(i\omega)|$ und $|G(i\omega)|^2$):

$$\frac{d}{d\omega} |G(i\omega)|^2_{\omega=\omega_m} = 0 \rightarrow \frac{-(K\omega_0^2)^2 \cdot [2(\omega_0^2 - \omega_m^2)(-2\omega_m) + 8D^2\omega_0^2\omega_m]}{[(\omega_0^2 - \omega_m^2)^2 + (2D\omega_0\omega_m)^2]^2} = 0$$

$$-\omega_0^2 + \omega_m^2 + 2D^2\omega_0^2 = 0 \rightarrow \omega_m^2 = \omega_0^2(1 - 2D^2) \rightarrow \boxed{\omega_m = \omega_0\sqrt{1 - 2D^2}}$$

Amplitude am Maximum:

$$|G(i\omega_m)| = \frac{K\omega_0^2}{\sqrt{(\omega_0^2 - \omega_0^2 + 2D^2\omega_0^2)^2 + 4D^2\omega_0^2\omega_0^2(1 - 2D^2)}} = \frac{K\omega_0^2}{\sqrt{4D^4\omega_0^4 + 4D^2\omega_0^4 - 8D^4\omega_0^4}}$$

7.2 Proportionales Verhalten

Amplitude am Maximum...:

$$|G(i\omega_m)| = \frac{K}{\sqrt{4D^2 - 4D^4}} = \frac{K}{2D\sqrt{1 - D^2}}$$

Überhöhung bei welcher Dämpfung D ?

$$\frac{K}{2D\sqrt{1 - D^2}} > K \rightarrow \frac{1}{2D} > \sqrt{1 - D^2} \rightarrow \frac{1}{4D^2} > 1 - D^2 \rightarrow \frac{1}{4} > D^2(1 - D^2)$$

$$\rightarrow \boxed{D < \frac{1}{\sqrt{2}}}$$

Achtung: Folgende Frequenzen sind unterschiedlich groß ($\omega_m \leq \omega_e \leq \omega_0$):

- Eckfrequenz ω_0 : Real- und Imaginärteil des Frequenzgangs sind gleich groß.
- Resonanzfrequenz ω_m : Maximale Überhöhung im Amplitudengang.
- Eigenfrequenz ω_e : Frequenz der Schwingung in der Sprungantwort o.ä.

Bei Dämpfung $D = 0$ fallen alle Werte wieder zusammen.

7.2 Proportionales Verhalten

Proportional verzögerndes Verhalten n . Ordnung (PT_n):

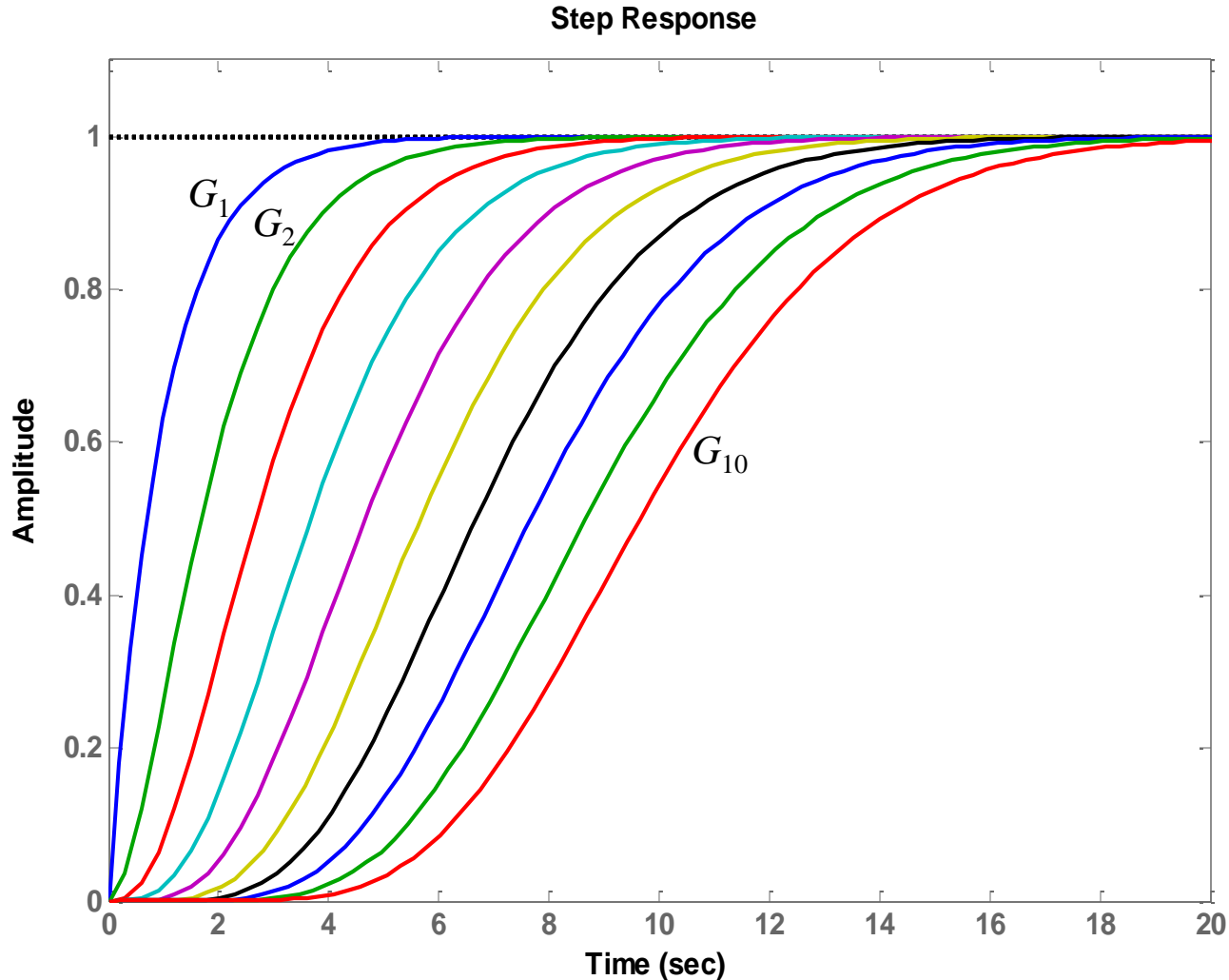
- Entsteht durch Hintereinanderschaltung von PT_1 - und/oder PT_2 -Gliedern.
- Abfall des Amplitudengangs: 0 bis $-n \cdot 20$ dB / Dekade.
- Phasenverzögerung: 0° bis $-n \cdot 90^\circ$.
- Ortskurve geht durch n Quadranten.
- Anteile mit sehr hohen Eckfrequenzen (bzw. sehr kurzen Zeitkonstanten) im Vergleich zu den anderen Anteilen, können oft vernachlässigt werden, um das Modell zu vereinfachen \rightarrow *Ordnungsreduktion*.
- Bei einer Parallelschaltung mehrerer Verzögerungselemente entstehen (wie bei der Hintereinanderschaltung) Systeme mit verzögertem Verhalten höherer Ordnung, allerdings ergibt sich auch eine Dynamik im Zähler (Nullstellen) der Übertragungsfunktion. *Beispiel:*

$$G(s) = G_1(s) + G_2(s) = \frac{K_1}{1 + T_1 s} + \frac{K_2}{1 + T_2 s} = \frac{K_1 + K_2 + (K_1 T_2 + K_2 T_1) s}{(1 + T_1 s)(1 + T_2 s)}$$

Diese Systeme werden im Abschnitt "Differenzierendes Verhalten" behandelt.

7.2 Proportionales Verhalten

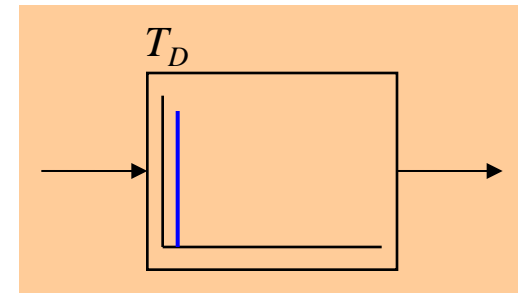
Beispiel: $G_n(s) = \frac{1}{(s + 1)^n}$



7.3 Differenzierendes Verhalten

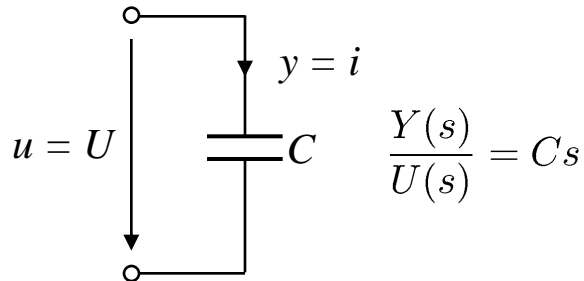
Ideales D-Verhalten:

- Übertragungsfunktion: $G(s) = T_D s$
- Dieses Verhalten ist nicht realisierbar.
- Nur ein Parameter: Differenzierbeiwert oder -zeit oder Vorhaltezeit T_D .

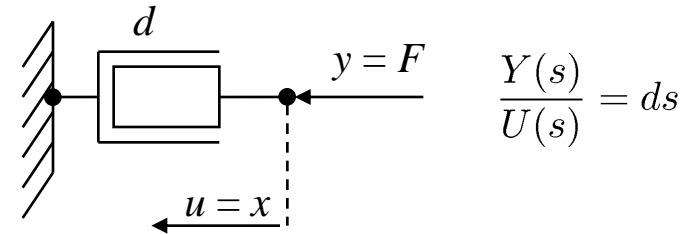


Physikalische Realisierungen:

Kondensator



Dämpfer



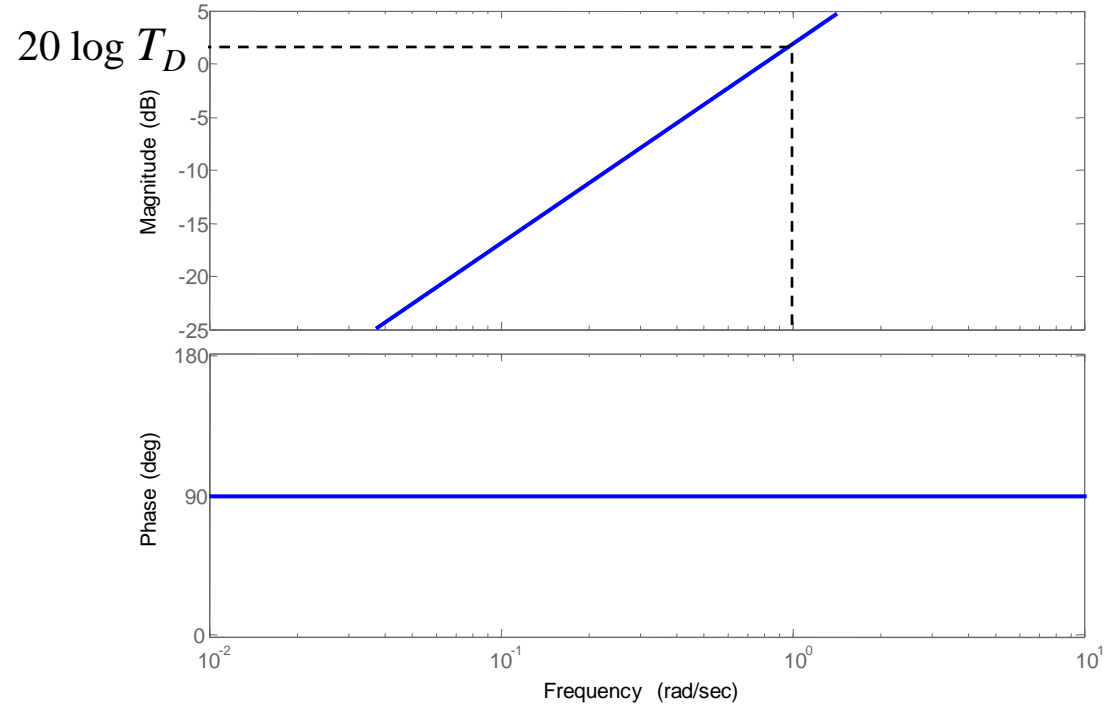
ACHTUNG: In Realität gibt es keine "reine" Kapazität ohne ohmschen Widerstandsanteil bzw. keinen "reinen" Dämpfer ohne federndes (kraftproportionales) Verhalten. Deshalb ist ein idealer Differenzierer nicht realisierbar.

7.3 Differenzierendes Verhalten

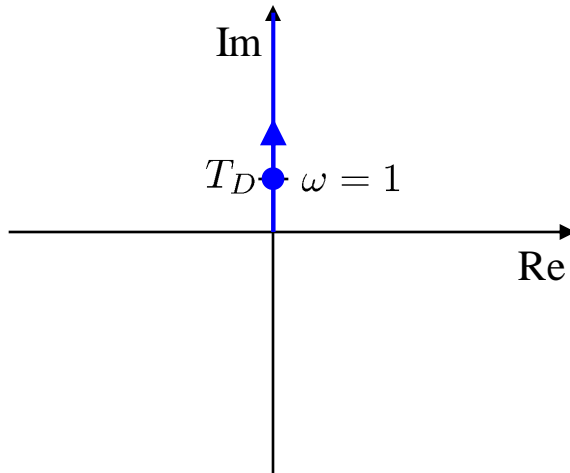
Sprungantwort eines D-Glieds



Bode-Diagramm eines D-Glieds



Ortskurve eines D-Glieds

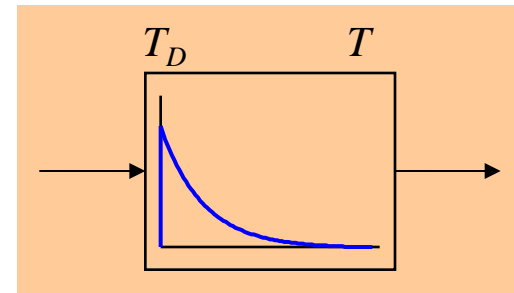


- Anstieg des Betrags: +20 dB/Dekade.
- Phasenverschiebung: +90°.

7.3 Differenzierendes Verhalten

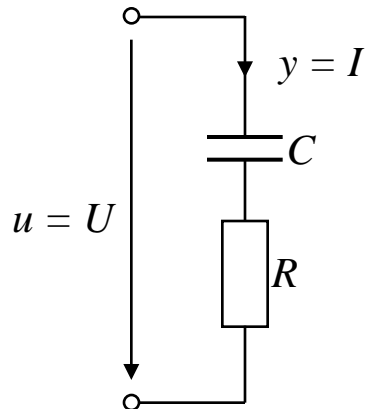
DT₁-Verhalten (reales D-Verhalten):

- Übertragungsfunktion: $G(s) = \frac{T_D s}{1 + T s}$
- Dieses Verhalten ist (im Gegensatz zu idealem D-Verhalten) realisierbar.
- Zwei Parameter: Differenzierbeiwert oder -zeit oder Vorhaltezeit T_D und Zeitkonstante T .
- Für $T \rightarrow 0$ nähert sich das Verhalten idealem D-Verhalten an.



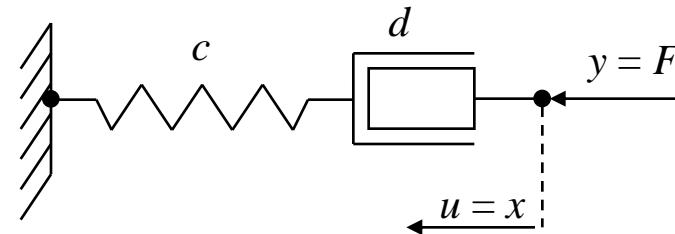
Physikalische Realisierungen:

Kondensator mit ohmschem Widerstandsanteil



$$\frac{Y(s)}{U(s)} = \frac{Cs}{1 + RCs}$$

Dämpfer mit Federanteil

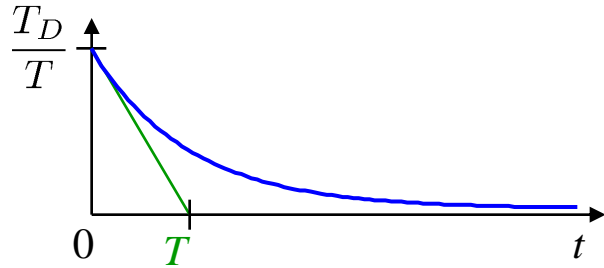


$$\frac{Y(s)}{U(s)} = \frac{ds}{1 + \frac{d}{c}s}$$

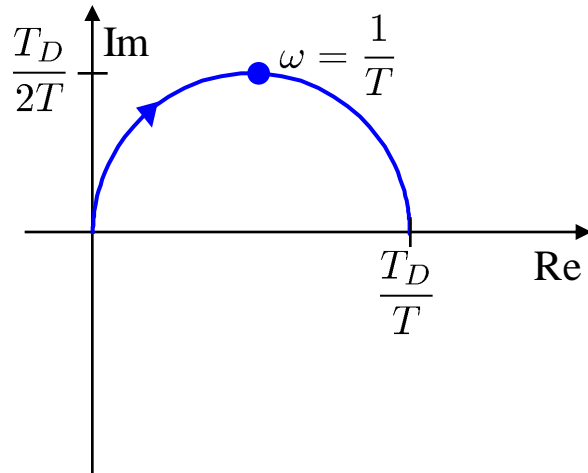
Für $R \rightarrow 0$ bzw. $c \rightarrow \infty$ ergibt sich wieder ideal differenzierendes Verhalten!

7.3 Differenzierendes Verhalten

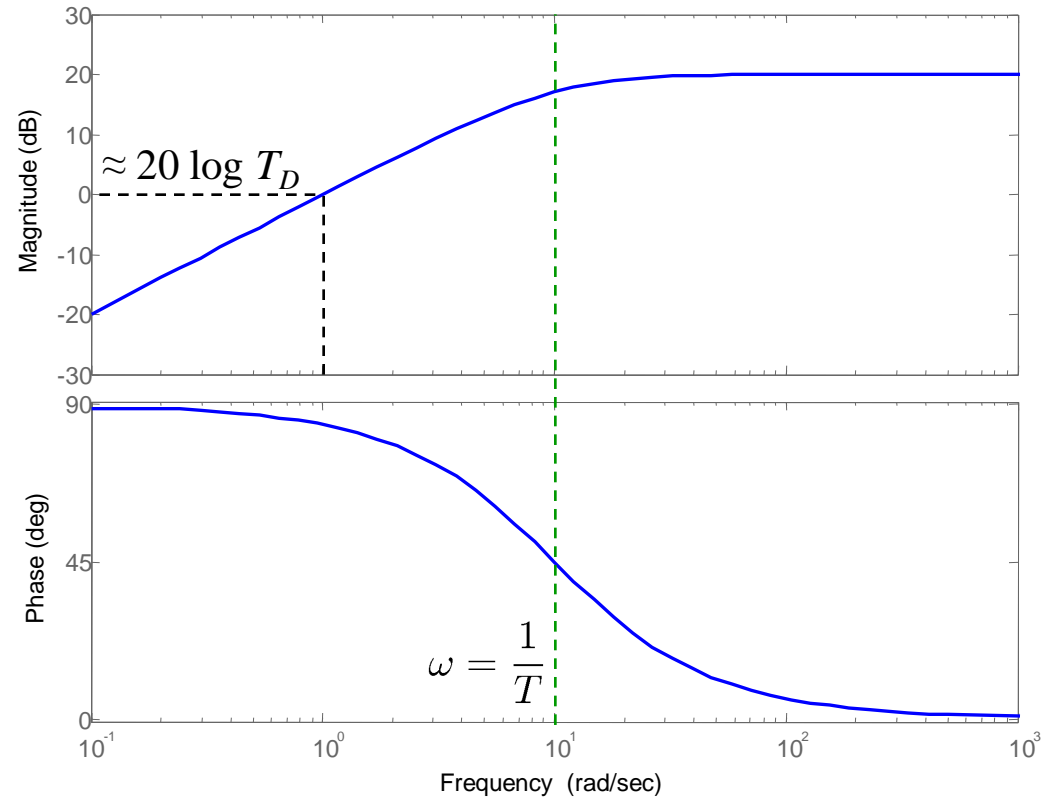
Sprungantwort eines DT₁-Glieds



Ortskurve eines DT₁-Glieds



Bode-Diagramm eines DT₁-Glieds



- Anstieg des Betrags: +20 bis 0 dB/Dekade.
- Phasenverschiebung: +90° bis 0°.

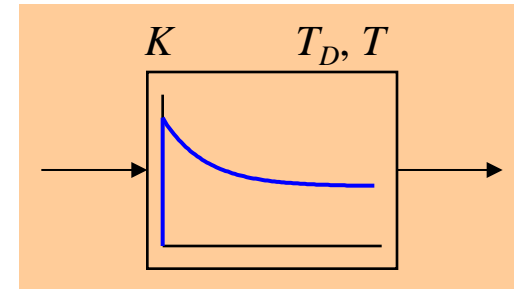
7.3 Differenzierendes Verhalten

PDT₁-Systeme können durch Parallelschaltung von P- und DT₁-Gliedern entstehen:

$$G(s) = K + \frac{T_D s}{1 + T_s} = \frac{K + (KT + T_D)s}{1 + T_s} = K \frac{1 + \tilde{T}_D s}{1 + T_s}$$

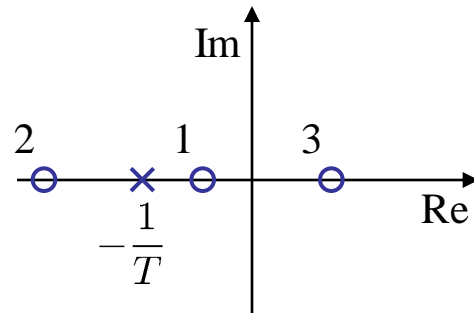
Oder durch Reihenschaltung idealer PD- mit PT₁-Gliedern:

$$G(s) = K(1 + T_D s) \cdot \frac{1}{1 + T_s} = K \frac{1 + T_D s}{1 + T_s}$$



Es können 3 Fälle unterschieden werden:

1. $T_D > T$: Pol $-1/T$ liegt links von der Nullstelle $-1/T_D$ → Lead-Glied (realer PD-Regler).
2. $0 < T_D < T$: Pol liegt rechts von der Nullstelle → Lag-Glied.
3. $T_D < 0$: Nullstelle ist positiv (instabil) → nicht phasenminimales Verhalten.

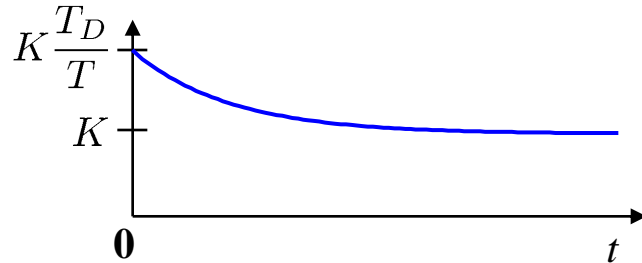


Fälle 1, 2 werden im Frequenzbereich in Kapitel 14.1 diskutiert.

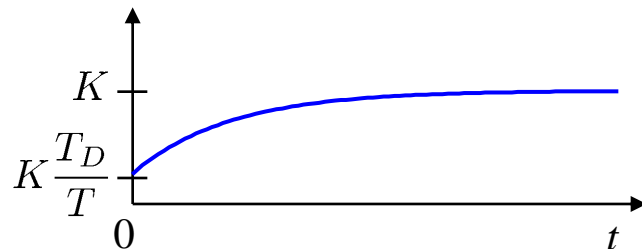
Fall 3 wird auch in Kapitel 7.5 diskutiert.

7.3 Differenzierendes Verhalten

Sprungantwort PDT₁-Glied mit dominanter Nullstelle ($T_D > T$) (realer PD-Regler):

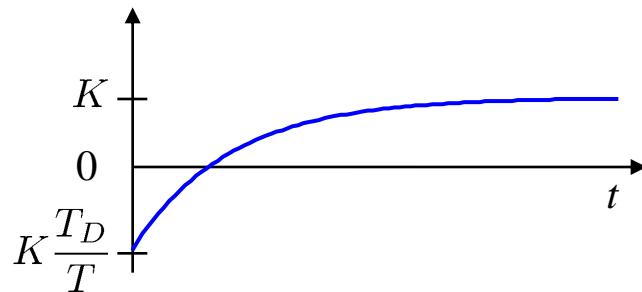


Sprungantwort PDT₁-Glied mit dominantem Pol ($0 < T_D < T$):



Für $T_D = T$ kürzt sich Zähler gegen Nenner und es ergibt sich ein P-Glied $G(s) = K$.

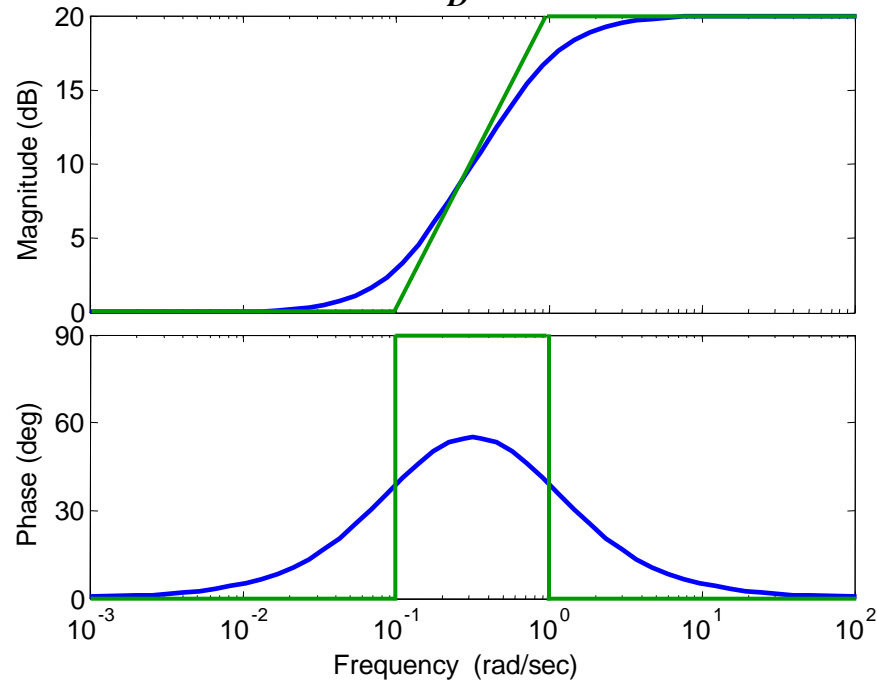
Sprungantwort PDT₁-Glied mit positiver Nullstelle ($T_D < 0$) (*nicht phasenminimal*):



Gegenläufige Anfangsreaktionen gibt es nur bei *nicht phasenminimalem* Verhalten!

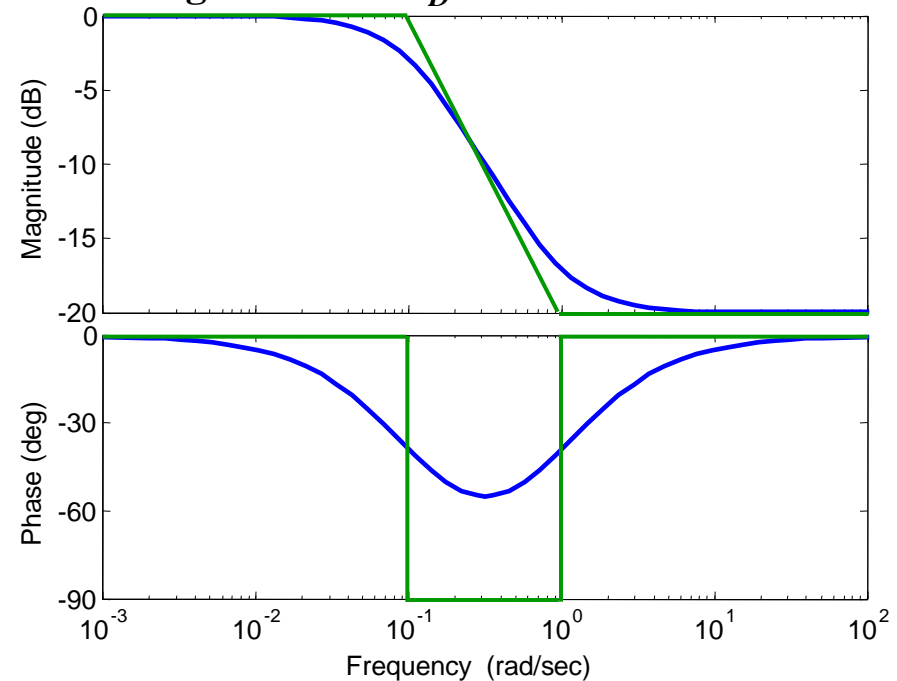
7.3 Differenzierendes Verhalten

1. Lead-Glied mit $T_D = 10$ sec und $T = 1$ sec



- Phasenhebung im Bereich mittlerer Frequenzen (Amplitudendurchtrittsfrequenz ω_D).
- Wirkt dadurch stabilisierend ähnlich wie ein D-Glied, hat aber geringer Rauschverstärkung bei hohen Frequenzen.

2. Lag-Glied mit $T_D = 1$ sec und $T = 10$ sec



- Amplitudenanhebung (zusammen mit vergrößertem Verstärkungsfaktor) bei niedrigen Frequenzen $\omega \ll \omega_D$
- Erhöht dadurch die stationäre Genauigkeit ähnlich wie ein I-Glied, hat aber nicht dessen destabilisierende Wirkung.

7.3 Differenzierendes Verhalten

Einfluss von Nullstellen auf die Systemdynamik:

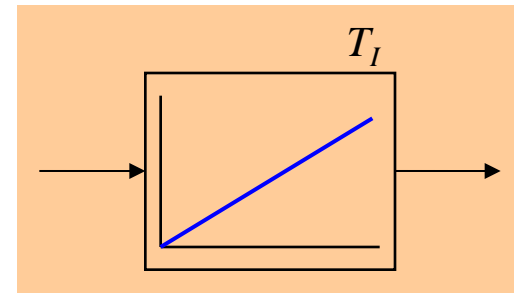
- Die Exponenten der e-Funktionen und die Frequenzen in den sin- und cos-Funktionen der Zeitlösung werden ausschließlich von den *Polen* bestimmt.
- Daher sind nur die *Pole* für *Stabilität* ausschlaggebend.
- Die Nullstellen beeinflussen aber entscheidend die Vorfaktoren vor den e-, sin- und cos-Funktionen und bestimmen daher mit, welches dynamische Verhalten sich wie stark auswirkt.
- Nullstellen filtern (blockieren) bestimmte Dynamiken im Eingangssignal heraus:
 1. Nullstellen bei $s = \pm i\omega_0$ vernichten alle Signalkomponenten mit Frequenz ω_0 .
 2. Nullstellen bei $s = 0$ vernichten den Gleichanteil (entspricht Frequenz $\omega_0 = 0$).
 3. Nullstellen bei $s = -a$ vernichten Anregungssignale der Form $\exp(-at)$.
- *Positive* (instabile) *Nullstellen* können dem Eingangssignal *gegenläufige Anfangsreaktionen* erzeugen. *Beispiel*: Anfangswert der Sprungantwort (Nullstelle bei $s = +1/4$):

$$h(+0) = \lim_{s \rightarrow \infty} sG(s) \frac{1}{s} = \lim_{s \rightarrow \infty} \frac{-4s + 1}{3s + 1} = -\frac{4}{3} < 0$$

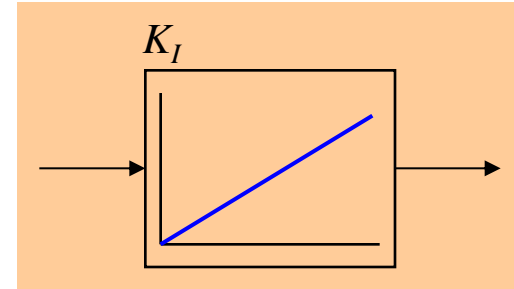
7.4 Integrierendes Verhalten

Reines I-Verhalten:

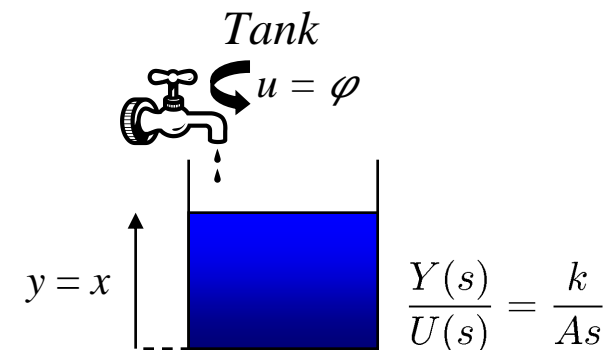
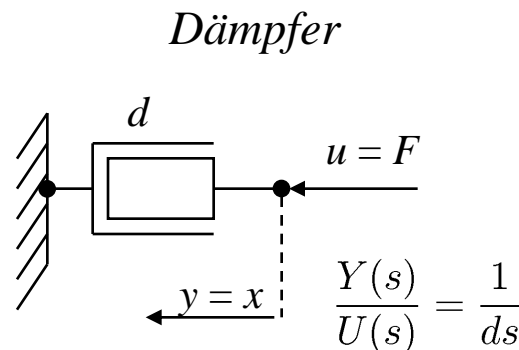
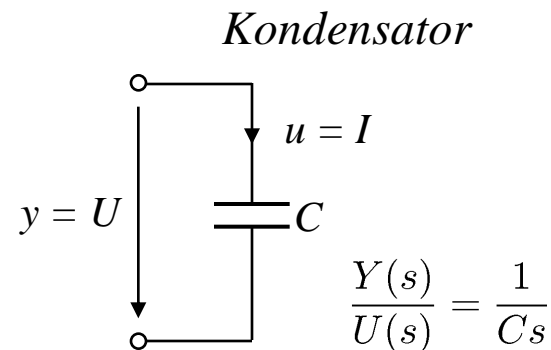
- Übertragungsfunktion: $G(s) = \frac{1}{T_I s} = \frac{K_I}{s}$
- Nur ein Parameter: Integrierzeit oder Nachstellzeit T_I bzw. Integrierbeiwert K_I .
- Häufig in Regelstrecken vorhanden:
Beschleunigung \rightarrow Drehzahl, Drehzahl \rightarrow Position.
- Häufig in Reglern enthalten zur Vermeidung bleibender Regelabweichungen (dazu später mehr...).
- Grenzstabiles Verhalten.



bzw.

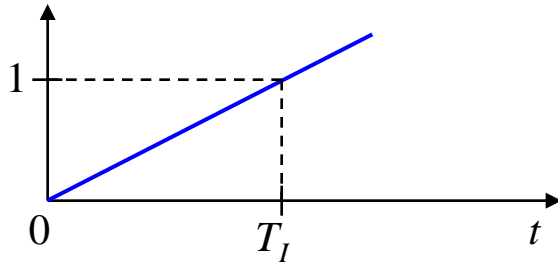


Physikalische Realisierungen:

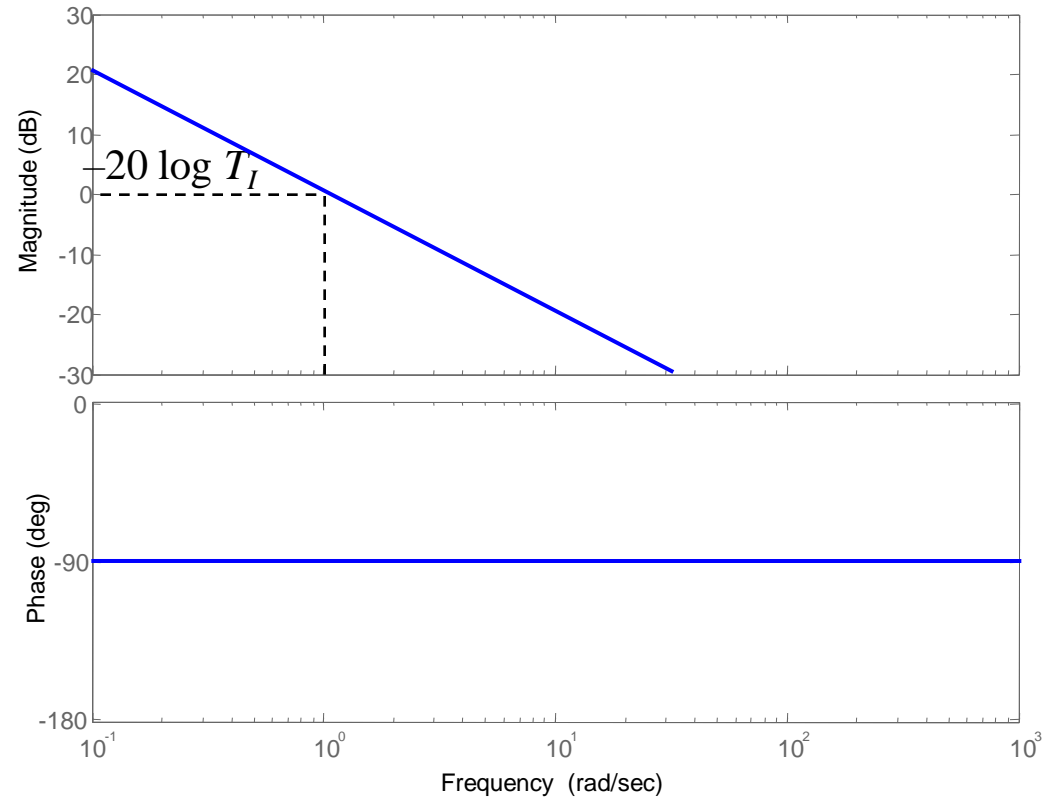


7.4 Integrierendes Verhalten

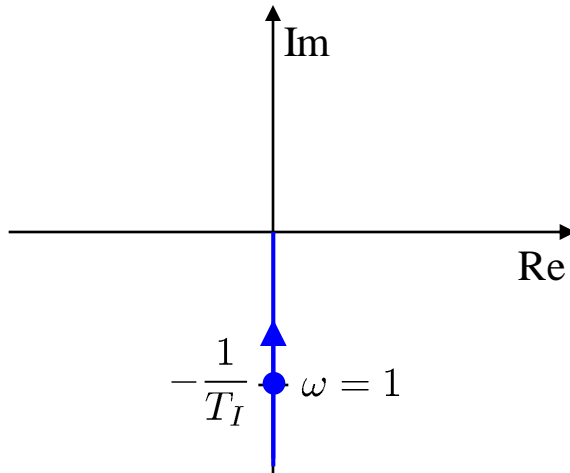
Sprungantwort eines I-Glieds



Bode-Diagramm eines I-Glieds



Ortskurve eines I-Glieds

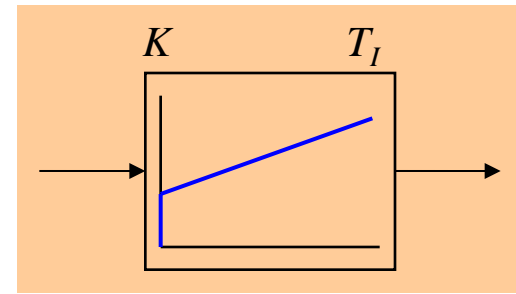


- Abfall des Betrags: -20 dB/Dekade.
- Phasenverschiebung: -90° .

7.4 Integrierendes Verhalten

PI-Verhalten:

- Übertragungsfunktion: $G(s) = K \left(1 + \frac{1}{T_I s} \right) = K \frac{1 + T_I s}{T_I s}$
- Zwei Parameter: Verstärkung K und Integrierzeit oder Nachstellzeit T_I .
- Wird häufig als Regler eingesetzt.
- Pol bei $s = 0$, Nullstelle bei $s = -1/T_I$.

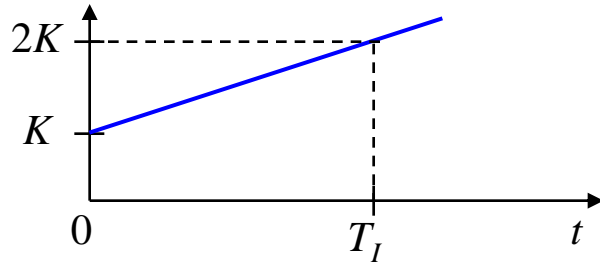


Physikalische Realisierungen:

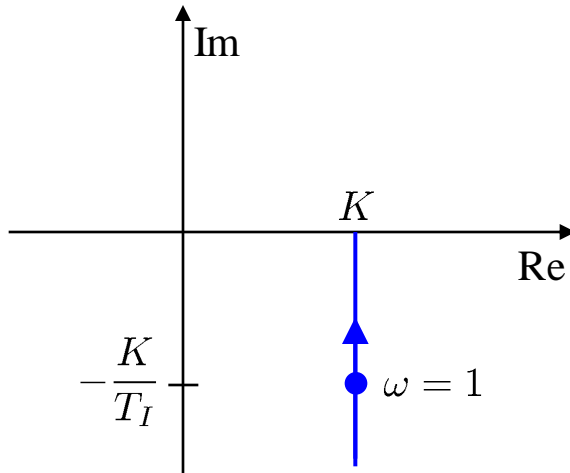
Siehe Kapitel 10 über Regler.

7.4 Integrierendes Verhalten

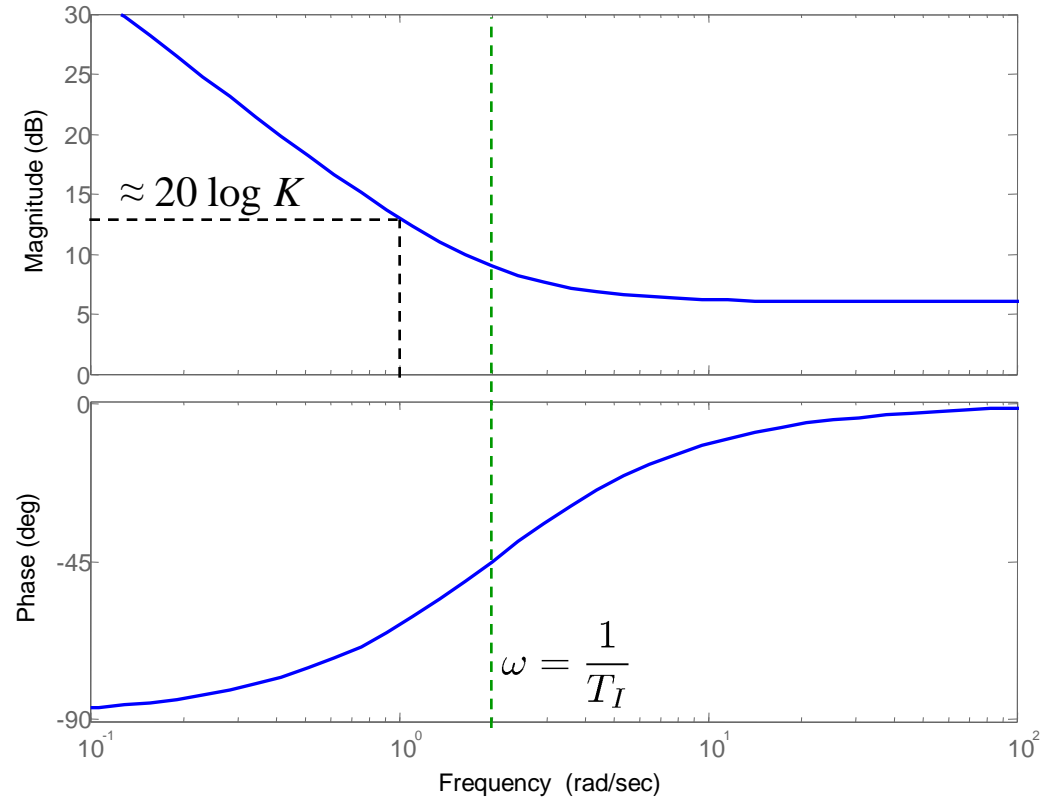
Sprungantwort eines PI-Glieds



Ortskurve eines PI-Glieds



Bode-Diagramm eines PI-Glieds

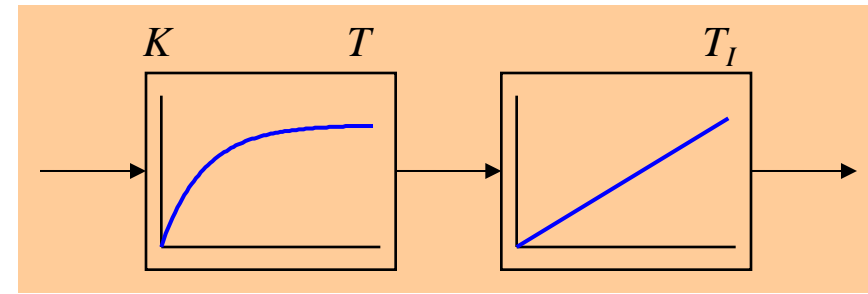


- Abfall des Betrags: -20 bis 0 dB/Dekade.
- Phasenverschiebung: -90° bis 0° .

7.4 Integrierendes Verhalten

IT₁-Verhalten:

- Übertragungsfunktion: $G(s) = \frac{1}{T_I s} \frac{1}{1 + T s}$
- Reihenschaltung aus I- und PT₁-Glied.
- Kommt häufig als Regelstrecke vor.
- Pole bei $s = 0$ und $s = -1/T$.



Physikalische Realisierungen:

Gleichstrommotor zur Positionsregelung des Tisches einer Werkzeugmaschine:

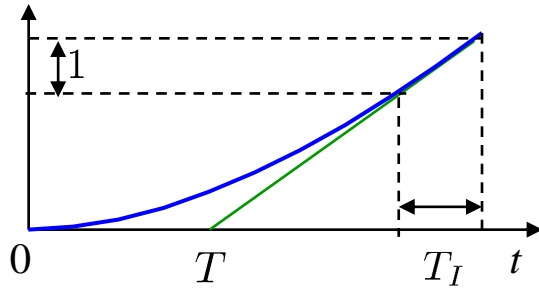
- Eingang: El. Spannung U des Motors. Ausgang: Position der Welle φ / des Tisches.
- Elektrisches System viel schneller als mechanisches. → Vernachlässigung der Dynamik.

Mechanisches System (Spannung → Drehzahl): $\Theta \frac{d\omega}{dt} + c\omega = k_1 U$ (PT₁-Verhalten)

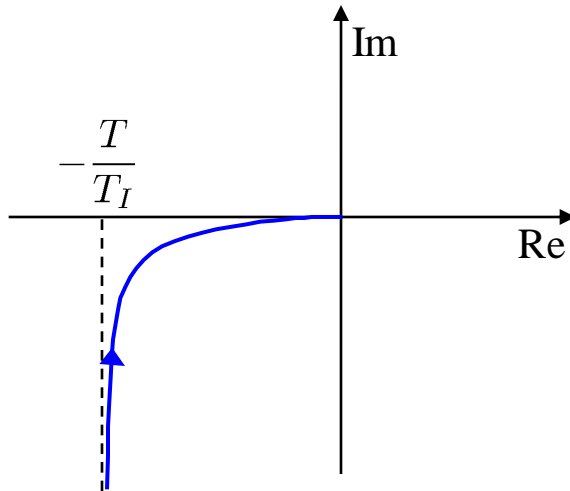
Drehzahl → Winkel/Position: $\frac{d\varphi}{dt} = k_2 \omega$ (I-Verhalten)

7.4 Integrierendes Verhalten

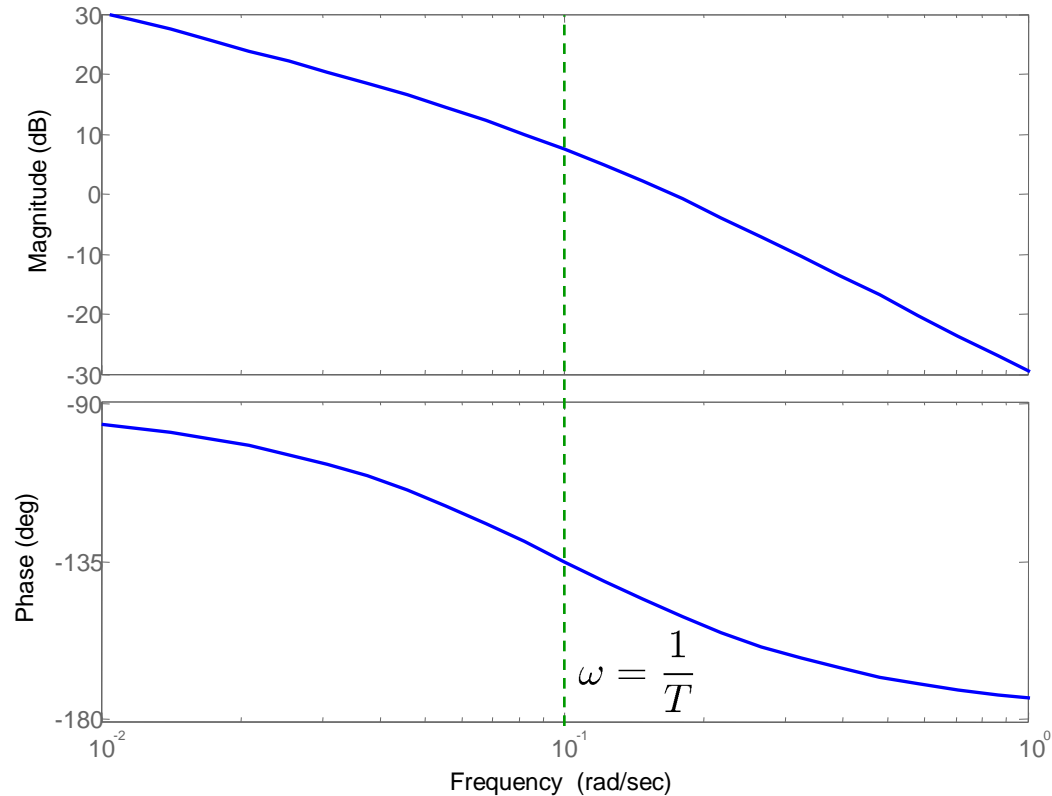
Sprungantwort eines IT₁-Glieds



Ortskurve eines IT₁-Glieds



Bode-Diagramm eines IT₁-Glieds



- Abfall des Betrags: -20 bis -40 dB/Dekade.
- Phasenverschiebung: -90° bis -180° .

7.5 Allpässe und nicht phasenminimale Systeme

Phasenminimale Systeme:

- Weisen zu gegebenem Amplitudengang die minimale Phase auf.
- Es gibt eine 1:1-Beziehung zwischen Amplitudengang und Phasengang
- ACHTUNG: Bei Abtastung (digitale Regelung) kann Phasenminimalität verloren gehen!

Nichtphasenminimale Systeme:

- Haben mindestens eine (instabile) Nullstelle mit positivem Realteil oder eine Totzeit.
- Ihre Inverse ist instabil. Das macht das Steuern und Regeln sehr viel schwieriger als für phasenminimale Systeme.
- Ihre Reaktion auf Änderungen des Eingangssignals sind oft *gegenläufig*, d.h. wenn das Eingangssignal ansteigt, fällt der Systemausgang kurzzeitig ab um dann doch wieder anzusteigen (im Unterschied zu Systemen mit negativer Verstärkung!). Dieses Verhalten lässt sich z.B. für die Sprungantwort eines Systems mit einer instabilen Nullstelle ($T_{ni} < 0$) und Zählergrad = Nennergrad ($n = m$) leicht zeigen:

$$h(+0) = \lim_{s \rightarrow \infty} sG(s) \frac{1}{s} = \lim_{s \rightarrow \infty} K \frac{(T_{n1}s + 1) \cdot \dots \cdot (T_{nm}s + 1)}{(T_{p1}s + 1) \cdot \dots \cdot (T_{pn}s + 1)} = K \frac{T_{n1} \cdot \dots \cdot T_{nn}}{T_{p1} \cdot \dots \cdot T_{pn}} < 0$$

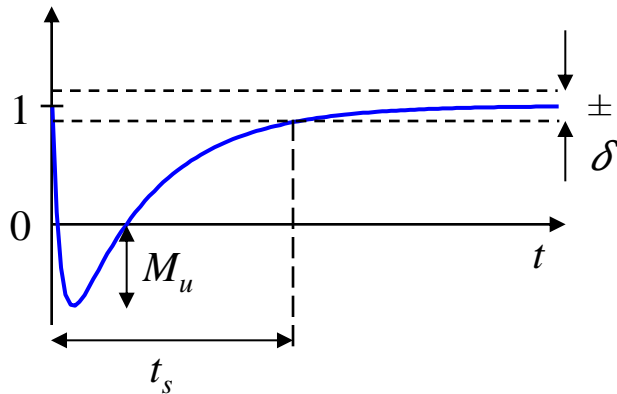
7.5 Allpässe und nicht phasenminimale Systeme

Unterschwingen der Sprungantwort nicht phasenminimaler Systeme:

Für ein nicht phasenminimales System mit Verstärkung 1 und einer instabilen Nullstelle bei $s = a > 0$ gilt für die Unterschwingweite der Sprungantwort:

$$M_u \geq \frac{1 - \delta}{e^{at_s} - 1}$$

Die e-Funktion nähert sich dem Wert 1, wenn $a \rightarrow 0$. Dann geht der Nenner gegen 0 und damit muss M_u sehr groß werden. Außerdem existiert ein Zielkonflikt zwischen Einschwinggeschwindigkeit und Unterschwingweite. Will man ein schnelles Einschwingen erreichen, muss man ein großes Unterschwingen in Kauf nehmen!



Einschwingzeit (Zeit bis das Signal im Band $1 \pm \delta$ bleibt): t_s

Unterschwingweite: M_u

7.5 Allpässe und nicht phasenminimale Systeme

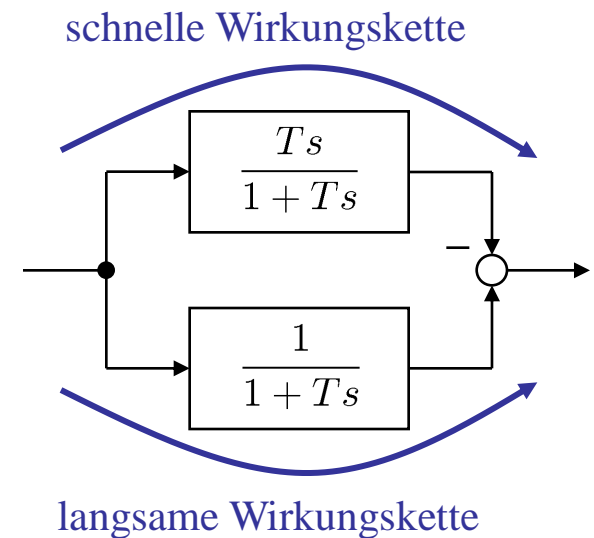
Allpässe: Nicht phasenminimale Systeme ohne Totzeit deren Amplitudengang konstant 1 ist.

- Übertragungsfunktion eines Allpasses n . Ordnung:

$$G(s) = \frac{(1 - T_1 s)(1 - T_2 s) \cdot \dots \cdot (1 - T_n s)}{(1 + T_1 s)(1 + T_2 s) \cdot \dots \cdot (1 + T_n s)}$$

- Zählergrad = Nennergrad.
- Die Nullstellen entsprechen den an der Imaginärachse gespiegelten Polen.
- Alle Pole sind stabil. \rightarrow Alle Nullstellen sind instabil.
- Für Allpass n . Ordnung: Phasenverschiebung wächst von 0° auf $-n \cdot 180^\circ$.

Beispiel: Allpass 1. Ordnung: $G(s) = \frac{1 - Ts}{1 + Ts}$



7.5 Allpässe und nicht phasenminimale Systeme

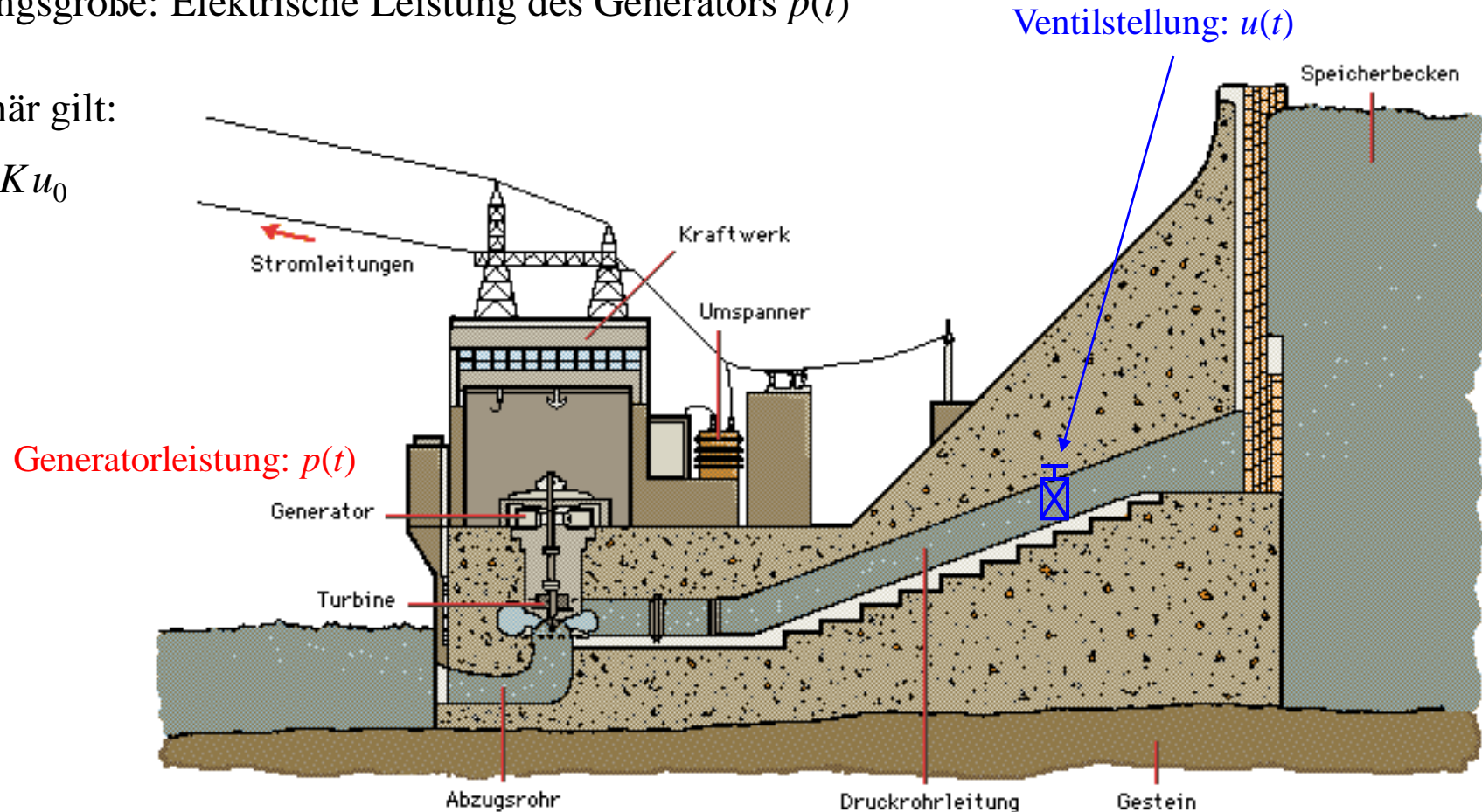
Beispiel: Nicht phasenminimales System: Leistungsregelung bei Wasserkraftwerken

Eingangsgröße: Ventilstellung $u(t)$ zur Einstellung der Wassermenge

Ausgangsgröße: Elektrische Leistung des Generators $p(t)$

Stationär gilt:

$$p_0 = K u_0$$



7.5 Allpässe und nicht phasenminimale Systeme

Beispiel: Nicht phasenminimales System: Leistungsregelung bei Wasserkraftwerken

- *Langfristig* (langsame Wirkungskette) folgt die Generatorleistung $p(t)$ proportional der durch das Ventil $u(t)$ eingestellten Wassermenge.
- *Kurzfristig* (schnelle Wirkungskette) sinkt die Generatorleistung aber beim Öffnen des Ventils, da sich die Fließgeschwindigkeit reduziert (Impuls muss erhalten bleiben!).

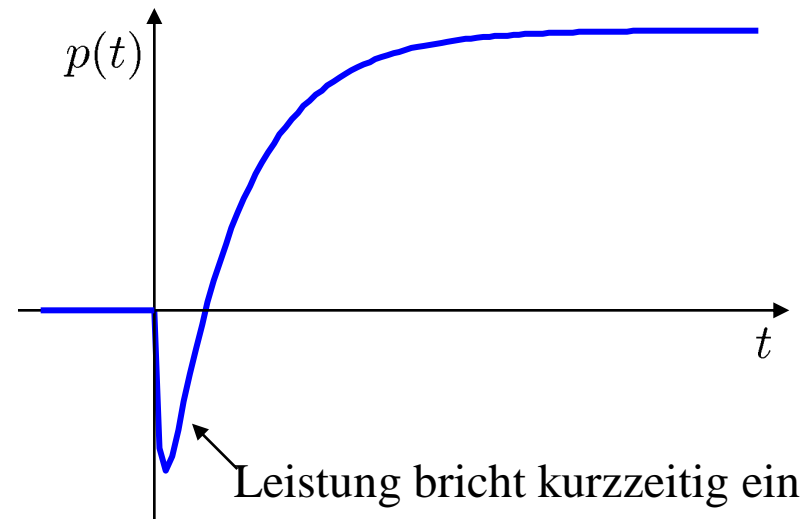
$$\text{Leistung: } p = \frac{d}{dt} \frac{1}{2} m v^2 = \frac{1}{2} \dot{m} v^2$$

Kurzfristige Reaktion nach Öffnen des Ventils:

- Massenstrom steigt proportional an: $\dot{m} \sim u$
 - Fließgeschwindigkeit nimmt um den selben Faktor ab (Impulserhaltung!): $v \sim \frac{1}{u}$
- Leistung bricht kurzfristig ein:

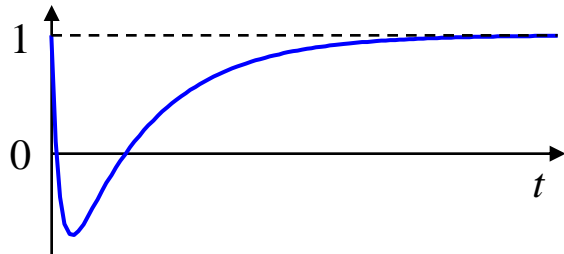
$$p \sim u \left(\frac{1}{u} \right)^2 = \frac{1}{u}$$

Antwort auf sprungförmiges Öffnen des Ventils $u(t)$

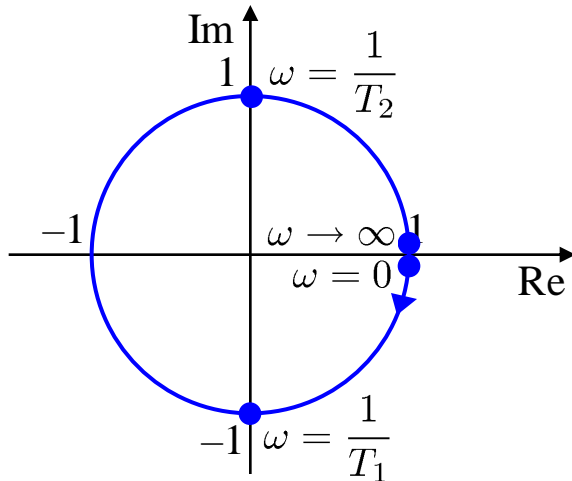


7.5 Allpässe und nicht phasenminimale Systeme

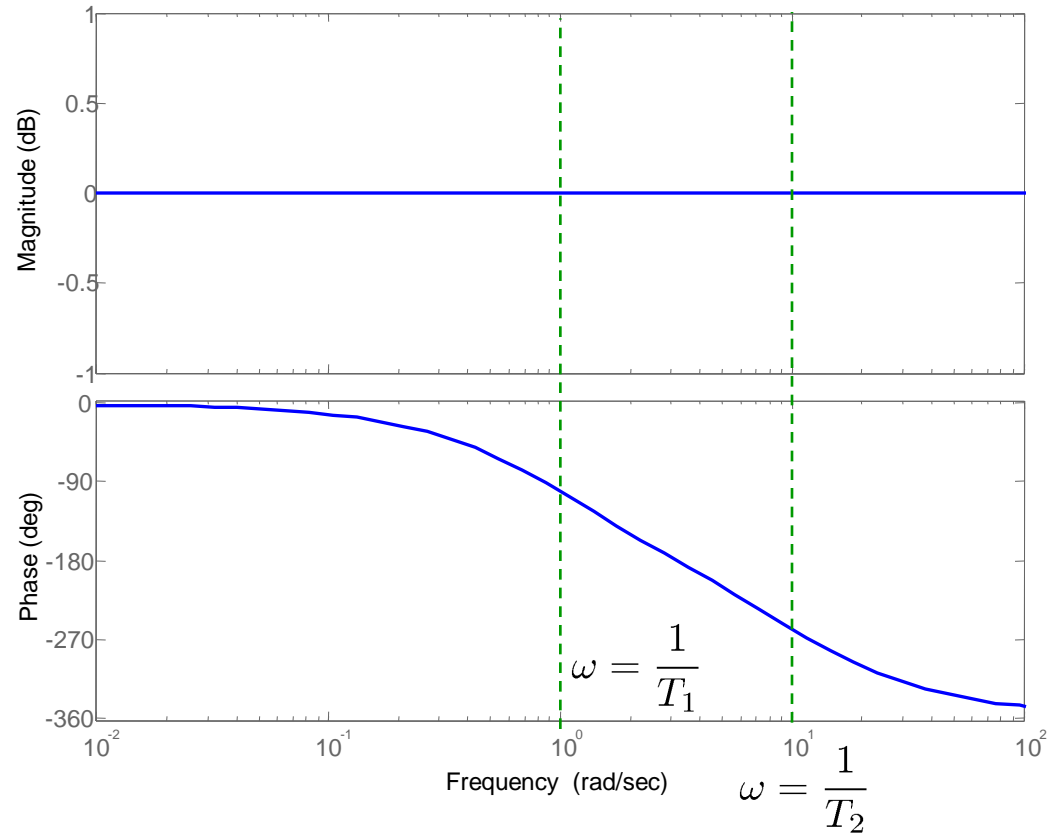
Sprungantwort eines Allpasses 2. Ordnung



Ortskurve eines Allpasses 2. Ordnung



Bode-Diagramm eines Allpasses 2. Ordnung



- Konstanter Betrag von 0 dB.
- Phasenverschiebung: 0° bis -360° .

7.5 Allpässe und nicht phasenminimale Systeme

Zerlegung eines nicht phasenminimalen Systems G_{NPM} in ein phasenminimales System G_{PM} und einen Allpass G_{AP} :

$$G_{\text{NPM}}(s) = G_{\text{PM}}(s) \cdot G_{\text{AP}}(s)$$

Eine solche Zerlegung lässt sich für jedes nicht phasenminimale System finden (ausgenommen Totzeiten). Dies sieht man leicht in der Produkt-Standardform. Die mv stabilen Nullstellen seien mit v_i bezeichnet, die mw instabilen Nullstellen mit w_i , also gilt $v_i < 0$ und $w_i \geq 0$:

$$G_{\text{NPM}}(s) = k \frac{(s - v_1) \dots (s - v_{mv})(s - w_1) \dots (s - w_{mw})}{(s - p_1) \dots (s - p_n)}$$

$$G_{\text{PM}}(s) = k \frac{(s - v_1) \dots (s - v_{mv})(s + w_1) \dots (s + w_{mw})}{(s - p_1) \dots (s - p_n)}$$

$$G_{\text{AP}}(s) = \frac{(s - w_1) \dots (s - w_{mw})}{(s + w_1) \dots (s + w_{mw})}$$

Umdrehen des Vorzeichens macht aus instabilen, stabile Nullstellen bzw. Pole.

7.5 Allpässe und nicht phasenminimale Systeme

Beispiel: $G(s) = \frac{(s + 4)(s - 5)}{(s + 1)(s + 2)(s + 3)}$

Phasenminimaler Anteil:

$$G_{\text{PM}}(s) = \frac{(s + 4)(s + 5)}{(s + 1)(s + 2)(s + 3)}$$

Allpass-Anteil:

$$G_{\text{AP}}(s) = \frac{(s - 5)}{(s + 5)}$$

Somit ergibt sich:

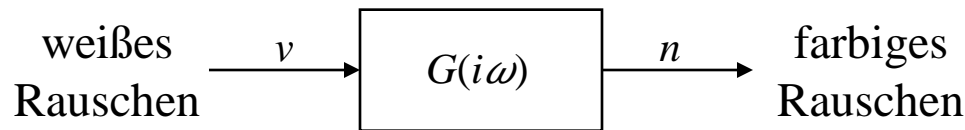
$$G_{\text{PM}}(s)G_{\text{AP}}(s) = \frac{(s + 4)\cancel{(s + 5)}}{(s + 1)(s + 2)(s + 3)} \frac{(s - 5)}{\cancel{(s + 5)}} = G(s)$$

Da Allpässe den Amplitudengang nicht verändern aber die Phase absenken, liegt ein *phasenminimales System* vor, wenn sich kein Allpass in oben gezeigter Form abspalten lässt, also alle Nullstellen stabil sind!

7.5 Allpässe und nicht phasenminimale Systeme

Spektralzerlegung:

Für manche Anwendungen ist nicht der Frequenzgang $G(i\omega)$ selbst wichtig, sondern dessen Betrag bzw. Betragsquadrat $|G(i\omega)|^2$. Dann geht die Information über die Phase verloren und nur der Amplitudengang ist relevant. Eine typische Anwendung ist die Veränderung des Leistungsdichtespektrums eines Rauschsignals, um seine Eigenschaften im Frequenzbereich zu modellieren:



Aus dem gemessenen bzw. berechneten Leistungsdichtespektrum $|G(i\omega)|^2$ muss dann der Frequenzgang $G(i\omega)$ rekonstruiert werden. Da die Phaseninformation fehlt, ist diese Rekonstruktion nicht eindeutig. Dabei ist es *immer* möglich $G(i\omega)$ so zu wählen, dass es stabil und phasenminimal ist:

$$|G(i\omega)|^2 = G(i\omega)G^*(i\omega) = G(i\omega)G(-i\omega)$$

konjugiert komplex

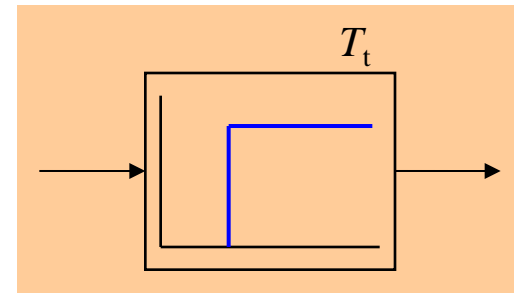
alle Pole und Nullstellen negativ:
stabil und phasenminimal

alle Pole und Nullstellen positiv:
instabil und nicht phasenminimal

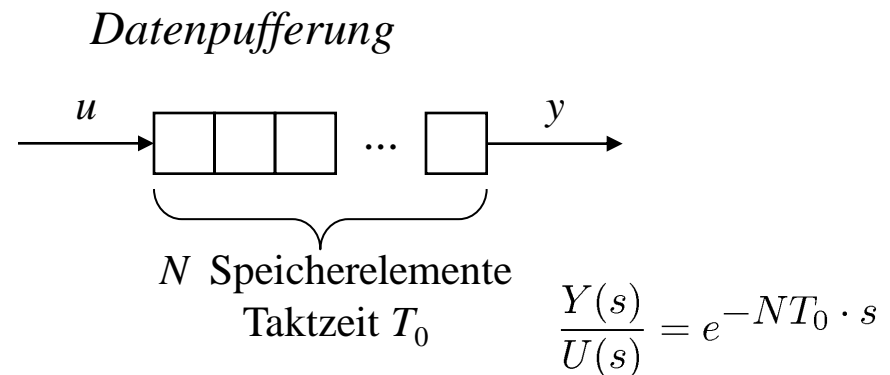
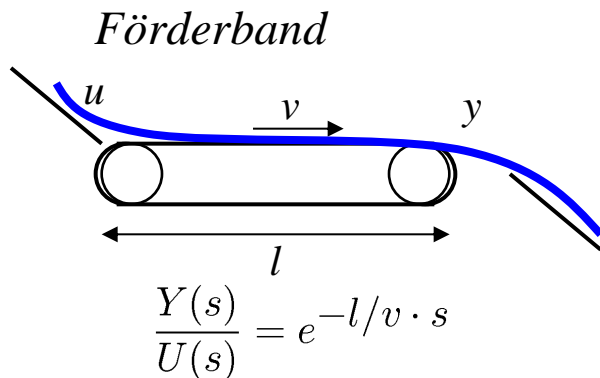
7.6 Totzeiten

Reines Totzeit-Verhalten:

- Übertragungsfunktion: $G(s) = e^{-T_t s}$
- Entsteht z.B. durch Laufzeiten von Material oder Signalen.
- Nur ein Parameter: Totzeit T_t .
- Je größer die Totzeit einer Regelstrecke ist, desto schwieriger ist sie zu regeln.
- Meist wird die Totzeit separat behandelt, denn ihre Laplace-Transformierte ist *keine* gebrochen rationale Funktion, wie bei allen anderen linearen Übertragungsgliedern.
- Totzeiten können über Allpass-Glieder mittels *Padé*-Approximation angenähert werden.

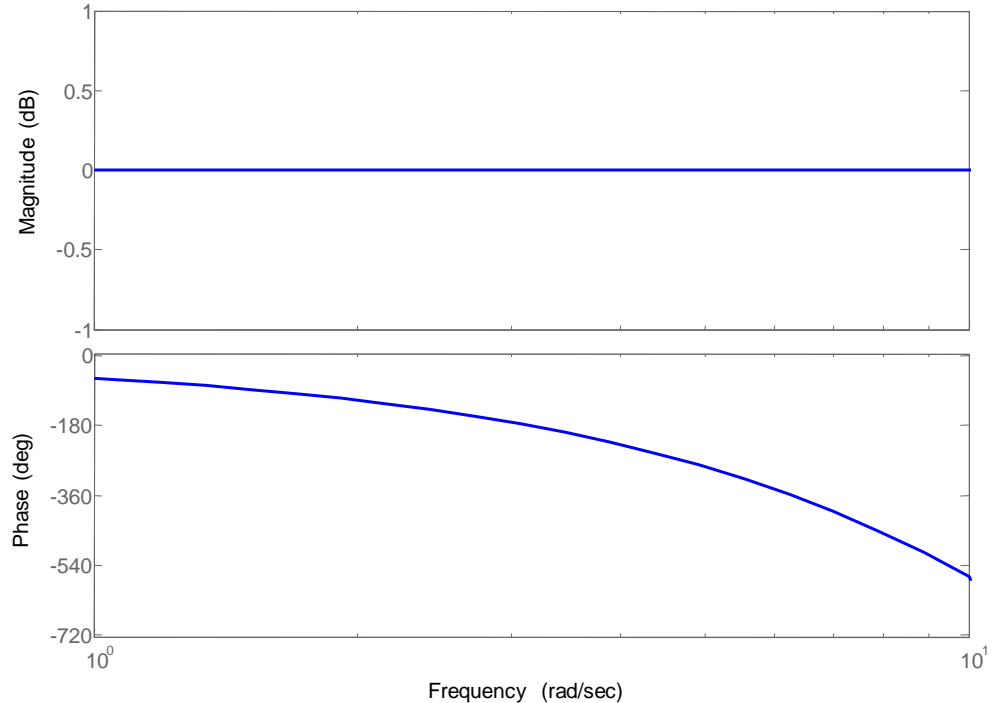
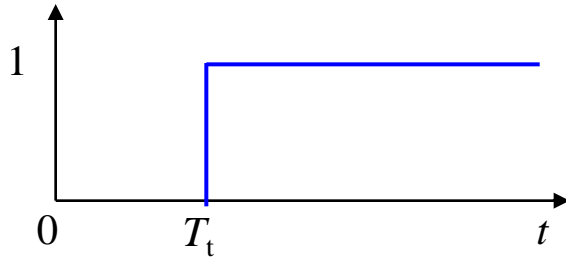


Physikalische Realisierungen:

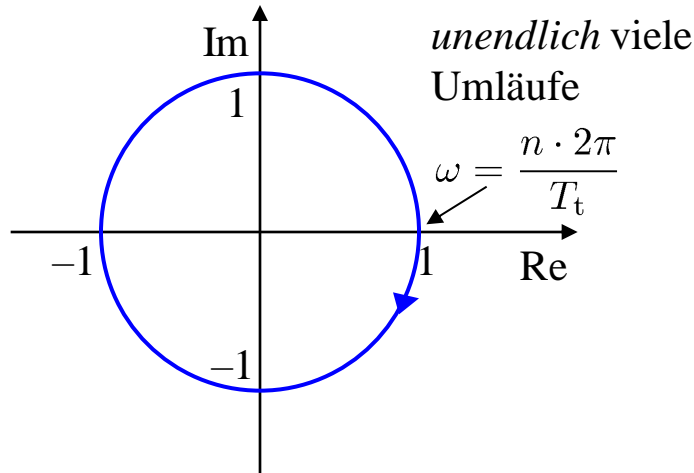


7.6 Totzeiten

Sprungantwort eines Totzeit-Glieds Bode-Diagramm eines Totzeit-Glieds



Ortskurve eines Totzeit-Glieds



- Betrag ist konstant 1.
- Phasenverschiebung wächst proportional zur Totzeit und exponentiell mit der Frequenz.

7.6 Totzeiten

Beispiel: Integrator mit Totzeiten I- T_t :

- Der Integrator lässt den Betrag mit $1/\omega$ gegen Null gehen
- Die Totzeit dreht die Phase immer weiter.
- Kommt von $-T_t - i\infty$ aus $\omega = 0$.
- Geht $\rightarrow 0$ für $\omega \rightarrow \infty$.

$$\text{Info: } \lim_{x \rightarrow 0} \frac{\sin x}{x} = 1$$

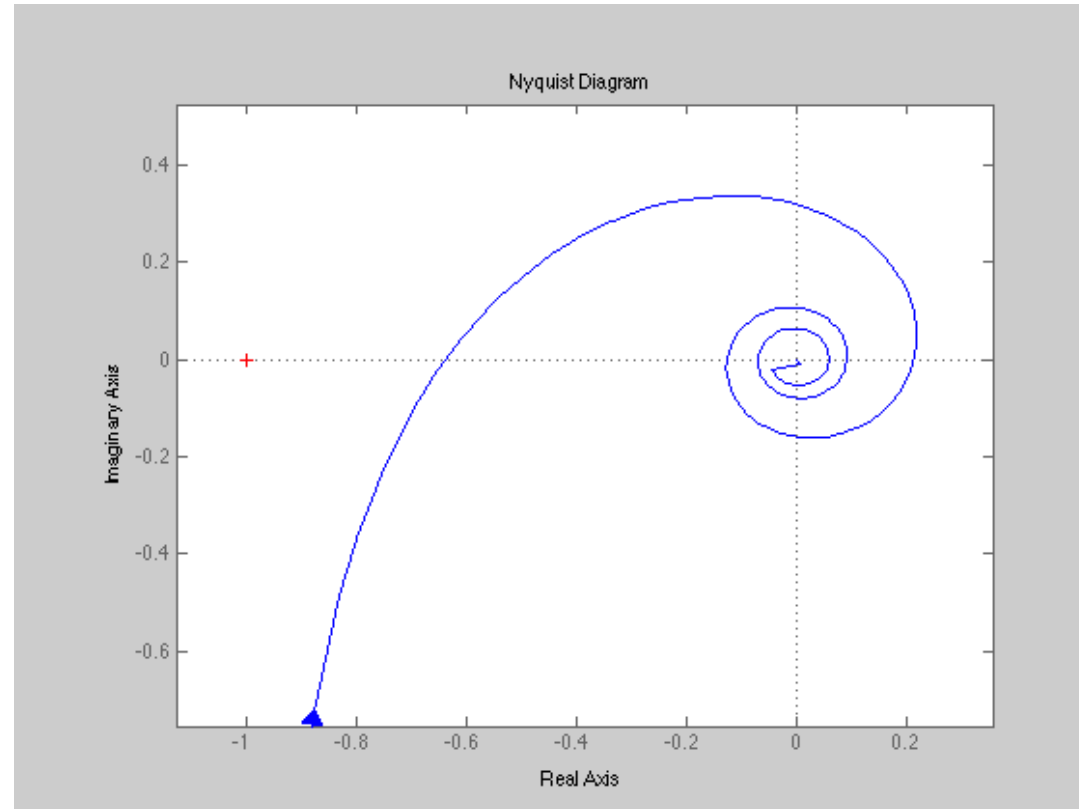
für $T_t = 1$ sec

$$G(s) = \frac{1}{s} e^{-sT_t}$$

$$G(i\omega) = \frac{1}{i\omega} e^{-i\omega T_t}$$

$$= \frac{1}{i\omega} (\cos(-\omega T_t) + i \sin(-\omega T_t))$$

$$= -\frac{1}{\omega} \sin(\omega T_t) - i \frac{1}{\omega} \cos(\omega T_t)$$



7.6 Totzeiten

Padé-Approximation von Totzeiten:

- Totzeit-Glied wird mittels Allpass approximiert.
- Amplitudengang kann exakt nachgebildet werden (konstant 1).
- Approximationsfehler im Phasengang wird mit steigender Frequenz größer.
- Je höher die Ordnung der Padé-Approximation, desto größerer Frequenzbereich kann gut beschrieben werden.

Totzeit

$$G(s) = e^{-T_t s}$$

*Padé-Approximation
1. Ordnung*

$$\hat{G}_1(s) = \frac{-\frac{T_t}{2}s + 1}{\frac{T_t}{2}s + 1}$$

*Padé-Approximation
2. Ordnung*

$$\hat{G}_2(s) = \frac{\frac{T_t}{12}s^2 - \frac{T_t}{2}s + 1}{\frac{T_t}{12}s^2 + \frac{T_t}{2}s + 1}$$

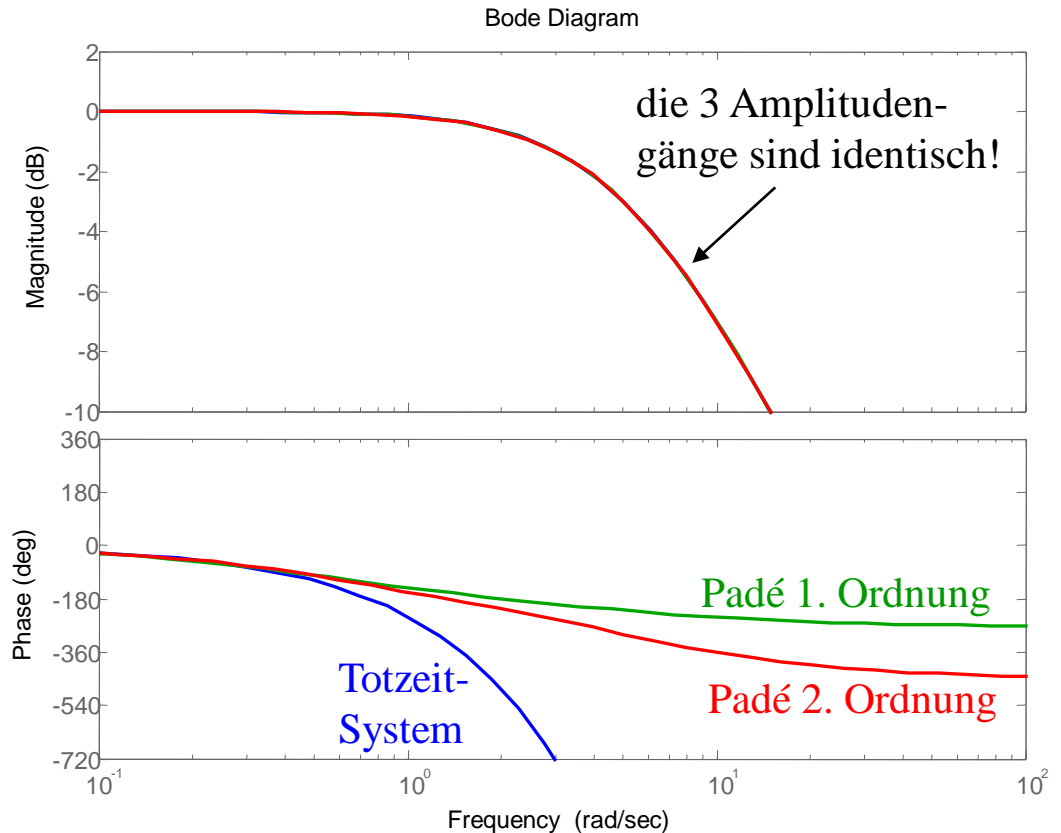
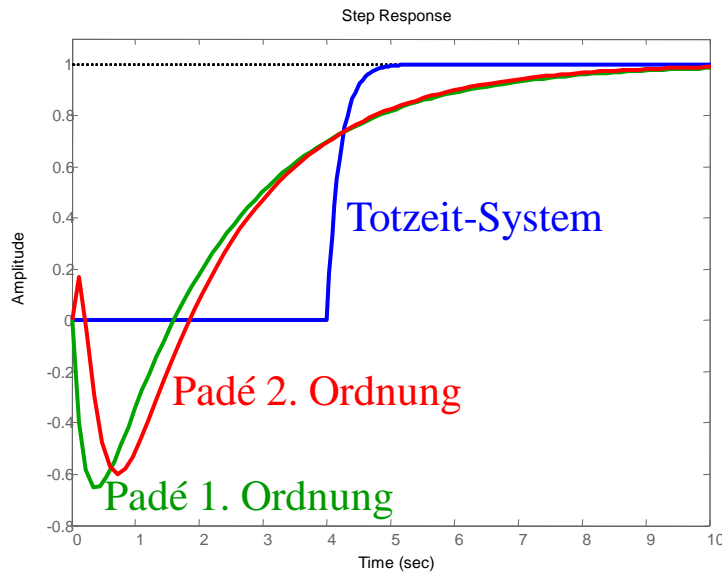
7.6 Totzeiten

Beispiel: $G(s) = \frac{5}{s + 5} e^{-4s}$

$$\hat{G}_1(s) = 5 \frac{-2s + 1}{(s + 5)(2s + 1)}$$

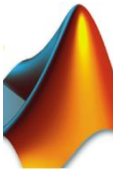
Padé-Approximation 1./2. Ordnung:

$$\hat{G}_2(s) = 5 \frac{\frac{1}{3}s^2 - 2s + 1}{(s + 5)(\frac{1}{3}s^2 + 2s + 1)}$$



7.7 MATLAB-Funktionen

MATLAB



Zeichnen der Spung- und Impulsantwort eines Systems:

```
s = tf('s'); % s als Laplace-Variable definieren
G = 1/(s+1); % PT1-Übertragungsfunktion

step(G); % Zeichnet die Sprungantwort von G
impulse(G); % Zeichnet die Impulsantwort von G
% Simulationszeit wird automatisch
% von MATLAB festgelegt.

step(G, T); % T ist ein Zeitvektor, der die
% Simulationszeit enthält

[h, t] = step(G); % Vektor h enthält die Sprungantwort
% Vektor t enthält die zugeh. Zeitpunkte

[g, t] = impulse(G); % Vektor g enthält die Impulsantwort
```



8. Stabilität linearer Systeme

Inhalt Kapitel 8

8 Stabilität linearer Systeme

8.1 Stabilität im Regelkreis

8.2 Hurwitz-Kriterium

8.3 Nyquist-Kriterium

8.1 Stabilität im Regelkreis

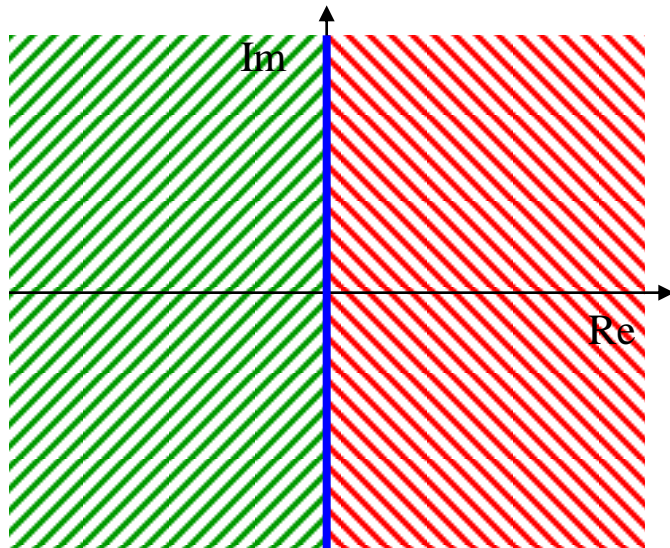
Definition der Stabilität:

Ein lineares dynamisches System $G(s)$ heißt *stabil*, wenn seine Gewichtsfunktion $g(t)$ gegen Null abklingt: $\lim_{t \rightarrow \infty} g(t) = 0$.

Strebt die Gewichtsfunktion betragsmäßig gegen unendlich, heißt das System *instabil*.

Überschreitet die Gewichtsfunktion mit wachsendem t einen endlichen Wert nicht, heißt das System *grenzstabil*.

Wir wissen bereits: Für den Verlauf von $g(t)$ sind die *Pole* von $G(s)$ entscheidend:



- Alle Pole haben Realteil $< 0 \rightarrow$ *stabil*.
- Einfache Pole oder Polpaare haben Realteil $= 0$, alle anderen Pole Realteil $< 0 \rightarrow$ *grenzstabil*.
- Ein oder mehrere Pole haben Realteil > 0 oder ein oder mehrere Mehrfachpole haben Realteil $= 0 \rightarrow$ *instabil*.

8.1 Stabilität im Regelkreis

Beispiele:

$$G(s) = \frac{5}{s+1} \rightarrow \text{stabil}$$

$$G(s) = \frac{5}{s-1} \rightarrow \text{instabil}$$

$$G(s) = \frac{5}{s} \rightarrow \text{grenzstabil}$$

$$G(s) = \frac{3}{s^2-1} \rightarrow \text{instabil}$$

$$G(s) = \frac{3}{s^2+1} \rightarrow \text{grenzstabil}$$

$$G(s) = \frac{3}{s^2} \rightarrow \text{instabil}$$

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

\rightarrow grenzstabil

$$G(s) = \frac{1-s}{(s+1)(s+2)} \rightarrow \text{stabil}$$

$$G(s) = \frac{s-1}{s^2-1} \rightarrow ??$$

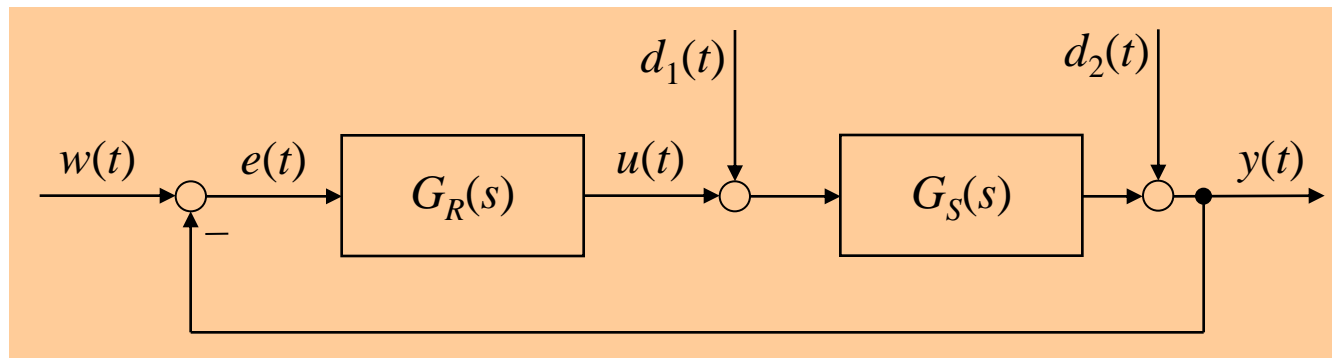
8.1 Stabilität im Regelkreis

Stabilität im Regelkreis:

Ein *Regelkreis* heißt *stabil*, wenn nach einer endlichen Erregung durch Führungs- oder Störsignale seine Regelgröße endlich bleibt. Verschwindet diese Erregung, dann klingt die Regelgröße gegen Null ab: $\lim_{t \rightarrow \infty} y(t) = 0$.

Dies bedeutet, dass die Übertragungsfunktionen von jedem möglichen Eingangssignal zur Regelgröße stabil sein müssen, d.h. insbesondere die Führungsübertragungsfunktion $G_w(s)$ und die Störübertragungsfunktionen $G_{d1}(s)$ und $G_{d2}(s)$.

$$G_w(s) = \frac{G_R(s)G_S(s)}{1 + G_R(s)G_S(s)} \quad G_{d1}(s) = \frac{G_S(s)}{1 + G_R(s)G_S(s)} \quad G_{d2}(s) = \frac{1}{1 + G_R(s)G_S(s)}$$



8.1 Stabilität im Regelkreis

Nullstellen und Pole des geschlossenen Regelkreises:

- Regler: $G_R(s) = \frac{Q(s)}{P(s)}$
- Strecke: $G_S(s) = \frac{B(s)}{A(s)}$
- Offener Regelkreis: $G_0(s) = G_R(s)G_S(s) = \frac{Q(s)}{P(s)} \frac{B(s)}{A(s)}$

Führungsübertragungsfunktion: $G_w(s) = \frac{Q(s)B(s)}{P(s)A(s) + Q(s)B(s)}$ Nullstellen des offenen Kreises bleiben erhalten!
Pole hängen von Nullstellen und Polen des offenen Kreises ab

Störübertragungsfunktion 1: $G_{d1}(s) = \frac{P(s)B(s)}{P(s)A(s) + Q(s)B(s)}$ Andere Nullstellen
Pole identisch mit $G_w(s)$

Störübertragungsfunktion 2: $G_{d2}(s) = \frac{P(s)A(s)}{P(s)A(s) + Q(s)B(s)}$ Andere Nullstellen
Pole identisch mit $G_w(s)$

8.1 Stabilität im Regelkreis

Beobachtung: $G_w(s)$, $G_{d1}(s)$ und $G_{d2}(s)$ haben den selben Nenner!

Im Normalfall (d.h. wenn keine instabile Pol/Nullstellen-Kürzung vorkommt) ist ein Regelkreis also stabil, wenn dieser Nenner

$$1 + G_0(s) = 1 + G_R(s)G_S(s) = 1 + \frac{Q(s)}{P(s)} \frac{B(s)}{A(s)}$$

ausschließlich stabile Nullstellen (= Pole des geschlossenen Kreises) hat.

Vereinfacht können wir also die Nullstellen des folgenden Polynoms zur Stabilitätsprüfung heranziehen:

$$P(s)A(s) + Q(s)B(s)$$

Sonderfall: Kürzung instabiler Pole!

Enthält die Regelstrecke instabile Pole (also $A(s)$ instabile Nullstellen), die durch identische Nullstellen des Reglers (also instabile Nullstellen in $Q(s)$) gekürzt werden, dann ist der geschlossene Regelkreis *instabil!*

Bemerkung: Es gibt aber andere Möglichkeiten eine instabile Regelstrecke zu stabilisieren!

8.1 Stabilität im Regelkreis

→ Ein geschlossener Regelkreis ist *instabil*, wenn instabile Pole der Regelstrecke mit entsprechenden Nullstellen des Reglers gekürzt werden.

Warum ist das so?

Annahme: Regelstrecke hat instabilen Pol p : $A(s) = \tilde{A}(s)(s - p)$

Dieser Pol soll durch den Regler gekürzt werden, also: $Q(s) = \tilde{Q}(s)(s - p)$

$$G_w(s) = \frac{\frac{Q(s) B(s)}{P(s) A(s)}}{1 + \frac{Q(s) B(s)}{P(s) A(s)}} = \frac{\frac{\tilde{Q}(s)(s-p) B(s)}{P(s) \tilde{A}(s)(s-p)}}{1 + \frac{\tilde{Q}(s)(s-p) B(s)}{P(s) \tilde{A}(s)(s-p)}} = \frac{\tilde{Q}(s) B(s)}{P(s) \tilde{A}(s) + \tilde{Q}(s) B(s)}$$

$$G_{d1}(s) = \frac{\frac{B(s)}{A(s)}}{1 + \frac{Q(s) B(s)}{P(s) A(s)}} = \frac{\frac{B(s)}{\tilde{A}(s)(s-p)}}{1 + \frac{\tilde{Q}(s)(s-p) B(s)}{P(s) \tilde{A}(s)(s-p)}} = \frac{1}{(s-p)} \cdot \frac{P(s) B(s)}{P(s) \tilde{A}(s) + \tilde{Q}(s) B(s)}$$

instabil!

$$G_{d2}(s) = \frac{1}{1 + \frac{Q(s) B(s)}{P(s) A(s)}} = \frac{1}{1 + \frac{\tilde{Q}(s)(s-p) B(s)}{P(s) \tilde{A}(s)(s-p)}} = \frac{P(s) \tilde{A}(s)}{P(s) \tilde{A}(s) + \tilde{Q}(s) B(s)}$$

8.1 Stabilität im Regelkreis

Innere Stabilität:

Selbst wenn $G_w(s)$, $G_{d1}(s)$ und $G_{d2}(s)$ stabil sind, kann es innerhalb des Regelkreises zu unendlich großen Signalen kommen, ohne dass dies (theoretisch) von Außen, also an der Regelgröße, bemerkt werden könnte. Eine solche Situation liegt vor, wenn ein instabiler Pol der Reglers gegen eine entsprechende Nullstelle der Strecke gekürzt werden würde.

Betrachten wir dazu die Übertragungsfunktion von der Führungsgröße w zur Stellgröße u :

Annahme: Regler hat instabilen Pol p : $P(s) = \tilde{P}(s)(s - p)$

Dieser Pol wird durch eine Nullstelle in der Strecke gekürzt: $B(s) = \tilde{B}(s)(s - p)$

$$G_{wu}(s) = \frac{G_R(s)}{1 + G_R(s)G_S(s)} = \frac{\frac{Q(s)}{P(s)}}{1 + \frac{Q(s)}{P(s)} \frac{B(s)}{A(s)}} = \frac{\frac{Q(s)}{\tilde{P}(s)(s-p)}}{1 + \frac{Q(s)}{\tilde{P}(s)(s-p)} \frac{\tilde{B}(s)(s-p)}{A(s)}}$$

instabil! $\frac{1}{(s-p)} \cdot \frac{Q(s)A(s)}{\tilde{P}(s)A(s) + Q(s)\tilde{B}(s)}$, d.h. die Stellgröße u wächst über alle Grenzen!

Innere Stabilität des Regelkreises liegt also nur dann vor, wenn auch alle inneren Signale endlich bleiben. Daher müssen wir fordern, dass $G_w(s)$, $G_{d1}(s)$, $G_{d2}(s)$ und $G_{wu}(s)$ stabil sind.

8.1 Stabilität im Regelkreis

Kürzung instabiler Pole:

- Führt selbst bei exakter Kürzung zu einer instabilen Störübertragungsfunktion $G_{d1}(s)$.
- Eine exakte Kürzung ist in der Praxis gar nicht zu erreichen, weil
 1. das Modell die Regelstrecke immer nur ungefähr widerspiegelt (Modellfehler),
 2. die Streckendynamik selbst (zumindest kleinen) zeitlichen Veränderungen unterworfen ist (Alterung, Abhängigkeit von im Modell vernachlässigten Einflussgrößen, ...).

In der Regel ergibt sich durch eine inexakte Kürzung instabiler Pole ein instabiles Gesamtsystem. Siehe Kapitel 13.

Kürzung instabiler Nullstellen:

- Aus den o.g. Gründen führt auch ein Kürzen *instabiler Nullstellen* der Regelstrecke mit entsprechenden Polen im Regler zu einem instabilen Regelkreis.
- Daher dürfen instabile Nullstellen in $B(s)$ nicht gekürzt werden und der geschlossene Regelkreis $G_w(s)$ hat zwangsläufig nicht phasenminimales Verhalten, falls die Strecke nicht phasenminimales Verhalten hat!

8.1 Stabilität im Regelkreis

Stabilität ist die wichtigste Forderung an einen Regelkreis. Deshalb wurden viele Kriterien entwickelt, um die Stabilität zu überprüfen:

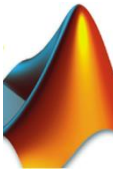
- *Direkte Berechnung* der Pole des *geschlossenen* Regelkreises, also der Nullstellen des charakteristischen Polynoms $P(s)A(s) + Q(s)B(s)$:
 1. Erfordert numerische Suchverfahren.
 2. War früher sehr aufwendig, heute mit MATLAB leicht möglich.
 3. Pole geben neben Stabilität auch Auskunft zur Entfernung von der Stabilitätsgrenze und der Art und Geschwindigkeit der Dynamik des geschlossen Regelkreises.
- Analyse der Koeffizienten des charakteristischen Polynoms des *geschlossenen* Regelkreises $P(s)A(s) + Q(s)B(s)$:
 1. Verfahren nach *Hurwitz* oder nach *Routh*.
 2. Liefern nur eine Ja/Nein-Aussage.
 3. Schnell und einfach für Systeme bis 3. Ordnung.

8.1 Stabilität im Regelkreis

- Analyse der *Ortskurve* des *offenen* Regelkreises $G_0(s)$:
 1. *Nyquist*-Kriterium.
 2. Amplituden- und Phasenrand geben auch Auskunft über Nähe zur Stabilitätsgrenze.
 3. Kann auch für den Reglerentwurf genutzt werden.
- Zeichnen der *Wurzelortskurve (WOK)*, d.h. Zeichnen der Pollagen des geschlossenen Regelkreises in der komplexen s -Ebene in Abhängigkeit der Reglerverstärkung (oder eines anderen Parameters). Zum Zeichnen der WOK werden aber lediglich die Pole und Nullstellen des *offenen* Regelkreises benötigt!
 1. Auskunft über Nähe zur Stabilitätsgrenze.
 2. Vermittelt Einsichten zur optimalen Wahl der Reglerverstärkung.
 3. Liefert Hinweise zur Auswahl der Reglerstruktur.
- *Q-Parametrierung* bzw. *Internal Model Control*:
 1. Neuartige Parametrierung des Reglers bzw. neuartige Regelkreisstruktur.
 2. Dadurch ist der Regelkreis *automatisch stabil*, völlig unabhängig von der Wahl des (neuen) Reglers!

8.1 Stabilität im Regelkreis

MATLAB



Berechnung der Nullstellen des charakteristischen Polynoms:

```
Polynom = [2 0 1 3];           % Koeffizienten des charakteristischen
                                % Polynoms in absteigender Reihenfolge
                                %  $2s^3 + s + 3$ 

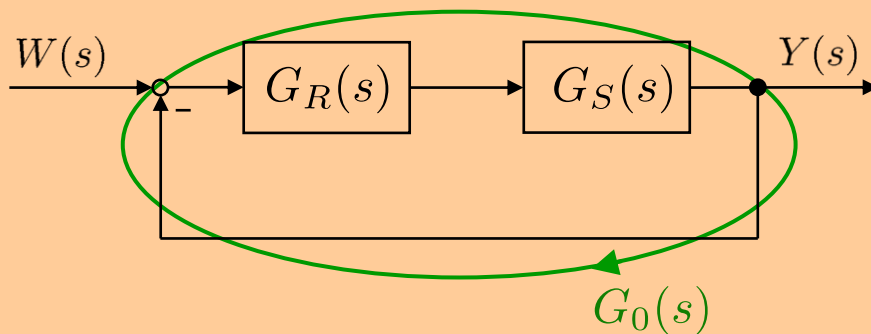
Nullstellen = roots(Polynom);  % Berechnet die Nullstellen des Polynoms
                                % Liefert:
                                %  $0.5000 + 1.1180i$ 
                                %  $0.5000 - 1.1180i$ 
                                %  $-1.0000$ 
```

8.2 Hurwitz-Kriterium

Zur Erinnerung:

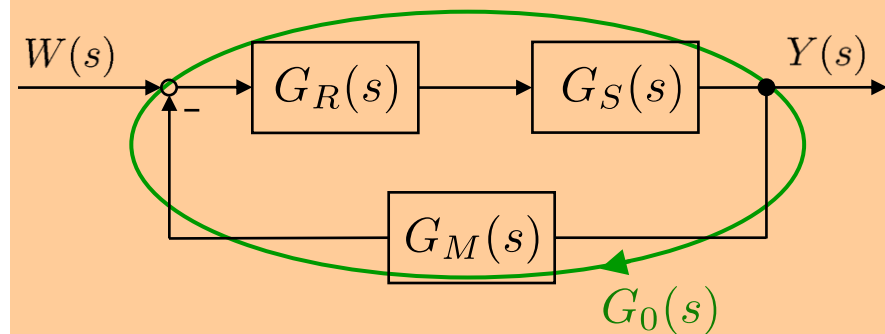
- Offener Regelkreis: $G_0(s) = \begin{cases} G_R(s)G_S(s) & \text{ohne Sensordynamik } G_M(s) = 1 \\ G_R(s)G_S(s)G_M(s) & \text{mit Sensordynamik } G_M(s) \end{cases}$
- Geschlossener Regelkreis: $G_w(s) = \frac{\text{Vorwärtszweig}}{1 + G_0(s)} = \frac{G_R(s)G_S(s)}{1 + G_0(s)}$

Ohne Sensordynamik



$$G_w(s) = \frac{G_R(s)G_S(s)}{1 + G_R(s)G_S(s)}$$

Mit Sensordynamik



$$G_w(s) = \frac{G_R(s)G_S(s)}{1 + G_R(s)G_S(s)G_M(s)}$$

8.2 Hurwitz-Kriterium

Eigenschaften des *Hurwitz-Kriteriums*:

- Überprüft anhand der Koeffizienten eines Polynoms, ob die Nullstellen dieses Polynoms alle einen negativen Realteil aufweisen (also stabile Nullstellen sind).
- Liefert nur eine Ja/Nein-Aussage.
- Bequem für Polynome niedriger Ordnung; für höhere Ordnungen (>3–4) heute nicht mehr besonders hilfreich, da sich die Nullstellen mit MATLAB schnell numerisch berechnen lassen.
- Stabilität lässt sich in Abhängigkeit von Koeffizienten bestimmen (geht numerisch nicht).

→ **Ausgangspunkt:** Charakteristische Gleichung (= Nenner der Übertragungsfunktion des geschlossenen Regelkreises):

$$C(s) = P(s)A(s) + Q(s)B(s) = c_n s^n + c_{n-1} s^{n-1} + \dots + c_1 s + c_0 = 0$$

Es werden bestimmte Bedingung für die Koeffizienten c_i ($i = 0, 1, \dots, n$) aufgestellt. Sind alle diese Bedingungen erfüllt, dann haben alle Nullstellen der charakteristischen Gleichung einen negativen Realteil und damit sind alle Pole des geschlossenen Regelkreises stabil!

8.2 Hurwitz-Kriterium

Hurwitz-Matrix:

$$\underline{H} = \begin{pmatrix} c_1 & c_3 & c_5 & c_7 & \cdots \\ c_0 & c_2 & c_4 & c_6 & \cdots \\ 0 & c_1 & c_3 & c_5 & \cdots \\ 0 & c_0 & c_2 & c_4 & \cdots \\ 0 & 0 & c_1 & c_3 & \cdots \\ 0 & 0 & c_0 & c_2 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Obere linke Determinanten der Hurwitz-Matrix:

$$\begin{pmatrix} \boxed{c_1} & \boxed{c_3} & \boxed{c_5} & \boxed{c_7} & \cdots \\ \boxed{c_0} & \boxed{c_2} & \boxed{c_4} & \boxed{c_6} & \cdots \\ \boxed{0} & \boxed{c_1} & \boxed{c_3} & \boxed{c_5} & \cdots \\ \boxed{0} & \boxed{c_0} & \boxed{c_2} & \boxed{c_4} & \cdots \\ 0 & 0 & c_1 & c_3 & \cdots \\ 0 & 0 & c_0 & c_2 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad \begin{array}{l} \boxed{D_1 = c_1} \\ \boxed{D_2 = \det \begin{pmatrix} c_1 & c_3 \\ c_0 & c_2 \end{pmatrix}} \\ \boxed{D_3 = \det \begin{pmatrix} c_1 & c_3 & c_5 \\ c_0 & c_2 & c_4 \\ 0 & c_1 & c_3 \end{pmatrix}} \end{array}$$

$$D_4 = \det \begin{pmatrix} c_1 & c_3 & c_5 & c_7 \\ c_0 & c_2 & c_4 & c_6 \\ 0 & c_1 & c_3 & c_5 \\ 0 & c_0 & c_2 & c_4 \end{pmatrix}$$

8.2 Hurwitz-Kriterium

Hurwitz-Kriterium:

Alle Nullstellen des Polynoms $c_n s^n + c_{n-1} s^{n-1} + \dots + c_1 s + c_0$ haben genau dann einen negativen Realteil, wenn die beiden folgenden Bedingungen erfüllt sind:

1. Alle Koeffizienten sind positiv:

$$c_i > 0 \quad (i = 0, 1, \dots, n).$$

2. Die $n-1$ führenden Hauptabschnittsdeterminanten D_i der Hurwitz-Matrix \underline{H} sind positiv:

$$D_i > 0 \quad (i = 1, 2, \dots, n - 1).$$

Die c_i mit $i > n$ sind zu Null zu setzen.

8.2 Hurwitz-Kriterium

Hurwitz-Kriterium für Polynome 1. – 4. Ordnung:

1. Ordnung:

$$c_0 > 0, c_1 > 0.$$

2. Ordnung:

$$c_0 > 0, c_1 > 0, c_2 > 0.$$

3. Ordnung:

$$c_0 > 0, c_1 > 0, c_2 > 0, c_3 > 0, \\ c_1 c_2 - c_0 c_3 > 0.$$

ab hier muss das Stabilitätsgebiet nicht mehr *konvex* sein!

4. Ordnung:

$$c_0 > 0, c_1 > 0, c_2 > 0, c_3 > 0, c_4 > 0, \\ c_1 c_2 - c_0 c_3 > 0, \\ c_1 c_2 c_3 - c_0 c_3^2 - c_1^2 c_4 > 0.$$

8.2 Hurwitz-Kriterium

Beispiele:

Regler: $G_R(s) = K$ Strecke: $G_S(s) = \frac{1}{(s+2)(s+5)} = \frac{1}{s^2 + 7s + 10}$ **Stabile Strecke!**

Charakteristische Gleichung: $1 + G_0(s) = 1 + \frac{K}{s^2 + 7s + 10} = 0 \rightarrow \boxed{s^2 + 7s + 10 + K = 0}$

→ Stabil für $K > -10$. D.h. für alle positiven Reglerverstärkung ist geschlossener Kreis stabil.

Regler: $G_R(s) = K$ Strecke: $G_S(s) = \frac{1}{(s-2)(s+5)} = \frac{1}{s^2 + 3s - 10}$ **Instabile Strecke! Langsamer instabiler Pol.**

Charakteristische Gleichung: $1 + G_0(s) = 1 + \frac{K}{s^2 + 3s - 10} = 0 \rightarrow \boxed{s^2 + 3s - 10 + K = 0}$

→ Stabil für $K > 10$.

8.2 Hurwitz-Kriterium

Beispiele:

Instabile Strecke! Schneller instabiler Pol.

Regler: $G_R(s) = K$ Strecke: $G_S(s) = \frac{1}{(s+2)(s-5)} = \frac{1}{s^2 - 3s - 10}$

Charakteristische Gleichung: $1 + G_0(s) = 1 + \frac{K}{s^2 - 3s - 10} = 0 \rightarrow \boxed{s^2 - 3s - 10 + K = 0}$

→ **Immer instabil!** Koeffizient "-3" lässt sich durch P-Regler nicht beeinflussen!

PDT₁-Regler Instabile Strecke! Schneller instabiler Pol.

Regler: $G_R(s) = K \frac{s+1}{s+10}$ Strecke: $G_S(s) = \frac{1}{(s+2)(s-5)} = \frac{1}{s^2 - 3s - 10}$

Charakteristische Gleichung: $1 + G_0(s) = 1 + \frac{K(s+1)}{s^3 + 7s^2 - 40s - 100} = 0$

→ $\boxed{s^3 + 7s^2 + (K - 40)s + (K - 100) = 0}$

→ Stabil für **$K > 100$** .

$(K - 40) \cdot 7 - (K - 100) \cdot 1 > 0 \rightarrow K > 30$ und $K > 40$ und $K > 100$

8.3 Nyquist-Kriterium

Eigenschaften des *Nyquist-Kriteriums*:

- Überprüft die Stabilität des geschlossenen Regelkreises anhand der Ortskurve des *offenen* Regelkreises.
- Erlaubt quantitative Aussagen über die Stabilität, d.h. Entfernung zur Stabilitätsgrenze.
- Unterstützt den Reglerentwurf. Der Kompromiss zwischen Reglerverstärkung und Entfernung von der Stabilitätsgrenze wird veranschaulicht.
- Ist einfach durchzuführen, auch für *komplexe* Systeme, auch für Systeme mit *Totzeit*.
- Benötigt kein mathematisches Modell der Strecke (Übertragungsfunktion). Die Ortskurve des offenen Kreises kann auch experimentell bestimmt worden sein.

Also, welcher Zusammenhang existiert zwischen der Stabilität des geschlossenen Kreises und der Ortskurve des offenen Kreises?

8.3 Nyquist-Kriterium

Vorbereitungen zum Nyquist-Kriterium:

Offener Regelkreis: $G_0(s) = \frac{Z_0(s)}{N_0(s)}$ enthält Nullstellen des offenen Kreises
enthält Pole des offenen Kreises

Für reale Systeme ist Nennerordnung größer als Zählerordnung ($n > m$):

$$1 + G_0(s) = 1 + \frac{Z_0(s)}{N_0(s)} = \frac{N_0(s) + Z_0(s)}{N_0(s)} = k \frac{(s - p_{g1}) \cdot \dots \cdot (s - p_{gn})}{(s - p_{o1}) \cdot \dots \cdot (s - p_{on})}$$

Pole des geschlossenen Kreises
Pole des offenen Kreises

Wir untersuchen im Folgenden die Ortskurve von $1 + G_0(s)$. Später führen wir diese auf die Ortskurve von $G_0(s)$ zurück.

Berechnen wir den Winkel obiger komplexer Gleichung:

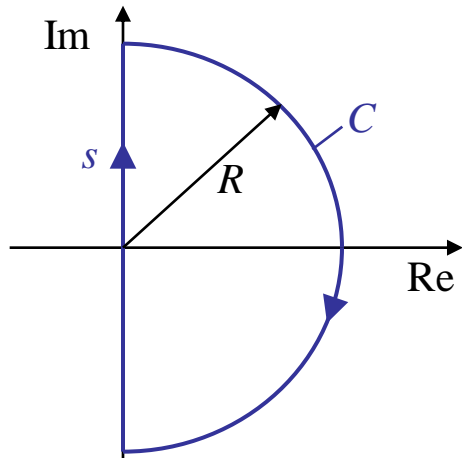
$$\angle(1 + G_0(s)) = \angle \left(k \frac{(s - p_{g1}) \cdot \dots \cdot (s - p_{gn})}{(s - p_{o1}) \cdot \dots \cdot (s - p_{on})} \right) = \angle k + \sum_{i=1}^n \angle(s - p_{gi}) - \sum_{i=1}^n \angle(s - p_{oi})$$

↑
bekannt, wenn man die Ortskurve des offenen Kreises kennt
↑
unabhängig von s
↑
unbekannt
↑
bekannt

8.3 Nyquist-Kriterium

$$\sum_{i=1}^n \angle(s - p_{gi}) = \angle(1 + G_0(s)) + \sum_{i=1}^n \angle(s - p_{oi}) - \angle k$$

Jetzt lassen wir s entlang der Imaginärachse von $-iR$ bis iR laufen und entlang eines Halbkreises mit Radius R durch die rechte s -Halbebene zurück und nennen diese Kurve C :



Für $R \rightarrow \infty$ umschließt die Kurve C alle instabilen Pole, weil diese in der positiven s -Halbebene liegen.

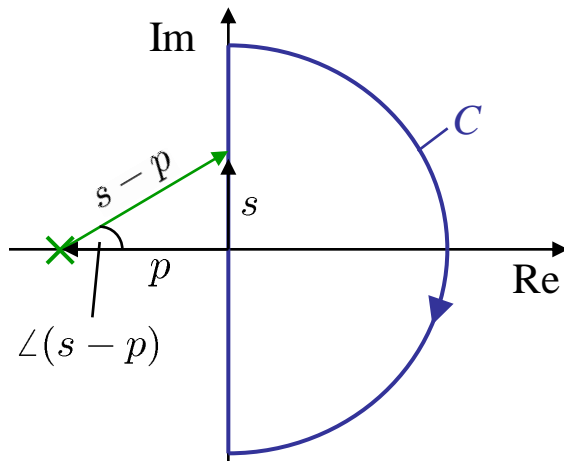
Welche Winkeländerung ergibt sich, wenn s einmal obige Kurve C durchläuft?

$$\sum_{i=1}^n \Delta_C \angle(s - p_{gi}) = \Delta_C \angle(1 + G_0(s)) + \sum_{i=1}^n \Delta_C \angle(s - p_{oi}) - \Delta_C \angle k$$

8.3 Nyquist-Kriterium

Berechnung der Winkeländerung für die Terme vom Typ $\Delta_C \angle(s - p)$

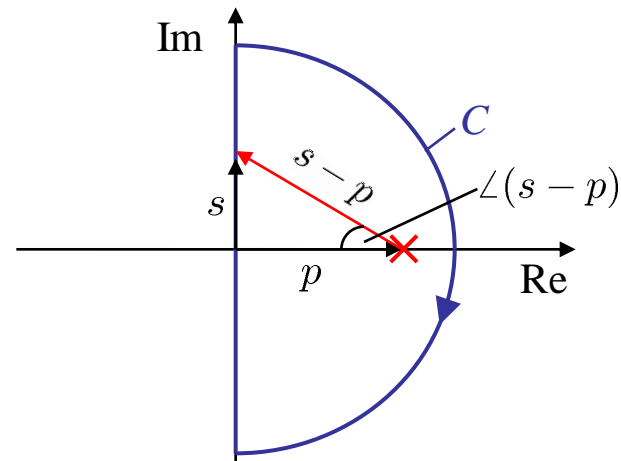
Stabiler Pol



Beim Durchlaufen der Kurve C "wackelt" der grüne Zeiger $s-p$ nach oben und unten aber insgesamt wird kein Winkel verändert:

$$\angle(s - p) = 0$$

Instabiler Pol



Beim Durchlaufen der Kurve C dreht sich der rote Zeiger $s-p$ einmal um volle 360° bzw. 2π im Uhrzeigersinn:

$$\angle(s - p) = 2\pi$$

8.3 Nyquist-Kriterium

Damit ergibt sich für die Winkeländerungen:

$\Delta_C \angle k = 0$, (für $k > 0$) da k ja gar nicht von s abhängt.

$\sum_{i=1}^n \Delta_C \angle (s - p_{gi}) = 2\pi n_g^+$, mit n_g^+ als Anzahl der instabilen Pole des *geschlossenen* Kreises.

$\sum_{i=1}^n \Delta_C \angle (s - p_{oi}) = 2\pi n_o^+$, mit n_o^+ als Anzahl der instabilen Pole des *offenen* Kreises.

$$\rightarrow \boxed{2\pi n_g^+ = \Delta_C \angle (1 + G_0(s)) + 2\pi n_o^+}$$

Für nicht sprungfähige Systeme gilt in der rechten s -Halbebene (da Nennergrad $n >$ Zählergrad m): $\lim_{s \rightarrow \infty} G(s) = 0$.

Daher brauchen wir für die Winkeländerung von $1 + G_0(s)$ entlang der Kurve C nur noch die Imaginärachse zu betrachten. Der Anteil des (unendlich großen) Halbkreises fällt weg:

$$\Delta_C \angle (1 + G_0(s)) = \angle (1 + G_0(i\omega))|_{\omega=-\infty \dots \infty}$$

8.3 Nyquist-Kriterium

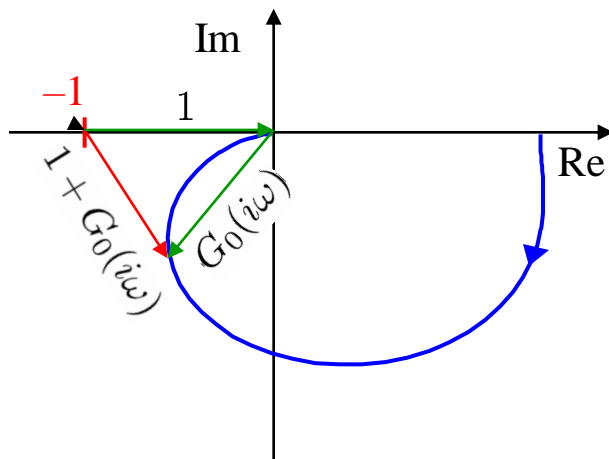
Damit ergibt sich:

$$2\pi n_g^+ = \angle (1 + G_0(i\omega)|_{\omega=-\infty \dots \infty}) + 2\pi n_o^+$$

Wir sind gewohnt Ortskurven nur für positive Kreisfrequenzen zu zeichnen und den symmetrischen negativen Teil wegzulassen (bringt keine Zusatzinformationen). Wenn wir also die Ortskurve nur für positive ω zeichnen ergibt sich:

$$\pi n_g^+ = \angle (1 + G_0(i\omega)|_{\omega=0 \dots \infty}) + \pi n_o^+$$

Geometrische Bedeutung der Ortskurve $1 + G_0(s)$:



Zur Auswertung von $\angle (1 + G_0(i\omega)|_{\omega=0 \dots \infty})$ müssen wir also die Drehung des Zeigers vom Punkt -1 an die Ortskurve des offenen Kreises $G_0(i\omega)$ bestimmen.

8.3 Nyquist-Kriterium

Für Stabilität des geschlossenen Regelkreises fordern wir, dass der geschlossene Kreis keine instabilen Pole aufweist: $n_g^+ = 0$. Damit ergibt sich:

$$\angle (1 + G_0(i\omega)|_{\omega=0\dots\infty}) = -\pi n_o^+$$

$$\text{bzw.} = -\pi n_o^+ - \frac{\pi}{2} n_o^g$$

wenn der offene Kreis n_o^g grenzstabile Pole hat.

Allgemeines Nyquist-Kriterium:

Weist der offene Regelkreis n_o^+ instabile Pole auf (grenzstabile Pole zählen halb!), dann muss die Ortskurve (für positive ω) des offenen Regelkreises den Punkt -1 der komplexen Ebene $\frac{1}{2}n_o^+$ -mal im Gegenuhrzeigersinn umschlingen bzw. der Zeiger vom Punkt -1 an die Ortskurve muss einen Winkel von $-\pi n_o^+$ (oder $-180^\circ \cdot n_o^+$) überstreichen, damit der geschlossen Regelkreis stabil ist.

Uhrzeigersinn = positive Winkel
Gegenuhrzeigersinn = negative Winkel

Einfaches Nyquist-Kriterium:

Ist der offene Regelkreis stabil (mit Ausnahme eines evtl. Pols bei $s = 0$), dann darf die Ortskurve (für positive ω) des offenen Regelkreises den Punkt -1 der komplexen Ebene nicht umschlingen, damit der geschlossen Regelkreis stabil ist.

8.3 Nyquist-Kriterium

Erweiterungen und Ergänzungen:

- *Grenzstabile* Pole des offenen und des geschlossenen Regelkreises zählen als *halbe* instabile Pole.
- Neben einer Stabilitätsaussage liefert das Nyquist-Kriterium auch Informationen über die *Entfernung zur Stabilitätsgrenze*. Dies erlaubt auch eine Aussage darüber, wie stark sich Modell und realer Prozess für verschiedene Frequenzen voneinander unterscheiden dürfen, um dennoch Stabilität garantieren zu können. Dies führt auf die Definition der *robusten Stabilität*, siehe Kapitel 9.
- Das einfache Nyquist-Kriterium wird manchmal auch als "*Linke-Hand-Regel*" formuliert:
Linke-Hand-Regel: Ist der offene Regelkreis stabil, dann muss der Punkt -1 links von der in Richtung wachsender Frequenzen durchlaufenen Ortskurve des offenen Regelkreises liegen, damit der geschlossenen Regelkreis stabil ist.
- Das Nyquist-Kriterium ist eines der wenigen Stabilitätskriterium, das auch für Systeme mit *Totzeit* anwendbar ist.

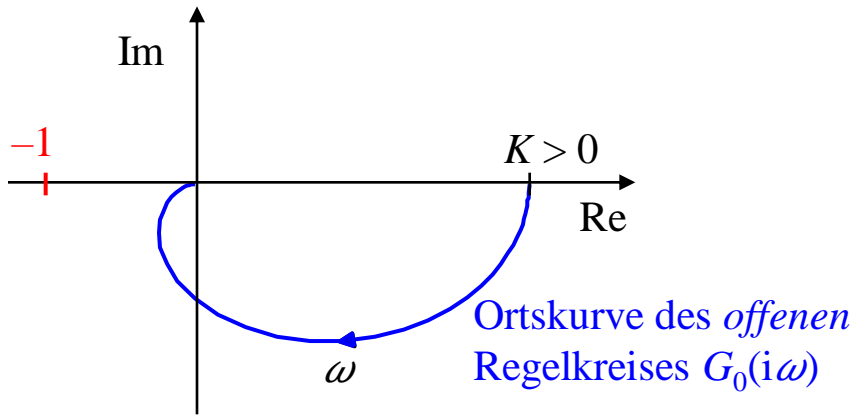
8.3 Nyquist-Kriterium

Beispiel: Stabiler offener Regelkreis

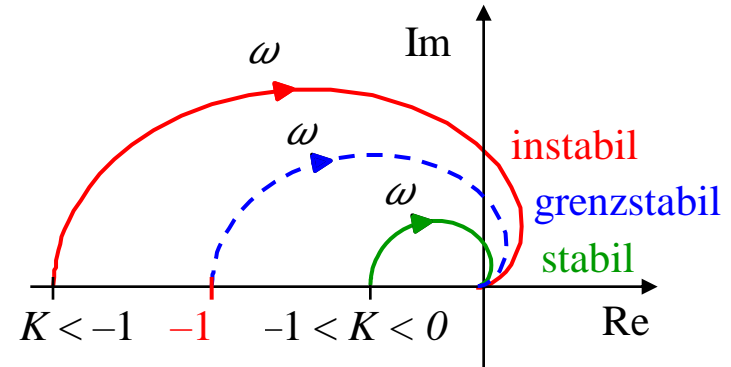
Regler: $G_R(s) = K$ Strecke: $G_S(s) = \frac{10}{s^2 + 7s + 10}$

Offener Regelkreis: $G_0(s) = \frac{10K}{s^2 + 7s + 10}$

Nach Hurwitz ist $G_0(s)$ stabil $\rightarrow n_o^+ = 0$.



Negative Reglerverstärkungen:



Linke-Hand-Regel: Punkt -1 liegt links der Ortskurve \rightarrow geschlossener Regelkreis ist stabil!

Diese Stabilitätsaussage gilt unabhängig von der Reglerverstärkung $K > 0$. Ein solches Verhalten nennt man *strukturstabil*!

Für negative Reglerverstärkungen kleiner -1 wird der Regelkreis allerdings instabil!

8.3 Nyquist-Kriterium

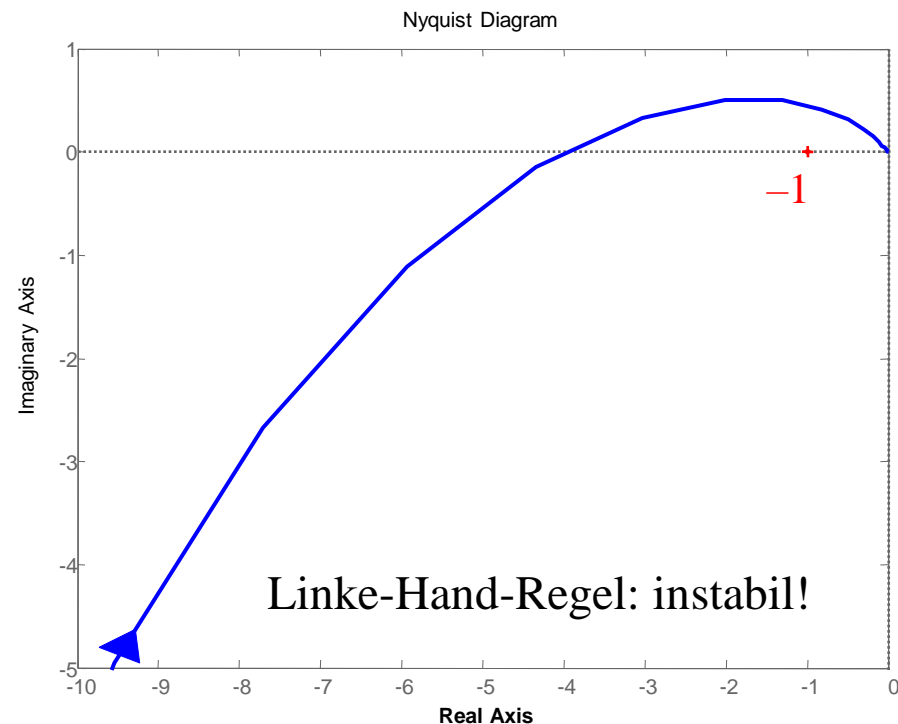
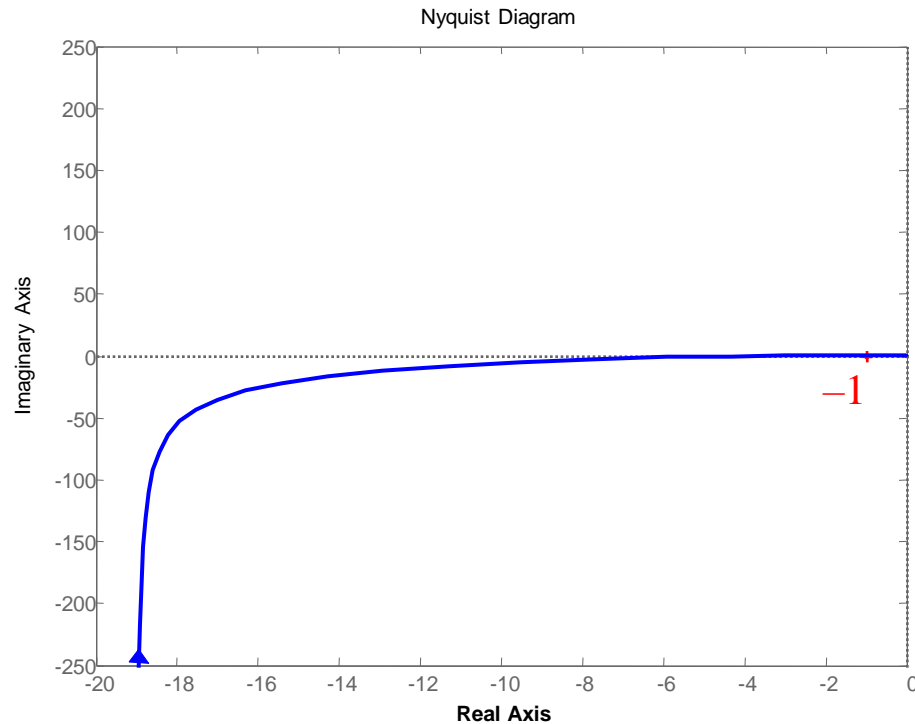
Beispiel: Grenzstabiler offener Regelkreis

Regler: $G_R(s) = 1 + \frac{10}{s} = \frac{s + 10}{s}$

Strecke: $G_S(s) = \frac{1}{s^2 + 2s + 1}$

Offener Regelkreis: $G_0(s) = \frac{s + 10}{s^3 + 2s^2 + s}$

$\rightarrow n_o^+ = 1/2.$



8.3 Nyquist-Kriterium

Beispiel: Instabiler offener Regelkreis

Regler: $G_R(s) = K$

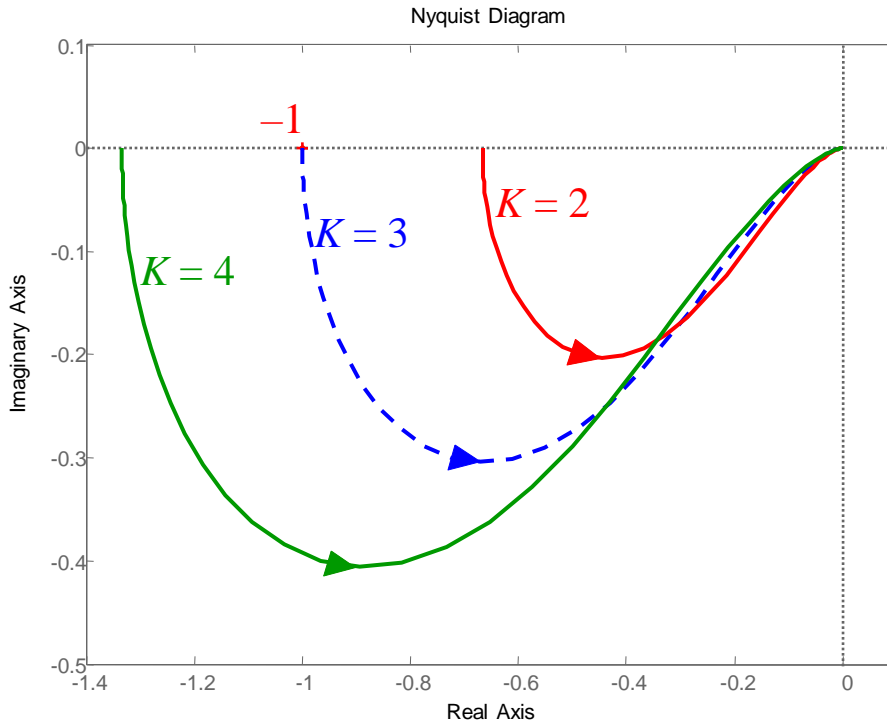
Strecke: $G_S(s) = \frac{1}{(s+3)(s-1)} = \frac{1}{s^2 + 2s - 3}$

Offener Regelkreis: $G_0(s) = \frac{K}{s^2 + 2s - 3}$

Zur Erinnerung:

$$\pi n_g^+ = \angle (1 + G_0(i\omega))|_{\omega=0 \dots \infty} + \pi n_o^+$$

$\rightarrow n_o^+ = 1.$



Drehwinkel um den Punkt -1:

$K > 3:$ -180° bzw. $-\pi$

$0 < K < 3:$ 0° bzw. 0

Für $K > 3:$ $\pi n_g^+ = -\pi + \pi$

$n_g^+ = 0$ instabile Pole im geschl. Kreis

Für $0 < K < 3:$ $\pi n_g^+ = 0 + \pi$

$n_g^+ = 1$ instabiler Pol im geschl. Kreis

\rightarrow Geschlossener Kreis für $K > 3$ stabil!

8.3 Nyquist-Kriterium

Beispiel: Instabiler offener Regelkreis

Regler: $G_R(s) = K$

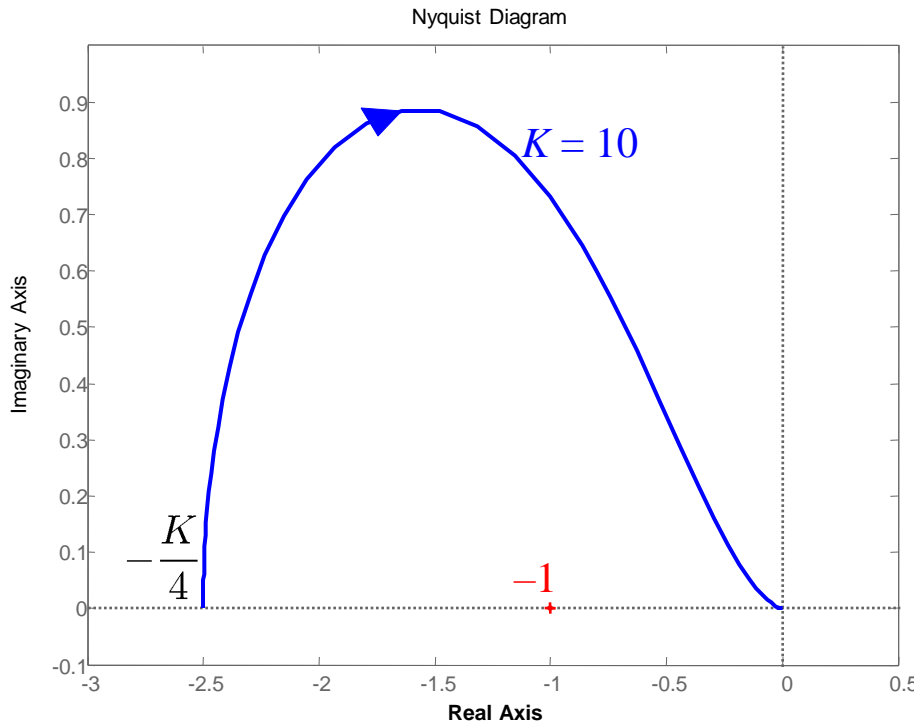
Strecke: $G_S(s) = \frac{1}{(s+1)(s-4)} = \frac{1}{s^2 - 3s - 4}$

Offener Regelkreis: $G_0(s) = \frac{K}{s^2 - 3s - 4}$

Zur Erinnerung:

$$\pi n_g^+ = \angle (1 + G_0(i\omega))|_{\omega=0 \dots \infty} + \pi n_o^+$$

$\rightarrow n_o^+ = 1.$



Drehwinkel um den Punkt -1 :

$K > 4$: 180° bzw. π

$0 < K < 4$: 0° bzw. 0

Für $K > 4$: $\pi n_g^+ = \pi + \pi$

$n_g^+ = 2$ instabile Pole

Für $0 < K < 4$: $\pi n_g^+ = 0 + \pi$

$n_g^+ = 1$ instabiler Pol

\rightarrow Mit P-Regler nicht stabilisierbar!

8.3 Nyquist-Kriterium

Beobachtungen:

- Es gibt Regelstrecken, die sich mit einem bestimmten Reglertyp (im einfachsten Fall einem P-Regler) für jede Reglerverstärkung stabilisieren lassen. Solche Regelstrecken nennt man *strukturstabil*.

ACHTUNG: Reale Prozesse sind meist viel komplizierter als die zum Reglerentwurf verwendeten Prozessmodelle. Oft werden schnelle dynamische Eigenschaften und kleine Totzeiten bei der Modellierung vernachlässigt. Dies führt insbesondere bei großen Frequenzen zu erheblichen Abweichungen zwischen realem Prozess und Modell. Deshalb kann es passieren, dass der Regelkreis mit dem Modell zwar stabil ist, am realen Prozess aber instabil. Ein *robuster* Regelungsentwurf berücksichtigt diese Zusammenhänge.

- Es gibt Regelstrecken, die sich nicht mit einem P-Regler stabilisieren lassen. Mit komplexeren Reglern ist dies aber möglich.
- Bei *stabilen Regelstrecken* muss die *Reglerverstärkung K unter einer Schwelle* bleiben ($K < K_{krit}$), um Stabilität des geschlossenen Kreises zu sichern.
- Bei *instabilen Regelstrecken* muss die *Reglerverstärkung K über einer Schwelle* bleiben ($K > K_{krit}$), um Stabilität des geschlossenen Kreises zu sichern. Warum ist das so?

9. Quantitative Stabilitätskriterien

Inhalt Kapitel 9

9 Quantitative Stabilitätskriterien

9.1 Pollage

9.2 Amplitudenrand

9.3 Phasenrand

9.4 Die Rolle des Punktes $(-1, 0)$

9.5 Robuste Stabilität

9 Qualitative Stabilitätskriterien

Wir wollen eine Aussage darüber machen, *wie* stabil ein Regelkreis ist, d.h. *wie weit* er von der *Stabilitätsgrenze entfernt* ist.

Da es verschiedene Möglichkeiten gibt die "Entfernung" zur Stabilitätsgrenze zu messen, gibt es verschiedene solcher quantitative Stabilitätskriterien.

Nur wenn die Entfernung von der Stabilitätsgrenze hinreichend groß ist, können wir für ein System auch in der Praxis Stabilität sicher stellen. Denn unser Modell stimmt nicht mit der realen (wahren) Strecke exakt überein, weil:

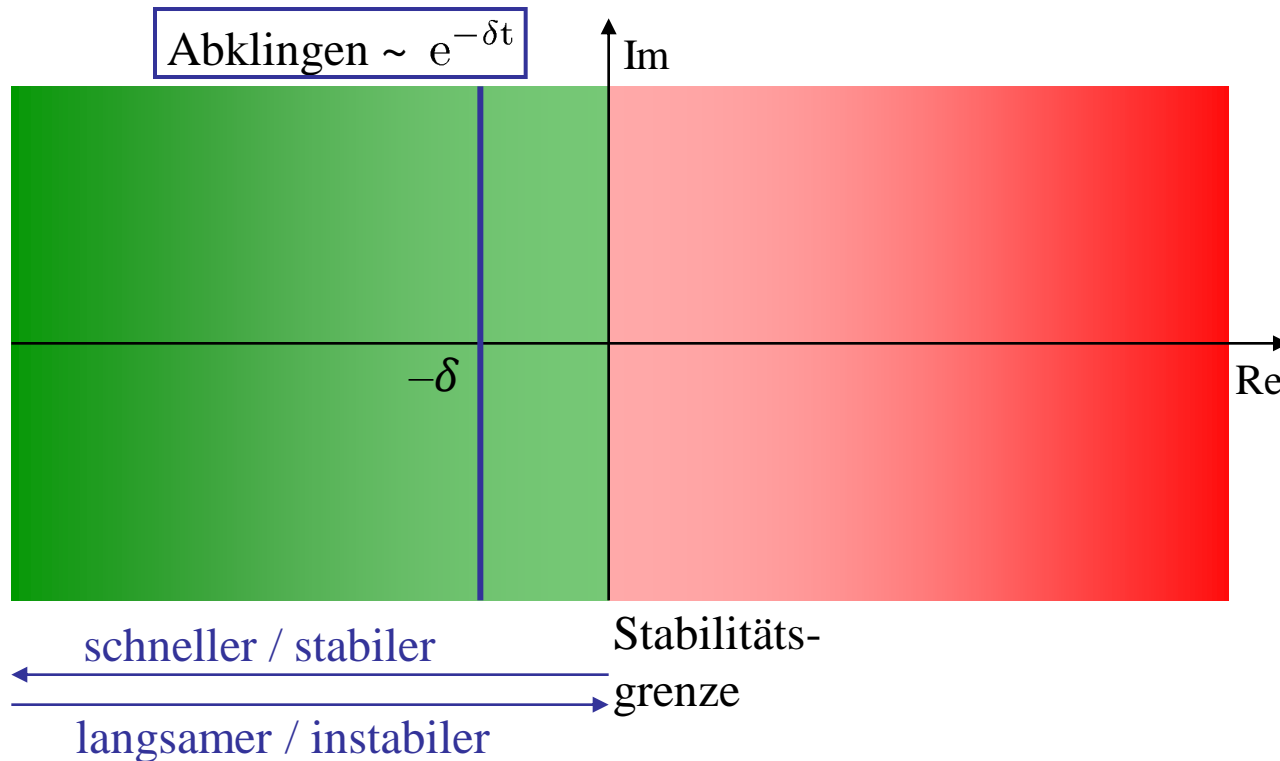
- bei der Modellierung bestimmte Effekte nicht erfasst oder beschrieben werden konnten,
- bei der Modellierung bestimmte Effekte zur Vereinfachung absichtlich vernachlässigt wurden,
- die Strecke mit der Zeit ihr Verhalten verändert (Alterung, Verschleiß),
- äußere Einflüsse das Verhalten der Strecke verändern (Teile werden ausgetauscht)...

→ **Wir brauchen genug "Spielraum", d.h. Abstand zur Stabilitätsgrenze, damit nicht nur unser Modell sondern auch die reale Strecke stabil ist!**

9.1 Pollage

Die Realteile der *Pole des geschlossenen Regelkreises* bestimmen das Abklingverhalten der zeitlichen DGL-Lösung. Daher sollten alle Pole "genügend weit" von der imaginären Achse entfernt sein (z.B. links der blauen Linie bei $-\delta$).

Was "genügend weit" konkret bedeutet, muss in jedem Einzelfall individuell eingeschätzt werden. Bei manchen Prozessen geht es um *msec* bei anderen um *h*.



Bemerkung:

Neben Stabilität spielen beim Reglerentwurf noch andere Kriterien eine wichtige Rolle, z.B. nicht zu starkes Überschwingen. Dies schränkt die erwünschten Pollagen weiter ein! Dazu später mehr...

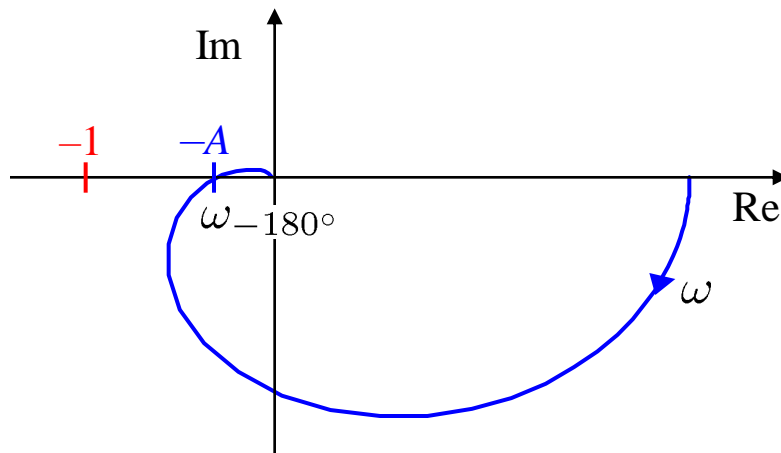
9.2 Amplitudenrand

Wir wissen aus dem einfachen *Nyquist-Kriterium*: Die Ortskurve des offenen Kreises darf den Punkt -1 nicht umschlingen, damit der geschlossene Kreis stabil ist. Die Frequenz bei der die Ortskurve die negative reelle Achse schneidet nennt man *Phasendurchtrittsfrequenz* ω_{-180° .

D.h. der Schnittpunkt der Ortskurve mit der negativen reellen Achse (bei $-A$) muss rechts von der -1 liegen. Das ist um so mehr erfüllt, je kleiner A ist.

$$\text{Amplitudenrand: } k_R = \frac{1}{|-A|} = \frac{1}{|G_0(i\omega_{-180^\circ})|}$$

Gilt nur für stabile offene Regelkreise!



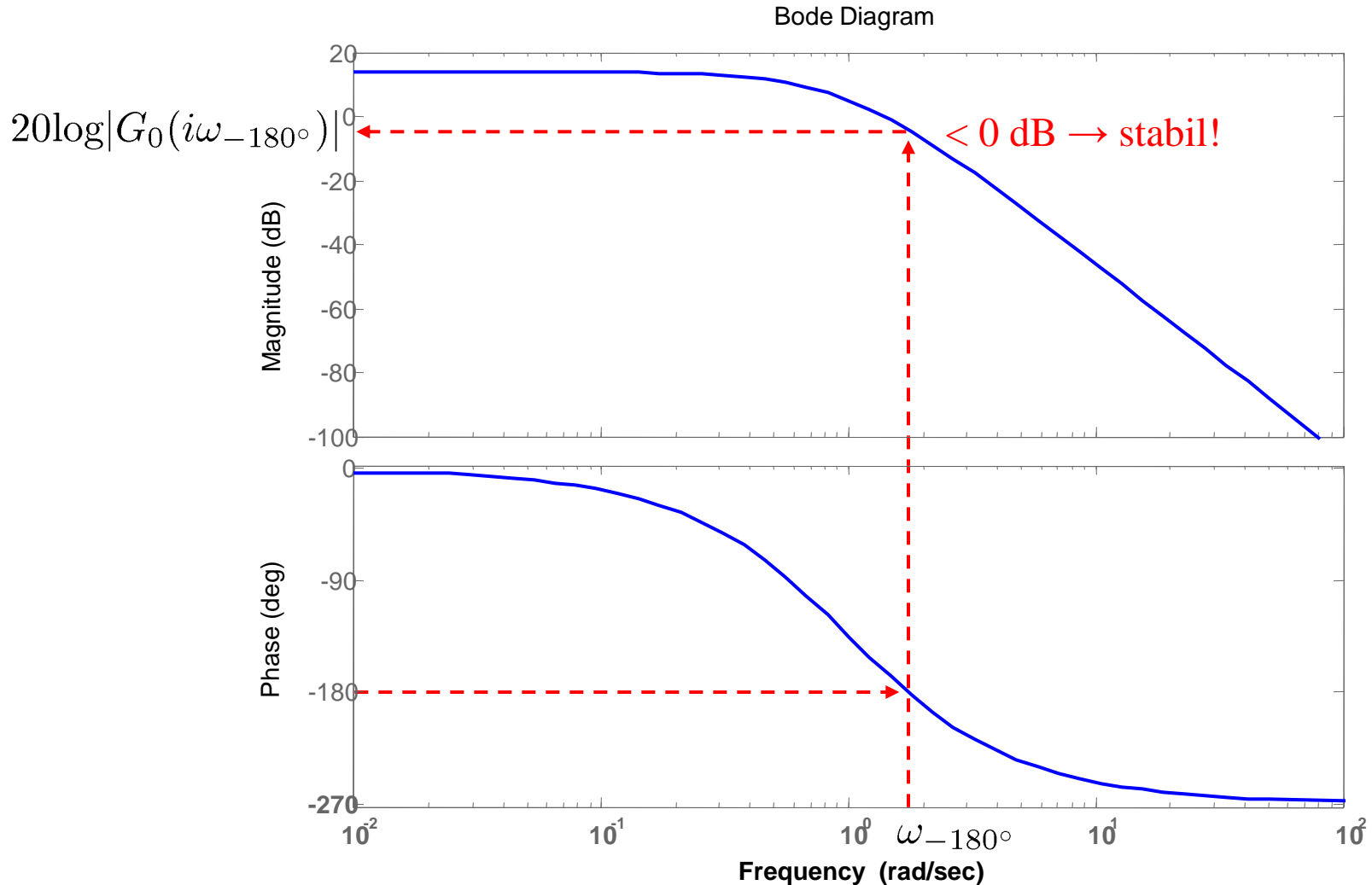
Stabilität für: $k_R > 1$.

Je größer k_R , desto stabiler!

Amplitudenrand wird auch *Amplitudenreserve* genannt, weil man die Kreisverstärkung um diesen Faktor erhöhen kann bis die Stabilitätsgrenze erreicht wird.

9.2 Amplitudenrand

Gleiche Betrachtungen im Bode-Diagramm möglich:

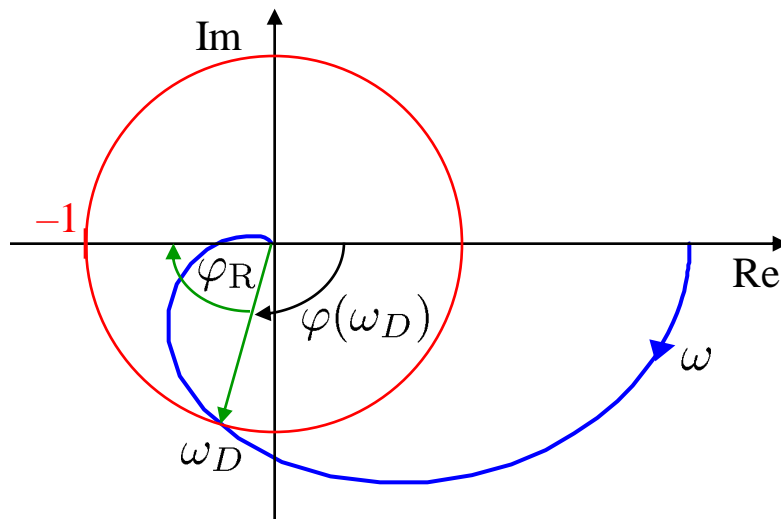


9.3 Phasenrand

Alternativ können wir uns auch anschauen, bei welchem Winkel die Ortskurve des offenen Regelkreises den Kreis mit Radius 1 um den Ursprung schneidet. Die Frequenz bei der dieser Schnittpunkt liegt nennt man *Amplitudendurchtrittsfrequenz* ω_D .

Je mehr Phasenwinkel bis zum Erreichen der negativen reellen Achse, also bei $\varphi = -180^\circ$, übrig bleibt, desto weiter ist man von der Stabilitätsgrenze entfernt.

Phasenrand: $\varphi_R = 180^\circ - |\varphi(\omega_D)|$ **Gilt nur für stabile offene Regelkreise!**



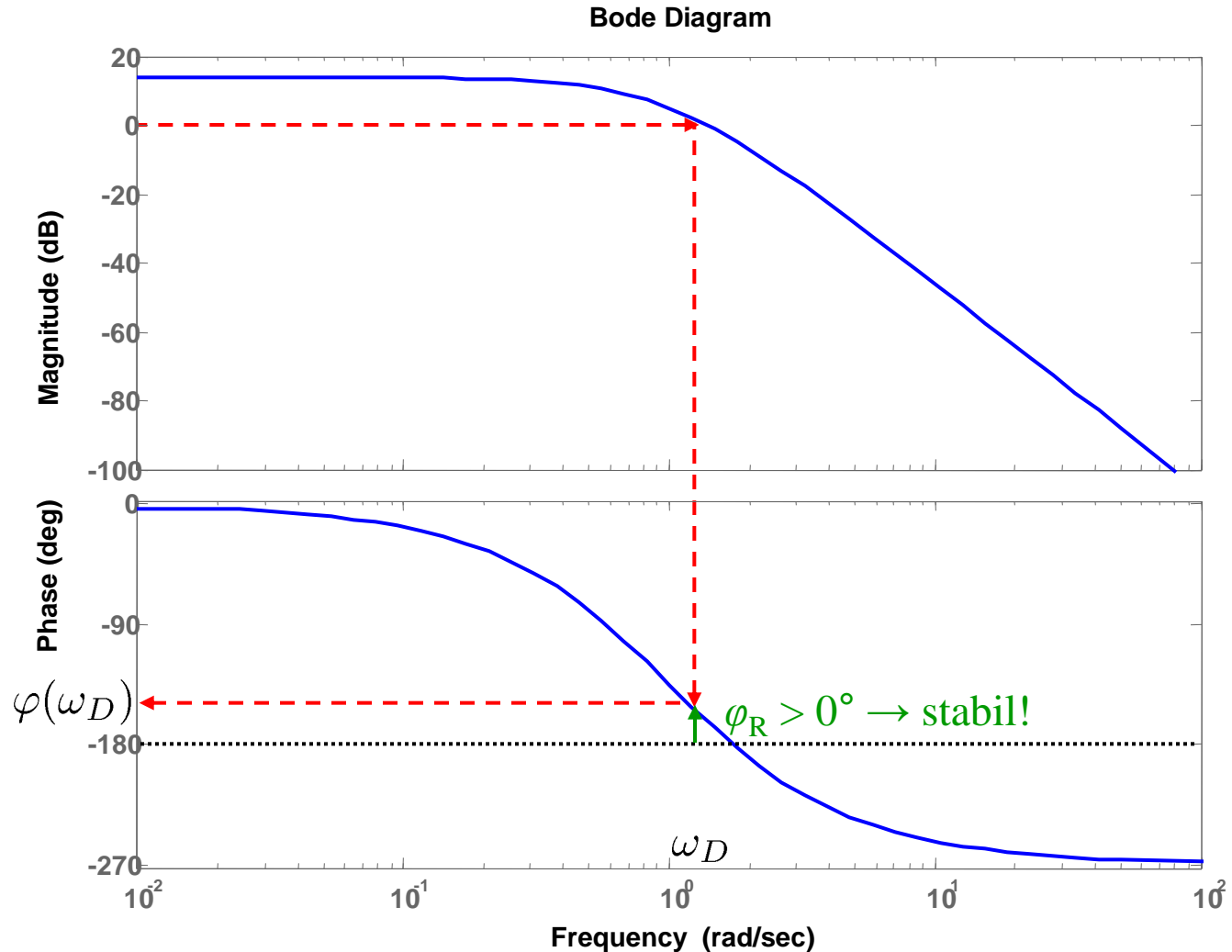
Stabilität für: $\varphi_R > 0^\circ$.

Je größer φ_R , desto stabiler!

Phasenrand wird auch Phasenreserve genannt, weil man die Phasenverschiebung um diesen Winkel erhöhen kann bis die Stabilitätsgrenze erreicht wird.

9.3 Phasenrand

Gleiche Betrachtungen im Bode-Diagramm möglich:



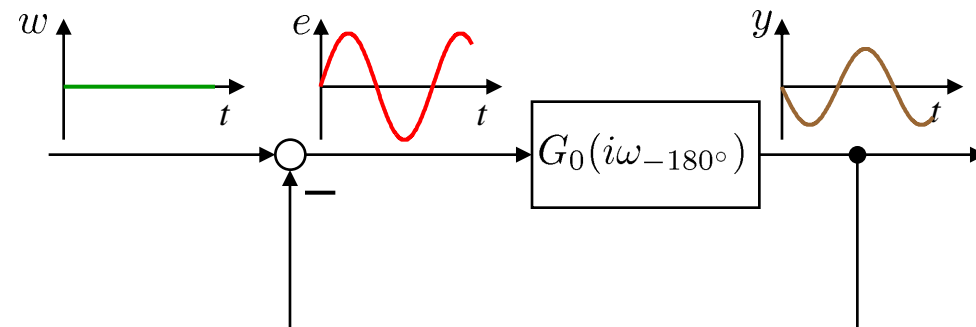
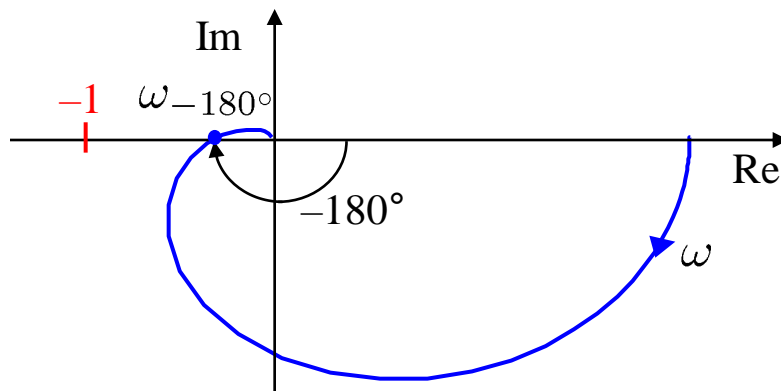
9.4 Die Rolle des Punktes $(-1, 0)$

Wenn die Ortskurve $G_0(i\omega)$ die negative reelle Achse schneidet, dann weist der offene Regelkreis eine Phasenverschiebung von -180° auf, d.h. für eine sinusförmige Schwingung der Frequenz ω_{-180° im eingeschwungenen Zustand ist die Regelgröße y gegenüber der Regelabweichung e genau um eine Halbwelle verschoben, hat also entgegengesetztes Vorzeichen!

Damit wird für diese Frequenz die negative Rückkopplung (Gegenkopplung) zu einer positiven Rückkopplung (Mitkopplung). Der Ausgang des offenen Kreises kommt also phasensynchron wieder an dessen Eingang an. Stabilität erhält man folglich dann, wenn die Verstärkung des offenen Kreises bei ω_{-180° kleiner als Eins ist:

$$|G_0(i\omega_{-180^\circ})| < 1 \rightarrow \text{stabil}$$

Das ist der Fall, wenn der Punkt $(-1, 0)$ links von der Ortskurve des offenen Regelkreises liegt!



9.5 Robuste Stabilität

Amplituden- und Phasenrand geben Auskunft darüber, wie stabil ein Regelkreis ist oder genauer: wie weit er von der Stabilitätsgrenze entfernt ist. Je größer Amplituden- und Phasenrand sind, desto sicherer können wir sein, dass Stabilität nicht nur für das Modell sondern auch für die (davon abweichende) reale Strecke herrscht. Aber wie können wir Stabilität für die reale Strecke *garantieren*?

Dazu benötigen wir zunächst eine Beschreibung der *maximal* möglichen Modellfehler (Abweichungen zwischen Modell und realer Strecke):

- Parameterunsicherheiten: Das Modell enthält Parameter, die nur in einem Intervall bekannt sind. Beispiele: Masse eines Autos [$m_{\text{leer}}, m_{\text{voll beladen}}$]; Dämpfung eines Stoßdämpfers (abgelesen aus dessen Kennlinie) [$d_{\text{min}}, d_{\text{max}}$]; ...
- Bereich des Arbeitspunkts: Die Linearisierung eines nichtlinearen Modells liefert arbeitspunktabhängige Parameter. Aus dem Bereich des Arbeitspunkts ergeben sich die Parameterschwankungen: (Flughöhe: [0 m, 10 000 m]; Fahrzeuggeschwindigkeit: [0 km/h, 250 km/h])
- Unmodellerte Dynamik: Vernachlässigte kleine Totzeiten; höherfrequente dynamische Eigenschaften; ...

9.5 Robuste Stabilität

Am einfachsten lassen sich diese Modellunsicherheiten im Frequenzbereich darstellen und weiterverarbeiten. Zwei alternative Ansätze:

1. Additive (absolute) Modellunsicherheiten

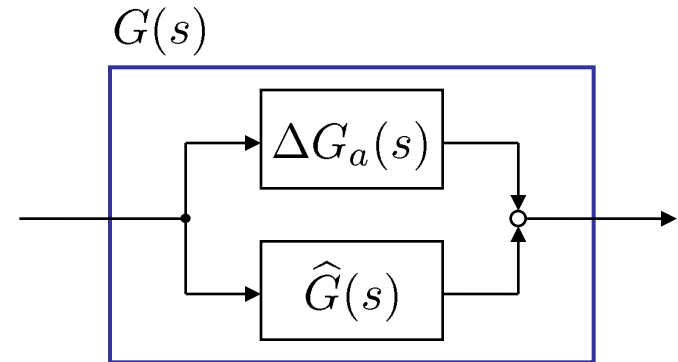
$$G(s) = \hat{G}(s) + \Delta G_a(s)$$

unbekannte
reale Strecke

Modell

Unsicherheit

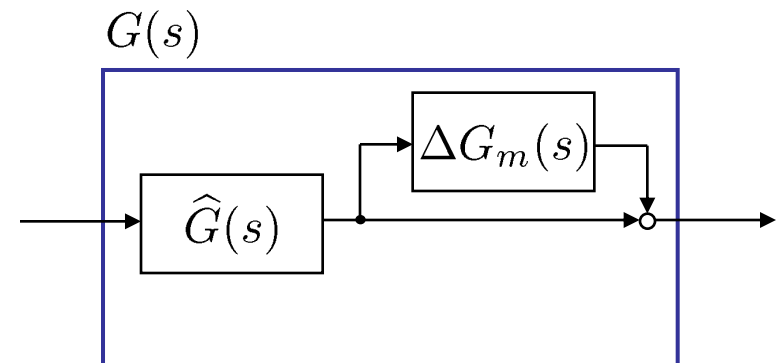
$$|\Delta G_a(s)| \leq G_{a,\max}(s)$$



2. Multiplikative (relative) Modellunsicherheiten

$$G(s) = \hat{G}(s) (1 + \Delta G_m(s))$$

$$|\Delta G_m(s)| \leq G_{m,\max}(s)$$

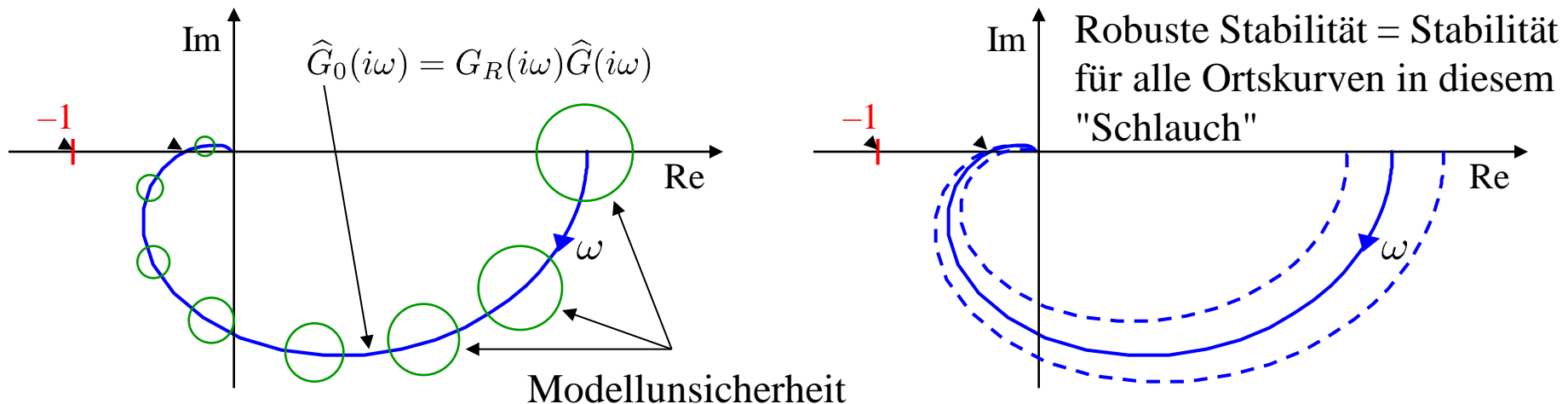


9.5 Robuste Stabilität

Wenn ein Regelkreis für *alle* möglichen $G(s)$ stabil ist, dann nennt man ihn *robust stabil*.

- Wichtig ist hierfür, dass wir mit $\Delta G_a(s)$ bzw. $\Delta G_m(s)$ eine *maximale* Modellunsicherheit definieren.
- In der Praxis sind maximale Modellunsicherheiten teilweise schwer abzuschätzen.
- Zu große Modellunsicherheiten führen dazu, dass der Regler sehr konservativ (langsam) eingestellt werden muss, um robuste Stabilität zu garantieren. Deshalb sollte die Unsicherheitsabschätzung nicht zu großzügig sein.

Beispiel: Konstanter multiplikativer (relativer) Fehler.



9.5 Robuste Stabilität

Stabilität ohne Berücksichtigung der Modellfehler lässt sich leicht anhand der Ortskurve des offenen Kreises zeigen:

$$\widehat{G}_0(i\omega) = G_R(i\omega)\widehat{G}(i\omega)$$

Robuste Stabilität (also mit Berücksichtigung der Modellfehler) muss für die Ortskurven *alle möglichen* offenen Kreise gezeigt werden:

$$G_0(i\omega) = G_R(i\omega)G(i\omega)$$

Es könnte passieren, dass zwar $\widehat{G}_0(i\omega)$ die -1 nicht umschlingt, aber Teile des "Schlauchs" $G_0(i\omega)$ die -1 umschlingen. Für robuste Stabilität darf der gesamte "Schlauch" die -1 nicht umschlingen.

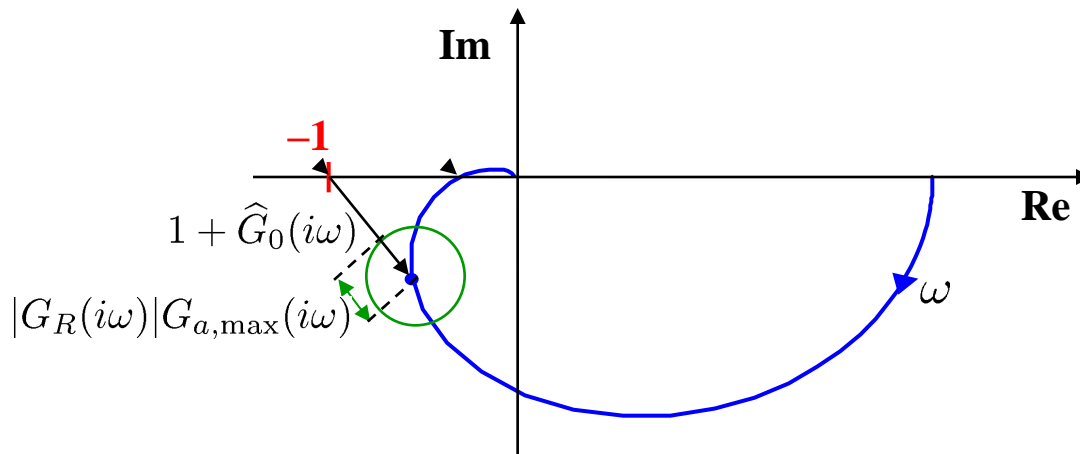
Bei *additiven* Modellunsicherheiten gilt für die möglichen Abweichungen zwischen Modell und Strecke (**grünen Kreise**):

$$|G_0(i\omega) - \widehat{G}_0(i\omega)| = |G_R(i\omega)[G(i\omega) - \widehat{G}(i\omega)]| = |G_R(i\omega)\Delta G_a(i\omega)| \leq |G_R(i\omega)|G_{a,\max}(i\omega)$$

(Gleichheit gilt auf dem Rand des grünen Kreises. Das "<" gilt im Inneren.)

Die Formel für multiplikative Modellunsicherheiten lässt sich in ähnlicher Weise herleiten...

9.5 Robuste Stabilität



Wir können robuste Stabilität garantieren, wenn kein Punkt innerhalb der grünen Kreise die -1 umschlingt. Dies ist genau dann der Fall, wenn gilt:

$$|1 + \hat{G}_0(i\omega)| = |1 + G_R(i\omega)\hat{G}(i\omega)| > |G_R(i\omega)|G_{a,\max}(i\omega)$$

Robuste Stabilität
bei additiven Modellunsicherheiten:

$$\left| \frac{G_R(i\omega)}{1 + G_R(i\omega)\hat{G}(i\omega)} \right| G_{a,\max}(i\omega) < 1$$

Robuste Stabilität
bei multiplikativen Modellunsicherheiten:

$$\left| \frac{G_R(i\omega)\hat{G}(i\omega)}{1 + G_R(i\omega)\hat{G}(i\omega)} \right| G_{m,\max}(i\omega) < 1$$

9.5 Robuste Stabilität

Beispiel: Regelung eines Systems 2. Ordnung / Vernachlässigung der Sensordynamik.

Modell 2. Ordnung:

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

P-Regler:

$$G_R(s) = K$$

Vernachlässigte Sensordynamik:

$$G_{\text{Sensor}}(s) = \frac{1}{Ts + 1}$$

Berechnung der multiplikativen Modellunsicherheit aus $G(s) = \hat{G}(s) (1 + \Delta G_m(s))$:

$$\frac{1}{s^2 + s + 1} \frac{1}{Ts + 1} = \frac{1}{s^2 + s + 1} (1 + \Delta G_m(s)) \quad \rightarrow \quad \boxed{\Delta G_m(s) = \frac{1}{Ts + 1} - 1 = \frac{-Ts}{Ts + 1}}$$

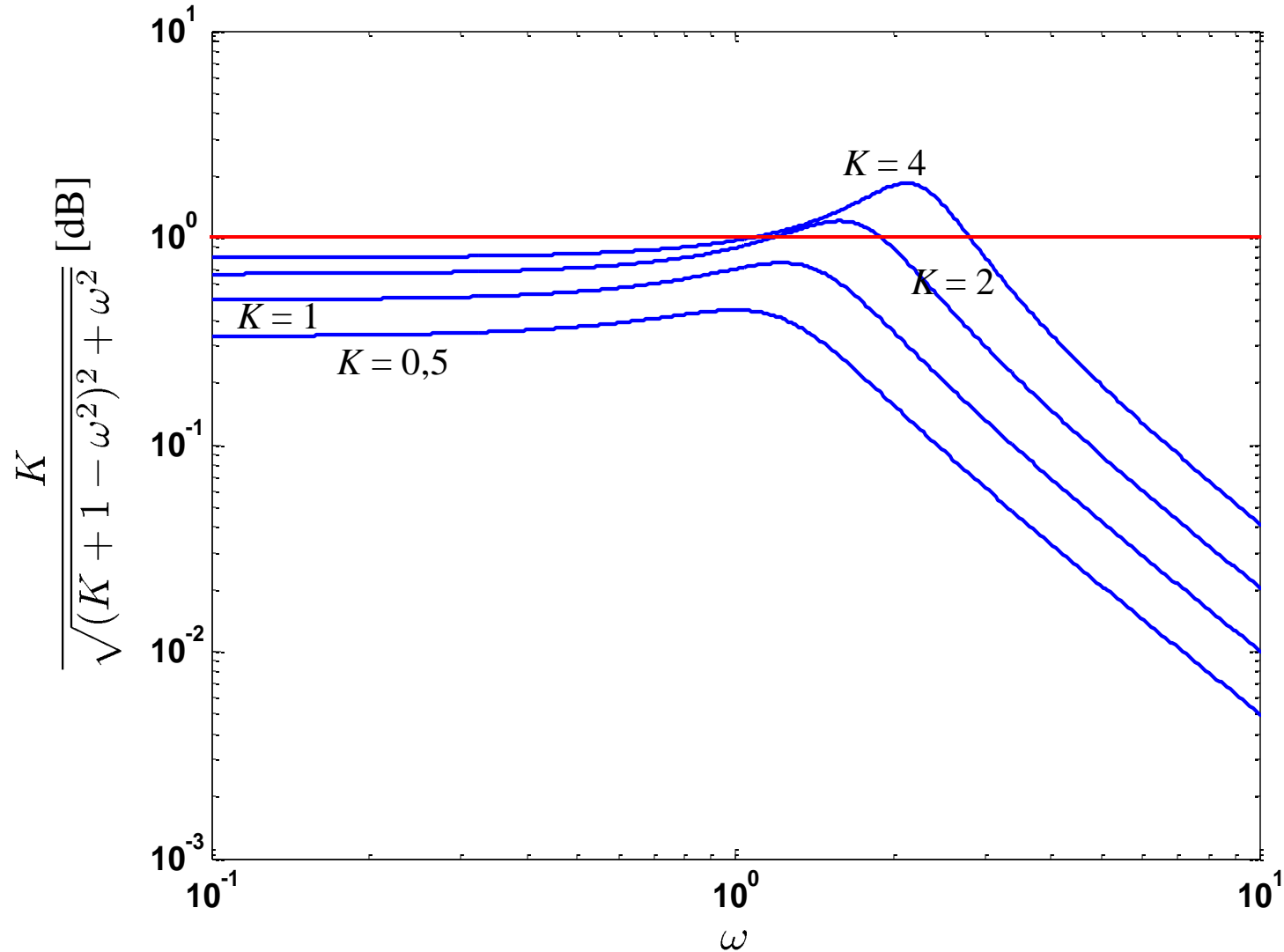
Obere Grenze für den Betrag der Modellunsicherheit, wenn T unbekannt ist:

$$|\Delta G_m(i\omega)| = \frac{T\omega}{\sqrt{T^2\omega^2 + 1}} \leq \boxed{1 = G_{m,\max}(i\omega)}$$

Bedingung für robuste Stabilität $\left| \frac{G_R(i\omega)\hat{G}(i\omega)}{1 + G_R(i\omega)\hat{G}(i\omega)} \right| G_{m,\max}(i\omega) < 1$:

$$\rightarrow \left| \frac{K}{(i\omega)^2 + (i\omega) + 1 + K} \right| = \frac{K}{\sqrt{(K + 1 - \omega^2)^2 + \omega^2}} < \frac{1}{G_{m,\max}(i\omega)}$$

9.5 Robuste Stabilität

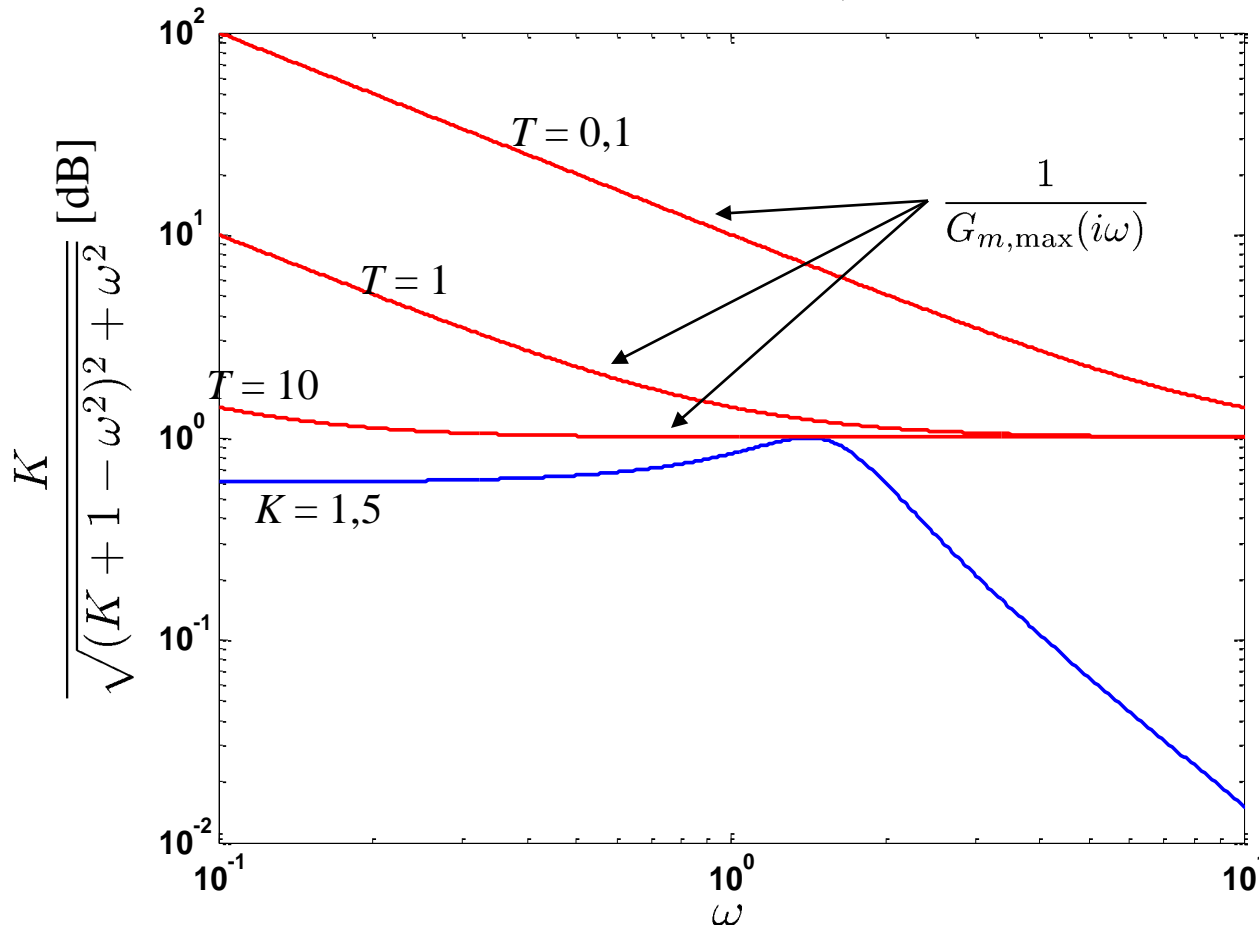


Robuste Stabilität bis ca. $K < 1,5$.

9.5 Robuste Stabilität

Für ein festes T können wir auch die frequenzabhängige und damit weniger konservative obere Schranke angeben:

$$G_{m,\max}(i\omega) = \frac{T\omega}{\sqrt{T^2\omega^2 + 1}}$$



Eine realistische Abschätzung der Sensordynamik mit $T = 0,1$ sec (d.h. ca. 10 mal schneller als die Strecke) erlaubt wesentlich höhere Verstärkungen als $K = 1,5$ bis zum Erreichen der robusten Stabilitätsgrenze!

9.5 Robuste Stabilität

Robuste Stabilität garantiert die Stabilität des Regelkreises für alle möglichen Strecken, die durch die Modellunsicherheit beschrieben werden.

Neben der Stabilität spielt auch die Güte (Performance) der Regelung eine wichtige Rolle. Sie kann auf verschiedenste Arten gemessen werden.

Entsprechend der Stabilität lässt sich auch die Güte (Performance) eines Regelkreises auf den Begriff der *robusten Güte* bzw. *robusten Performance* erweitern:

Robuste Güte (Performance) garantiert, dass vom Regelkreis ein Gütemaß (Performancemaß) bestimmter Größe erfüllt wird, und zwar für alle möglichen Strecken, die durch die Modellunsicherheit beschrieben werden.

→ Reglerentwurfverfahren, die robuste Stabilität und evtl. auch robuste Güte (Performance) des Regelkreises sicher stellen (oder das Fehlschlagen des Entwurfs melden) werden unter dem Schlagwort "robuste Regelungen" behandelt. Dies soll aber nicht darüber hinweg täuschen, dass alle (vernünftig entworfenen) Regelkreise durch die prinzipiellen Vorteile der Rückkopplung (siehe Kapitel 1) robust sind.

10. Einfache lineare Regler

Inhalt Kapitel 10

10 Einfache lineare Regler

- 10.1 Aufgaben einer Regelung
- 10.2 Statische Anforderungen
- 10.3 Dynamische Anforderungen
- 10.4 PID-Regler
- 10.5 Reglerrealisierungen

10.1 Aufgaben einer Regelung

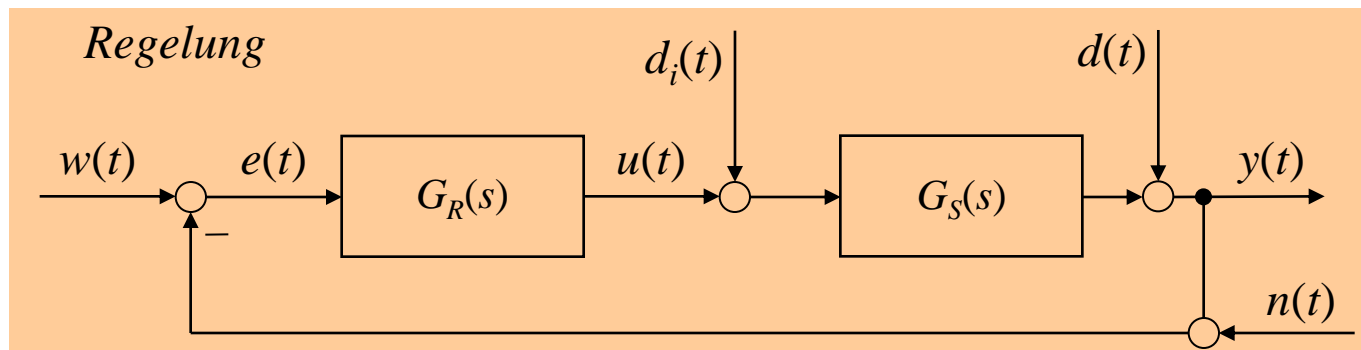
- Regelkreis soll stabil sein.
- Regelgröße y soll von den Störungen d_i und d möglichst wenig beeinflusst werden, d.h. die *Störübertragungsfunktionen* sollen möglichst klein sein:

$$G_{d_i}(s) = \frac{G_S(s)}{1 + G_R(s)G_S(s)} \ll 1 \quad G_d(s) = \frac{1}{1 + G_R(s)G_S(s)} \ll 1$$

- Regelgröße y soll der Führungsgröße w möglichst gut folgen, d.h. die *Führungsübertragungsfunktion* soll möglichst gleich 1 sein:

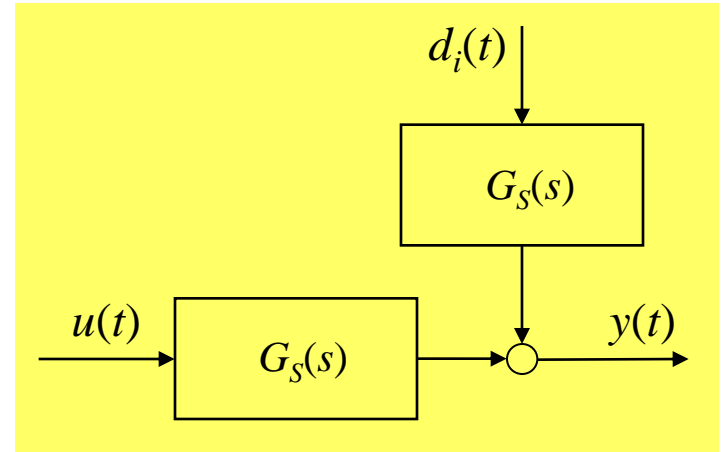
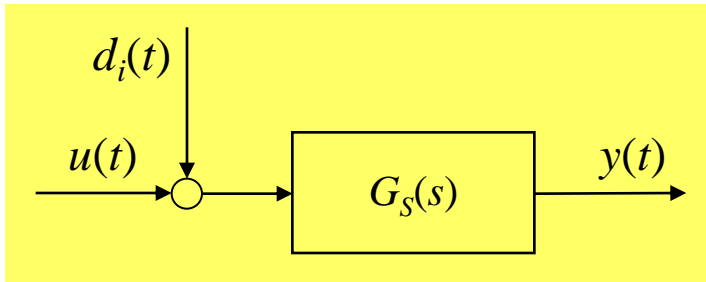
$$G_w(s) = \frac{G_R(s)G_S(s)}{1 + G_R(s)G_S(s)} \approx 1 \quad G_n(s) = -G_w(s)$$

→ Regler $G_R(s)$ sollte so gewählt werden, dass diese Forderungen erfüllt werden können.



10.1 Aufgaben einer Regelung

Eine Störung am Streckeneingang kann auf eine Störung am Streckenausgang umgerechnet werden:



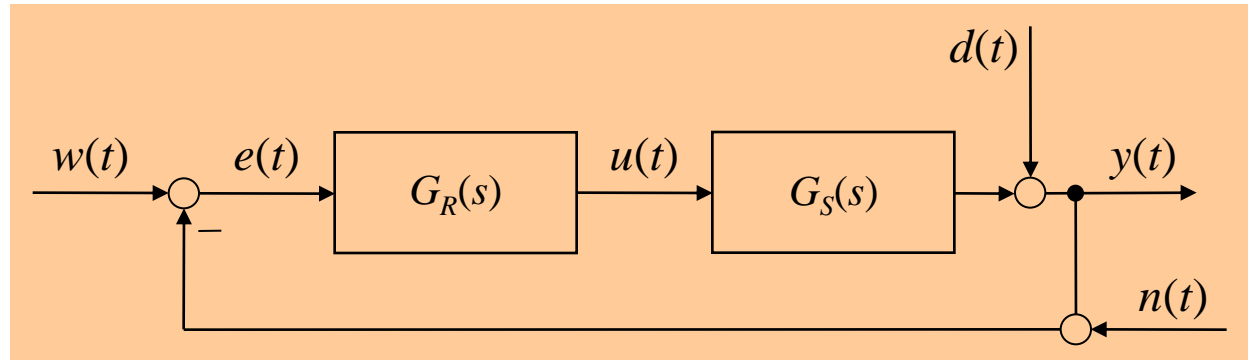
→ Die beiden wichtigsten Übertragungsfunktionen für die Güte einer Regelung sind:

Ausgangsstörung → Regelgröße: $G_d(s) = \frac{1}{1 + G_R(s)G_S(s)} = S(s)$ Empfindlichkeitsfunktion

Führungsgröße → Regelgröße: $G_w(s) = \frac{G_R(s)G_S(s)}{1 + G_R(s)G_S(s)} = T(s)$ komplementäre Empfindlichkeitsfkt.

10.1 Aufgaben einer Regelung

Die verschiedenen Ziele einer Regelung können sich widersprechen. Dann muss entweder ein Kompromiss gefunden werden oder die Struktur des Regelkreises muss erweitert werden (Vorsteuerung, Filterung der Regelgröße), siehe Kap. 12.



$$S(s) = \frac{1}{1 + G_R(s)G_S(s)} : d \rightarrow y, w \rightarrow e \quad T(s) = \frac{G_R(s)G_S(s)}{1 + G_R(s)G_S(s)} : w \rightarrow y, -n \rightarrow y$$

Konflikte treten auf, wenn Führungsgröße w und Störgröße d in unterschiedlichen Frequenzbereichen liegen.

Und wenn Führungsgröße w und Messrauschen n in überlappenden Frequenzbereichen liegen.

Weitere Betrachtungen im Frequenzbereich werden in Kapitel 14 näher behandelt. Für moderne Reglerentwurfsverfahren spielen S und T eine entscheidende Rolle. Die beiden Forderungen $S(s) \ll 1$ und $T(s) \approx 1$ können aber immer nur für bestimmte Frequenzbereiche erfüllt werden.

10.1 Aufgaben einer Regelung

Nun aber zunächst zu den etwas einfacheren Betrachtungen im Zeitbereich.

Die wichtigsten Forderungen sind:

1. Stabilität → Stabilitätskriterien
2. Stationäre Kompensation der Störung und Folge der Führungsgröße → Kapitel 10.2
3. Dynamische Anforderungen → Kapitel 10.3
4. Robustheit → quantitative Stabilitätskriterien (robuste Stabilität und Performance)

Schwierigkeit: Bei der Einstellung bzw. Optimierung des Reglers ist die *Stabilität* des Regelkreises immer eine *notwendige Nebenbedingung*. Das macht den Reglerentwurf schwierig.

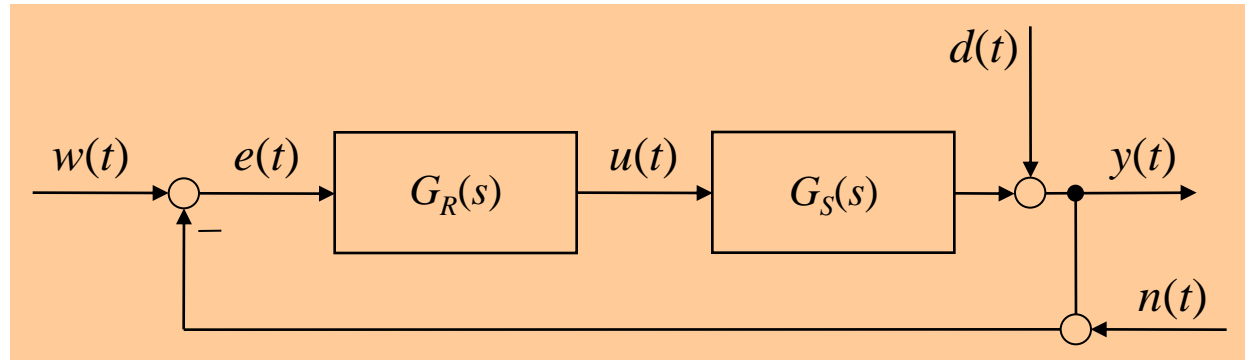
Ausweg: Es gibt eine erweiterte Regelkreisstruktur (IMC = *Internal Model Control* bzw. *Q-Parametrierung* oder *Youla-Parametrierung*), welche die Stabilität des Regelkreises für alle stabilen Regler sicherstellt. Damit wird sowohl die manuelle Einstellung als auch die automatische Optimierung des Reglers wesentlich vereinfacht. → Kap. 15.

Sehr viele moderne regelungstechnische Verfahren basieren darauf!

10.2 Statische Anforderungen

$$w \rightarrow y : \frac{G_R(s)G_S(s)}{1 + G_R(s)G_S(s)}$$

$$d \rightarrow y : \frac{1}{1 + G_R(s)G_S(s)}$$



$$Y(s) = \underbrace{\frac{G_R(s)G_S(s)}{1 + G_R(s)G_S(s)}}_{G_w(s)} W(s) + \underbrace{\frac{1}{1 + G_R(s)G_S(s)}}_{G_d(s)} D(s)$$

Forderung: Im eingeschwungenen Zustand soll die Regelgröße $y(t)$ der Führungsgröße $w(t)$ folgen: $\lim_{t \rightarrow \infty} (w(t) - y(t)) = 0$. Das bedeutet im Laplace-Bereich (Endwertsatz!):

$$\begin{aligned} \lim_{s \rightarrow 0} s (W(s) - Y(s)) &= \lim_{s \rightarrow 0} s \left(W(s) - \frac{G_R(s)G_S(s)}{1 + G_R(s)G_S(s)} W(s) - \frac{1}{1 + G_R(s)G_S(s)} D(s) \right) \\ &= \lim_{s \rightarrow 0} \left(\frac{s}{1 + G_R(s)G_S(s)} W(s) - \frac{s}{1 + G_R(s)G_S(s)} D(s) \right) \stackrel{!}{=} 0 \end{aligned}$$

10.2 Statische Anforderungen

Diese Forderung soll sowohl für $w(t) = 0$ (also auch $W(s) = 0$) als auch für $d(t) = 0$ (also auch $D(s) = 0$) gelten. Daher müssen die Summanden auch einzeln $= 0$ sein:

$$\lim_{s \rightarrow 0} \frac{s}{1 + G_R(s)G_S(s)} W(s) \stackrel{!}{=} 0$$

$$\lim_{s \rightarrow 0} \frac{s}{1 + G_R(s)G_S(s)} D(s) \stackrel{!}{=} 0$$

Die Forderungen sind also für Führungsgröße und Störgröße (am Streckenausgang) identisch!
Die folgenden Beispiele gelten also für beides...

Beispiel: Impuls $\delta(t)$

$$\lim_{s \rightarrow 0} \frac{s}{1 + G_R(s)G_S(s)} \cdot \overset{\mathcal{L}\{\delta(t)\}}{1} = 0 \rightarrow \text{Wird geregelt!}$$

Beispiel: Sprung $\sigma(t)$

$$\lim_{s \rightarrow 0} \frac{s}{1 + G_R(s)G_S(s)} \cdot \overset{\mathcal{L}\{\sigma(t)\}}{\frac{1}{s}} = \lim_{s \rightarrow 0} \frac{1}{1 + G_R(s)G_S(s)}$$

$$= \begin{cases} 0 & \text{falls } G_R G_S \text{ einen I-Anteil aufweist, also } \rightarrow \infty \text{ geht} \\ \frac{1}{1 + G_R(0)G_S(0)} & \text{sonst} \end{cases}$$

10.2 Statische Anforderungen

Beispiel: Rampe $t\sigma(t)$ $\xrightarrow{\mathcal{L}\{t\sigma(t)\}}$

$$\lim_{s \rightarrow 0} \frac{s}{1 + G_R(s)G_S(s)} \cdot \frac{1}{s^2} = \lim_{s \rightarrow 0} \frac{1}{s + sG_R(s)G_S(s)}$$

$$= \begin{cases} 0 & \text{falls } G_R G_S \text{ einen doppelten I-Anteil aufweist} \\ \text{konst.} & \text{falls } G_R G_S \text{ einen einfachen I-Anteil aufweist} \\ \infty & \text{sonst} \end{cases}$$

Beispiel: Sinusförmig

Lässt sich mit dem Endwertsatz nicht analysieren, da für $t \rightarrow \infty$ kein stationärer Zustand sondern Dauerschwingungen auftreten!

Beobachtung: Es fällt auf, dass für eine stationär (also für $t \rightarrow \infty$) exakte Regelung der offene Regelkreis $G_0(s) = G_R(s)G_S(s)$ als Faktor die Laplace-Transformierte der Führungs- bzw. Störgröße enthalten muss. Der offene Regelkreis muss also ein Modell der Signale enthalten. Dieses *"Innere Modell"-Prinzip* lässt sich auf beliebige Signaltypen erweitern.

Wenn das Signalmodell nicht in der Strecke enthalten ist, muss man es im Regler realisieren!

10.2 Statische Anforderungen

"Inneres Modell"-Prinzip

Zur Vereinfachung nehmen wir an $w(t) = 0$ und betrachten nur die Störgröße $d(t)$. Umgekehrt ($d(t) = 0$) gilt dann entsprechendes.

$$G_0(s) = G_R(s)G_S(s) \qquad Y(s) = G_d(s)D(s) = \frac{1}{1 + G_0(s)}D(s)$$

Ein Störsignal der Form $d(t) = e^{-at}$ entspricht im Frequenzbereich $D(s) = \frac{1}{s + a}$.

Ein Störsignal der Form $d(t) = \sin\omega_0 t$ entspricht im Frequenzbereich $D(s) = \frac{\omega_0}{s^2 + \omega_0^2}$.

Aus Kapitel 7.3 wissen wir, dass eine Nullstelle im System den Pol (die Pole) des Signals wegekürzt und somit das Signal nicht durchlässt. Genau das ist hier unser Ziel, denn wir wollen verhindern, dass die Störung (dauerhaft) die Regelgröße beeinflusst. Deshalb fordern wir, dass die Störübertragungsfunktion Nullstellen aufweist, wo das Störsignal Pole hat:

$$D(s) = \frac{1}{s + a} \rightarrow G_d(s) = (s + a)\tilde{G}_d(s) \qquad D(s) = \frac{\omega_0}{s^2 + \omega_0^2} \rightarrow G_d(s) = (s^2 + \omega_0^2)\tilde{G}_d(s)$$

10.2 Statische Anforderungen

Allgemein fordern wir einen Zähler in der Störübertragungsfunktion, der dem Nenner $N_d(s)$ des Signals $D(s)$ entspricht:

$$D(s) = \frac{Z_d(s)}{N_d(s)} \rightarrow G_d(s) = N_d(s) \cdot \tilde{G}_d(s) = \frac{N_d(s)\tilde{Z}(s)}{\tilde{N}(s)} = \frac{1}{1 + G_0(s)}$$

Auflösen nach $G_0(s)$ liefert:

$$G_0(s) = \frac{\tilde{N}(s) - N_d(s)\tilde{Z}(s)}{N_d(s)\tilde{Z}(s)}$$

Nenner von $D(s)$ ist auch im Nenner von $G_0(s)$ vorhanden!

Gleiches kann man auch für das Führungsverhalten zeigen. Daraus folgt das sehr allgemeine "Innere Modell"-Prinzip:

"Inneres Modell"-Prinzip:

Ein stabiler Regelkreis kann nur dann eine Störgröße vollständig unterdrücken bzw. einer Führungsgröße ohne bleibende Regelabweichung folgen, wenn er (d.h. $G_0(s)$) ein inneres Modell der Störsignale bzw. der Führungssignale enthält.

10.2 Statische Anforderungen

Zusammenfassung

- Um einem Führungsgrößensprung ohne bleibende Regelabweichung zu folgen bzw. eine sprungförmige Störung komplett auszuregeln, muss der offene Regelkreis $G_0(s)$ einen I-Anteil ($1/s$) enthalten. Weist die Strecke keinen I-Anteil auf, muss dieser folglich im Regler eingebaut werden.

- Wenn weder Strecke noch Regler einen I-Anteil enthalten, dann ergibt sich eine

$$\text{bleibende Regelabweichung} = \frac{1}{1 + G_R(0)G_S(0)} = \frac{1}{1 + K_R K_S} .$$

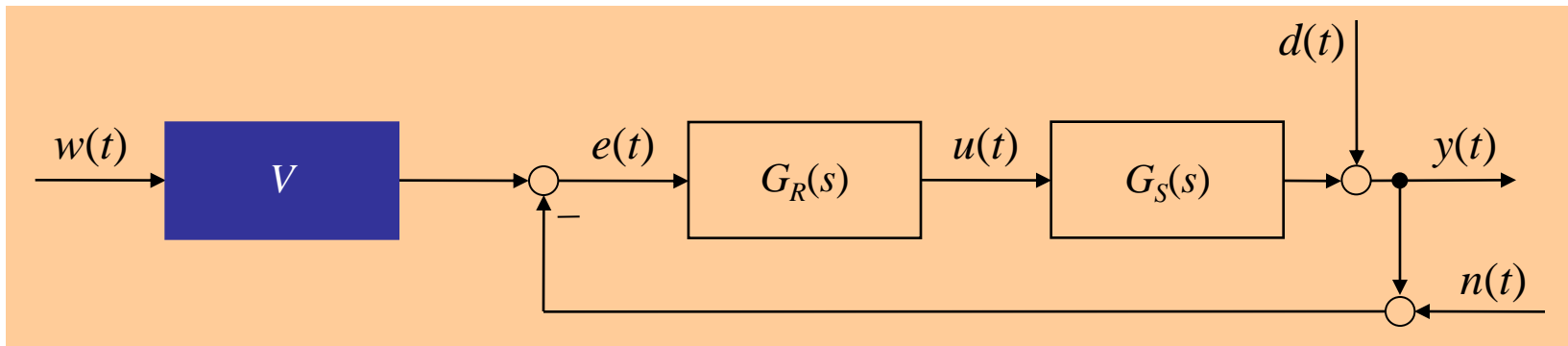
Diese Regelabweichung muss um so kleiner sein, desto größer die Kreisverstärkung ist (also das Produkt aus Regler- und Streckenverstärkung). Geht z.B. $K_R \rightarrow \infty$, dann verschwindet die bleibende Regelabweichung. Das ist konsistent damit, dass ein I-Glied eine unendlich hohe statische Verstärkung hat (Amplitude im Bode-Diagramm für $\omega \rightarrow 0$). Außerdem stimmt es mit unserer Beobachtung aus dem 1. Kapitel überein, dass sich theoretisch bei unendlicher Reglerverstärkung eine perfekte Regelung ergibt (praktisch ist natürlich die Stabilitätsgrenze eine obere Schranke für die Reglerverstärkung).

- Für *rampenförmige* Signale wird entsprechend ein doppelter I-Anteil im offenen Kreis benötigt. Ansonsten ergibt sich ein sog. *Schleppfehler*.

10.2 Statische Anforderungen

Zwischenbemerkung: Stationäre Genauigkeit ohne I-Anteil

Wir haben gerade gelernt, dass ohne einen I-Anteil in $G_0(s)$ eine bleibende Regelabweichung bei einem Führungsgrößenprung auftritt. Es gibt eine 2. Möglichkeit, wie dies vermieden werden kann: Die Erweiterung des Standardregelkreises um einen statischen **Vorfilter**.



Der Quotient aus $y(t \rightarrow \infty) / w(t \rightarrow \infty)$ ohne Vorfilter ist $\frac{K_R K_S}{1 + K_R K_S}$. Durch geeignete Wahl des Vorfilters lässt sich dies kompensieren:

$$V = \frac{1 + K_R K_S}{K_R K_S}$$

Dieser Vorfilter beseitigt die bleibende Regelabweichung auch ohne I-Anteil in $G_0(s)$.

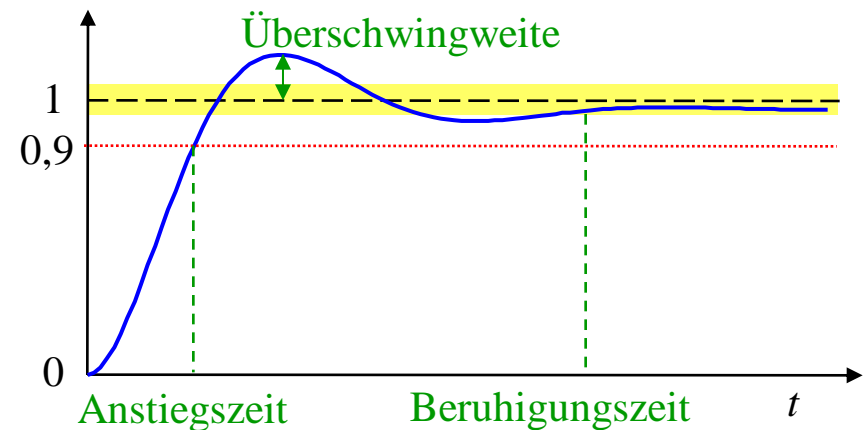
- ABER:**
- Im Gegensatz zum I-Anteil, müssen hier die Verstärkungen *exakt* bekannt sein!
 - Vorfilter kann nicht für die (nicht messbaren) Störgrößen eingesetzt werden!

10.3 Dynamische Anforderungen

Neben dem eben betrachteten stationären Verhalten spielt natürlich auch das zeitliche Verhalten des Regelkreises eine Rolle. Die Forderungen werden meist auf sprungförmige Veränderung der Führungs- bzw. Störgröße bezogen. Bezogen auf einen Einheitssprung der Führungsgröße liegen folgende Forderungen nahe:

1. Möglichst geringes bzw. kein Unterschwingen (für nicht phasenminimale Systeme).
2. Möglichst kurze Anstiegszeit.
3. Möglichst geringes Überschwingen.
4. Möglichst kurze Beruhigungszeit.

Diese Forderungen stehen teilweise in Konflikt zueinander und können nicht alle gleichzeitig optimal erfüllt werden.



In vielen Anwendungen werden spezifische Forderungen ausgestellt, die den dortigen charakteristischen Gegebenheiten Rechnung tragen (Rampenförmige Führungsgröße, Sinusförmige Störung einer bestimmten Frequenz, Überschwingen ganz verboten, ...).

10.3 Dynamische Anforderungen

Prinzipiell kann zwischen zwei Aufgaben einer Regelung unterschieden werden:

1. Folgen der Führungsgröße.
2. Ausregeln der Störgröße.

Dominiert bei einer Regelung eine dieser beiden Aufgaben, dann unterscheidet man in:

1. **Folgeregelung** oder **Nachlaufregelung** (*tracking*). (Störgröße ist meist $d(t) \approx 0$.)
2. **Festwertregelung** (*regulation*). (Führungsgröße ist meist $w(t) = \text{konst.}$)

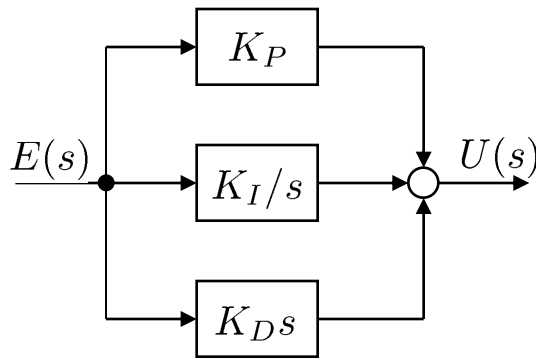
Typische Beispiele sind:

1. Roboterbahnregelung, Positionsregelung CNC-Maschinen, Lenkwaffen, ...
2. Raumtemperaturregelung, Tempomat beim Auto, Magnetlager, Synchronmaschine, ...

Wenn beide Aspekte (Führungsgrößenfolge und Störgrößenunterdrückung) sehr wichtig sind, reicht der Standardregelkreis meist nicht aus, um beide Ziele gleichzeitig zu erfüllen. Dann muss der Regelkreis um eine *Vorsteuerung* erweitert werden, siehe Kapitel 12.

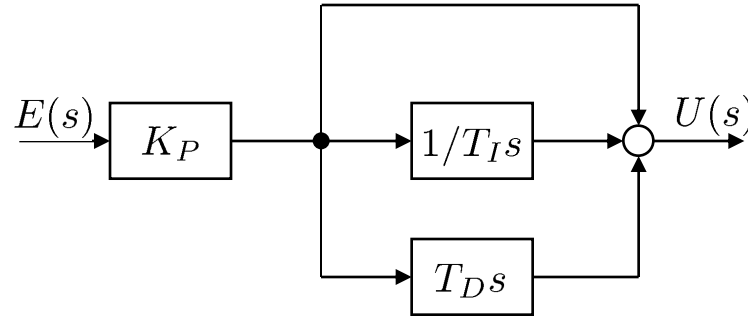
10.4 PID-Regler

Die in der Praxis am häufigsten eingesetzte Reglerstruktur ist der *PID-Regler*. Es gibt 2 alternative Darstellungsmöglichkeiten, die inhaltlich identisch sind:



$$G_R(s) = K_P + K_I \frac{1}{s} + K_D s$$

↑ ↑ ↑
 P-Anteil I-Anteil D-Anteil



$$G_R(s) = K_P \left(1 + \frac{1}{T_I s} + T_D s \right)$$

↑ ↑ ↑
 proportionale Nachstellzeit Vorhaltezeit
 Verstärkung (Integrierzeit) (Differenzierzeit)

Im Zeitbereich:

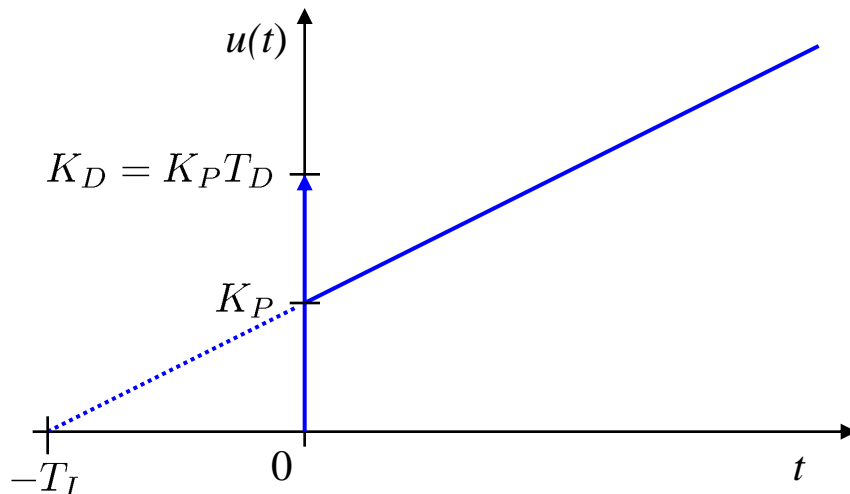
$$u(t) = K_P e(t) + K_I \int_0^t e(\tau) d\tau + K_D \frac{d}{dt} e(t) \qquad u(t) = K_P \left(e(t) + \frac{1}{T_I} \int_0^t e(\tau) d\tau + T_D \frac{d}{dt} e(t) \right)$$

10.4 PID-Regler

Als vereinfachte Sonderfälle des PID-Reglers entstehen durch Nullsetzen der K_P -, K_I -, K_D -Faktoren P -, PI -, PD - oder I -Regler (alle anderen Kombinationsmöglichkeiten sind von untergeordneter praktischer Bedeutung).

Typischerweise fällt beim Reglerentwurf erst die Entscheidung für eine bestimmte *Reglerstruktur* (z.B. PI - oder PID -Regler) und anschließend werden die *Reglerparameter* (z.B. K_P und K_I) bestimmt.

Sprungantwort eines PID-Reglers



D-Anteil: Driac-Impuls der Höhe $K_P T_D$

P-Anteil: Auf diesen Wert (K_P) fällt die Sprungantwort nach dem Abklingen des D-Anteils wieder zurück.

I-Anteil: Linearer Anstieg der Steigung K_P/T_I

10.4 PID-Regler

P-Anteil: $u(t) = K_P e(t)$

- Stellgröße ist proportional zur Regelabweichung.
- "Je größer die Regelabweichung ist, desto größer muss auch die Stellgröße sein!"
- Schneller Abbau der Regelabweichung; aber sie wird evtl. nicht vollständig abgebaut.

I-Anteil: $u(t) = K_I \int_0^t e(\tau) d\tau$ oder $\frac{du(t)}{dt} = K_I e(t)$

- Stellgröße entspricht dem Integral (Summe) der vorangegangenen Regelabweichungen.
- Änderung der Stellgröße entspricht der Regelabweichung.
- "Solange eine Regelabweichung auftritt, muss die Stellgröße verändert werden!"
- Langsamer aber vollständiger Abbau der Regelabweichung.

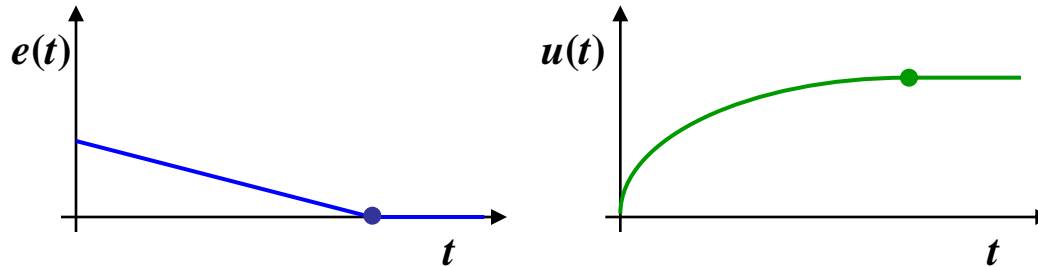
D-Anteil: $u(t) = K_D \frac{d}{dt} e(t)$

- Stellgröße entspricht der Änderung (Steigung, Trend) der Regelabweichung.
- "Je stärker sich die Regelabweichung ändert, umso stärker muss die Regelung eingreifen!"
- Sehr schnelle Reaktion. Auch bei kleinen Regelabweichungen mit falscher Tendenz aktiv.

10.4 PID-Regler

Veranschaulichung der Wirkung des I-Anteils

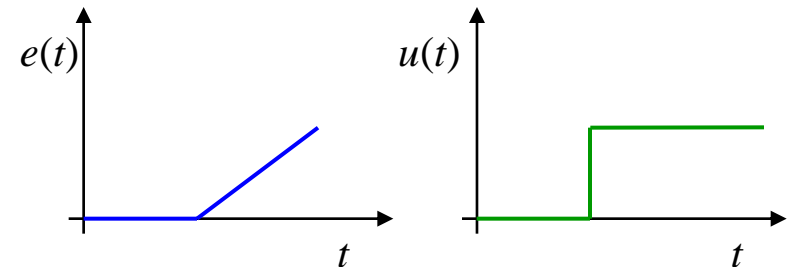
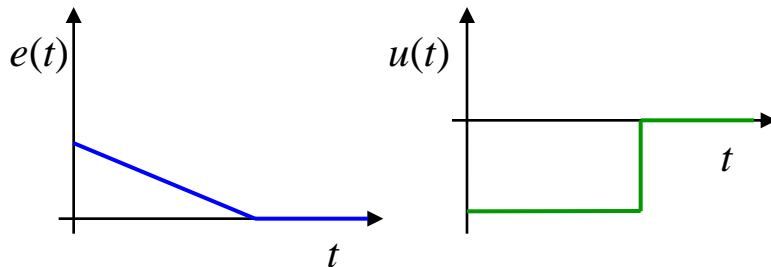
Die Stellgröße wird so lange erhöht bis die Regelabweichung bei 0 angekommen ist. Dann wächst das Integral nicht mehr und die Stellgröße bleibt konstant.



Veranschaulichung der Wirkung des D-Anteils (im Zusammenspiel mit einem P-Anteil)

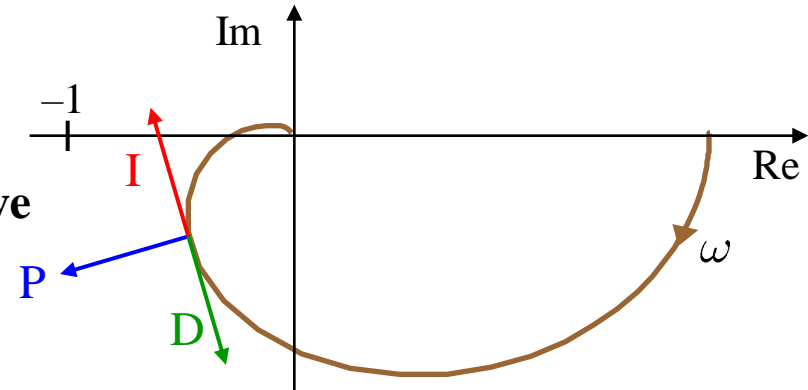
Wenn die Regelabweichung $\rightarrow 0$ abnimmt, wird eine negative Komponente erzeugt, die den P-Anteil abschwächt.

Wenn die Regelabweichung $\rightarrow \infty$ zunimmt, wird eine positive Komponente erzeugt, die den P-Anteil unterstützt.



10.4 PID-Regler

Einfluss des P-, I- und D-Anteils auf die Ortskurve



P-Anteil

Die Ortskurve wird proportional zum P-Anteil (Reglerverstärkung) skaliert, da der Betrag mit dem P-Anteil multipliziert wird und die Phase unverändert bleibt. Eine Vergrößerung des P-Anteils bringt die Ortskurve näher an den Punkt $(-1, 0)$ heran und wirkt daher destabilisierend.

I-Anteil

Die Phase wird je nach Stärke des I-Anteils zwischen 0° und -90° verschoben. D.h. Punkte gleicher Frequenz erfahren eine entsprechende zusätzliche Drehung im Uhrzeigersinn. Eine Vergrößerung des I-Anteils dreht die Ortskurve näher an den Punkt $(-1, 0)$ heran und wirkt daher destabilisierend.

D-Anteil

Die Phase wird je nach Stärke des D-Anteils zwischen 0° und $+90^\circ$ verschoben. D.h. Punkte gleicher Frequenz erfahren eine entsprechende zusätzliche Drehung gegen den Uhrzeigersinn. Eine Vergrößerung des D-Anteils dreht die Ortskurve weiter vom Punkt $(-1, 0)$ weg und wirkt daher stabilisierend.

10.4 PID-Regler

Zusammenfassung

- P-Anteil: Bewertet die Gegenwart, d.h. die aktuellen Regelabweichung
- I-Anteil: Bewertet die Vergangenheit, d.h. die akkumulierte (integrierte) Regelabweichung
→ Nachstellzeit T_I
- D-Anteil: Bewertet die Zukunft; mit der Berechnung des Trends (Ableitung) wird der zukünftige Verlauf der Regelabweichung vorhergesagt.
→ Vorhaltezeit T_D

Bemerkung: Wenn der Mensch selbst regelt, dann tut er dies oft als I- oder PI-Regler, d.h. die Änderung der Stellgröße hängt von der Regelabweichung ab. Absolute Werte sind oft nicht so entscheidend.

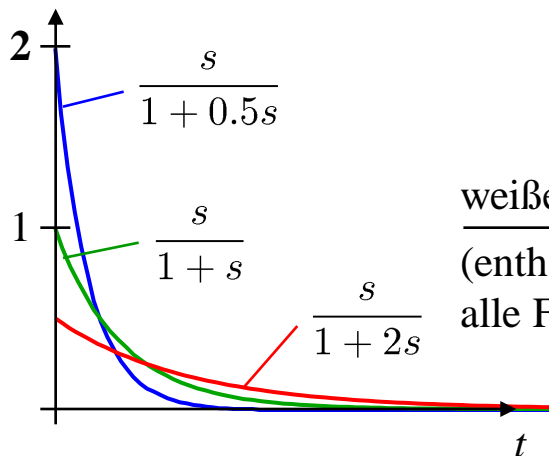
Beispiel: *Temperaturregelung unter der Dusche.* Typischerweise dreht man den Temperaturstellknopf nach rechts oder links mit einer Drehwinkelgeschwindigkeit proportional zur Regelabweichung (Differenz zwischen gewünschter und tatsächlicher Wassertemperatur). Dabei ist die absolute Skala ziemlich egal. Es sei denn, der Temperaturstellknopf ist am Anfang total verstellt. Dann wird wahrscheinlich noch eine Steuerungskomponente aktiv.

10.4 PID-Regler

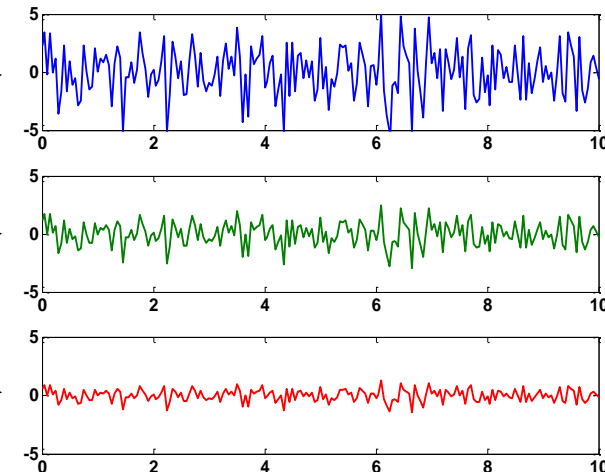
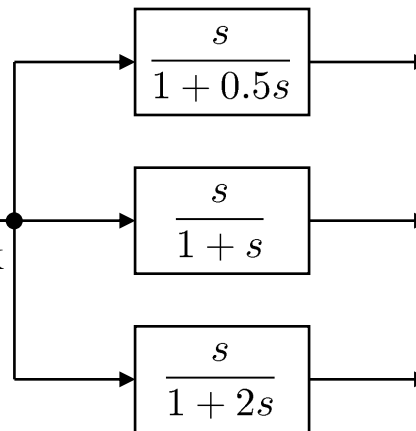
PID-Regler ist so nicht realisierbar! → Reiner D-Anteil, siehe Kapitel 7.

- D-Anteil wird als DT_1 -Glieder realisiert.
- Zeitkonstante des DT_1 -Glieds bestimmt die "Stärke" der Differentiation.
- DT_1 -Glieder: $K_D \frac{s}{1 + Ts}$
- Für den Grenzfall $T \rightarrow 0$ ergibt sich wieder ein reines D-Verhalten.
- Für den Grenzfall $T \rightarrow \infty$ ergibt sich ein P-Verhalten.

DT_1 -Sprungantworten: **Rauschverstärkung:**



weißes Rauschen
(enthält gleich stark
alle Frequenzen)



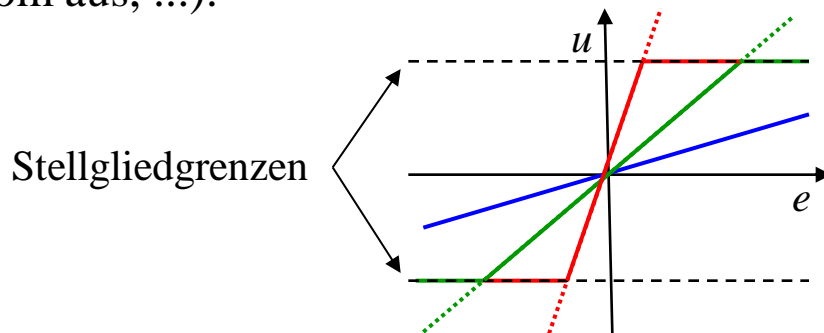
10.4 PID-Regler

Prinzipien bei der Auslegung von PID-Reglern, P-Anteil:

Der P-Anteil ist fast immer vorhanden (Ausnahme: reiner I-Regler). Der P-Anteil bestimmt die Kreisverstärkung des Regelkreises und ist somit für die Stabilität sehr wichtig. Man muss einen Kompromiss finden zwischen schneller Dynamik und kleiner bleibender Regelabweichung $1/(1+K)$ (ohne I-Anteil) einerseits und genügendem Abstand zur Stabilitätsgrenze und guter Dämpfung andererseits.

Für einfache Regelaufgaben reicht oft ein P-Regler schon aus, wenn die bleibende Regelabweichung vernachlässigt oder toleriert werden kann (oder = 0 ist, weil die Strecke einen I-Anteil besitzt).

Für sehr hohe Verstärkungen nähert sich der P-Regler einem *Zweipunktregler* an, der zwischen den Stellgliedergrenzen hin- und herschaltet (z.B. Ventil auf / Ventil zu, Strom an / Strom aus, ...):



P-Regler mit sehr großer Verstärkung

P-Regler mit großer Verstärkung

P-Regler mit kleiner Verstärkung

10.4 PID-Regler

Prinzipien bei der Auslegung von PID-Reglern, I-Anteil:

Hauptzweck eines I-Anteils ist die Vermeidung bleibender Regelabweichungen. Daher ist ein I-Anteil normalerweise nicht nötig, wenn die Strecke schon einen I-Anteil besitzt.

Ausnahme: Es wird ein doppelter I-Anteil zur Vermeidung von Schleppfehlern benötigt.

Oder ein sehr vorsichtiger Regler soll entworfen werden: Ein reiner I-Regler erzeugt keine Sprünge in der Stellgröße, ist also immer recht langsam. Das ist z.B. für Prozesse mit großer Totzeit ratsam.

Wird ein I-Anteil im Regler nicht aus wichtigen Gründen benötigt, sollte er weggelassen werden (P- oder PD-Regler), da seine Phasenabsenkung um 90° destabilisierend wirkt und den Phasenrand entsprechend verkleinert.

10.4 PID-Regler

Prinzipien bei der Auslegung von PID-Reglern, **D-Anteil**:

Der D-Anteil wird oft nicht verwendet (P- oder PI-Regler). Zum Einsatz kommt er hauptsächlich, wenn schnelle Dynamik gefragt ist oder die Strecke instabil ist (häufig kommt I- bzw. doppeltes I-Verhalten in der Strecke vor; z.B. Drehzahl \rightarrow Position oder Beschleunigung \rightarrow Drehzahl bzw. Beschleunigung \rightarrow Position).

Mit dem D-Anteil sollte man sehr vorsichtig umgehen. Eine gute Strategie ist es, ohne D-Anteil zu starten und ihn schrittweise zu vergrößern. Sich schnell ändernde Führungs- oder Störgrößen (z.B. Sprünge) können wegen der starken Reaktion des D-Anteils leicht zu Überschwingern führen. Dies kann man für die Führungsgröße abschwächen, indem man diese mit einem sog. Führungsgrößenfilter/Vorfilter tiefpassfiltert.

Hochfrequente Rauschanteile werden durch die Differentiation verstärkt (um so mehr, je kleiner die Zeitkonstante des realen DT_1 -Glieds ist).

Generell hat der D-Anteil einen positiven Einfluss auf die Dynamik: Er macht den Regelkreis schneller und hebt die Phase um 90° an.

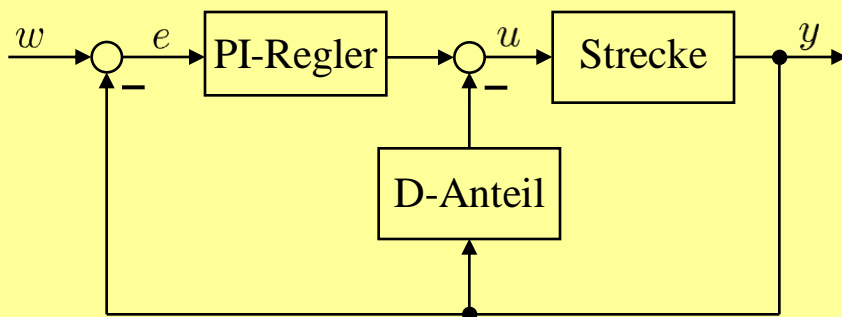
10.4 PID-Regler

PID-Regler mit D-Anteil nur in der Rückführung (in der Industrie weit verbreitet!):

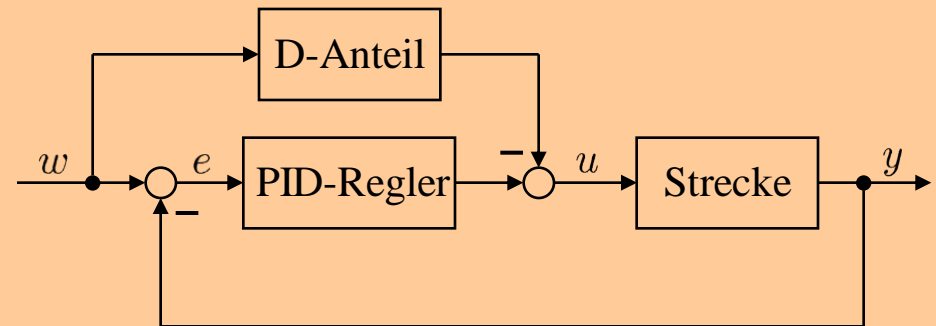
Der D-Anteil soll auf den Trend (Änderung) der Regelabweichung reagieren. Das funktioniert sehr gut solange sich die Führungsgröße nicht ändert (*Festwertregelung*) bzw. zumindest nicht sehr schnell ändert. Bei einem sprungförmigen Verlauf der Führungsgröße w wird eine extreme Reaktion (theoretisch unendlich) des D-Anteils provoziert, die normalerweise nicht gewollt ist.

Daher wird in **industriellen PID-Reglern** der D-Anteil oft nur in die Rückführung integriert. D.h. es wird die Ableitung von $-y$ und nicht von $w-y$ verwendet. Alternativ (rechtes Bild) kann man den D-Anteil von w über eine negative Vorsteuerung wieder eliminieren.

PID-Regler ohne D-Anteil im Pfad $w \rightarrow y$



PID-Regler ohne D-Anteil im Pfad $w \rightarrow y$



10.4 PID-Regler

Übertragungsfunktion eines PID-Regler mit D-Anteil nur in der Rückführung:

Die Stellgröße berechnet sich auf Führungs- und Regelgröße wie folgt:

$$\begin{aligned} U(s) &= \left(K_P + \frac{K_I}{s} \right) (W(s) - Y(s)) - K_D s Y(s) \\ &= \left(K_P + \frac{K_I}{s} + K_D s \right) (W(s) - Y(s)) - K_D s W(s) \end{aligned}$$

Die Übertragungsfunktion des offenen Regelkreises $G_0(s)$ hat sich im Vergleich zum "normalen" PID-Regler nicht verändert:

$$G_0(s) = \left(K_P + \frac{K_I}{s} + K_D s \right) G_S(s) \quad \text{(identisch zu } G_0(s) \text{ mit normalem PID-Regler)}$$

wobei $G_S(s)$ die Übertragungsfunktion der Strecke ist. D.h. wir können mit dem industriellen PID-Regler genauso die Pole des geschlossenen Regelkreises beeinflussen wie mit dem "normalen" PID-Regler. Unterschiedlich ist aber der Vorwärtszweig, der nur aus dem PI-Anteil besteht. Daher ergibt sich ein verändertes Führungsverhalten des Regelkreises:

$$G_w(s) = \frac{\left(K_P + \frac{K_I}{s} \right) G_S(s)}{1 + \left(K_P + \frac{K_I}{s} + K_D s \right) G_S(s)} \quad \text{(anders als } G_w(s) \text{ mit normalem PID-Regler)}$$

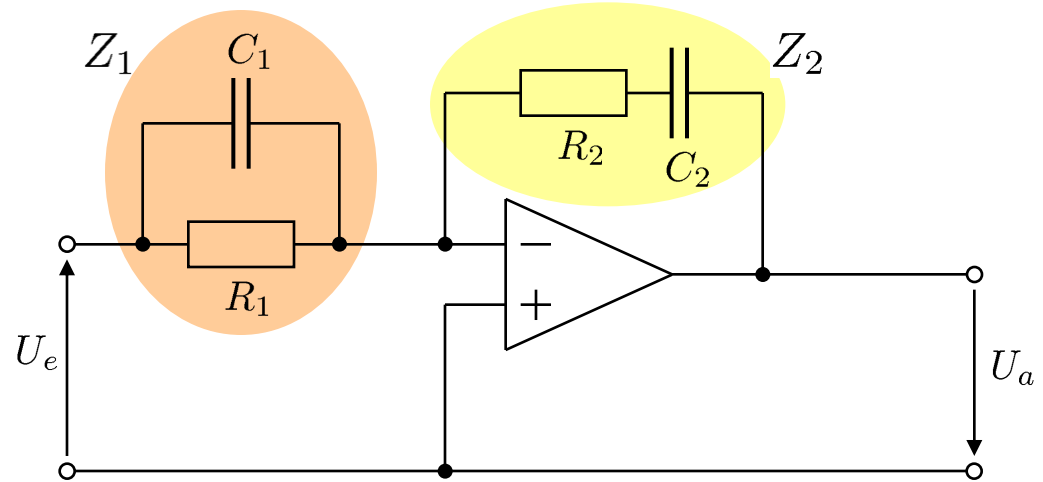
10.5 Reglerrealisierungen

Elektronische Realisierung eines PID-Reglers mittels Operationsverstärkerschaltung

Eigenschaften eines idealen negativ rückgekoppelten Operationsverstärkers:

- Verstärkung $\rightarrow \infty$.
- Eingangsspannung ≈ 0 .
- Eingangsstrom ≈ 0 .

$$G_R(s) = \frac{Z_2}{Z_1}$$



Übertragungsfunktion:

$$\frac{U_a(s)}{U_e(s)} = G_R(s) = \frac{R_2 + \frac{1}{sC_2}}{\frac{R_1}{sC_1}} = \frac{(R_1 + \frac{1}{sC_1})(R_2 + \frac{1}{sC_2})}{\frac{R_1}{sC_1}} = \frac{(1 + sR_1C_1)(1 + sR_2C_2)}{sR_1C_2}$$

$$G_R(s) = \frac{1}{R_1C_2s} + \frac{R_1C_1 + R_2C_2}{R_1C_2} + R_2C_1s$$

I-Anteil
P-Anteil
D-Anteil

4 Schaltungsparameter zum Einstellen von 3 Reglerparametern
 → Jeder beliebige PID-Regler möglich!

11. Reglerentwurf mittels Optimierung oder Einstellregeln

Inhalt Kapitel 11

11 Reglerentwurf mittels Optimierung und Einstellregeln

11.1 Gütekriterien

11.2 Parameteroptimierung

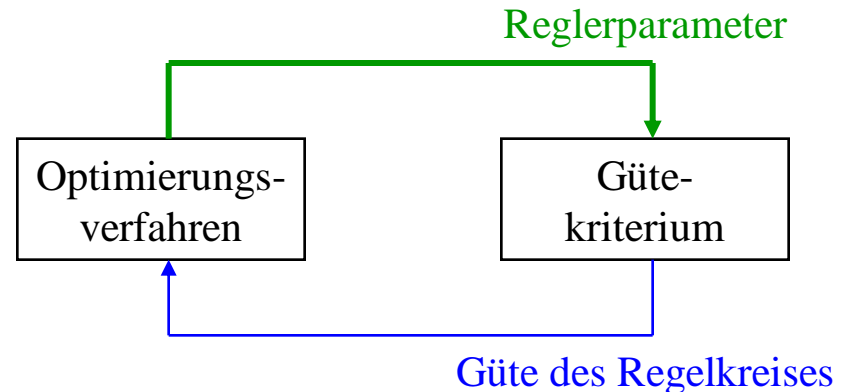
11.3 Einstellregeln nach Ziegler-Nichols

11.4 Einstellregeln nach Summenzeitkonstante

11.1 Gütekriterien

Eine nahe liegende Idee für den Reglerentwurf ist folgende:

1. Lege die Reglerstruktur fest (z.B. P-Regler oder PI-Regler, ...).
2. Überlege ob das Störgrößen- oder Führungsgrößenverhalten wichtiger ist und welcher Signaltyp als Störgröße bzw. Führungsgröße zu erwarten ist (z.B. Einheitssprung von $w(t)$).
3. Stelle ein Gütekriterium auf, das die Qualität (Performance) der Regelung beurteilt.
4. Programmiere eine Funktion, die dieses Gütekriterium ausrechnet in Abhängigkeit der noch zu ermittelnden Reglerparameter (z.B. K_P und K_I).
5. Setze ein Optimierungsverfahren ein, um die Reglerparameter so lange zu verändern, bis das Optimum des Gütekriteriums gefunden ist.
6. Verwende diese optimalen Parameter.



11.1 Gütekriterien

Vorbemerkungen

- Man kann neutral von einer Zielfunktion (*objective function*) sprechen, die optimiert werden soll oder von einer Gütefunktion (*quality function*), die maximiert werden soll oder von Verlust- (*loss*) bzw. Kosten- (*cost*) bzw. Fehlerfunktion (*error function*), die minimiert werden soll. Das ist Alles äquivalent, und lässt sich per Vorzeichenumkehr ineinander überführen.

Wir sprechen hier meist von einer zu minimierenden **Verlustfunktion J** .

- Man unterscheidet *lineare* und *nichtlineare* Optimierungsprobleme.
- Wenn alle zu optimierenden Parameter den Fehler linear beeinflussen und die Verlustfunktion als Summe (oder Integral) der Fehlerquadrate gewählt wird, dann entsteht ein lineares Optimierungsproblem.

Beispiel: $y = c_0 + c_1u + c_2u^2 + c_3u^3$. $J = \sum_{i=1}^N [y_d(i) - y(i)]^2 \rightarrow \min.$

Anzahl der Datenpunkte

gewünschte Funktionswerte

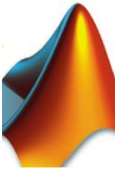
Polynomwerte

11.1 Gütekriterien

Lineare Optimierungsprobleme haben folgende sehr "schöne" Eigenschaften:

- Eindeutige Lösung.
- Geschlossene Lösung.
- Ausgereifte, schnelle, numerisch stabile Lösungsverfahren.
- MATLAB-Funktionen: `pinv`, `inv`, `"\"`.

MATLAB



Nichtlineare Optimierungsprobleme haben folgende weniger "schöne" Eigenschaften:

- Lokale Optima können neben dem globalen Optimum existieren.
- Keine geschlossene Lösung, nicht mal für lokale Optima.
- Iterative, langsamere aber numerisch stabile und ausgereifte *lokale* Lösungsverfahren. D.h. deren Lösung hängt vom Anfangswert/Startwert der Parameter ab.
- Iterative, sehr langsame *globale* Lösungsverfahren, die versuchen, das globale bzw. zumindest ein "gutes" lokales Optimum zu finden.
- MATLAB-Funktionen: `fminsearch`, `fminunc`, `lsqnonlin` (lokale Verfahren in der Optimierungs-Toolbox).

MATLAB



11.2 Parameteroptimierung

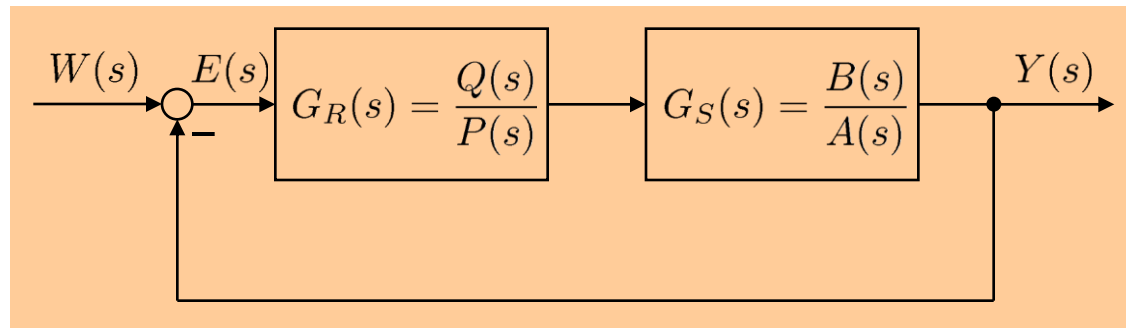
Regleroptimierung ist ein nichtlineares Optimierungsproblem

Im Standardregelkreis ist die Regelabweichung im Frequenzbereich gegeben durch:

$$\begin{aligned} E(s) &= W(s) - Y(s) = W(s) - G_w(s)W(s) = (1 - G_w(s))W(s) \\ &= \left(1 - \frac{G_R(s)G_S(s)}{1 + G_R(s)G_S(s)}\right) W(s) = \left(1 - \frac{Q(s)B(s)}{P(s)A(s) + Q(s)B(s)}\right) W(s) \\ &= \frac{P(s)A(s)}{P(s)A(s) + Q(s)B(s)} W(s) \end{aligned}$$

Selbst für einfache Regler, wie z.B. einen P-Regler ($Q(s) = K_R$, $P(s) = 1$) ergibt sich ein nichtlinearer Zusammenhang zwischen dem/den Reglerparameter(n) und der Regelabweichung:

$$E(s) = \frac{A(s)}{A(s) + K_R B(s)} W(s)$$



11.1 Gütekriterien

Mögliche Gütekriterien

$$e(t) = w(t) - y(t)$$

- **Aufintegrierter Fehlerbetrag**

$$J = \int_0^{T_{\text{sim}}} |e(t)| dt$$

Positive und negative Fehleranteile sollen sich nicht aufheben.

- **Aufintegriertes Fehlerquadrat**

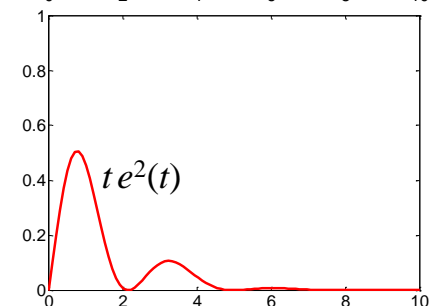
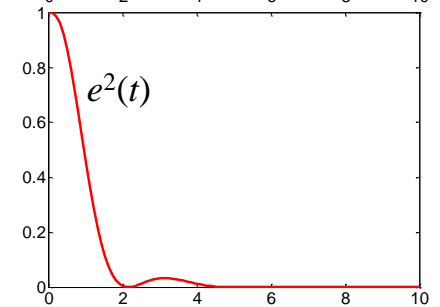
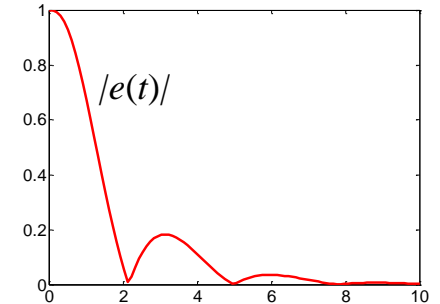
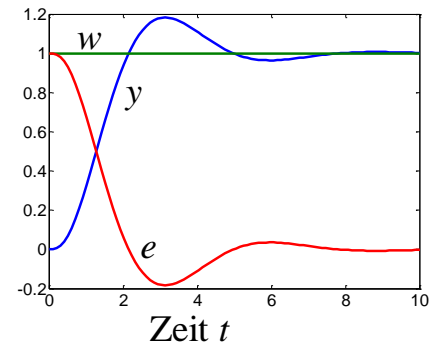
$$J = \int_0^{T_{\text{sim}}} e^2(t) dt$$

Bietet viele rechnerische Vorteile bei der Optimierung.

- **Aufintegriertes zeitgewichtetes Fehlerquadrat**

$$J = \int_0^{T_{\text{sim}}} t e^2(t) dt$$

Eine spätere Abweichung wird stärker bestraft.



11.1 Gütekriterien

- **Aufintegriertes Fehlerquadrat mit Bestrafung der quadratischen Stellgröße**

$$J = \int_0^{T_{\text{sim}}} [e^2(t) + \alpha u^2(t)] dt$$

Das aufintegrierte Quadrat der Stellgröße entspricht der eingesetzten Stellenergie. In einigen Anwendungen wird auch direkt eine Energieeinsparung damit bezweckt (Luft- und Raumfahrt). Typischerweise ist dieser Strafterm aber nur ein Vehikel, um die Stellgrößenauslässe niedriger (im Vergleich zu $\alpha = 0$) zu halten, damit

- a) ihre Begrenzungen/Anschläge nicht oder nicht so oft/lange überschritten werden (bzw. $u(t)$ nicht gegen unendlich strebt),
- b) das Stellglied nicht so schnell verschleißt,
- c) durch große Stellgrößenamplituden keine so hohen Frequenzen angeregt werden, dass das zugrunde liegende Streckenmodell zu ungenau wird (Vernachlässigung schneller Dynamik bei der Modellierung),
- d) die Regelung weniger aggressiv (d.h. langsamer) und damit robuster wird.

11.1 Gütekriterien

- **Aufintegriertes Fehlerquadrat mit Bestrafung der quadratischen *Änderung* der Stellgröße**

$$J = \int_0^{T_{\text{sim}}} [e^2(t) + \alpha \Delta u^2(t)] dt$$

Mit $\Delta u(t)$ ist die Änderung der Stellgröße pro Zeiteinheit gemeint, also eigentlich die Ableitung von $u(t)$. Wenn die Schrittweite Δt des Integrationsverfahrens aber gleich bleibt, kann man vereinfacht auch $\Delta u(t) = u(t) - u(t - \Delta t)$ schreiben (genauso wie bei abgetasteten Systemen).

Die Bestrafung von $\Delta u^2(t)$ statt $u^2(t)$ hat folgende Vorteile:

1. In vielen Anwendungen verbraucht nicht das Halten eines Absolutwerts der Stellgröße Energie sondern die *Änderung* der Stellgröße.
2. Wenn die Strecke keinen I-Anteil besitzt, dann gilt für den Regelkreis im eingeschwungenen Zustand: Falls $y_0 \neq 0$ dann folgt daraus auch $u_0 \neq 0$. D.h. der Strafterm $\int \alpha u^2(t) dt$ wird mit steigender Simulationszeit T_{sim} immer größer.
3. Eine Bestrafung von $u^2(t)$ ist besonders dann sinnvoll, wenn die Stellgröße im eingeschwungenen Zustand wieder auf 0 zurück geht (Ausregelung von Impuls-Störungen).

11.1 Gütekriterien

Rolle des *Hyperparameters* α

- Der Name "Hyper-" oder "Metaparameter" rührt daher, dass dieser Parameter die Verlustfunktion bestimmt und damit indirekt alle Reglerparameter steuert.
- $\alpha \rightarrow 0$: Stellgröße wird nicht berücksichtigt. Regelkreis wird sehr (zu) schnell. In der Praxis ergeben sich häufig Stellgrößen, die vom Stellglied gar nicht realisiert werden können (z.B. ein Ventil kann nicht mehr als "auf" und weniger als "zu" sein; ein Motor hat eine Maximalleistung, ...). Solche Begrenzungen sind nichtlineare Effekte und wir berücksichtigen diese in RT nicht! Es ist klar, dass sich dann große Abweichungen zwischen Simulation und Realität ergeben können.
- $\alpha \rightarrow \infty$: Regelabweichung wird nicht berücksichtigt. Regelkreis wird beliebig langsam.

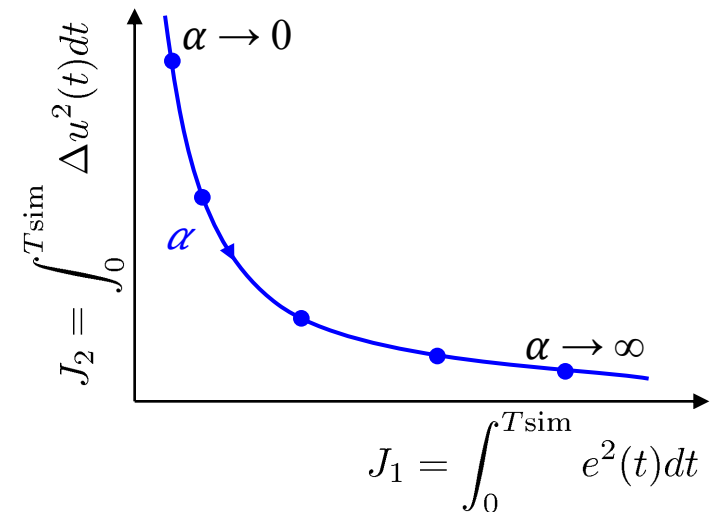
Menge der **pareto-optimalen Lösungen** von

$$J = \int_0^{T_{\text{sim}}} [e^2(t) + \alpha \Delta u^2(t)] dt$$

für $\alpha = 0 \dots \infty$.

J_1

J_2



11.1 Gütekriterien

Pareto-optimale Lösungen

- Alle möglichen Punkte auf der blauen Kurve sind pareto-optimal, d.h. sie lassen sich nicht in einer Zielfunktion verbessern ohne gleichzeitig die andere zu verschlechtern.
- Alle pareto-optimalen Lösungen sind prinzipiell gleichwertig. Welche den besten Kompromiss (*trade-off*) zwischen den beiden Zielen darstellt, muss der Benutzer (nach anderen Kriterien) entscheiden. Meist lassen sich diese "anderen Kriterien" schwer mathematisch fassen, sonst hätte man sie von vorn herein in die Verlustfunktion mit aufgenommen und das Optimierungsverfahren die insgesamt beste Lösung suchen lassen.

Mehrzieloptimierung (*multi-objective optimization*)

- Die Bildung einer gewichteten Summe aus 2 oder mehreren Zielfunktionen ist eine sehr typische Vorgehensweise, um aus einem Optimierungsproblem mit mehreren Zielen (z.B. minimaler Regelfehler und minimale Stellgrößenaktivität) ein Optimierungsproblem mit einem Ziel zu machen. Dafür muss man meist an dem Hyperparameter α "drehen" bis man einen guten Kompromiss zwischen den Zielen gefunden hat.
- Es gibt spezielle Mehrzieloptimierungsverfahren, die effizient viele Lösungen in der pareto-optimalen Menge finden und damit das "Drehen" an α automatisch durchführen.

11.2 Parameteroptimierung

Beispiel: Optimierung eines P-Reglers für die Strecke $G_S(s) = \frac{6}{(s+1)(s+2)(s+3)}$
auf einen Einheitssprung der Führungsgröße.

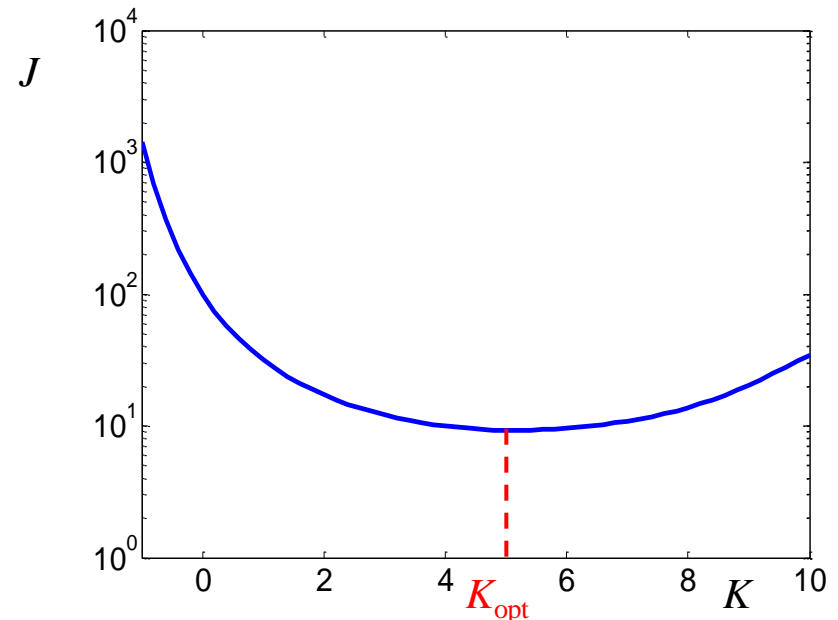
Hurwitz-Kriterium liefert folgende Stabilitätsgrenzen:

Charakteristische Gleichung: $s^3 + 6s^2 + 11s + 6 + 6K$

Bedingungen: $6 + 6K > 0 \rightarrow K > -1$

$11 \cdot 6 - (6 + 6K) \cdot 1 > 0 \rightarrow K < 10$

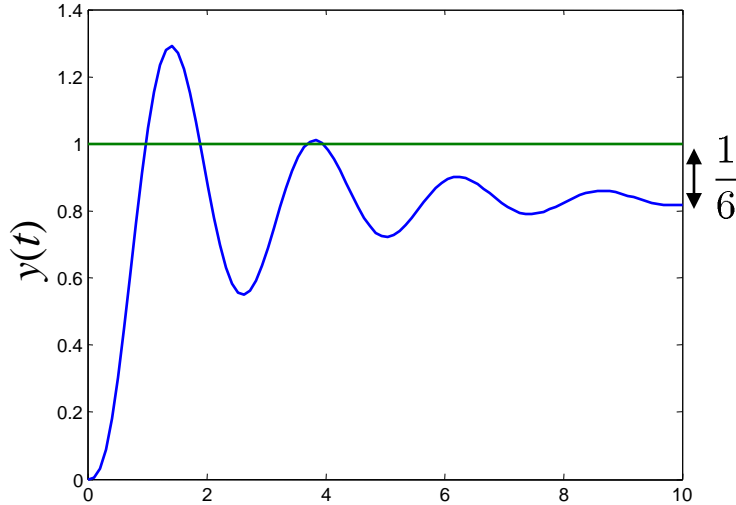
- Eindimensionales Optimierungsproblem
- Simulationszeit: $T_{\text{sim}} = 10 \text{ sec.}$
- Verlustfunktion: $J = \int_0^{T_{\text{sim}}} e^2(t) dt$
- $G_w(s) = \frac{6K}{s^3 + 6s^2 + 11s + 6 + 6K}$
- Bleibende Regelabw.: $1 - G_w(0) = \frac{1}{1 + K}$



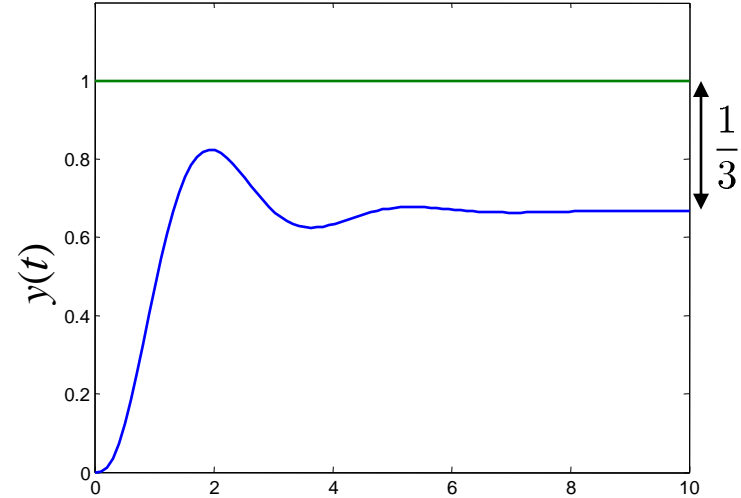
11.2 Parameteroptimierung

Optimierungsergebnisse

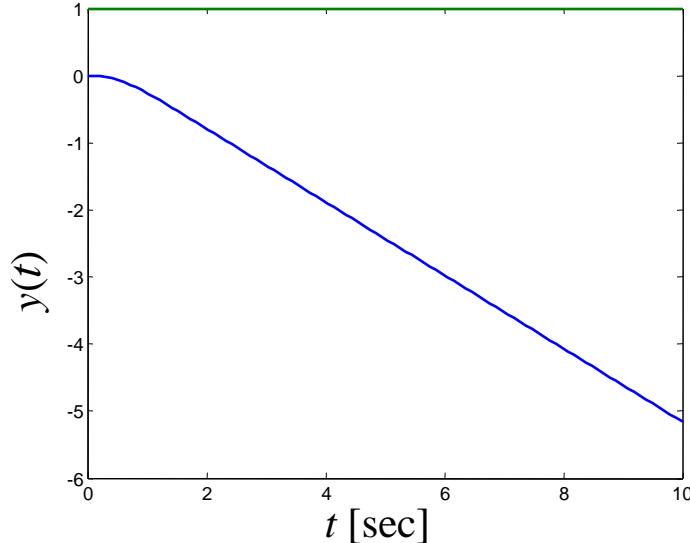
$K_{\text{opt}} = 5$



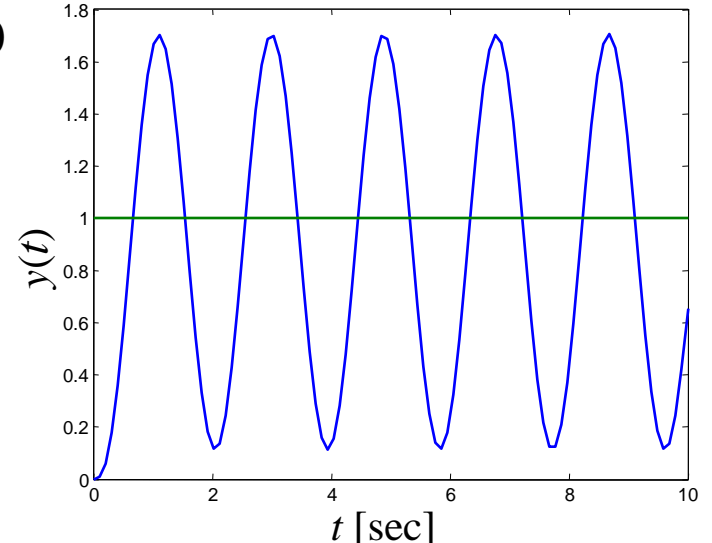
$K = 2$



$K = -1$



$K = 10$



11.2 Parameteroptimierung

Analyse der Ergebnisse

- Da es sich um einen P-Regler handelt und die Strecke keinen I-Anteil aufweist, ergibt sich eine bleibende Regelabweichung, die um so kleiner ist, je größer die Verstärkung gewählt wird.
- Der optimale P-Regler mit $K = 5$ erreicht kein befriedigendes Regelverhalten. Das stark oszillatorische Verhalten musste offensichtlich in Kauf genommen werden, um die bleibende Regelabweichung zu reduzieren, vergleiche Ergebnis für $K = 2$.
- Der Kompromiss zwischen wenig oszillatorischem Verhalten und geringer Regelabweichung hängt auch von der Wahl der Simulationszeit T_{sim} ab. Bei einer längeren Simulationszeit wird die bleibende Regelabweichung größeren Einfluss in der Verlustfunktion bekommen. Dann wird die optimale Reglerverstärkung ansteigen, und die Stärke der Oszillationen wird zunehmen.
- Die beiden Fälle $K = -1$ und $K = 10$ zeigen das Verhalten an der Stabilitätsgrenze, also für den grenzstabilen Fall. Offensichtlich hat der geschlossene Regelkreis für $K = -1$ einen Pol im Ursprung (integrales Verhalten) und für $K = 10$ einen imaginären Doppelpol bei $\pm i\omega_0$ (Dauerschwingung mit Kreisfrequenz ω_0).

11.2 Parameteroptimierung

Wahl der Simulationszeit T_{sim}

Im Regelfall wird der offene Regelkreis $G_0(s)$ einen Integrator enthalten und somit eine bleibende Regelabweichung vermeiden. Dann kann die Simulationszeit gegen unendlich gehen ($T_{\text{sim}} \rightarrow \infty$), um sicher zu stellen, dass alle Einschwingvorgänge abgeschlossen sind.

Die sehr häufig verwendete Verlustfunktion der aufintegrierten Fehlerquadrate kann dann mittels des folgenden Satzes in den Frequenzbereich transformiert werden:

Parsevalsches Theorem:

$$\int_0^\infty e^2(t) dt = \frac{1}{2\pi} \int_0^\infty |E(i\omega)|^2 d\omega$$

Energie eines Signals im Zeitbereich = Energie des transformierten Signals im Frequenzbereich

Wir müssen also $E(s)$ berechnen und können dann mit Hilfe des Parsevalschen Theorems eine geschlossene Formel (abhängig von den Reglerparametern) für die Verlustfunktion herleiten. Diese lässt sich dann viel schneller berechnen als eine Simulation im Zeitbereich.

11.2 Parameteroptimierung

Beispiel: Optimierung eines PI-Reglers für die Strecke $G_S(s) = \frac{6}{(s+1)(s+2)(s+3)}$ **auf einen Einheitssprung der Führungsgröße.**

Zunächst vereinfachen wir das Problem zu einem 1-Parameter-Optimierungsproblem. Später wird dann die allgemeine 2-Parameter-Optimierung (P- und I-Anteil gleichzeitig) behandelt.

Wir wählen die Zählerzeitkonstante des PI-Reglers gleich der langsamsten Nennerzeitkonstante der Strecke, so dass sich die langsamste Dynamik wegekürzt:

$$G_R(s) = K_P \left(1 + \frac{1}{T_I s} \right) = K_P \frac{s + 1/T_I}{s} \rightarrow \text{Wähle } T_I = 1 \text{ sec.}$$

Durch diese sinnvolle (aber nicht notwendigerweise optimale) Wahl des einen Reglerparameters reduziert sich das Problem auf die Optimierung von K_P .

Führungsübertragungsfunktion:

$$G_w(s) = \frac{\frac{6K_P}{s(s+2)(s+3)}}{1 + \frac{6K_P}{s(s+2)(s+3)}} = \frac{6K_P}{s^3 + 5s^2 + 6s + 6K_P}$$

Bleibende Reglabweichung ist nun durch I-Anteil beseitigt: $1 - G_w(0) = 1 - \frac{6K_P}{6K_P} = 0$

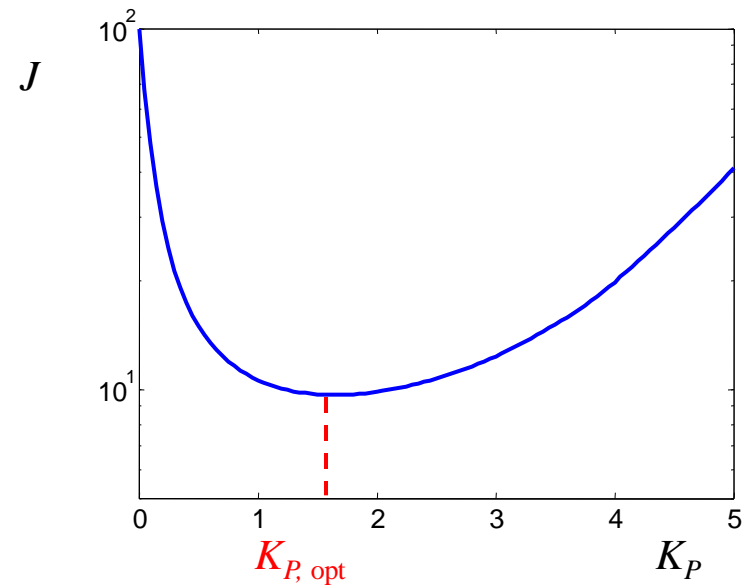
11.2 Parameteroptimierung

Hurwitz-Kriterium liefert folgende Stabilitätsgrenzen:

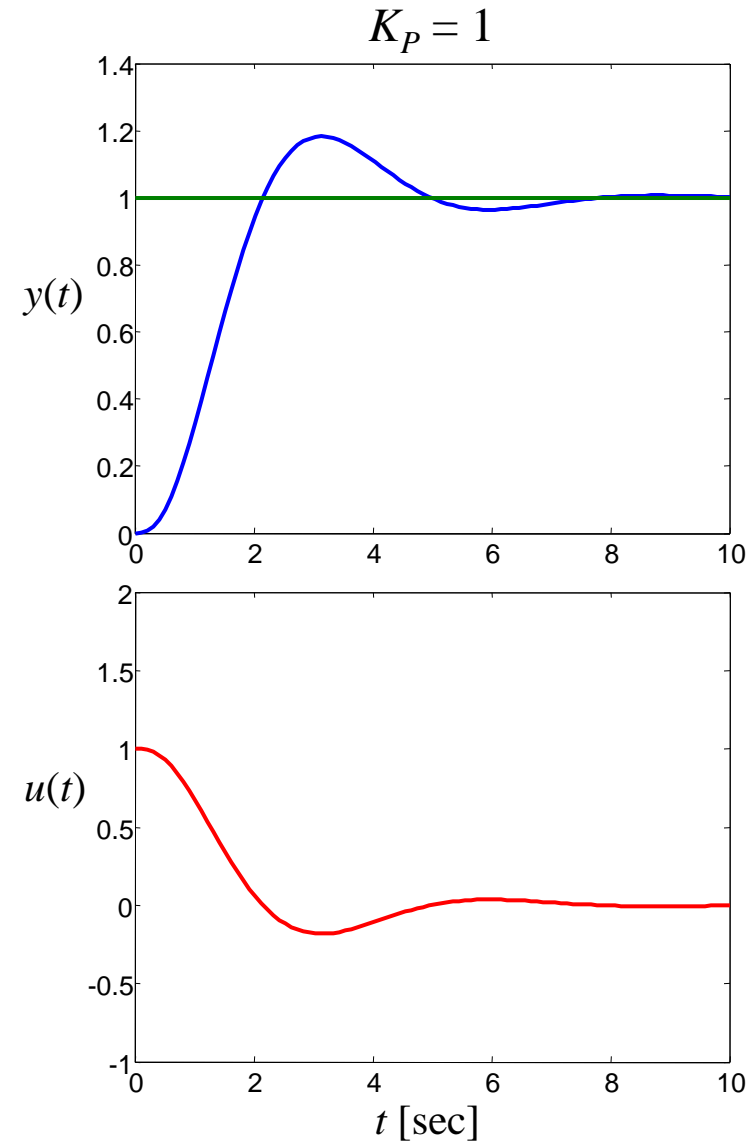
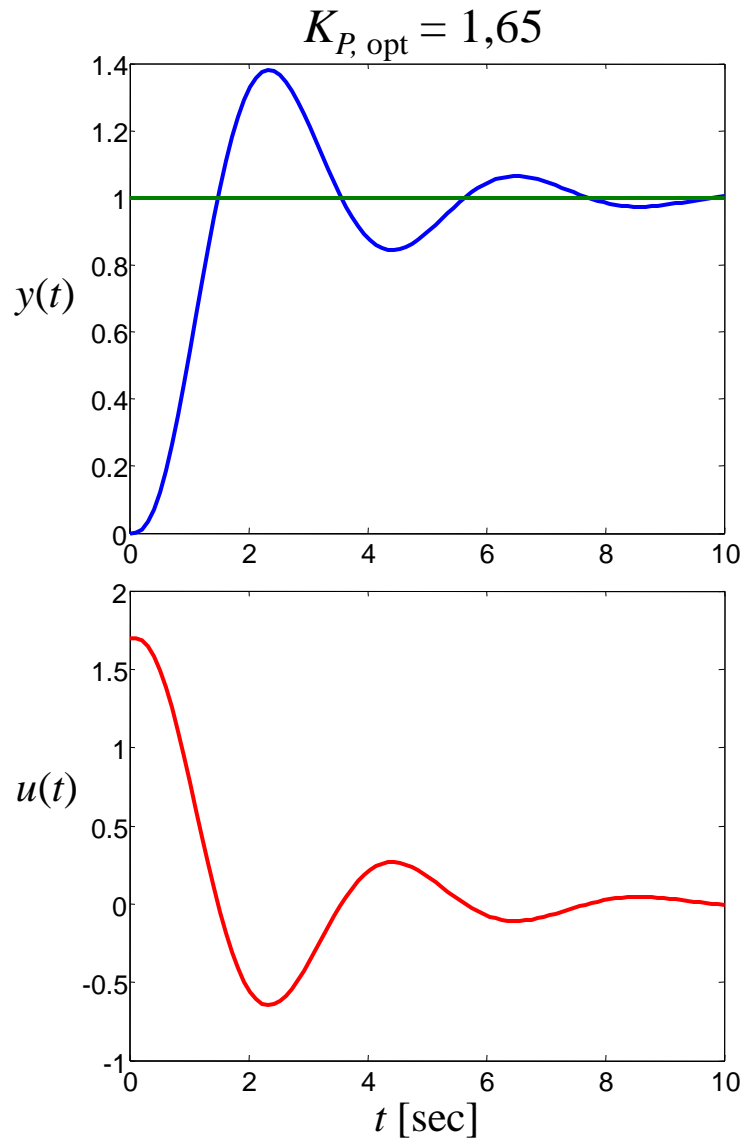
$$6K_P > 0 \rightarrow K_P > 0$$

$$6 \cdot 5 - 6K_P \cdot 1 > 0 \rightarrow K_P < 5$$

- Eindimensionales Optimierungsproblem
- Simulationszeit: $T_{\text{sim}} = 10 \text{ sec.}$
- Verlustfunktion: $J = \int_0^{T_{\text{sim}}} e^2(t) dt$
- $K_{P, \text{opt}} = 1,65$



11.2 Parameteroptimierung



11.2 Parameteroptimierung

Analyse der Ergebnisse

- Für das optimale Regelergebnis werden deutlich höhere (nach oben und nach unten) Stellgrößen benötigt als für die langsamere Variante mit $K_p = 1$.
- Die langsamere Variante scheint einige Vorteile für sich verbuchen zu können:
 1. Geringere Stellgrößenausschläge.
 2. Weniger Überschwingen.
 3. Kürzere Beruhigungszeit, d.h. schnelleres Einschwingen innerhalb eines schmalen Bandes um den Sollwert.

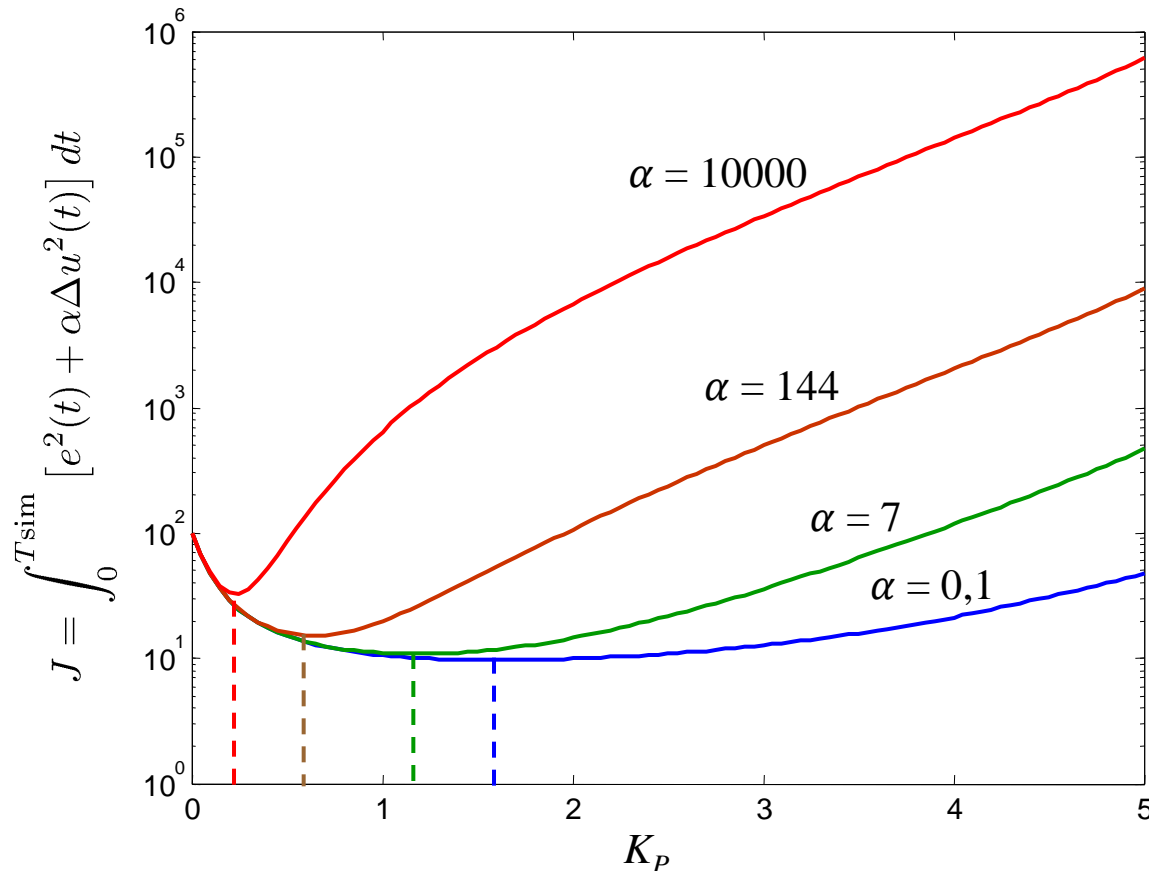
Als Nachteil im Vergleich zur optimalen Lösung steht dem "nur" gegenüber:

1. Längere Anstiegszeit.
- Wenn wir nun eine etwas langsamere Anstiegszeit in Kauf nehmen möchten, dann müssen wir die Verlustfunktion ändern. Z.B. könnten wir nur die Beträge –nicht die Quadrate– des Fehlers bestrafen, dann würden die großen Fehler (nahe $t = 0$ sec) nicht so stark gewichtet. Alternativ können wir die Stellgrößenaktivität bestrafen. Dies hat den Vorteil, dass wir mit dem Hyperparameter α unsere Zielvorstellung ausdrücken können.

11.2 Parameteroptimierung

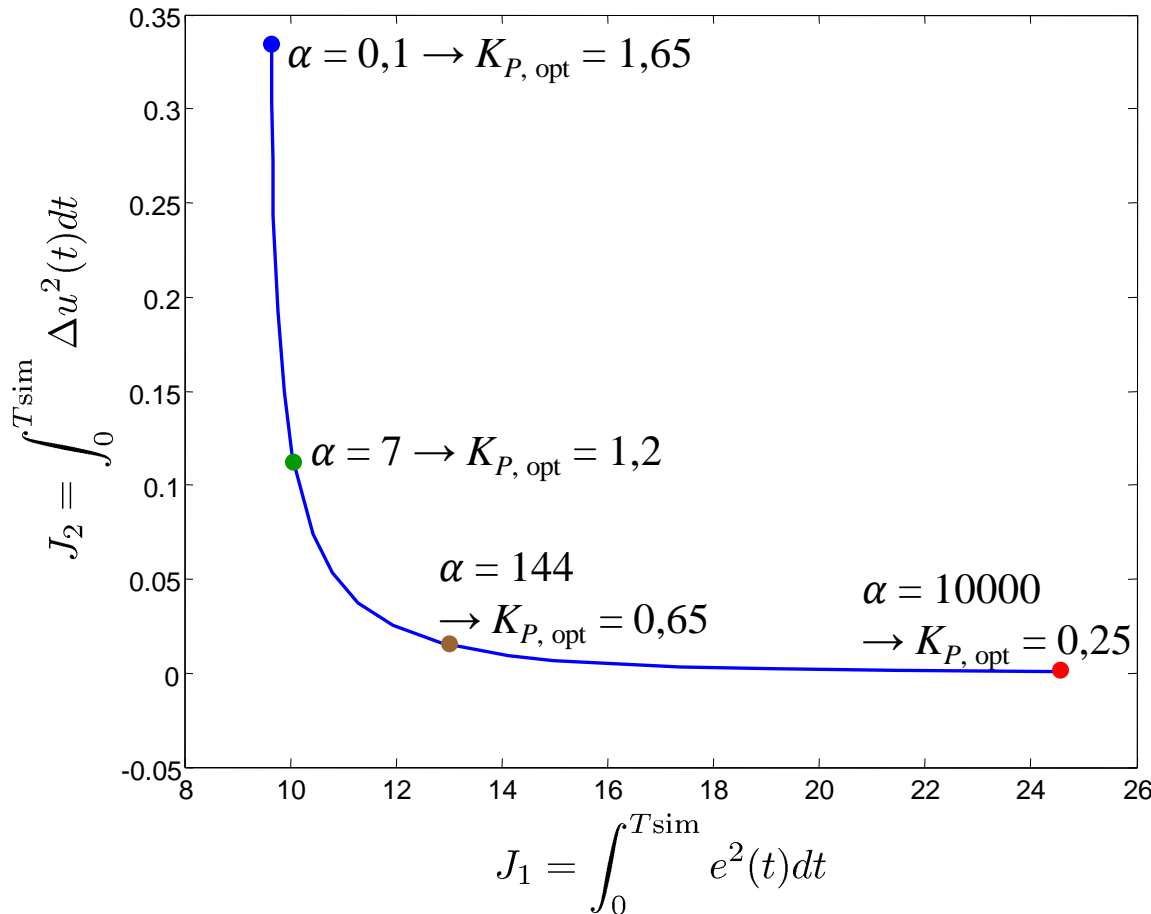
Verlauf der Verlustfunktion für verschiedenen Werte von α :

- Je größer α wird, desto weiter rückt das Minimum nach links, d.h. in Richtung kleinerer Verstärkungen, weil die Stellgrößenänderung verringert werden muss.



11.2 Parameteroptimierung

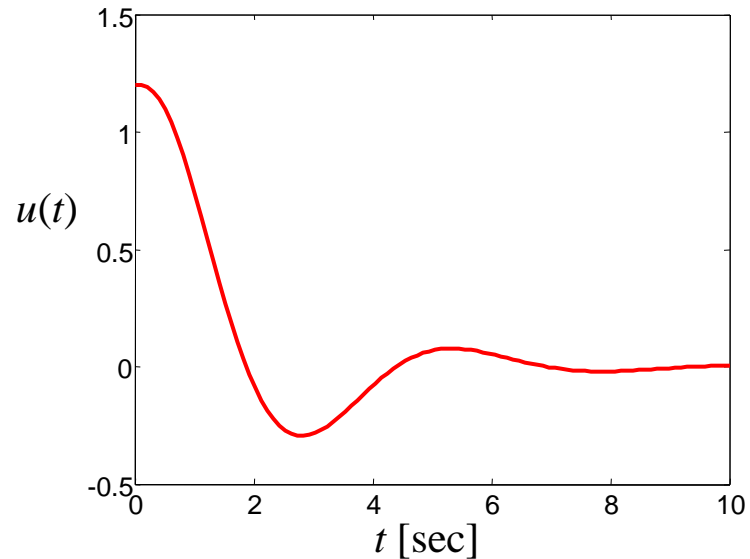
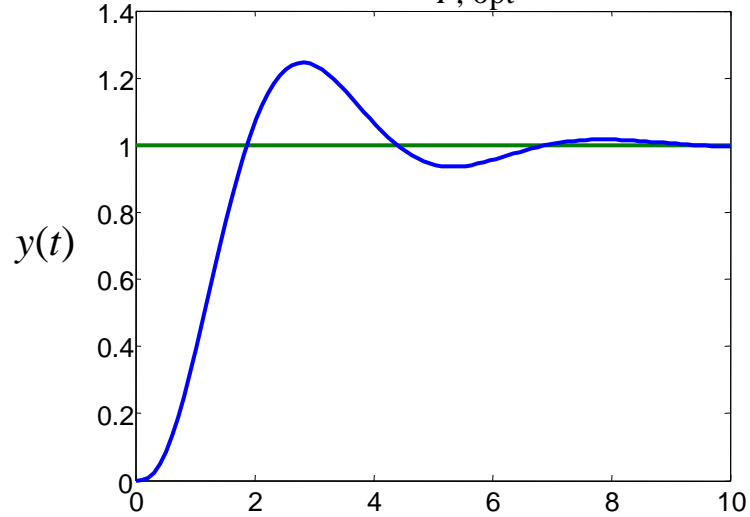
Also wählen wir als Verlustfunktion: $J = \int_0^{T_{\text{sim}}} [e^2(t) + \alpha \Delta u^2(t)] dt$



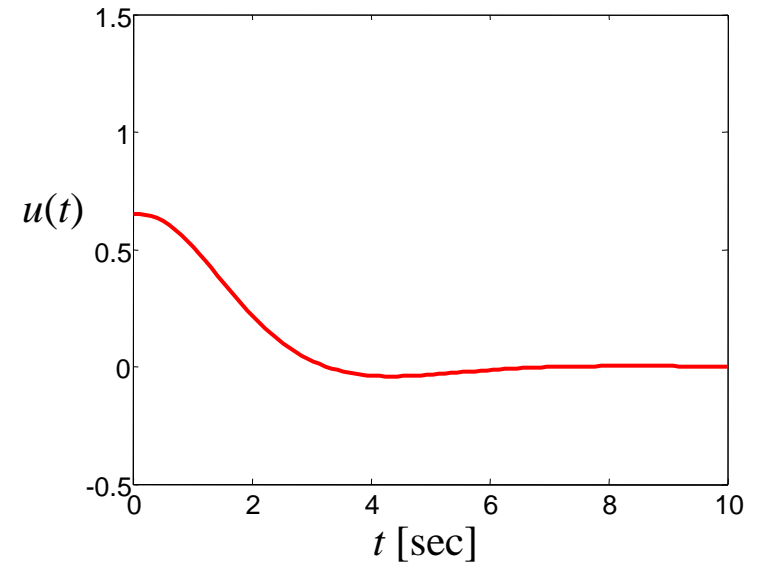
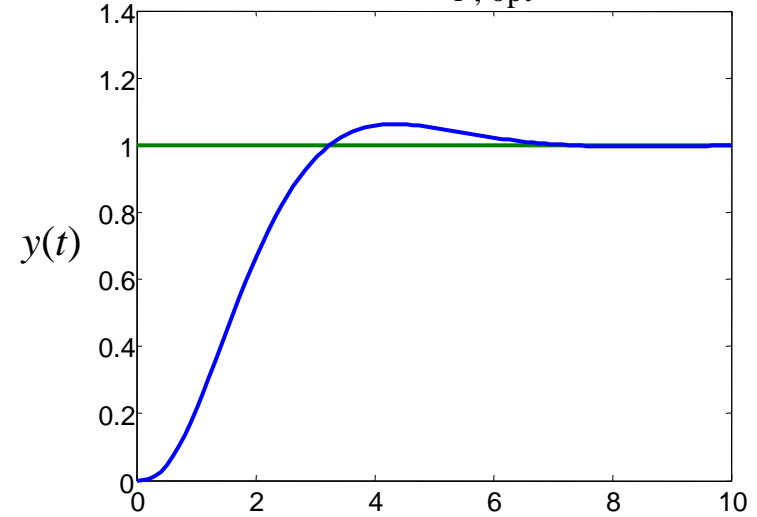
Die absolute Skalierung von α hängt von den absoluten Skalen von y und Δu ab. Eine Normierung dieser Größen kann die Wahl des sinnvollen Bereichs für α erleichtern. Auch ist es zweckmäßig α entsprechend einer logarithmischen Skala zu verändern, da große Bereiche abgedeckt werden müssen.

11.2 Parameteroptimierung

$$\alpha = 7 \rightarrow K_{P, \text{opt}} = 1,2$$



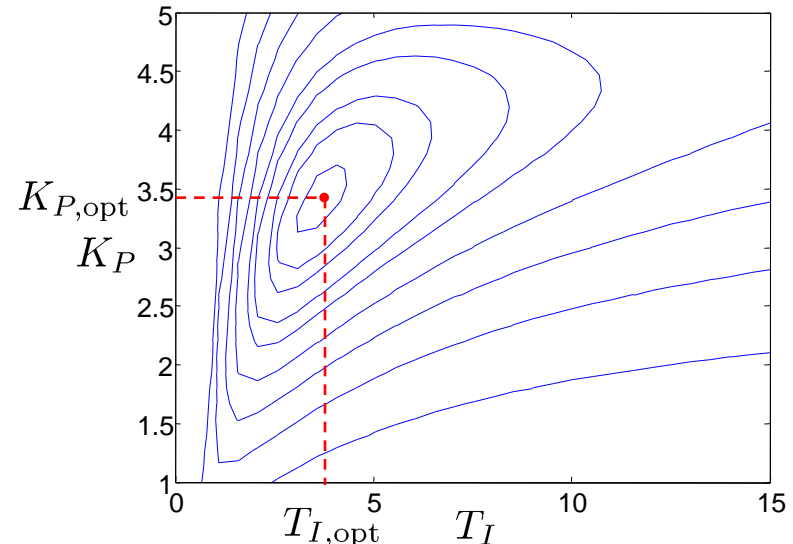
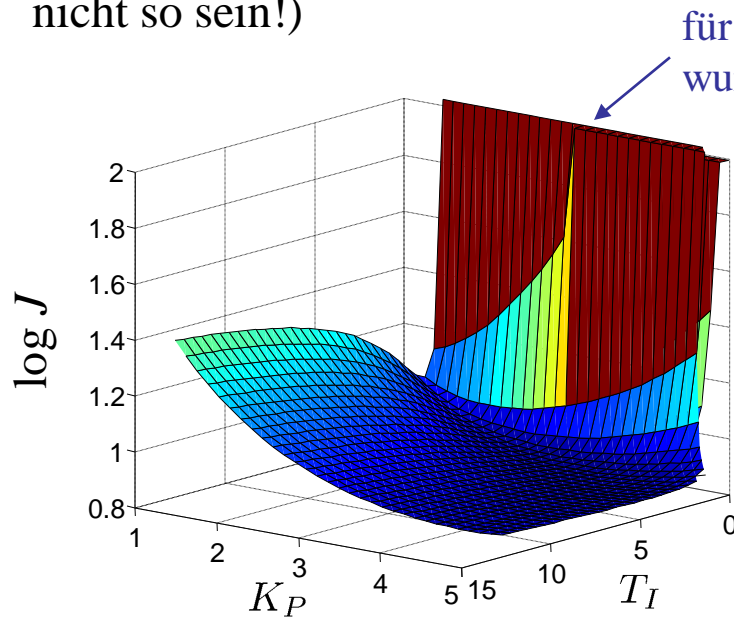
$$\alpha = 144 \rightarrow K_{P, \text{opt}} = 0,65$$



11.2 Parameteroptimierung

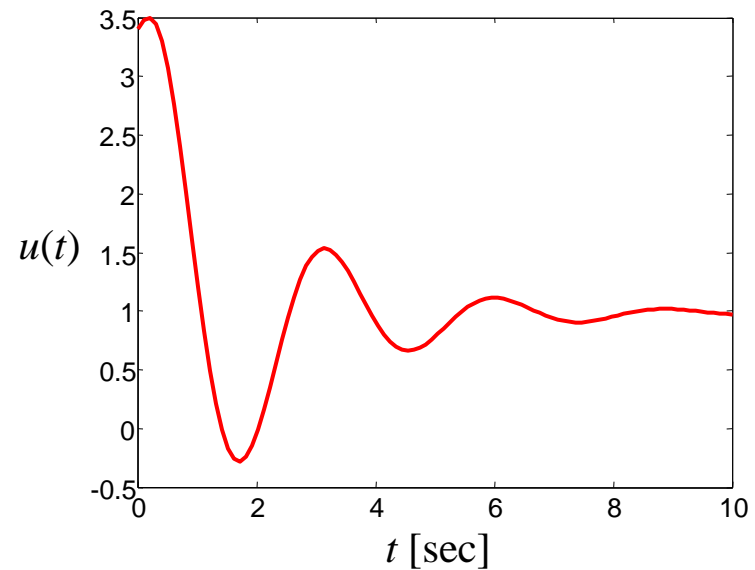
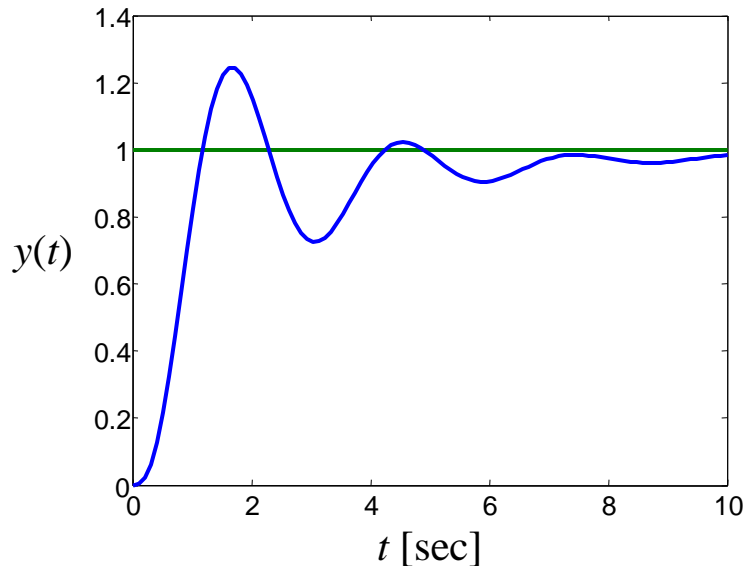
Gleichzeitige Optimierung des P- und I-Anteils des PI-Reglers

- Verlustfunktion wurde für instabile Fälle auf 100 begrenzt.
- Stark unterschiedlicher P- und I-Anteil im Vergleich zur vorherigen Strategie bestehend aus Faustregel für I-Anteil und Optimierung des P-Anteils.
- Trotzdem relativ ähnlicher zeitlicher Verlauf.
- Verlustfunktion stark nichtlinear aber mit nur einem eindeutigen Minimum. (Das muss nicht so sein!)



11.2 Parameteroptimierung

Für $K_{P, \text{opt}}$ und $T_{I, \text{opt}}$



Simulationszeit von 0 bis

10 sec / ∞

↓ ↓

$J = 9,3$ / ∞ (bleibende Regelabweichung)

$J = 10,6$ / $10,7$

$J = 7,4$ / $7,6$

Zusammenfassung:

- Optimaler P-Regler:
- PI-Regler mit I-Anteil nach Faustregel:
- Optimaler PI-Regler:

11.3 Einstellregeln nach Ziegler-Nichols

Wenn ein (explizites) Modell der Regelstrecke nicht vorliegt bzw. dessen Entwicklung unverhältnismäßigen Aufwand bedeuten würde oder wenn die Rechnerunterstützung zur Optimierung nicht kurzfristig verfügbar ist, dann können heuristische Einstellregeln helfen einen Standardregler (P, PI, PID) mit minimalem Aufwand sinnvoll einzustellen, um ein ineffektives langwieriges Ausprobieren zu vermeiden bzw. zu verkürzen.

Vorteile von Einstellregeln

- Ein (explizites) Modell wird nicht benötigt.
- Schnell und ohne Rechneinsatz verwendbar.

Nachteile von Einstellregeln

- Nur für die Standardreglerstrukturen (P, PI, PID) zu gebrauchen.
- Regelkreis muss stabil sein und Experimente (Sprungantwort, Schwingversuch) erlauben.
- Nur suboptimal.
- Bestimmte dynamische Anforderungen können nicht garantiert werden.

11.3 Einstellregeln nach Ziegler-Nichols

Die bekanntesten Einstellregeln kommen von *Ziegler* und *Nichols*. Es existieren 2 alternative Varianten:

Einstellregel 1 nach Ziegler-Nichols

Voraussetzungen: Die Regelstrecke ist stabil und die Sprungantwort kann gemessen werden und weist näherungsweise aperiodisches (d.h. nicht oder nur sehr schwach oszillatorisches) Verhalten auf.

1. Bestimme die Sprungantwort der Regelstrecke
2. Die Sprungantwort wird durch ein PT_1 - T_t -Glied approximiert (siehe nächste Seite).
3. Die Reglerparameter werden entsprechend der unten stehenden Tabelle gewählt.

$$\text{P-Regler: } K_P = \frac{1}{K} \frac{T}{T_t}$$

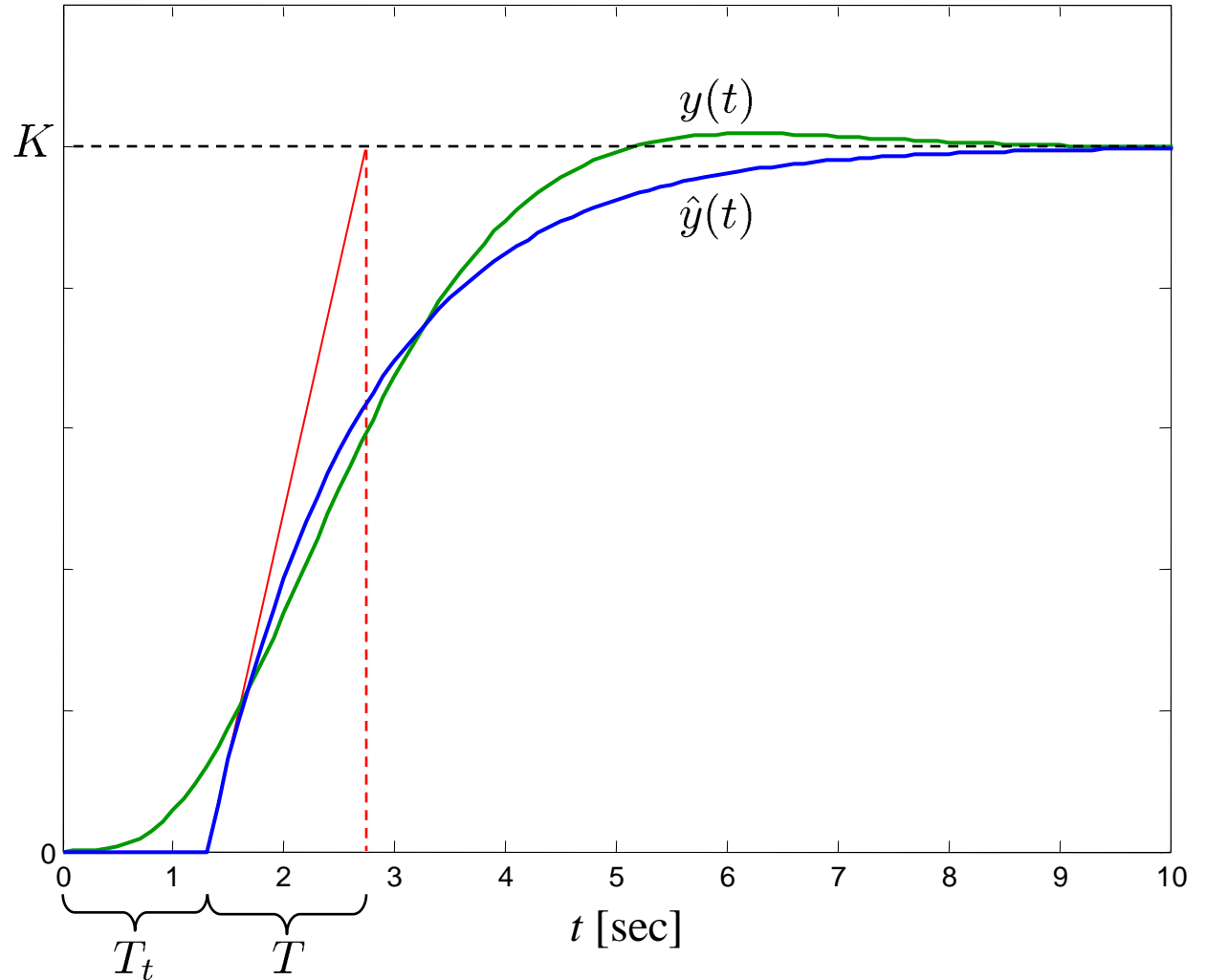
$$\text{PI-Regler: } K_P = \frac{0.9}{K} \frac{T}{T_t} \quad T_I = 3.33T_t$$

$$\text{PID-Regler: } K_P = \frac{1.2}{K} \frac{T}{T_t} \quad T_I = 2T_t \quad T_D = 0.5T_t$$

11.3 Einstellregeln nach Ziegler-Nichols

Die Sprungantwort muss durch ein PT_1T_t -Glied approximiert werden:

$$G_S(s) = \frac{K}{1 + Ts} e^{-T_t s}$$



11.3 Einstellregeln nach Ziegler-Nichols

Einstellregel 2 nach Ziegler-Nichols

Voraussetzungen: Die Regelstrecke ist stabil und darf in einem Dauerschwingversuch an ihre Stabilitätsgrenze gebracht werden (das ist oft nicht durchführbar).

1. Die Regelstrecke wird zunächst mit einem reinen P-Regler geregelt.
2. Die Verstärkung des P-Reglers wird im Experiment so lange erhöht bis die Stabilitätsgrenze erreicht ist, d.h. bis die Regelgröße eine Dauerschwingung konstanter Amplitude ausführt (Ortskurve von $G_0(i\omega)$ geht durch -1).
3. Die Verstärkung K_{krit} an der Stabilitätsgrenze und die Periodendauer T_{krit} der Dauerschwingung sind die Basis für die Bestimmung der Reglerparameter mittels der unten stehenden Tabelle.

P-Regler: $K_P = 0.5K_{\text{krit}}$ → erzeugt einen Amplitudenrand von $k_R = 2$

PI-Regler: $K_P = 0.45K_{\text{krit}}$ $T_I = 0.83T_{\text{krit}}$

PID-Regler: $K_P = 0.6K_{\text{krit}}$ $T_I = 0.5T_{\text{krit}}$ $T_D = 0.125T_{\text{krit}}$

11.4 Einstellregeln nach Summenzeitkonstante

Einstellregel mittels Summenzeitkonstante

Voraussetzungen: Die Regelstrecke ist stabil und die Sprungantwort kann gemessen werden und weist näherungsweise aperiodisches Verhalten auf.

1. Messung einer Sprungantwort der Regelstrecke.
2. Bestimmung der Summenzeitkonstante T_{Σ} und der Streckenverstärkung K aus der Sprungantwort (siehe nächste Folie).
3. Die Reglerparameter werden entsprechend der unten stehenden Tabelle gewählt.

PI-Regler: $K_P = \frac{0.5}{K} \quad T_I = 0.5T_{\Sigma}$

PID-Regler: $K_P = \frac{1}{K} \quad T_I = 0.667T_{\Sigma} \quad T_D = 0.167T_{\Sigma}$

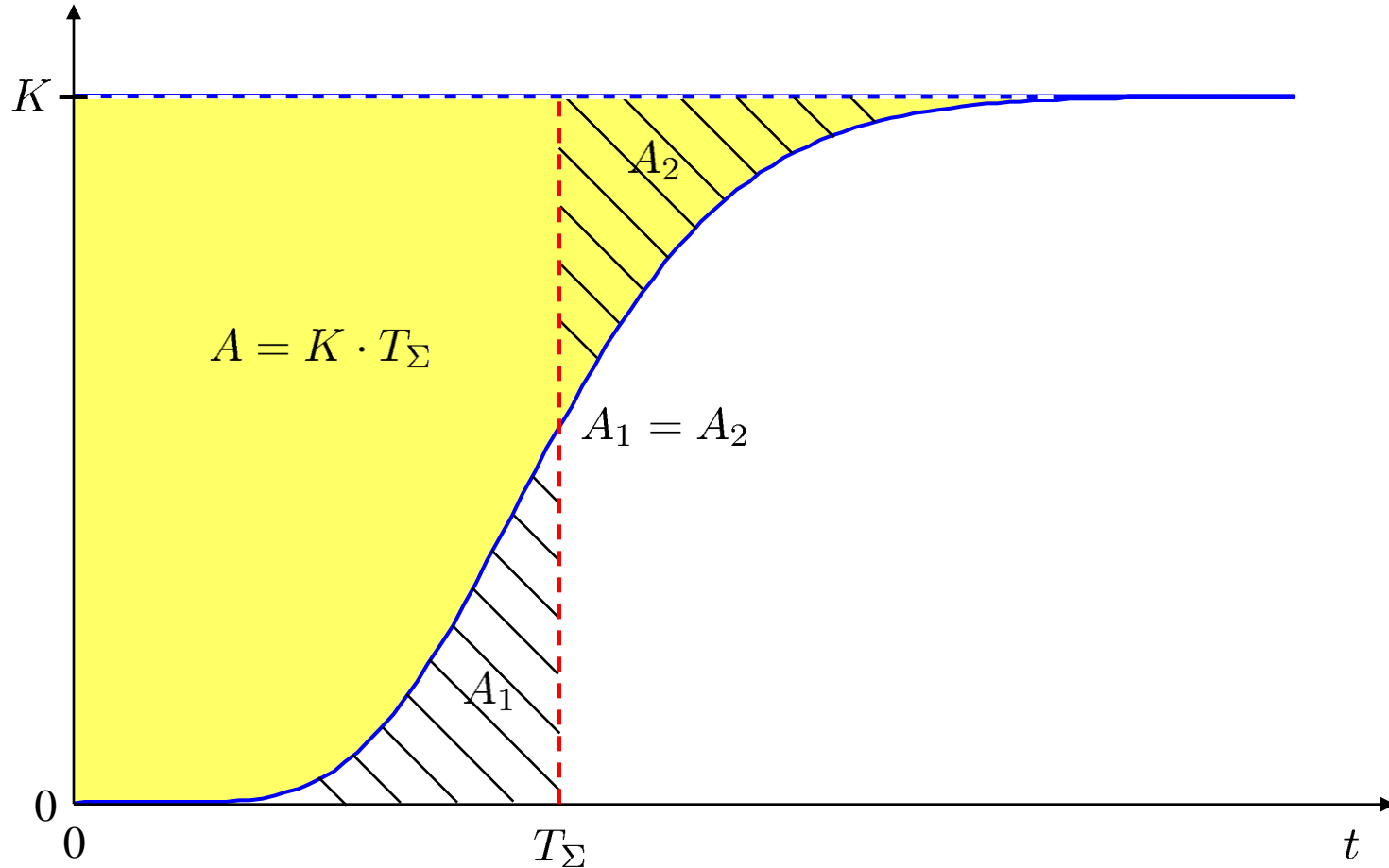
Bemerkung: Falls ein Modell der Strecke bekannt ist, dann kann daraus die Summenzeitkonstante berechnet werden (Bed.: alle $T_{Vi} < \max T_i$, damit kein Überschwingen):

$$G_S(s) = K \frac{(1 + T_{V1}s) \cdot \dots \cdot (1 + T_{Vm}s)}{(1 + T_1s) \cdot \dots \cdot (1 + T_ns)} e^{T_t s} \quad \rightarrow \quad T_{\Sigma} = T_t + \sum_{i=1}^n T_i - \sum_{i=1}^m T_{Vi}$$

11.4 Einstellregeln nach Summenzeitkonstante

Ermittlung der Summenzeitkonstante aus einer Sprungsantwort

- Die Flächen A_1 und A_2 sollen gleich groß sein!



12. Reglerentwurf mittels Kompensation

Inhalt Kapitel 12

12 Reglerentwurf mittels Kompensation

12.1 Steuerung phasenminimaler Systeme

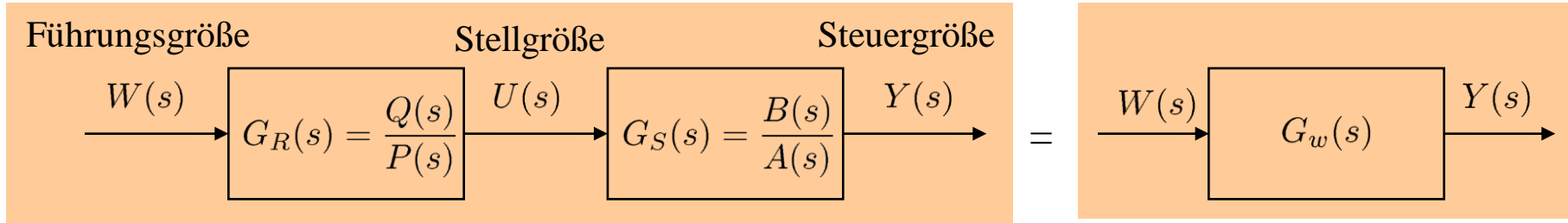
12.2 Steuerung nicht phasenminimaler Systeme

12.3 Kompensationsregler

12.4 Polvorgabe-Regler

12.1 Steuerung phasenminimaler Systeme

Betrachten wir zunächst einmal das einfachere Problem der Steuerung.



Zur Vereinfachung nehmen wir zunächst an, dass $G_S(s)$ stabil und phasenminimal ist, d.h. $B(s)$ und $A(s)$ haben nur stabile Nullstellen. Instabile Strecken können prinzipiell nicht gesteuert werden (sie müssen vorher über eine Rückkopplung, also eine Regelung, stabilisiert werden). Nicht phasenminimale Strecken \rightarrow nächster Abschnitt. Außerdem nehmen wir an, dass evtl. mögliche Pol/Nullstellen-Kürzungen an der Strecke vorgenommen worden sind.

Für die *optimale* Steuerung gilt:

$$y(t) = w(t) \quad \text{bzw.} \quad Y(s) = W(s) \quad \text{bzw.} \quad G_w(s) = \frac{Y(s)}{W(s)} = 1$$

Dies ist erfüllt für:

$$G_R(s)G_S(s) = 1 \quad \rightarrow \quad G_R(s) = G_S^{-1}(s) \quad \rightarrow \quad Q(s) = A(s) \quad \text{und} \quad P(s) = B(s)$$

12.1 Steuerung phasenminimaler Systeme

Die optimale Steuerung ist also die *Inverse* der Strecke (siehe Kapitel 1.4):

$$G_R^{(\text{opt})}(s) = G_S^{-1}(s) = \frac{A(s)}{B(s)}$$

Ist diese optimale Steuerung aber auch realisierbar?

(Fast) alle *realen Strecken* sind nicht sprungfähig, d.h. es gilt:

$$\text{Zählergrad } \deg(B) = m < \text{Nennergrad } \deg(A) = n$$

Für die Inverse müsste also gelten, dass der Zählergrad *größer* als der Nennergrad ist. Eine solche Steuerung ist aber nicht realisierbar (reine D-Anteile)!

Wir können die Realisierbarkeit aber erzwingen, indem wir (wie bei einem realen D-Glied) Verzögerungsglieder mit kleiner Zeitkonstante hinzu multiplizieren:

$$G_R^{(\text{opt, ideal})}(s) = G_S^{-1}(s)$$

$$G_R^{(\text{opt, real})}(s) = G_S^{-1}(s) \cdot \frac{1}{(1 + Ts)^{n-m}}$$

$$\text{Zählergrad} = n \quad \text{Nennergrad} = m$$

$$\text{Zählergrad} = n \quad \text{Nennergrad} = m + n - m = n$$

realisierbar!

12.1 Steuerung phasenminimaler Systeme

Beispiel: Steuerung der Strecke $G_S(s) = \frac{1 + 5s}{(1 + 4s)(1 + 7s)}$

$$m = 1, n = 2$$

$$G_R^{(\text{opt, ideal})}(s) = G_S^{-1}(s) = \frac{(1 + 4s)(1 + 7s)}{1 + 5s}$$

$$G_R^{(\text{opt, real})}(s) = G_S^{-1}(s) \cdot \frac{1}{(1 + Ts)^{n-m}} = \frac{(1 + 4s)(1 + 7s)}{(1 + 5s)(1 + Ts)}$$

Beispiel: Steuerung der Strecke $G_S(s) = \frac{1}{(1 + 4s)(1 + 7s)}$

$$m = 0, n = 2$$

$$G_R^{(\text{opt, ideal})}(s) = G_S^{-1}(s) = (1 + 4s)(1 + 7s)$$

$$G_R^{(\text{opt, real})}(s) = G_S^{-1}(s) \cdot \frac{1}{(1 + Ts)^{n-m}} = \frac{(1 + 4s)(1 + 7s)}{(1 + Ts)^2}$$

Die Wahl von T ist ein Kompromiss zwischen Rauschunempfindlichkeit und Schnelligkeit (Dynamik) der Steuerung!

mit z.B. $T = 0.1$ sec

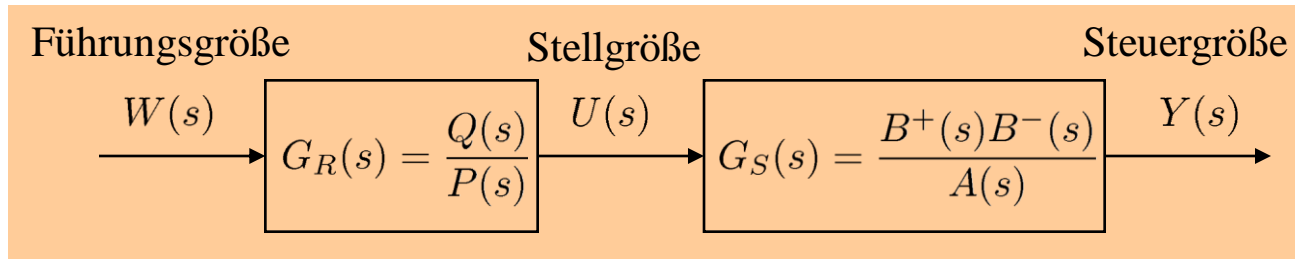
mit z.B. $T = 0.1$ sec

12.2 Steuerung nicht phasenminimaler Systeme

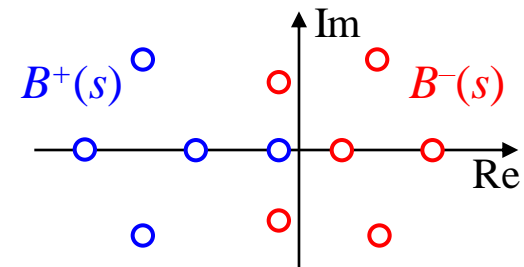
Instabile Nullstellen dürfen nicht weggekürzt werden, weil:

- Strecke und Modell der Strecke nie genau übereinstimmen und somit ein instabiler Pol in der Steuerung entsteht.
- Selbst wenn Strecke und Modell der Strecke exakt übereinstimmen würden, ist die Übertragungsfunktion $G_R(s)$ instabil, was zu unendlich großen Stellgrößen führen kann.

Also müssen stabile und instabile Nullstellen der Strecke offensichtlich unterschiedlich behandelt werden. Dazu spalten wir $B(s)$ in seinen stabilen $B^+(s)$ und seinen instabilen $B^-(s)$ Teil auf:



Bemerkung: In $B^-(s)$ sollten auch sehr schwach gedämpfte (aber noch stabile) Nullstellen einbezogen werden, weil es auch ratsam ist, diese nicht zu kürzen. Denn bei nicht exakter Kürzung können aufgrund der schwachen Dämpfung starke Oszillationen entstehen.



12.2 Steuerung nicht phasenminimaler Systeme

Es gibt im Wesentlichen 3 Strategien zur Behandlung der instabilen Nullstellen in $B^-(s)$:

1. **Ignorieren:** Die instabilen Nullstellen verbleiben unkompensiert im Zähler der Führungsübertragungsfunktion $G_w(s)$. Dies ist die einfachste Methode.
2. **Amplitudengang der Führungsübertragungsfunktion $|G_w(i\omega)| = 1$:** Allerdings erreicht man damit keine perfekte Steuerung, weil die Phase verzerrt wird. Der ideale Amplitudengang $|G_w(i\omega)| = 1$ wird erreicht, wenn jede instabile Nullstelle der Strecke mit ihrem (an der Im-Achse) gespiegelten Gegenpart im *Nenner* der Steuerung "kompensiert" wird.
Jedem Faktor $(1-T_i s)$ in $B^-(s)$ wird also ein Faktor $(1+T_i s)$ im *Nenner* $P(s)$ der Steuerung entgegen gesetzt.
3. **Phasengang der Führungsübertragungsfunktion $\angle G_w(i\omega) = 0$:** Allerdings erreicht man damit keine perfekte Steuerung, weil die Amplitude verzerrt wird. Der ideale Phasengang $\angle G_w(i\omega) = 0$ wird erreicht, wenn jede instabile Nullstelle der Strecke mit ihrem (an der Im-Achse) gespiegelten Gegenpart im *Zähler* der Steuerung "kompensiert" wird.
Jedem Faktor $(1-T_i s)$ in $B^-(s)$ wird also ein Faktor $(1+T_i s)$ im *Zähler* $Q(s)$ der Steuerung entgegen gesetzt.

12.2 Steuerung nicht phasenminimaler Systeme

Beispiel: Steuerung der Strecke $G_S(s) = \frac{1 - 5s}{(1 + 4s)(1 + 7s)}$

Methode 1:

$$G_R^{(\text{opt, real})}(s) = \frac{(1 + 4s)(1 + 7s)}{(1 + Ts)^2} \rightarrow G_w(s) = \frac{1 - 5s}{(1 + Ts)^2}$$

Methode 2:

Allpass
(Amplitudengang = 1)

$$G_R^{(\text{opt, real})}(s) = \frac{(1 + 4s)(1 + 7s)}{(1 + 5s)(1 + Ts)} \rightarrow G_w(s) = \frac{1 - 5s}{(1 + 5s)(1 + Ts)}$$

Methode 3:

Rein reeller Frequenzgang
(Phasengang = 0)

$$G_R^{(\text{opt, real})}(s) = \frac{(1 + 5s)(1 + 4s)(1 + 7s)}{(1 + Ts)^3} \rightarrow G_w(s) = \frac{1 - 25s^2}{(1 + Ts)^3}$$

12.2 Steuerung nicht phasenminimaler Systeme

Methode 2:

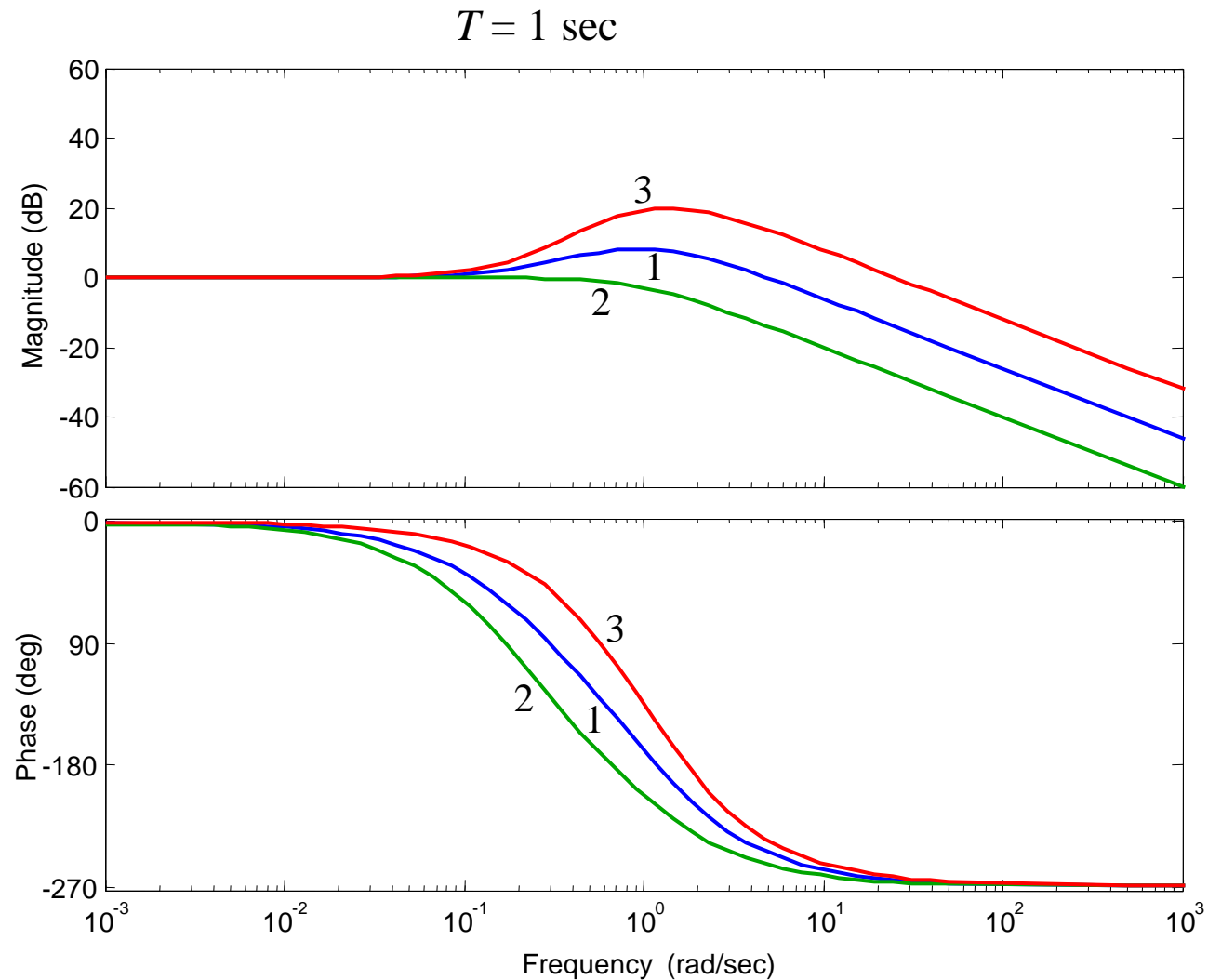
Beste Amplitudenwiedergabe (~ 0 dB)

Methode 3:

Beste Phasewiedergabe ($\sim 0^\circ$)

Methode 1:

Zwischenweg



12.2 Steuerung nicht phasenminimaler Systeme

Methode 2:

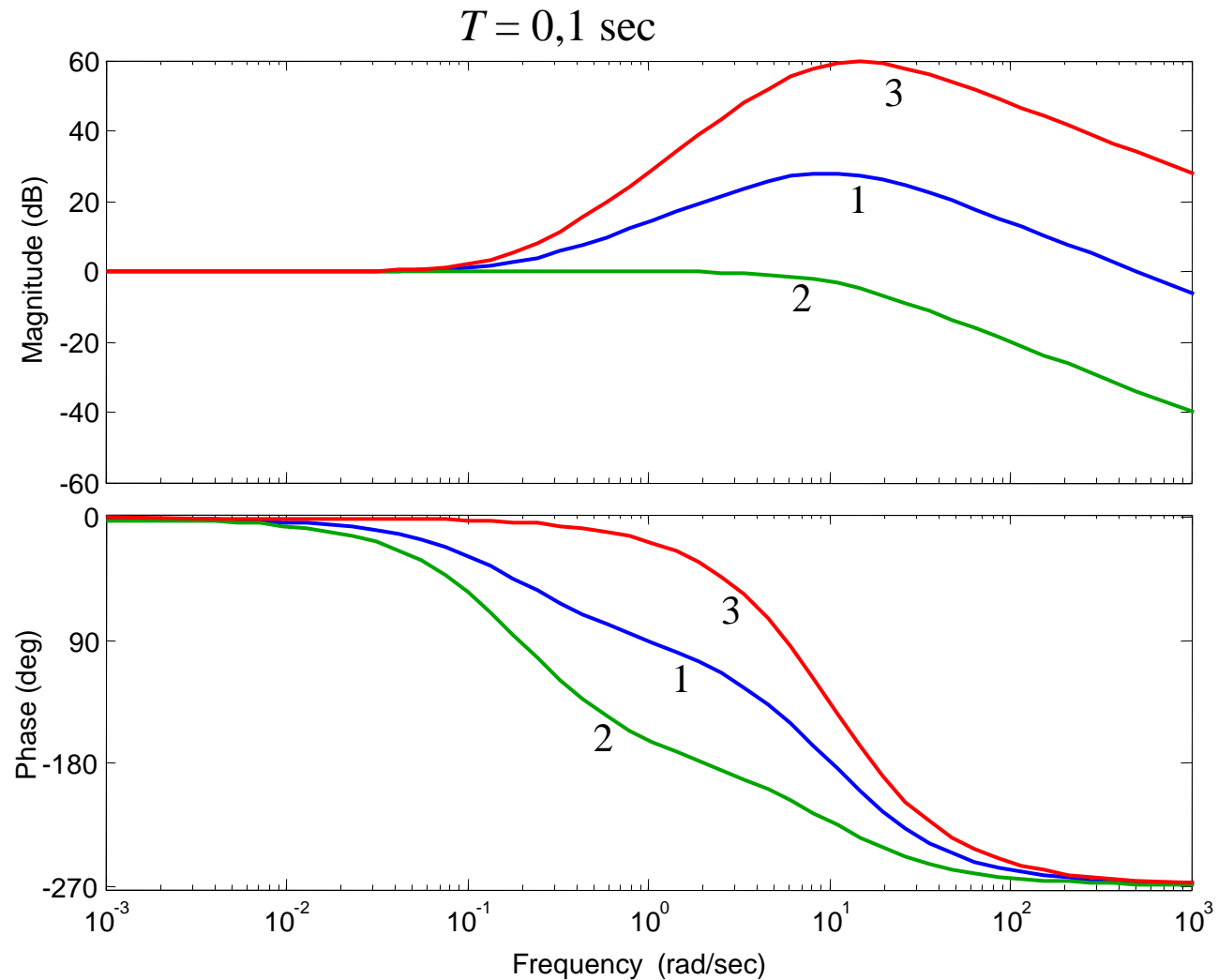
Beste Amplitudenwiedergabe (~ 0 dB)

Methode 3:

Beste Phasewiedergabe ($\sim 0^\circ$)

Methode 1:

Zwischenweg



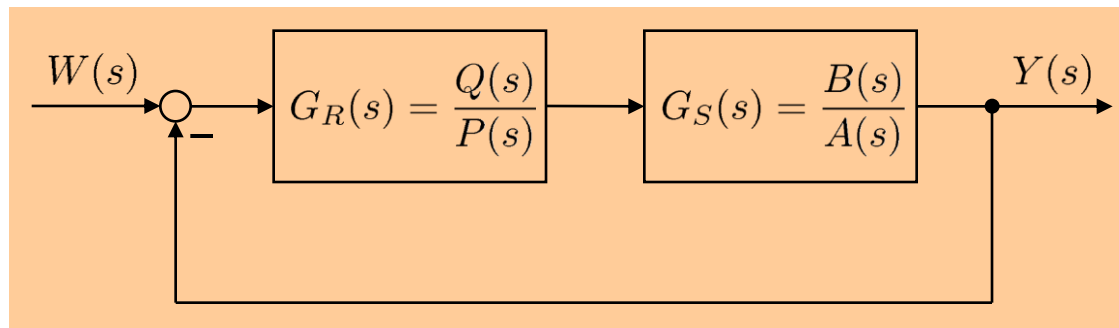
12.3 Kompensationsregler

Vorbemerkung: Nullstellen/Pole der Regelstrecke und des geschlossenen Regelkreises

Betrachten wir die Übertragungsfunktionen der Strecke und des geschlossenen Regelkreises:

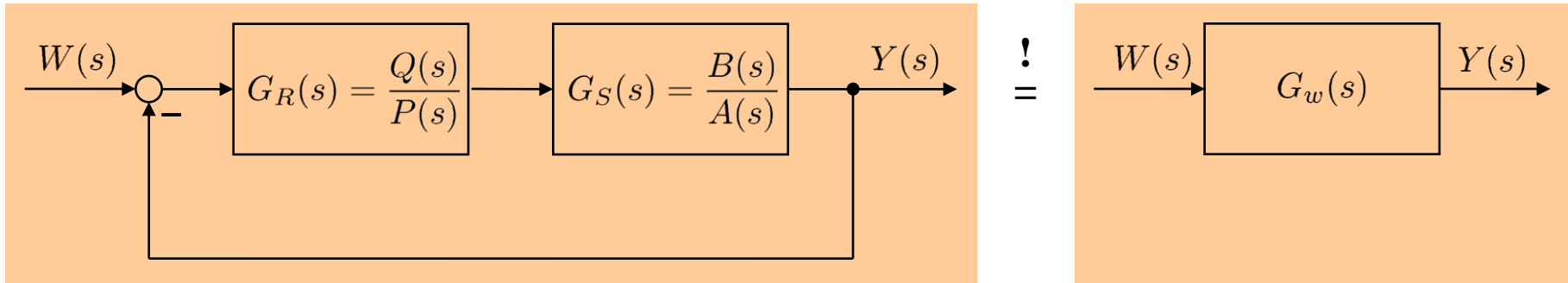
$$G_S(s) = \frac{B(s)}{A(s)} \quad G_w(s) = \frac{G_R(s)G_S(s)}{1 + G_R(s)G_S(s)} = \frac{Q(s)B(s)}{P(s)A(s) + Q(s)B(s)}$$

- Die *Pole* der Strecke können durch die Rückkopplung mittels $Q(s)$ und $P(s)$ verschoben werden.
- Stabile Pole der Strecke können mittels $Q(s)$ gekürzt werden.
- Die *Nullstellen* der Strecke bleiben durch die Rückkopplung erhalten.
- Stabile Nullstellen der Strecke können mittels $P(s)$ gekürzt werden.



12.3 Kompensationsregler

Wir wünschen uns ein bestimmtes Verhalten für die Führungsübertragungsfunktion $G_w(s)$ des geschlossenen Regelkreises. Daraus rechnen wir den Regler aus, der dieses $G_w(s)$ erzeugt.



$$G_w(s) = \frac{G_R(s)G_S(s)}{1 + G_R(s)G_S(s)} \rightarrow G_w(s) + G_w(s)G_R(s)G_S(s) = G_R(s)G_S(s)$$

$$\rightarrow G_w(s) = G_R(s)G_S(s)(1 - G_w(s)) \rightarrow G_R(s) = \frac{G_w(s)}{G_S(s)(1 - G_w(s))}$$

- Die Forderung $G_w(s) = 1$ ist nicht möglich, da sie einen Regler $G_R(s) \rightarrow \infty$ erfordern würde (perfekte Regelung).
- Die Strecke $G_S(s)$ wird invertiert (steht im Nenner), d.h. sie muss *stabil* und *phasenminimal* sein, damit keine instabilen Pol/Nullstellen-Kürzungen auftreten!

12.3 Kompensationsregler

Eine vernünftige Forderung für das Führungsgrößenverhalten ist z.B.: $G_w(s) = \frac{1}{1 + Ts}$.

Daraus ergibt sich folgender (streckenabhängiger) Regler:

$$G_R(s) = \frac{\frac{1}{1+Ts}}{G_S(s)\left(1 - \frac{1}{1+Ts}\right)} = \frac{1}{G_S(s)(1 + Ts - 1)} = \frac{1}{G_S(s)Ts}$$

Beispiel: PT₁-Strecke $G_S(s) = \frac{K}{1 + T_1s}$

$$\rightarrow G_R(s) = \frac{1 + T_1s}{KTs} = \frac{T_1}{KT} + \frac{1}{KTs} \rightarrow \text{PI-Regler}$$

Beispiel: PT₂-Strecke $G_S(s) = \frac{K}{(1 + T_1s)(1 + T_2s)}$

$$\rightarrow G_R(s) = \frac{(1 + T_1s)(1 + T_1s)}{KTs} = \frac{T_1 + T_2}{KT} + \frac{1}{KTs} + \frac{T_1T_2}{KT}s \rightarrow \text{PID-Regler}$$

12.3 Kompensationsregler

Die Forderung nach PT_1 -Führungsverhalten mit einem einzigen freien Parameter T führt automatisch sowohl auf die Reglerstruktur als auch auf alle Reglerparameter!

Die beiden Beispiele zeigen:

- PT_1 -Strecke: Regler mit Zählergrad = Nennergrad
- PT_2 -Strecke: Regler mit Zählergrad = Nennergrad + 1, also nicht mehr realisierbar!
- Für eine PT_3 -Strecke würden wir erhalten: Regler mit Zählergrad = Nennergrad + 2 ...

Um die Realisierbarkeit des Reglers sicherzustellen sollten wir für das Führungsverhalten z.B. fordern:

$$G_w(s) = \frac{1}{(1 + Ts)^{n-m}} \quad \rightarrow \quad G_R(s) = \frac{1}{G_S(s)((1 + Ts)^{n-m} - 1)}$$

Dann ergibt sich immer ein Regler mit Zählergrad = Nennergrad.

Interessant ist auch: Falls $G_S(s)$ keinen I-Anteil hat, ergibt sich immer ein Regler mit I-Anteil. (Die "1" in $(1+Ts)^{n-m}$ fällt gegen die "-1" weg und ein "s" lässt sich ausklammern.) Das muss auch so sein, da das vorgegebene Führungsverhalten keine bleibende Regelabweichung hat.

12.3 Kompensationsregler

Was können wir tun, wenn die Strecke nicht phasenminimal ist?

Aus der Steuerung nicht phasenminimaler Systeme wissen wir, dass man die instabilen Zählernullstellen nicht wegkürzen darf. Man muss sie im Führungsverhalten akzeptieren. Wir können den Kompensationsregler also auf nicht phasenminimale Strecken erweitern, wenn wir fordern, dass alle instabilen (in Praxis auch alle sehr schwach gedämpften) Nullstellen der Strecke im gewünschten Führungsverhalten $G_w(s)$ enthalten sein müssen.

Beispiel: Nicht phasenminimale Strecke $G_S(s) = \frac{1 - T_D s}{(1 + T_1 s)(1 + T_2 s)}$

Wir fordern: $G_w(s) = \frac{1 - T_D s}{(1 + T s)^2}$ ← instabile Nullstelle wird in das „gewünschte“ Verhalten übernommen

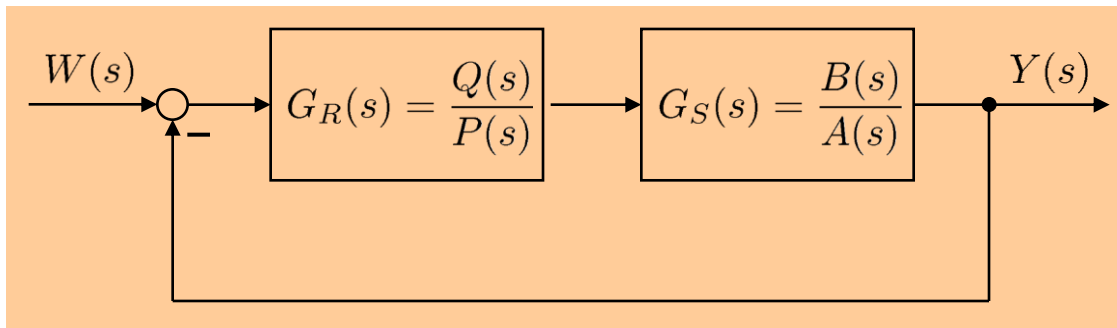
$$G_R(s) = \frac{\frac{1 - T_D s}{(1 + T s)^2}}{\frac{1 - T_D s}{(1 + T_1 s)(1 + T_2 s)} \left(1 - \frac{1 - T_D s}{(1 + T s)^2}\right)} = \frac{(1 + T_1 s)(1 + T_2 s)}{(1 + T s)^2 - (1 - T_D s)} = \frac{(1 + T_1 s)(1 + T_2 s)}{s(2T + T_D + T^2 s)}$$

→ Regler kürzt die instabile Nullstelle nicht!

12.4 Polvorgabe-Regler

Der Kompensationsregler kann nur für stabile Prozesse verwendet werden. Außerdem ist die Forderung, einem beliebig vorgegebenen Führungsverhalten zu folgen, recht "streng". Der Regler wird daher (zumindest in manchen Frequenzbereichen) aggressiv eingreifen müssen und nicht besonders robust ausfallen.

Im Folgenden gehen wir von dieser Maximalforderung einen Schritt zurück und verlangen "bloß", dass der geschlossene Regelkreis vorgegebene *Pole* aufweist. Es wird sich zeigen, dass diese Forderung (prinzipiell) für *jede beliebige* Strecke erfüllbar ist, und zwar unabhängig von der Wahl der gewünschten Pole! Bzgl. der Nullstellen und des sonstigen Führungsverhaltens stellen wir keine Forderungen.



$$G_w(s) = \frac{G_R(s)G_S(s)}{1 + G_R(s)G_S(s)} = \frac{Q(s)B(s)}{P(s)A(s) + Q(s)B(s)}$$

Diophantische Gleichung

Polvorgabe:

$$P(s)A(s) + Q(s)B(s) = N(s)$$

12.4 Polvorgabe-Regler

Beispiel: Phasenminimale Strecke 2. Ordnung $G_S(s) = \frac{1 + 5s}{(1 + 4s)(1 + 7s)}$

Frage: Wie viele Pole können/müssen wir vorgeben?

Charakteristische Gleichung: $P(s)(28s^2 + 11s + 1) + Q(s)(5s + 1) = \dots$

Regler 0. Ordnung: $P(s) = p_0$, $Q(s) = q_0$: 2 Unbekannte, 3 Gleichungen.

Regler 1. Ordnung: $P(s) = p_1s + p_0$, $Q(s) = q_1s + q_0$: 4 Unbekannte, 4 Gleichungen.

Regler 2. Ordnung: $P(s) = p_2s^2 + p_1s + p_0$, $Q(s) = q_2s^2 + q_1s + q_0$: 6 Unbekannte, 5 Gl.

- Ein Regler 0. Ordnung kann die Pole nicht wie gefordert verschieben.
- Ein Regler 1. Ordnung kann die Pole wie gefordert verschieben.
- Ein Regler 2. und höherer Ordnung kann die Pole wie gefordert verschieben und hat zusätzliche Freiheitsgrade, die für andere Zwecke genutzt werden können (z.B. zur Erzwingung eines I-Anteils, um bleibende Regelabweichungen zu vermeiden).

Wählen wir zunächst den Regler minimaler Ordnung um unsere Forderung zu erfüllen...

12.4 Polvorgabe-Regler

Charakteristische Gleichung mit Regler 1. Ordnung:

$$(p_1s + p_0)(28s^2 + 11s + 1) + (q_1s + q_0)(5s + 1) =$$

$$28p_1s^3 + (11p_1 + 28p_0 + 5q_1)s^2 + (p_1 + 11p_0 + q_1 + 5q_0)s + p_0 + q_0 =$$

Wir fordern einen Dreifachpol bei $-0,25$ bzw. PT_3 -Verhalten mit Zeitkonstante 4 sec.:

$$= (1 + 4s)^3 = 64s^3 + 48s^2 + 12s + 1$$

Koeffizientenvergleich ergibt:

$$s^3 : 28p_1 = 64$$

$$s^1 : p_1 + 11p_0 + q_1 + 5q_0 = 12$$

$$s^2 : 11p_1 + 28p_0 + 5q_1 = 48$$

$$s^0 : p_0 + q_0 = 1$$

In Matrix-Vektor-Form schreibt sich das:

$$\begin{pmatrix} 28 & 0 & 0 & 0 \\ 11 & 28 & 5 & 0 \\ 1 & 11 & 1 & 5 \\ 0 & 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} p_1 \\ p_0 \\ q_1 \\ q_0 \end{pmatrix} = \begin{pmatrix} 64 \\ 48 \\ 12 \\ 1 \end{pmatrix} \rightarrow \begin{pmatrix} p_1 \\ p_0 \\ q_1 \\ q_0 \end{pmatrix} = \begin{pmatrix} 28 & 0 & 0 & 0 \\ 11 & 28 & 5 & 0 \\ 1 & 11 & 1 & 5 \\ 0 & 1 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 64 \\ 48 \\ 12 \\ 1 \end{pmatrix}$$

12.4 Polvorgabe-Regler

Damit ergeben sich die Reglerparameter:

$$\begin{pmatrix} p_1 \\ p_0 \\ q_1 \\ q_0 \end{pmatrix} = \begin{pmatrix} 2,286 \\ 0,357 \\ 2,571 \\ 0,643 \end{pmatrix}$$

Und somit der Regler und die Reglerverstärkung: → PDT₁-Regler

$$G_R(s) = \frac{2,571s + 0,643}{2,286s + 0,357} = 0,643 \frac{4s + 1}{2,286s + 0,357} \quad K_R = G_R(0) = 1,8$$

Daraus berechnet sich die Führungsübertragungsfunktion zu:

$$G_w(s) = 0,643 \frac{(1 + 4s)(1 + 5s)}{(1 + 4s)^3} = 0,643 \frac{(1 + 5s)}{(1 + 4s)^2}$$

← vorgegebene Pole!

Auf einen Führungsgrößenprung der Höhe 1 haben wir also eine bleibende Regelabweichung von $1 - 0,643 = 0,357$. Es wurden ja auch nur die Pole vorgegeben und nichts für den Zähler bzw. die Verstärkung gefordert!

12.4 Polvorgabe-Regler

Beispiel: Nicht phasenminimale Strecke 2. Ordnung $G_S(s) = \frac{1 - 5s}{(1 + 4s)(1 + 7s)}$

Mit den selben Forderungen wie im vorherigen Beispiel führt dies auf ein ähnliches Gleichungssystem. Anstelle der "+5" tritt aber eine "-5":

$$\begin{pmatrix} p_1 \\ p_0 \\ q_1 \\ q_0 \end{pmatrix} = \begin{pmatrix} 28 & 0 & 0 & 0 \\ 11 & 28 & -5 & 0 \\ 1 & 11 & 1 & -5 \\ 0 & 1 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 64 \\ 48 \\ 12 \\ 1 \end{pmatrix}$$

Der Regler und die Reglerverstärkung berechnen sich zu: → PDT₁-Regler

$$G_R(s) = \frac{0,429s + 0,107}{2,286s + 0,893} \quad K_R = G_R(0) = 0,12$$

Dies ergibt folgende Führungsübertragungsfunktion:

$$G_w(s) = 0,107 \frac{(1 + 4s)(1 - 5s)}{(1 + 4s)^3} = 0,107 \frac{(1 - 5s)}{(1 + 4s)^2}$$

Sehr kleine Verstärkung!
(Regler muss bei nicht phasenminimaler Strecke vorsichtig sein. Stabilitätsgrenze ist wegen zusätzlicher Phase näher!)

Riesige bleibende Regelabweichung auf Einheitssprung von $1 - 0,107 = 0,893$.

12.4 Polvorgabe-Regler

Wie können wir Regler *höherer* Ordnung mit Polvorgabe entwerfen?

Beispiel: Strecke 2. Ordnung $G_S(s) = \frac{b_2s^2 + b_1s + b_0}{s^2 + a_1s + a_0}$

Charakteristische Gleichung: $P(s)(s^2 + a_1s + a_0) + Q(s)(b_2s^2 + b_1s + b_0) = \dots$

Regler 1. Ordnung kann alle Pole beliebig verschieben. Regler 2. Ordnung bietet einen überschüssigen Freiheitsgrad (5 Gleichungen für 6 unbekannte Parameter):

$$(p_2s^2 + p_1s + p_0)(s^2 + a_1s + a_0) + (q_2s^2 + q_1s + q_0)(b_2s^2 + b_1s + b_0) = \dots$$

Den freien Parameter können wir z.B. sinnvoll nutzen, um im Regler einen *I-Anteil* zu erzwingen, indem wir $p_0 = 0$ setzen. Dann lautet die Übertragungsfunktion des Reglers:

$$G_R(s) = \frac{Q(s)}{P(s)} = \frac{q_2s^2 + q_1s + q_0}{p_2s^2 + p_1s} = \frac{q_2s + q_1 + \frac{q_0}{s}}{p_2s + p_1} \rightarrow \text{PIDT}_1\text{-Regler}$$

Nun haben wir ein Problem mit 5 unbekanntem Parametern zu deren Bestimmung 5 Gleichungen zur Verfügung stehen. Das ist analog zu den vorherigen Beispielen lösbar.

12.4 Polvorgabe-Regler

Charakteristische Gleichung im allgemeinen Fall:

$$P(s)A(s) + Q(s)B(s) = N(s)$$

Ordnung: $r \quad n \quad r \quad m \quad n+r$ (da $n \geq m$)

Wie groß muss die Ordnung r des Reglers gewählt werden, damit das Polvorgabeproblem exakt gelöst werden kann?

- Anzahl der Gleichungen: $n + r + 1$ (höchste Polynomordnung + 1 für Gleichwert)
- Anzahl der unbekannt Parameter: $2(r + 1)$ (je $r + 1$ für Zähler und Nenner des Reglers)

$$\rightarrow n + r + 1 = 2r + 2 \quad \rightarrow \boxed{r = n - 1}$$

Ausgeschrieben ergibt sich dann also folgende charakteristische Gleichung:

$$\begin{aligned} & (p_{n-1}s^{n-1} + \dots + p_1s + p_0)(s^n + a_{n-1}s^{n-1} + \dots + a_1s + a_0) \\ & + (q_{n-1}s^{n-1} + \dots + q_1s + q_0)(b_n s^n + b_{n-1}s^{n-1} + \dots + b_1s + b_0) \\ = N(s) & = n_{2n-1}s^{2n-1} + \dots + n_1s + n_0 \end{aligned}$$

$= 0$, wenn $m < n - 1$
 $= 0$, wenn $m < n$

12.4 Polvorgabe-Regler

In Matrix-Vektor-Form lassen sich die $2n$ Gleichungen, die sich aus dem Koeffizientenvergleich ergeben wie folgt schreiben:

$$\underline{M} \cdot \underline{r} = \underline{n} \quad \rightarrow \quad \underline{r} = \underline{M}^{-1} \cdot \underline{n}$$

enthält die
enthält die
enthält die Koeffizienten des
Streckenparameter
Reglerparameter
Polynoms aus der Polvorgabe

$$\begin{matrix}
 \left. \begin{matrix} n \\ n \end{matrix} \right\} \\
 \left(\begin{array}{cccccc}
 a_n & & & & & b_n \\
 a_{n-1} & a_n & & & & b_{n-1} & b_n \\
 \vdots & & \ddots & & & \vdots & & \ddots \\
 a_1 & & & a_n & & b_1 & & b_n \\
 a_0 & a_1 & & a_{n-1} & b_0 & b_1 & & b_{n-1} \\
 & a_0 & & \vdots & & b_0 & & \vdots \\
 & & \ddots & \vdots & & & \ddots & \vdots \\
 & & & a_0 & & & & b_0
 \end{array} \right) \cdot \begin{pmatrix} p_{n-1} \\ \vdots \\ p_0 \\ q_{n-1} \\ \vdots \\ q_0 \end{pmatrix} = \begin{pmatrix} n_{2n-1} \\ n_{2n-2} \\ \vdots \\ \vdots \\ n_1 \\ n_0 \end{pmatrix}
 \end{matrix}$$

n
 n

12.4 Polvorgabe-Regler

Die Matrix \underline{M} nennt man *Sylvester*-Matrix. Sie spielt in vielen Bereichen der angewandten Mathematik eine große Rolle. Für uns sind folgende Eigenschaften wichtig:

- \underline{M} ist invertierbar, wenn die Polynome $A(s)$ und $B(s)$ keine gemeinsamen Nullstellen aufweisen. Dies lässt sich immer durch evtl. Pol/Nullstellen-Kürzungen in $B(s)/A(s)$ erreichen. Damit lassen sich die Reglerparameter eindeutig bestimmen.
- Wir normieren typischerweise $a_n = 1$.
- Fast alle realen Prozesse sind nicht sprungfähig; dann gilt $m < n$ (nicht nur $m \leq n$) und somit ist $b_n = 0$.

Wie sollen die gewünschten Pole eigentlich gewählt werden?

Diese Frage wird im nächsten Kapitel ausführlich behandelt. Die sinnvolle Wahl der gewünschten Pole ist schwierig. Man sollte auf jeden Fall aufpassen keine zu scharfen Forderungen an den Regelkreis zu stellen, wie z.B. sehr schnelles und gleichzeitig überschwingfreies Einschwingen. Denn obwohl alle Polkombinationen prinzipiell realisierbar sind, können zu scharfe Forderungen auf sehr unrobustes Verhalten in Bezug auf Modellungenauigkeiten und zu sehr aggressiven Stelleingriffen führen!

12.4 Polvorgabe-Regler

Bemerkung: Wir sind beim Polvorgaberegler stets davon ausgegangen, dass wir einen Regler mit Zählergrad = Nennergrad haben und dass alle Parameter unabhängig voneinander bestimmt werden können bzw. müssen. Diese Vorgehensweise ist deswegen richtig, weil wir auch keine Einschränkungen bei dem Polvorgabepolynom gemacht haben. So können wir die Forderung nach einem Dreifachpol bei $s = -2$ auf die verschiedensten Arten ausdrücken, z.B.:

$$N(s) = (s + 2)^3 \quad \text{oder} \quad N(s) = (1 + 0,5s)^3 \quad \text{oder} \quad \dots$$

Würden wir uns auf eine Form festlegen, z.B. auf den Typ $N(s) = (s+s_1) \cdot \dots \cdot (s+s_{2n-1})$, dann wäre automatisch der Koeffizient vor der höchsten Potenz s^{2n-1} gleich 1. Dann müssten wir auch $p_{n-1} = 1$ festlegen (und wie zuvor beschrieben: $a_n = 1, b_n = 0$), damit das Gleichungssystem noch lösbar ist. Die erste Gleichung lautet dann " $p_{n-1} = 1$ ", was keine neue Information bringt.

13. Wurzelortskurve (WOK)

Inhalt Kapitel 13

13 Wurzelortskurve (WOK)

13.1 Beziehung zwischen Polen und Zeitverhalten

13.2 Konstruktion der WOK

13.3 Reglerentwurf mittels WOK

13.1 Beziehung zwischen Polen und Zeitverhalten

Viele Entwurfsverfahren geben eine Wunschlage für die Pole des geschlossenen Regelkreises vor. **Welche Polvorgabe ist sinnvoll?** Zunächst einige quantitative Überlegungen:

- Alle Pole müssen stabil sein, also einen negativen Realteil aufweisen.
- Pole sehr nahe der imaginären Achse führen auf (zu) langsames Verhalten.
- Pole sehr weit entfernt von der imaginären Achse führen auf (zu) aggressives Stellverhalten und eine (zu) geringe Robustheit gegenüber Modellfehlern wegen der Anregung hoher Frequenzen.
- Pole mit großem Imaginärteil (im Vergleich zu ihrem Realteil) führen auf schwach gedämpftes Verhalten, d.h. starkes Überschwingen bzw. Oszillationen.

Bemerkung: Wenn in der Folge von schnellen und langsamen Polen bzw. Zeitkonstanten die Rede ist, dann verstehen sich diese Ausdrücke immer relativ zu den Polen bzw. Zeitkonstanten der Regelstrecke.

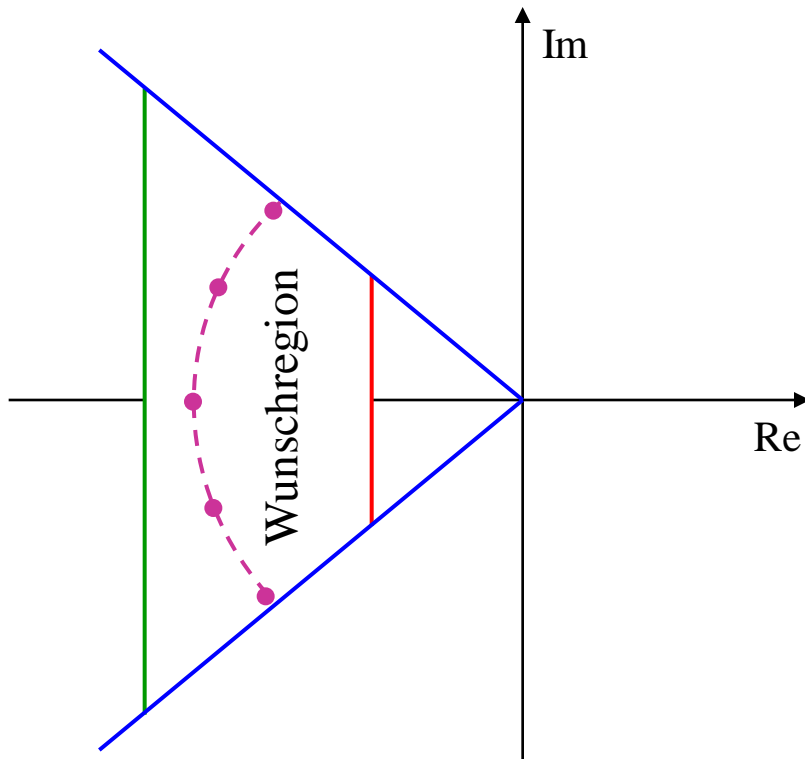
Für einen Kläranlagenprozess kann 10 min eine schnelle Zeitkonstante sein; für ein Magnetlager kann 0,1 sec eine langsame Zeitkonstante sein.

Daher sind die Re- und Im-Achsen in der Folge auch nicht absolut skaliert!

13.1 Beziehung zwischen Polen und Zeitverhalten

Die Überlegungen der vorherigen Folie legen folgendes Zielgebiet für die Wunschkpole nahe:

- Nicht zu langsam. → Links von der roten (weiter rechts liegenden senkrechten) Linie.
- Nicht zu schnell. → Rechts von der grünen (weiter links liegenden senkrechten) Linie.
- Keine zu niedrige Dämpfung. → Zwischen den blauen schrägen Linien.

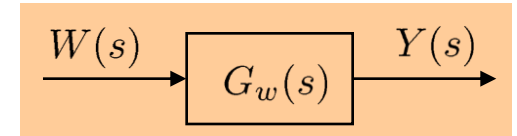


Bemerkung: Die Erfahrung zeigt, dass es ungünstig ist, die gewünschten Pole dicht beieinander zu platzieren. Es wird empfohlen, die Wunschkpole z.B. in gleichmäßigen Abständen auf einem Kreis um den Ursprung des Koordinatensystems innerhalb der Wunschregion zu legen (gestrichelt violett).

13.1 Beziehung zwischen Polen und Zeitverhalten

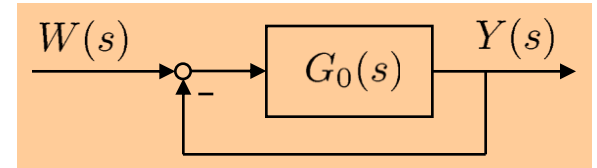
Zur weiteren Analyse nehmen wir vereinfachend an, dass sich der *geschlossene* Regelkreis durch ein schwingungsfähiges System 2. Ordnung mit Verstärkung $K_w=1$ beschreiben lässt:

$$G_w(s) = \frac{\omega_0^2}{s^2 + 2D\omega_0s + \omega_0^2}$$



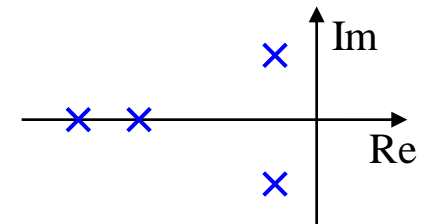
Dies ist z.B. dann der Fall, wenn der *offene* Regelkreis IT₁-Verhalten aufweist:

$$G_w(s) = \frac{G_0(s)}{1 + G_0(s)} \quad \rightarrow \quad G_0(s) = \frac{\omega_0^2}{s(s + 2D\omega_0)}$$



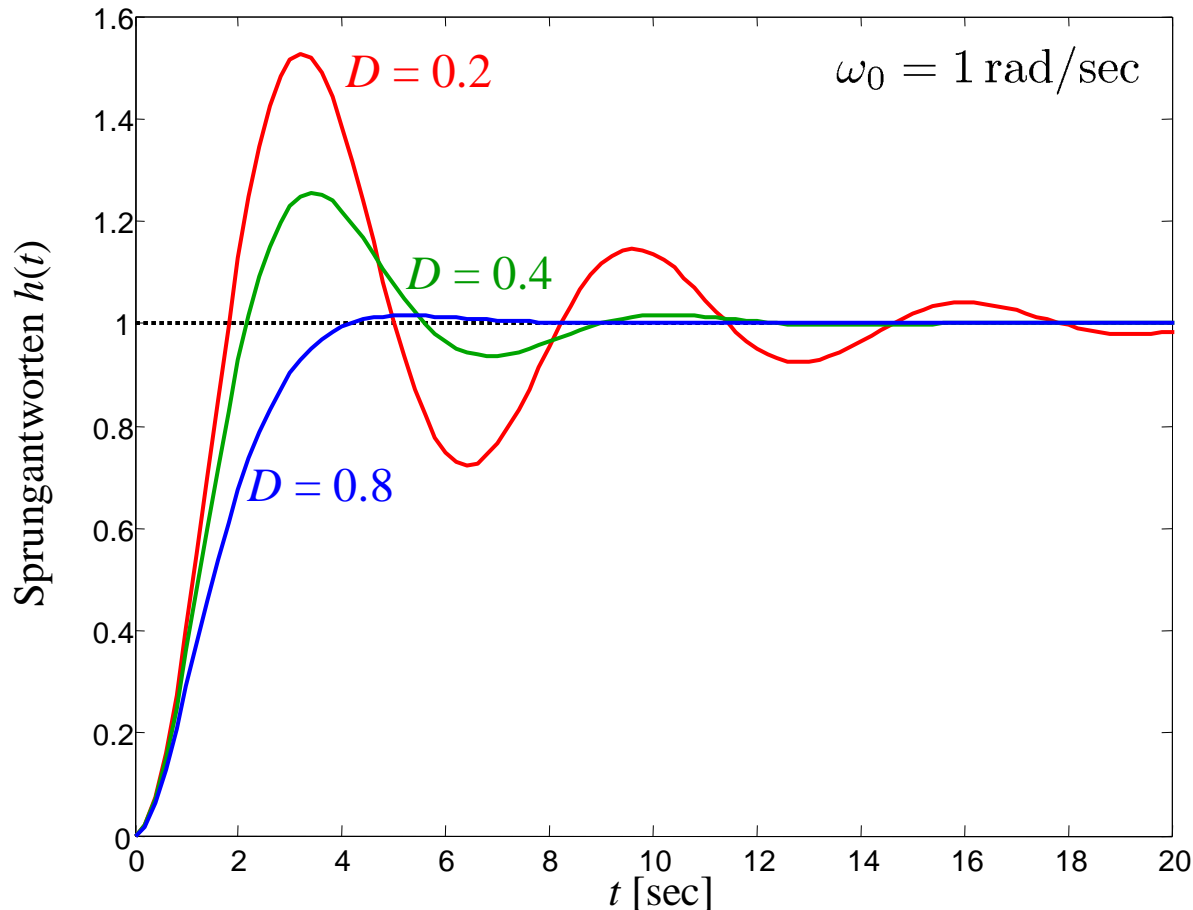
Ein solches IT₁-Verhalten kommt häufig vor. Es kann sich z.B. aus einem P-Regler mit IT₁-Strecke oder einem I-Regler mit PT₁-Strecke oder einem PI-Regler mit PT₂-Strecke (mit einer Pol/Nullstellen-Kürzung) ergeben.

Auch viele Systeme höherer Ordnung können so approximiert werden, wenn das konjugiert komplexe Polpaar in $G_w(s)$ *dominant* ist, d.h. weit rechts im Vergleich zu den restlichen Polen des geschlossenen Regelkreises liegt (also sehr langsam ist):

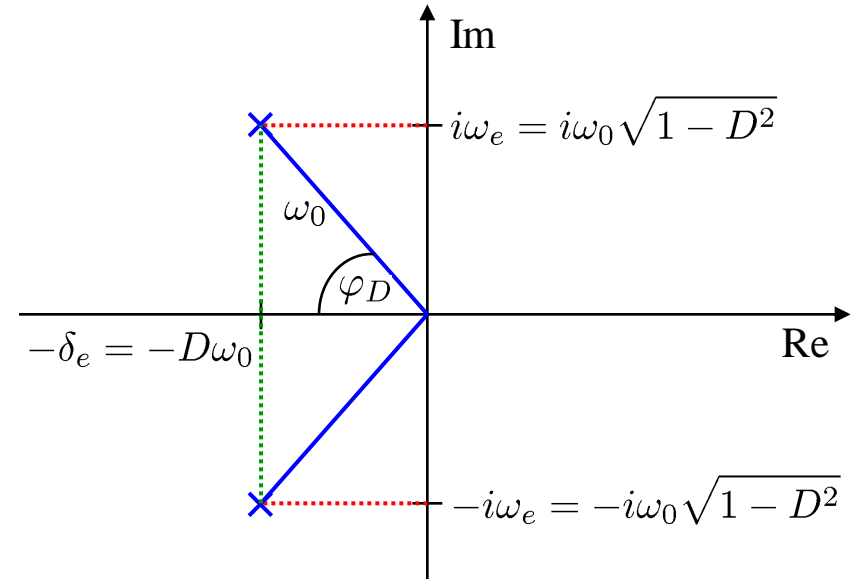
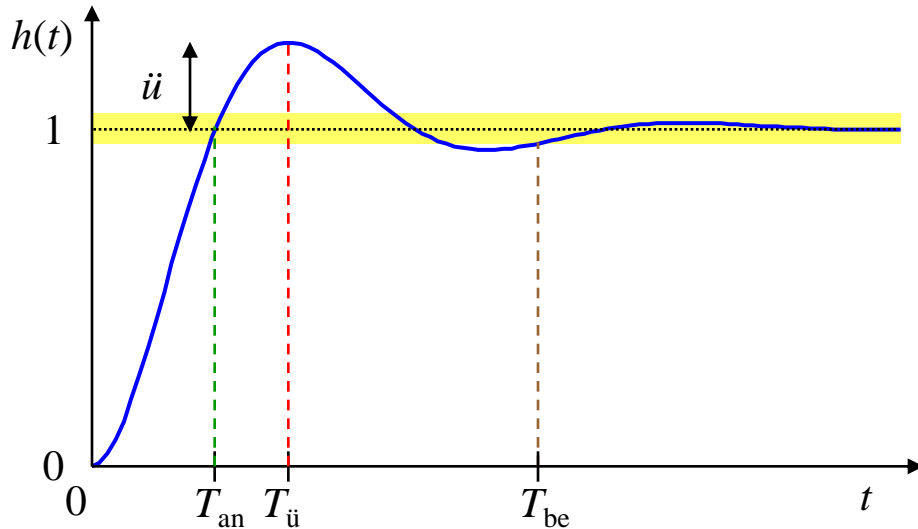


13.1 Beziehung zwischen Polen und Zeitverhalten

Nun untersuchen wir, wie charakteristische Werte der Sprungantwort mit der Lage der Pole zusammenhängen. Dann können wir von Forderungen im Zeitbereich auf Forderungen im Frequenzbereich (Pollage) umrechnen und somit eine sinnvolle Polvorgabe machen.



13.1 Beziehung zwischen Polen und Zeitverhalten



$$h(t) = 1 - \frac{1}{\sqrt{1-D^2}} e^{-D\omega_0 t} \sin(\omega_e t + \varphi_D)$$

$$\ddot{u} = e^{-\frac{\pi D}{\sqrt{1-D^2}}} = e^{-\frac{\delta_e}{\omega_e} \pi}$$

$$T_{\ddot{u}} = \frac{\pi}{\omega_e}$$

$$T_{an} = \frac{\pi - \varphi_D}{\omega_e}$$

$$T_{be}(\pm 5\%) = \frac{\ln(0.05\sqrt{1-D^2})}{-\delta_e}$$

Pole bei $s_{1/2} = -\delta_e \pm i\omega_e$

Re-Teil: $-\delta_e = -D\omega_0$

Im-Teil: $\pm\omega_e = \pm\omega_0\sqrt{1-D^2}$

Dämpfung: $D = \cos\varphi_D$

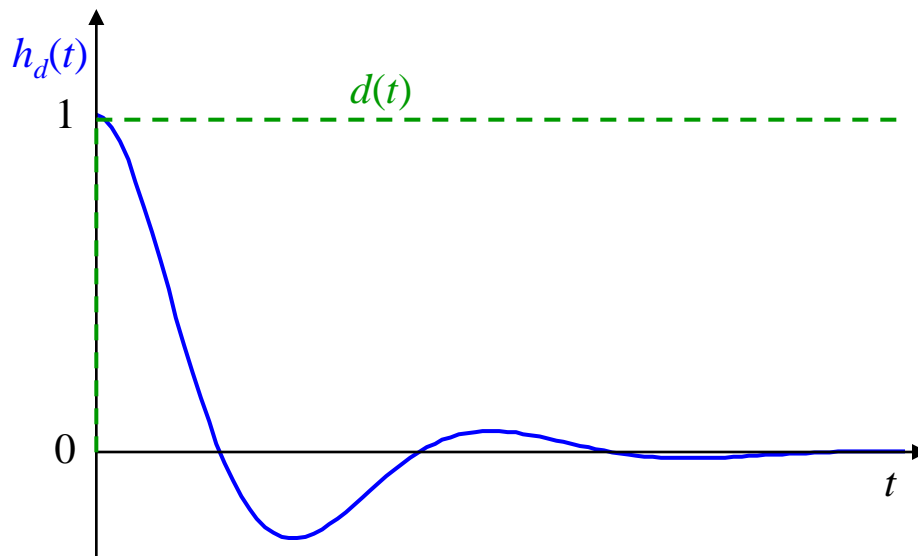
13.1 Beziehung zwischen Polen und Zeitverhalten

Lassen sich diese Überlegungen auch auf das Störverhalten des Regelkreises übertragen?

Ja, denn es gilt folgender Zusammenhang:

$$\text{Störgrößenübertragungsfunktion: } G_d(s) = \frac{1}{1 + G_0(s)} = \frac{1 + G_0(s) - G_0(s)}{1 + G_0(s)} = 1 - G_w(s)$$

Damit ist die Antwort auf einen Einheitssprung der Störgröße: $h_d(t) = 1 - h(t)$

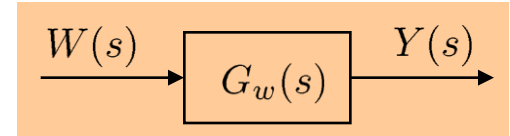


→ Aus *Überschwingen* wird *Unterschwingen*; alle sonstigen Größen sind 1:1 übertragbar!

13.1 Beziehung zwischen Polen und Zeitverhalten

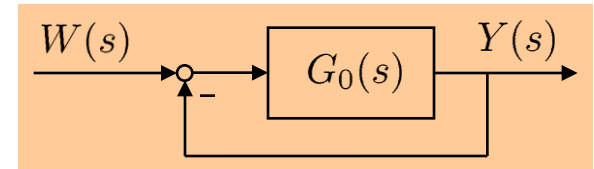
Geschlossener Regelkreis hat einen dominanten reellen Pol

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



Dieser Fall entsteht, wenn der offene Regelkreis integrales Verhalten aufweist bzw. als Integrator approximiert werden kann (d.h. alle anderen Pole sind sehr schnell und damit vernachlässigbar). Dies passiert insbesondere bei sehr langsam eingestellten Reglern.

$$G_0(s) = \frac{1}{Ts} \rightarrow G_w(s) = \frac{G_0(s)}{1 + G_0(s)}$$



Einzigste charakteristische Größe der Sprungantwort ist die Beruhigungszeit $T_{be}(5\%)$, d.h. die Zeit, in der das $\pm 5\%$ -Intervall um den Endwert erreicht und nicht mehr verlassen wird.

Ein PT_1 -Glied erreicht 95% seines Endwertes nach der ca. 3-fachen Zeitkonstante. Daher:

$$T_{be}(\pm 5\%) = 3T$$

13.2 Konstruktion der WOK

Wir wissen jetzt, wo wir die Pole des geschlossenen Regelkreises haben wollen. Aber wie wählen wir einen Regler und dessen Parameter aus, um diese Pole zu erzeugen?

Möglichkeit 1: Kompensationsregler oder Polvorgabe (→ Kapitel 12)

- Liefert schnell und bequem Reglerstruktur und -parameter.
- Keine Information über Robustheit oder Abstand zur Stabilitätsgrenze.
- Keine Einsichten oder "Gefühl" für Zusammenhänge.

Möglichkeit 2: Wurzelortskurve (WOK)

- Die WOK sind die Orte in der s -Ebene, wo die Pole des *geschlossenen* Regelkreises (in Abhängigkeit von Regler- oder auch Streckenparametern) liegen können.
- Meist wird zunächst die Struktur und die dynamischen Parameter (Pole und Nullstellen) des Reglers anhand qualitativer Überlegungen festgelegt; dann wird die WOK in Abhängigkeit der Reglerverstärkung gezeichnet.
- Die Pole des geschlossenen Regelkreises sind die Nullstellen oder *Wurzeln* des charakteristischen Polynoms. Daher kommt der Name *Wurzelortskurve*.

13.2 Konstruktion der WOK

Beispiel: Lage der Pole (Wurzeln der char. Gleichung) des geschlossenen Regelkreises

IT_I-Strecke: $G_S(s) = \frac{1}{s(s+4)}$

P-Regler: $G_R(s) = K_R$

$$G_0(s) = \frac{K_R}{s(s+4)}$$

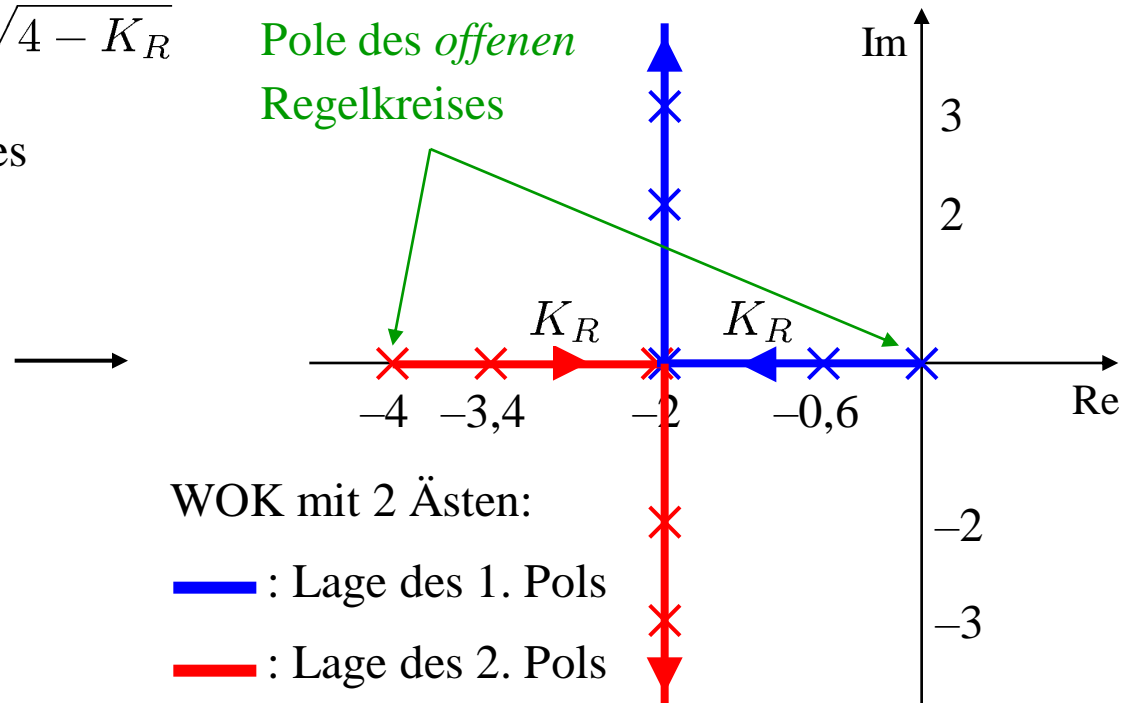
Charakteristische Gleichung: $1 + G_0(s) = 1 + \frac{K_R}{s^2 + 4s} = 0 \rightarrow s^2 + 4s + K_R = 0$

Pole (Wurzeln): $s_{1/2} = -2 \pm \sqrt{4 - K_R}$

Pole des offenen Regelkreises

Pole des geschlossenen Regelkreises

K_R	s_1	s_2
0	0	-4
2	-0,6	-3,4
4	-2	-2
8	-2+i2	-2-i2
13	-2+i3	-2-i3



13.2 Konstruktion der WOK

ACHTUNG: Positive Vorzeichen!

Für die Konstruktion der WOK gehen wir von der charakteristischen Gleichung aus:

$$1 + k\hat{G}_0(s) = 0 \quad \text{mit} \quad \hat{G}_0(s) = \frac{s^m + b_{m-1}s^{m-1} + \dots + b_1s + b_0}{s^n + a_{n-1}s^{n-1} + \dots + a_1s + a_0} \quad (m \leq n)$$

Evtl. zuvor vorhandene Koeffizienten b_m und a_n werden in den Faktor k mit herein gezogen. k ist *nicht* die Kreisverstärkung, d.h. das Produkt aus Regler- und Streckenverstärkung:

$$K_R K_S = k \frac{b_0}{a_0}$$

Für jeden Wert der Verstärkung k hat die charakteristische Gleichung genau n Lösungen für s , nämlich die n Pole des geschlossenen Regelkreises. Variiert man nun $k = 0, \dots, \infty$, dann ergeben sich n Kurven, die sog. Äste der WOK, welche den Verlauf der Pole des geschlossenen Regelkreises in Abhängigkeit der Verstärkung k repräsentieren. Am Verlauf der WOK können wir daher sehen, welche Pole sich für welche Verstärkungswerte ergeben.

Dann wählen wir auf den Ästen der WOK eine Punkt Kombination aus, die unseren Wunschpolen möglichst nahe kommt und rechnen die zugehörige Reglerverstärkung aus:

$$K_R = \frac{k}{K_S} \frac{b_0}{a_0}$$

13.2 Konstruktion der WOK

Annahme: $k \geq 0$, für $k < 0$
siehe weiter hinten

Eigenschaften der WOK

$$k\widehat{G}_0(s) = k \frac{s^m + b_{m-1}s^{m-1} + \dots + b_1s + b_0}{s^n + a_{n-1}s^{n-1} + \dots + a_1s + a_0} = k \frac{\prod_{i=1}^m (s - n_i)}{\prod_{i=1}^n (s - p_i)}$$

$$= k \frac{\prod_{i=1}^m |s - n_i|}{\prod_{i=1}^n |s - p_i|} e^{i \left(\sum_{i=1}^m \varphi_{ni} - \sum_{i=1}^n \varphi_{pi} \right)} = -1$$

Achtung: $i = \text{imaginäre Einheit!}$

Jeder Punkt auf der WOK muss also folgende Bedingungen einhalten:

→ **Amplitudenbedingung:** $\frac{\prod_{i=1}^m |s - n_i|}{\prod_{i=1}^n |s - p_i|} = \frac{1}{k}$

→ **Phasenbedingung:** $\sum_{i=1}^m \varphi_{ni} - \sum_{i=1}^n \varphi_{pi} = (2l + 1) \cdot 180^\circ$

$l = \dots, -2, -1, 0, 1, 2, \dots$

13.2 Konstruktion der WOK

Die Konstruktion der WOK direkt aus der Amplituden- und Phasenbedingung ist zu aufwändig. Statt dessen können Konstruktionsregeln abgeleitet werden, die eine einfache Skizzierung der WOK erlauben.

Den exakten Verlauf der WOK kann man mit MATLAB berechnen. Es ist trotzdem sehr wichtig, die wesentlichen Konstruktionsregeln zu kennen, denn nur so hat man später einen Anhaltspunkt für den Entwurf der Reglerstruktur!

Einige Eigenschaften der WOK lassen sich leicht ableiten:

- **Die WOK ist symmetrisch zur reellen Achse:** Die Pole des geschlossenen Kreises sind reell oder konjugiert komplex. Daraus folgt unmittelbar die Symmetrie der WOK.
- **Die n Äste der WOK beginnen in den Polen p_i des offenen Regelkreises:**
Das folgt aus der charakteristischen Gleichung für $k \rightarrow 0$:

$$1 + k\widehat{G}_0(s) = 0$$

$$k \prod_{i=1}^m (s - n_i) + \prod_{i=1}^n (s - p_i) = 0 \quad \rightarrow \quad \prod_{i=1}^n (s^{(k=0)} - p_i) = 0 \quad \rightarrow \quad \boxed{s_i^{(k=0)} = p_i}$$

13.2 Konstruktion der WOK

- **m Äste der WOK enden in den Nullstellen n_i des offenen Regelkreises:** Das folgt aus der charakteristischen Gleichung für $k \rightarrow \infty$:

$$\prod_{i=1}^m (s - n_i) + \frac{1}{k} \prod_{i=1}^n (s - p_i) = 0 \quad \rightarrow \quad \prod_{i=1}^m (s^{(k \rightarrow \infty)} - n_i) = 0 \quad \rightarrow \quad \boxed{s_i^{(k \rightarrow \infty)} = n_i}$$

- **$n-m$ Äste der WOK enden im Unendlichen:** Man kann sich die "fehlenden" $n-m$ Nullstellen als Nullstellen im Unendlichen vorstellen. Eine allgemeine Übertragungsfunktion mit Zählergrad m und Nennergrad n lässt sich schreiben als:

$$\frac{b_m s^m + b_{m-1} s^{m-1} + \dots + b_1 s + b_0}{a_n s^n + a_{n-1} s^{n-1} + \dots + a_1 s + a_0} = \frac{(1 + T_{Z1}s)(1 + T_{Z2}s) \cdot \dots \cdot (1 + T_{Zm}s)}{(1 + T_1s)(1 + T_2s) \cdot \dots \cdot (1 + T_ns)}$$

m Nullstellen bei: $-\frac{1}{T_{Zi}}$

$n-m$ Nullstellen bei: $-\frac{1}{T_{Zi}} \rightarrow -\infty$

$$\frac{(1 + T_{Z1}s)(1 + T_{Z2}s) \cdot \dots \cdot (1 + T_{Zm}s) \cdot (1 + T_{Z(m+1)}s) \cdot \dots \cdot (1 + T_{Zn}s)}{(1 + T_1s)(1 + T_2s) \cdot \dots \cdot (1 + T_ns)}$$

unter der Bedingung $T_{Z(m+1)} = \dots = T_{Zn} \rightarrow 0$.

13.2 Konstruktion der WOK

- **Jeder Punkt der reellen Achse, auf dessen rechter Seite die Anzahl der reellen Pole und Nullstellen (zusammen gezählt) ungerade ist, gehört zur WOK.**
- **Die Asymptotenwinkel der $n-m$ Äste, die ins Unendliche gehen, berechnen sich aus:**

$$\Psi_l = (1 + 2l) \frac{180^\circ}{n - m} \quad \text{mit } l = 0, 1, 2, \dots, n - m - 1$$

- **Der Schnittpunkt s_A dieser Asymptoten liegt auf der reellen Achse bei:**

$$s_A = \frac{\sum_{i=1}^n p_i - \sum_{i=1}^m n_i}{n - m} = \frac{b_{m-1} - a_{n-1}}{n - m}$$

- **Die Verzweigungspunkte s_v der WOK auf der reellen Achse genügen der Gleichung:**

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

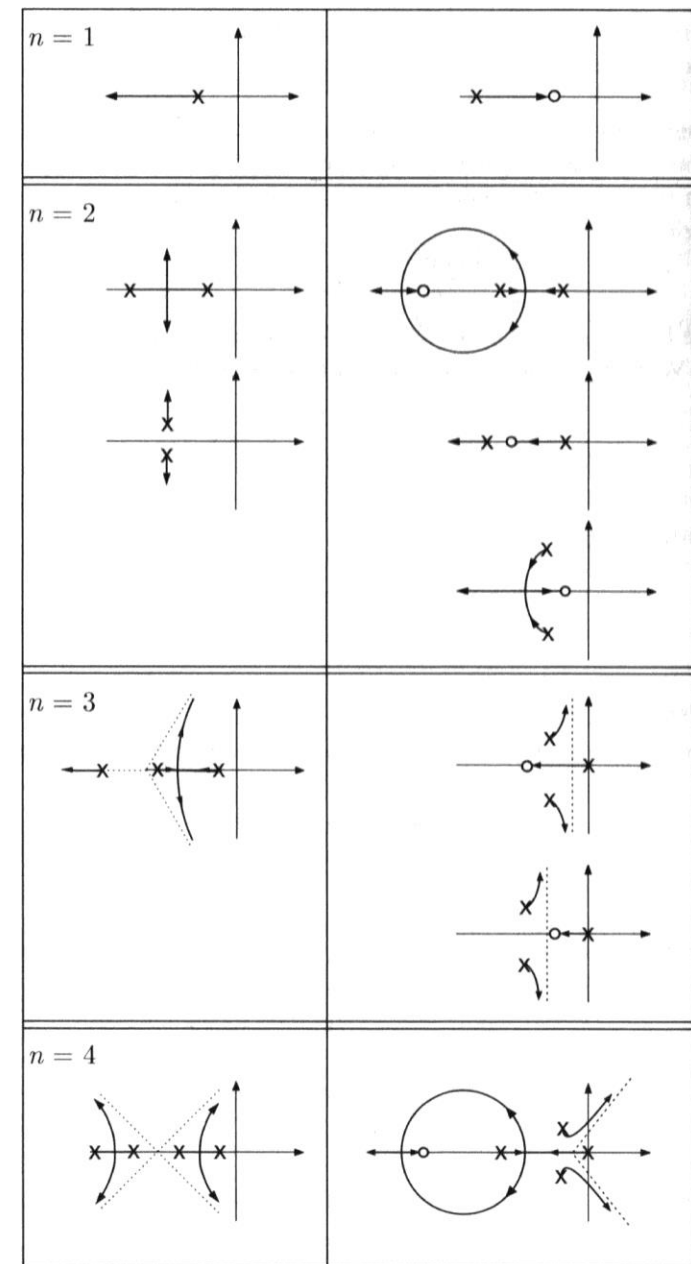
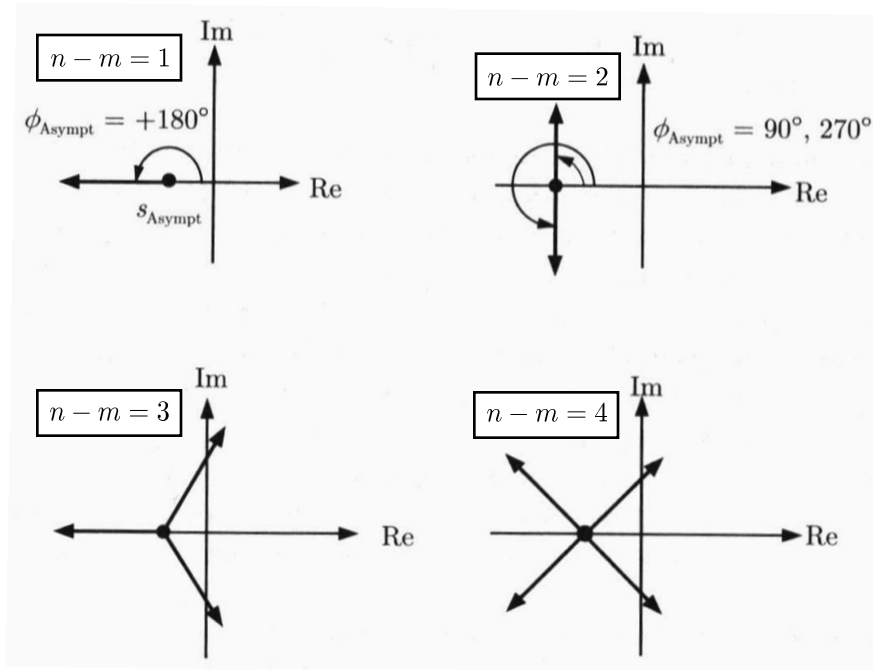
- **Schneiden sich in einem Verzweigungspunkt q Äste der WOK, so ist der Betrag des Schnittwinkels:**

$$\Delta\Psi_v = \frac{180^\circ}{q}$$

13.2 Konstruktion der WOK

Typischer Verlauf von WOKs [Lunze1]:

Asymptoten der WOK für unterschiedlichen Polüberschuss $n-m$ [Lunze1]:

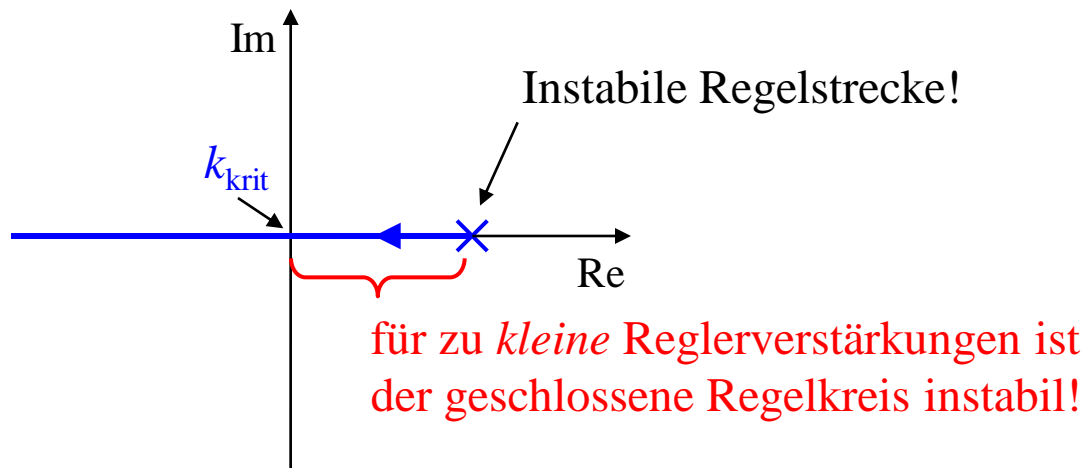


13.2 Konstruktion der WOK

Die WOK vermittelt einige tiefere Einsichten über das dynamische Verhalten von Regelkreisen, die weit über die Analyse einzelner spezieller Fälle hinausgehen:

1. Instabile Systeme

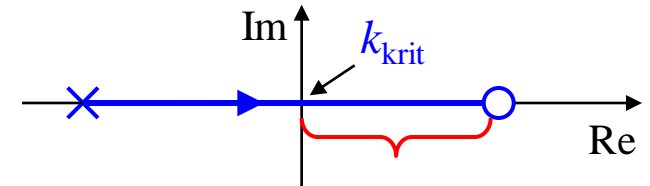
- Da die WOK bei den Polen des offenen Kreises startet, sind geschlossene Regelkreise mit instabiler Strecke (oder Regler) bei kleinen Verstärkungen auch instabil. Es ist eine gewisse Mindestverstärkung $k > k_{\text{krit}}$ notwendig, um den Regelkreis zu stabilisieren. D.h. die Rückkopplung muss genügend stark wirken.



13.2 Konstruktion der WOK

2. Nicht phasenminimale Systeme

- Da die WOK in den Nullstellen des offenen Kreises endet, werden geschlossene Regelkreise mit nicht phasenminimaler Strecke (oder Regler) bei großen Verstärkungen instabil. Es darf eine gewisse Maximalverstärkung $k < k_{\text{krit}}$ nicht überschritten werden, um den Regelkreis stabil zu halten. D.h. die Rückkopplung muss genügend schwach wirken.



für zu große Reglerverstärkungen ist der geschlossene Regelkreis instabil!

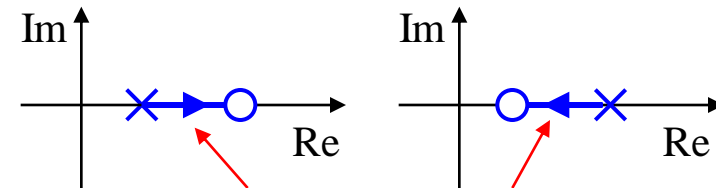
3. Systeme mit instabilen Polen und Nullstellen

- Um einen instabilen Pol der Strecke zu stabilisieren, muss die Regelung genügend stark eingreifen. Die Regelung muss schneller sein als der Pol.
Um mit einer instabilen Nullstelle der Strecke zurecht zu kommen, muss die Regelung genügend schwach eingreifen. Die Regelung muss langsamer sein als die Nullstelle.
Beide Bedingungen lassen sich nicht gleichzeitig erfüllen, wenn eine Strecke einen instabilen Pol und eine instabile Nullstelle aufweist, wobei der Pol schneller ist (weiter rechts liegt) als die Nullstelle. Eine solche Strecke wäre praktisch nicht regelbar!

13.2 Konstruktion der WOK

4. Kürzung instabiler Pol/Nullstellen

- Wir wissen bereits, dass wir aus Gründen der *inneren* Stabilität keine instabilen Pol/Nullstellen-Kürzungen durchführen dürfen. Die WOK zeigt, dass auch der gesamte geschlossene Regelkreis höchstwahrscheinlich instabil sein wird, wenn die Nullstelle den Pol nicht exakt trifft (was in der Praxis immer der Fall sein wird). Dann entsteht ein kurzes Stückchen WOK in der rechten s -Halbebene; egal ob die Nullstelle leicht rechts oder links vom Pol liegt.



Geschlossener Regelkreis hat einen instabilen Pol!

5. Kürzung stabiler Pol/Nullstellen

- Bei einer stabilen Pol/Nullstellen-Kürzung liegt dieses kurze Stückchen WOK aber in der linken s -Halbebene. Dort richtet es keinen Schaden an und beeinflusst den sonstigen Verlauf der WOK nur minimal (je weniger, desto näher sich Pol und Nullstelle sind). Daher können bei einer stabilen Pol/Nullstellen-Kürzung der Pol und die Nullstelle aus dem WOK-Diagramm entfernt werden.

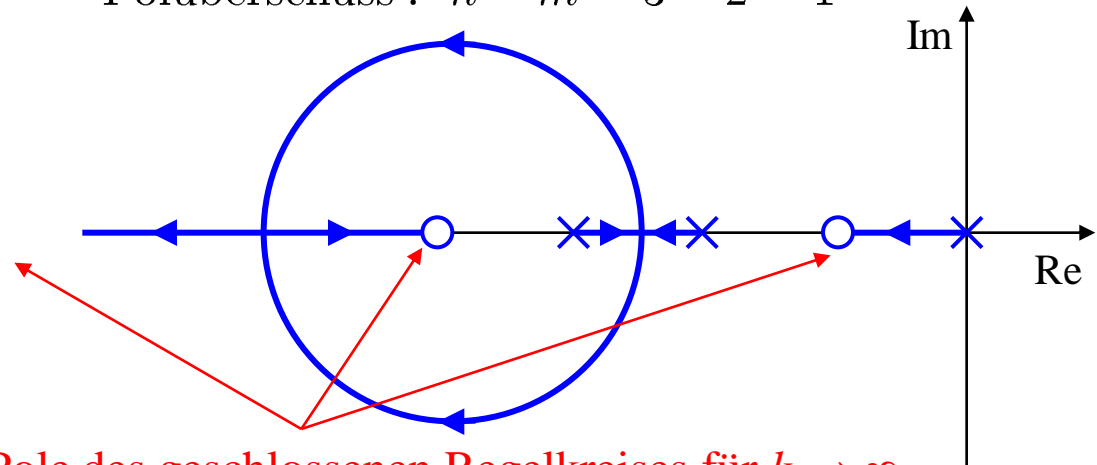
13.2 Konstruktion der WOK

6. Regelung mit hoher Kreisverstärkung (*high gain control*)

- Für phasenminimale offene Kreise mit Polüberschuss 1 enden alle bis auf einen der WOK-Äste in den Nullstellen. Dieser eine Ast strebt gegen $-\infty$. Da wegen der Phasenminimalität alle Nullstellen in der linken s -Halbebene liegen, erhält man für hinreichend große Kreisverstärkungen k *immer* einen stabilen Regelkreis. Die Stabilität ist gesichert, unabhängig von der genauen Lage der Nullstellen und Pole! Eine solche Regelstrategie nennt man "High Gain Control".

Leider weisen nahezu alle realen Prozesse einen Polüberschuss größer als 1 auf, so dass diese sehr einfache Regelstrategie meist nicht realisierbar ist.

$$\text{Polüberschuss : } n - m = 3 - 2 = 1$$



Endlage der Pole des geschlossenen Regelkreises für $k \rightarrow \infty$

13.3 Reglerentwurf mittels WOK

Ist die Reglerstruktur inkl. der dynamischen Parameter (Lage der Pole und Nullstellen) erst einmal festgelegt, kann man mittels WOK-Analyse leicht einen geeigneten Wert für die Reglerverstärkung finden. Dazu orientiert man sich an den Wunschregionen für die Pole des geschlossenen Regelkreises (siehe Anfang dieses Kapitels).

Wie soll man aber die Reglerstruktur und die Lage der Reglerpolstellen und -nullstellen festlegen? Dafür gibt es keine allgemein gültige, immer funktionierende Vorgehensweise. Etwas Ausprobieren wird dabei immer eine Rolle spielen. Folgende Gedanken sollten die Reglersynthese leiten:

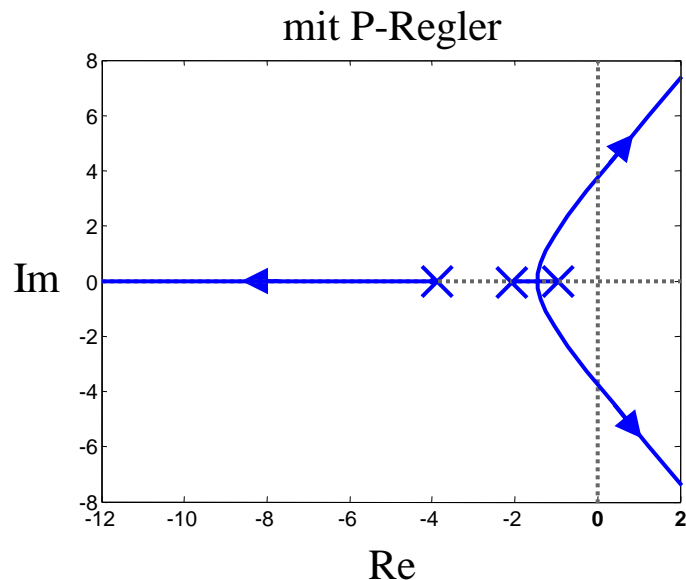
- Der Regler sollte einfach sein, also möglichst wenige Nullstellen und Pole besitzen.
- Vorwissen (z.B. Notwendigkeit für einen I-Anteil o.ä. aus dem "inneren Modell"-Prinzip) sollte von vorn herein in den Regler integriert werden.
- Standardüberlegungen sollten ausprobiert werden, wie z.B. das Kürzen des oder der dominanten Streckenpole mit Reglernullstellen.
- Nullstellen wirken auf die WOK-Äste "anziehend", Pole "abstoßend".
- Ein reiner D-Anteil reduziert die Anzahl der Asymptoten um 1, ein Polüberschuss von q im Regler erhöht die Anzahl der Asymptoten um q .

13.3 Reglerentwurf mittels WOK

Beispiel: Reglersynthese für die Strecke $G_S(s) = \frac{8}{(s+1)(s+2)(s+4)}$

1. P-Regler: $G_R(s) = K_R \rightarrow G_0(s) = k \frac{1}{(s+1)(s+2)(s+4)}$, $k = 8K_R$

Die Stabilitätsgrenze wird schon bei recht niedrigen Verstärkungen erreicht.



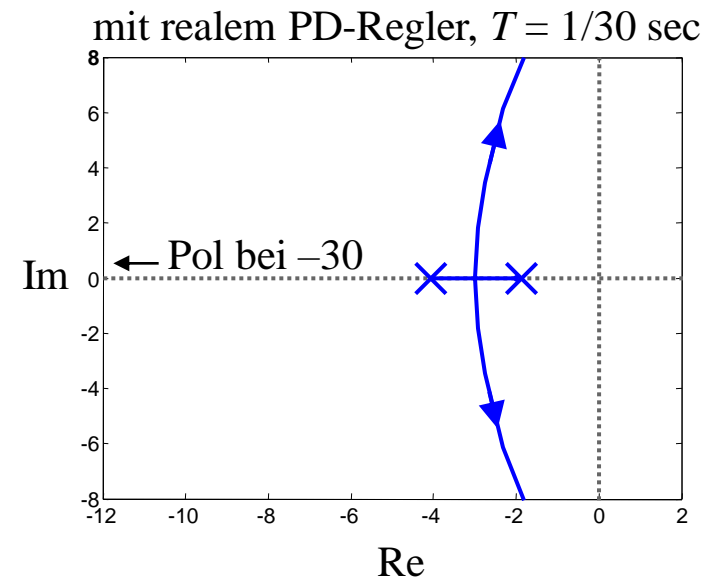
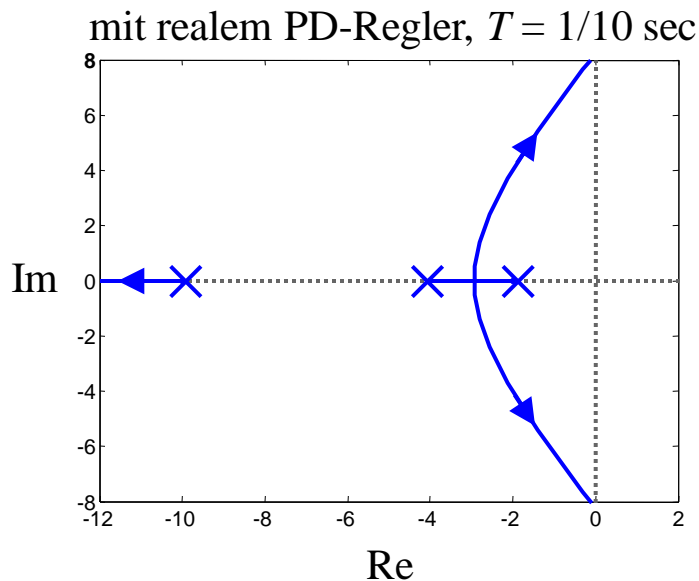
13.3 Reglerentwurf mittels WOK

2. Mehr Dynamik erzeugt ein realer PD-Regler (PDT₁): $G_R(s) = K_P \frac{1 + T_D s}{1 + T s}$

Kürzen des langsamsten Streckenpols mittels $T_D = 1$ sec und die Wahl $T = 1/10$ sec führen auf:

$$G_R(s) = 10K_P \frac{1 + s}{10 + s} \rightarrow G_0(s) = k \frac{1}{(s + 2)(s + 4)(s + 10)}, \quad k = 80K_P$$

Schnellere Regelung als mit P-Regler möglich. Weiter von Stabilitätsgrenze entfernt.



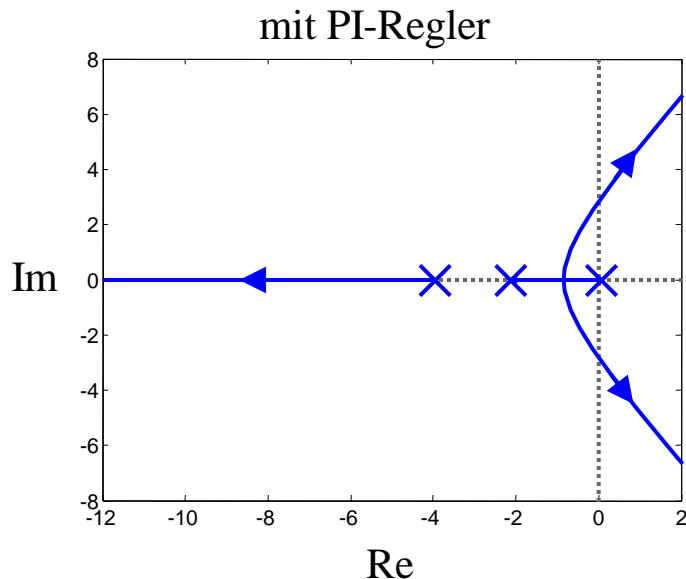
13.3 Reglerentwurf mittels WOK

3. Um bleibende Regelabweichungen zu vermeiden könnte man einen PI-Regler probieren:

$$G_R(s) = K_P \frac{1 + T_I s}{T_I s}$$

Kürzen des langsamsten Streckenpols mittels $T_I = 1$ sec führt auf:

$$G_R(s) = K_P \frac{1 + s}{s} \rightarrow G_0(s) = k \frac{1}{s(s+2)(s+4)}, \quad k = 8K_P$$



Der I-Anteil wirkt destabilisierend. Die WOK ist durch den Pol bei 0 näher an die Stabilitätsgrenze gerutscht.

Das ist der Preis, den man für die Vermeidung der bleibenden Regelabweichung zahlt, solange man bei einem 2-Parameter-Regler bleibt.

Ein PID-Regler kombiniert die Vorteile des PD- und des PI-Reglers...

13.3 Reglerentwurf mittels WOK

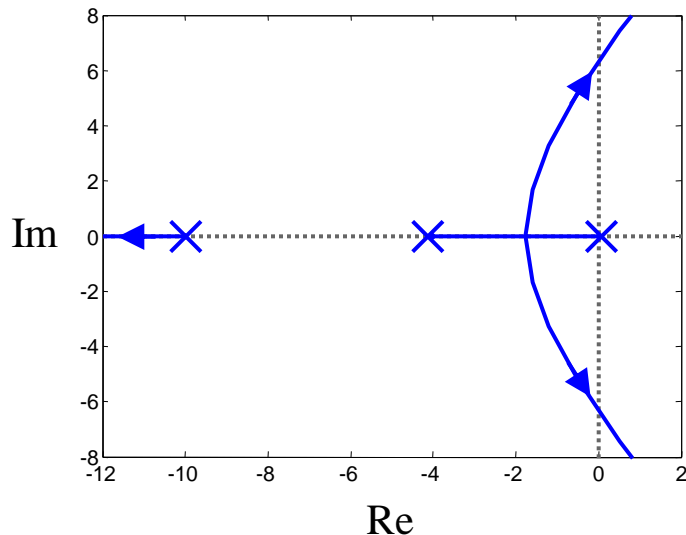
4. Mit einem PID-Regler können die Vorteile des PD- und des PI-Reglers kombiniert werden. Realistischerweise nehmen wir einen realen PID-Regler:

$$G_R(s) = K_P \frac{(1 + T_I s)(1 + T_D s)}{T_I s(1 + T s)}$$

Kürzen der 2 langsamen Streckenpole ($T_I = 1$ sec, $T_D = 1/2$ sec) und $T = 1/10$ sec:

$$G_R(s) = 5K_P \frac{(1 + s)(2 + s)}{s(10 + s)} \rightarrow G_0(s) = k \frac{1}{s(s + 4)(s + 10)}, \quad k = 40K_P$$

mit realem PID-Regler, $T = 1/10$ sec



Die Dynamik ist im Vergleich zum P-Regler deutlich verbessert. Zusätzlich verhindert der I-Anteil eine bleibende Regelabweichung.

Je weniger der D-Anteil durch die Nennerzeitkonstante T verwässert wird, desto weiter biegen sich die beiden rechten WOK-Äste nach links und desto größer wird der Abstand zur Stabilitätsgrenze.

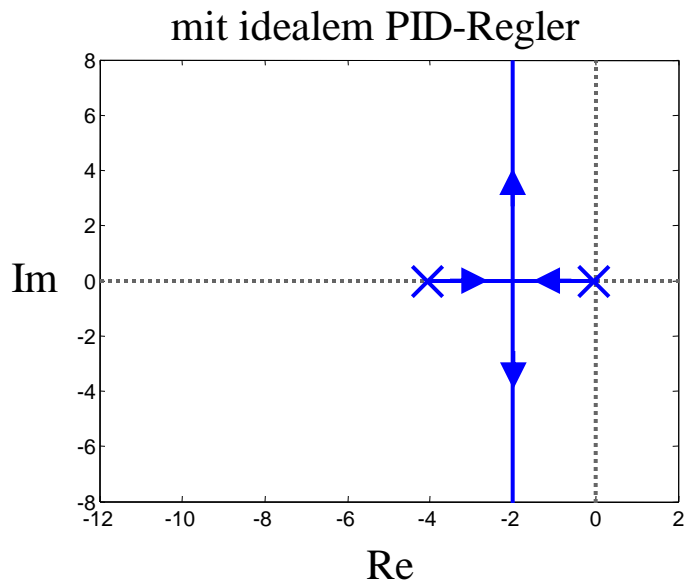
13.3 Reglerentwurf mittels WOK

5. Für den Grenzfall $T \rightarrow 0$ wird der reale PID-Regler zum idealen PID-Regler:

$$G_R(s) = K_P \frac{(1 + T_I s)(1 + T_D s)}{T_I s}$$

Gleicher Entwurf wie beim realen PID-Regler führt auf:

$$G_R(s) = \frac{1}{2} K_P \frac{(1 + s)(2 + s)}{s} \rightarrow G_0(s) = k \frac{1}{s(s + 4)}, \quad k = 4K_P$$



Durch den reinen D-Anteil reduziert sich die Anzahl der Asymptoten auf 2. Dadurch wird der Regelkreis *strukturstabil*, d.h. er bleibt für beliebig hohe Verstärkungen stabil; allerdings wird seine Dämpfung immer kleiner, er schwingt also immer stärker.

13.3 Reglerentwurf mittels WOK

WOK von nicht phasenminimalen Systemen

Bei der Herleitung der WOK-Regeln war vorausgesetzt worden, dass $k \geq 0$ ist und dass Zähler- und Nennerpolynom mit $+s^m$ bzw. $+s^n$ starten:

$$k\hat{G}_0(s) = k \frac{s^m + b_{m-1}s^{m-1} + \dots + b_1s + b_0}{s^n + a_{n-1}s^{n-1} + \dots + a_1s + a_0}$$

Bei nicht phasenminimalen offenen Regelkreisen *kann* es vorkommen, dass sich vor s^m ein Minuszeichen ergibt bzw. (alternativ) dieses Minus in k herein gezogen wird und dann $k < 0$ ist. Dies passiert genau dann, wenn es eine ungerade Anzahl instabiler Nullstellen gibt, z.B.:

$$k \frac{-s + 1}{(s + 3)(s + 5)} \quad \text{oder} \quad k \frac{(-s + 1)(s + 2)}{(s + 3)^2(s + 5)} \quad \text{oder} \quad k \frac{(-s + 1)^2(-s + 6)(s + 2)}{(s + 3)^2(s + 4)^2(s + 5)}$$

Instabile Pole sind kein Problem, solange sie in der üblichen Form $\frac{1}{(s - 3)}$ auftreten.

Hingegen zählen sie wie instabile Nullstellen, wenn sie als $\frac{1}{(-s + 3)}$ geschrieben sind.

In solchen Fällen ändern sich die WOK-Regeln! Sie entsprechen dann den WOK-Regeln "normaler" Systeme (ohne das Minuszeichen) für den Fall negativer Kreisverstärkungen.

13.3 Reglerentwurf mittels WOK

Für offene Regelkreise der Form

$$k\widehat{G}_0(s) = k \frac{-s^m + b_{m-1}s^{m-1} + \dots + b_1s + b_0}{s^n + a_{n-1}s^{n-1} + \dots + a_1s + a_0} \quad \text{mit } k \geq 0$$

bzw. der Form

$$k\widehat{G}_0(s) = k \frac{s^m + b_{m-1}s^{m-1} + \dots + b_1s + b_0}{s^n + a_{n-1}s^{n-1} + \dots + a_1s + a_0} \quad \text{mit } k \leq 0$$

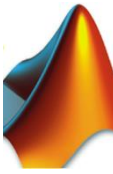
wenn interessiert, wie die Pole bei negativer Reglerverstärkung liegen.

ändern sich die WOK-Konstruktionsregeln wie folgt:

- Jeder Punkt der reellen Achse, auf dessen rechter Seite die Anzahl der reellen Pole und Nullstellen (zusammen gezählt) **gerade** ist, gehört zur WOK.
- Die Asymptotenwinkel der $n-m$ Äste, die ins Unendliche gehen, berechnen sich aus:

$$\Psi_l = \frac{l \cdot 360^\circ}{n - m} \quad \text{mit } l = 0, 1, 2, \dots, n - m - 1$$

Die Auswirkungen der Regeln werden sozusagen an der Imaginärachse "gespiegelt"!



WOK-Funktionen

```
s = tf('s');  
G0 = 8/((s+1)*(s+2)*(s+4));  
  
rlocus(G0);  
[k, Pole] = rlocfind(G0);  
  
sisotool
```

% s als Laplace-Variable definieren
% Übertragungsfunktion offener Kreis
% Zeichnet die WOK von G0
% Mit der Maus kann man auf einen Punkt
% eines WOK-Astes klicken und bekommt
% die zugehörige Kreisverstärkung sowie
% die Pole des geschlossenen Kreises
% zurückgegeben und angezeigt.

% Sehr komfortables GUI-Tool zur
% interaktiven Reglersynthese.
% Zunächst übergibt man nur die
% Übertragungsfunktion der
% Regelstrecke. Pole und Nullstellen
% des Reglers kann man nachher
% graphisch ergänzen und deren
% Auswirkungen direkt an der WOK
% beobachten.



14. *Reglerentwurf im Frequenzbereich

Inhalt Kapitel 14

14 *Reglerentwurf im Frequenzbereich

14.1 *Lead- und Lag-Glieder

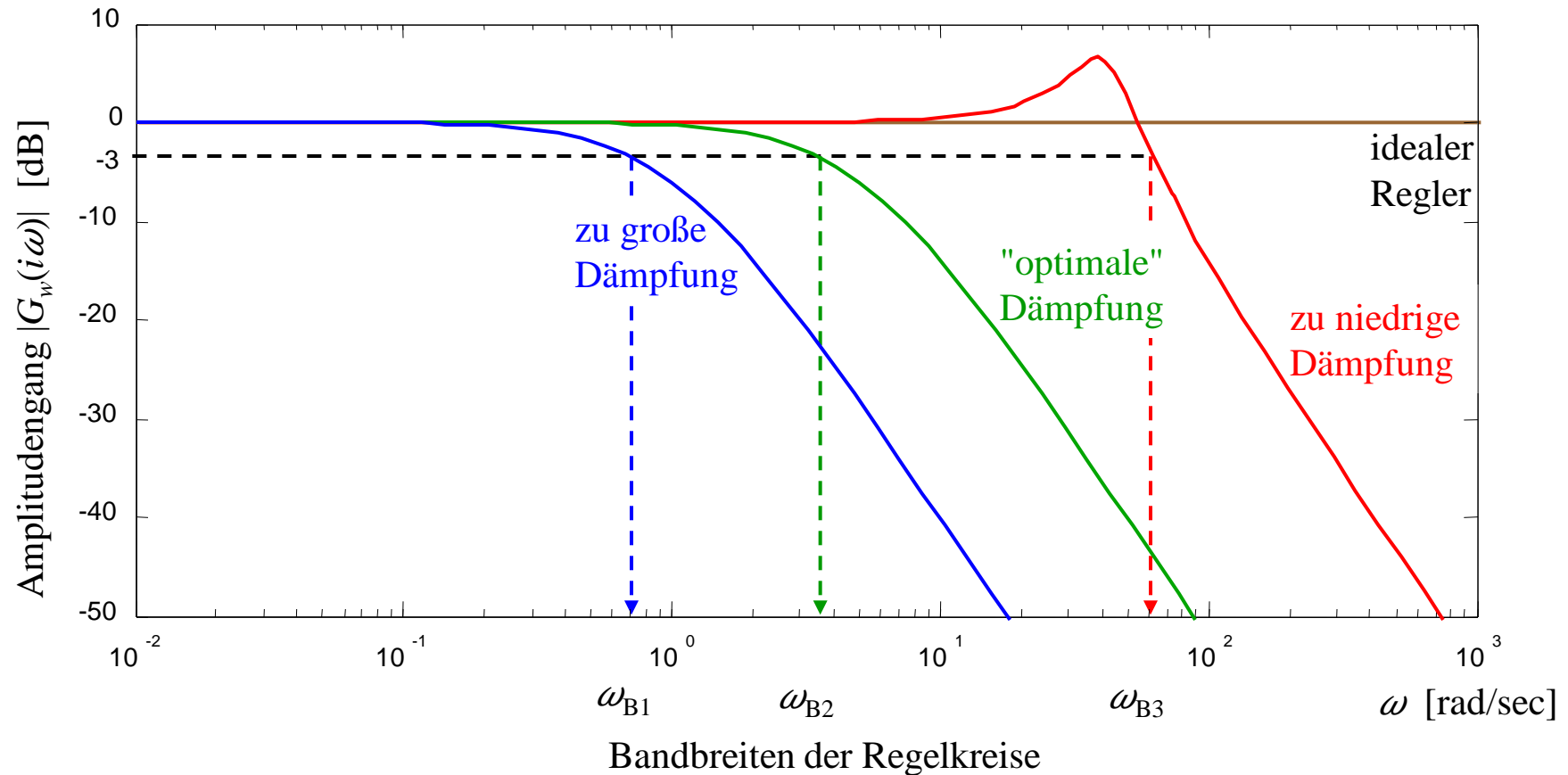
14.2 *Betragsoptimum

14.3 *Symmetrisches Optimum

14.1 *Lead- und Lag-Glieder

Der ideale Regler im Frequenzbereich: $G_w(i\omega) = 1$

In der Realität wird es im offenen Regelkreis immer einen Polüberschuss geben (Strecke inkl. Aktor und Sensor nicht sprungfähig) und $G_w(s)$ für hohe Frequenzen abfallen.



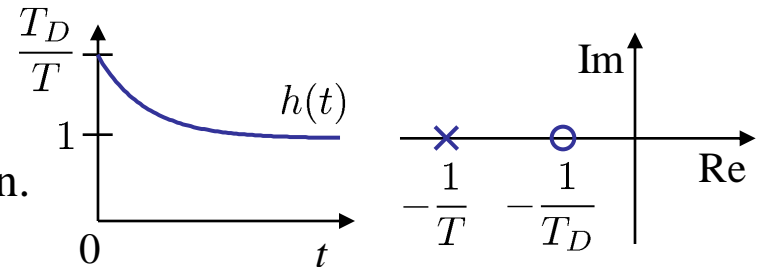
14.1 *Lead- und Lag-Glieder

Neben den klassischen Reglern (P, PI, PID, ...) gibt es noch zwei häufig als Regler bzw. als Teil des Reglers verwendete Korrekturglieder:

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

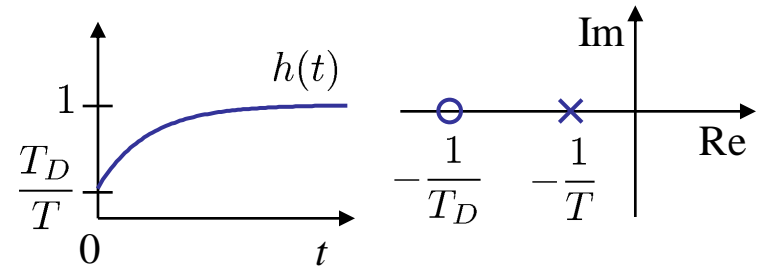
Lead-Glied oder phasenhebendes Glied ($T_D > T$):

- Reales PD-Glied.
- Abgeschwächte Eigenschaften eines D-Glieds.
- Wird verwendet, um den Phasenrand zu vergrößern.

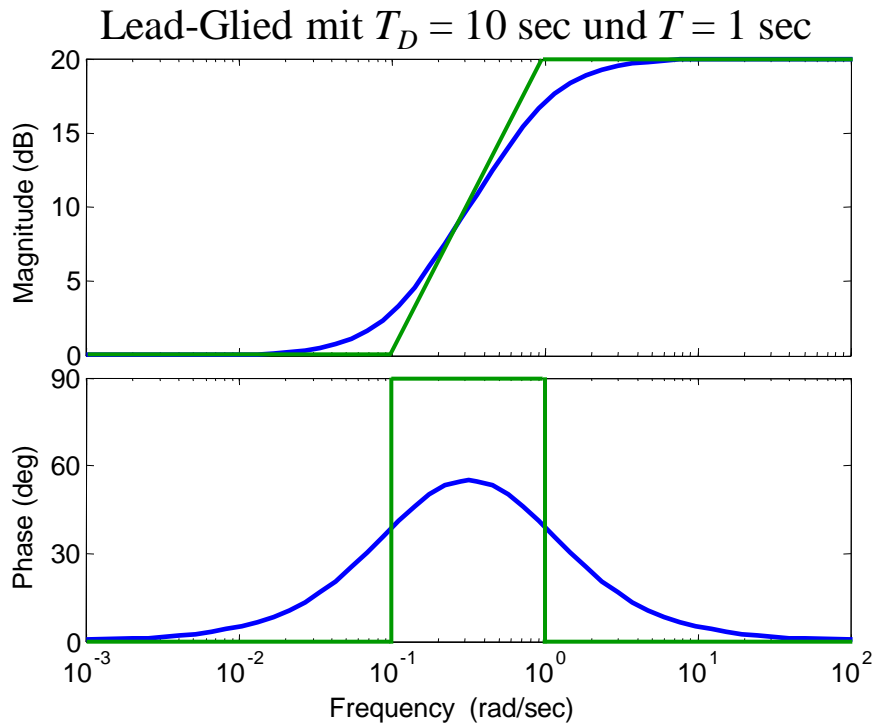


Lag-Glied oder phasenabsenkendes Glied ($T_D < T$):

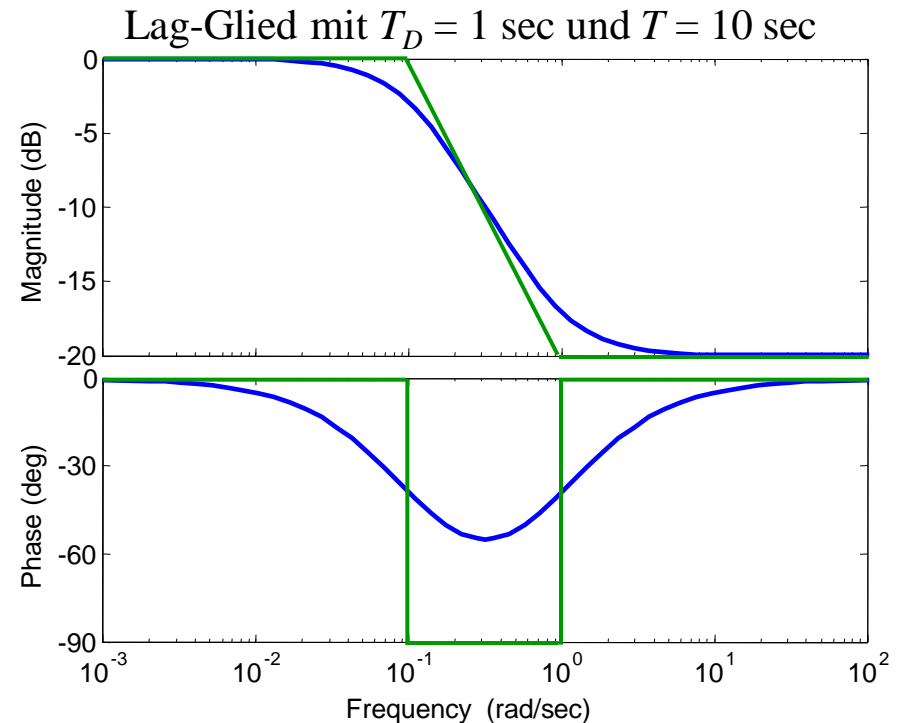
- Abgeschwächte Eigenschaften eines I-Glieds.
- Wird als Alternative zum I-Glied verwendet, um den Amplitudengang bei niedrigen Frequenzen anzuheben (Verbesserung der stationären Genauigkeit).



14.1 *Lead- und Lag-Glieder



- Phasenhebung im Bereich mittlerer Frequenzen Amplitudendurchtrittsfrequenz ω_D .
- Wirkt dadurch stabilisierend ähnlich wie ein D-Glied, hat aber geringer Rauschverstärkung bei hohen Frequenzen.



- Amplitudenanhebung (zusammen mit vergrößertem Verstärkungsfaktor) bei niedrigen Frequenzen $\omega \ll \omega_D$
- Erhöht dadurch die stationäre Genauigkeit ähnlich wie ein I-Glied, hat aber nicht dessen destabilisierende Wirkung.

14.1 *Lead- und Lag-Glieder

Beispiel für den Einsatz eines Lead-Glieds:

Regelstrecke: $G_S(s) = \frac{5}{(s + 1)(0.25s + 1)}$

Regler: P-Regler $G_R(s) = K_R$

1. Wo liegt die Amplitudendurchtrittsfrequenz ω_D bei einer Reglerverstärkung von $K_R = 1$?
Aus dem Bode-Diagramm lesen wir ab: $\omega_D = 3.5$ rad/sec.

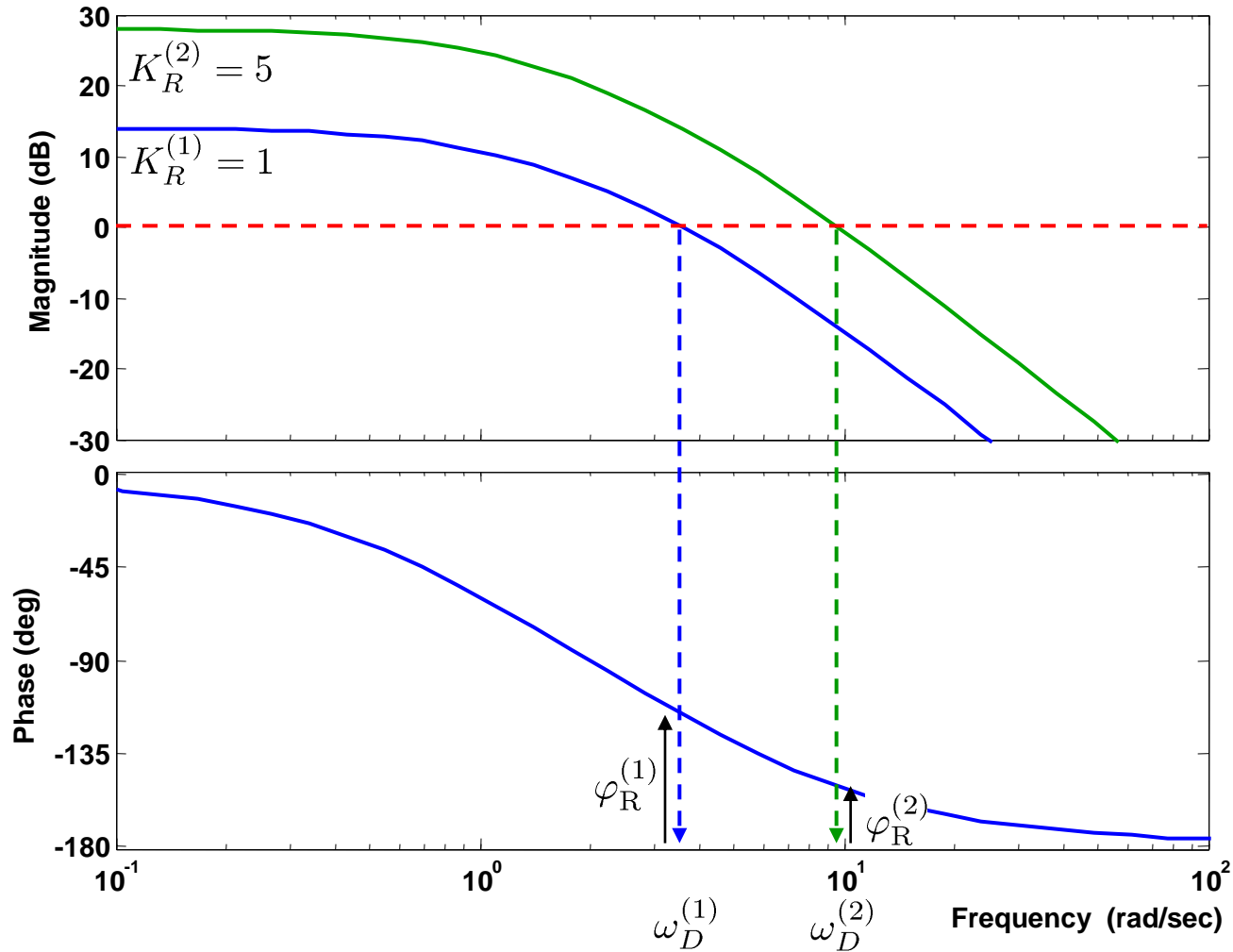
2. Dieser Regelkreis ist uns zu langsam. Wir erhöhen deshalb die Reglerverstärkung auf $K_R = 5$. Wo liegt nun die Amplitudendurchtrittsfrequenz?

Der Amplitudengang schiebt sich nach oben, die Amplitudendurchtrittsfrequenz wandert daher nach rechts auf $\omega_D^{(2)} = 9.5$ rad/sec.

3. Vergleiche den Phasenrand für Fall 1 und 2.

Fall 1: $\varphi_R = 60^\circ$. Fall 2: $\varphi_R = 30^\circ$. Der Abstand zu Stabilitätsgrenze hat sich durch die Erhöhung der Reglerverstärkung erheblich reduziert!

14.1 *Lead- und Lag-Glieder



14.1 *Lead- und Lag-Glieder

4. Nun fügen wir ein Lead-Glied in den Regler ein, um die Phase in der Gegend um $\omega_D = 9.5$ rad/sec anzuheben und damit den Regelkreis robuster zu machen. Der Regler ist also nun von der Form:

$$G_R(s) = K_R \frac{1 + T_D s}{1 + T s} \quad \text{mit } T_D > T$$

Die Knickfrequenzen für das Lead-Glied wählen wir symmetrisch um ω_D , also wegen der logarithmischen Skalierung bei 5 rad/sec und 20 rad/sec. Wie lautet dann der Regler?

Die Verstärkung bleibt bei $K_R = 5$, die Zeitkonstanten sind dann $T_D = 1/5$ sec, $T = 1/20$ sec.

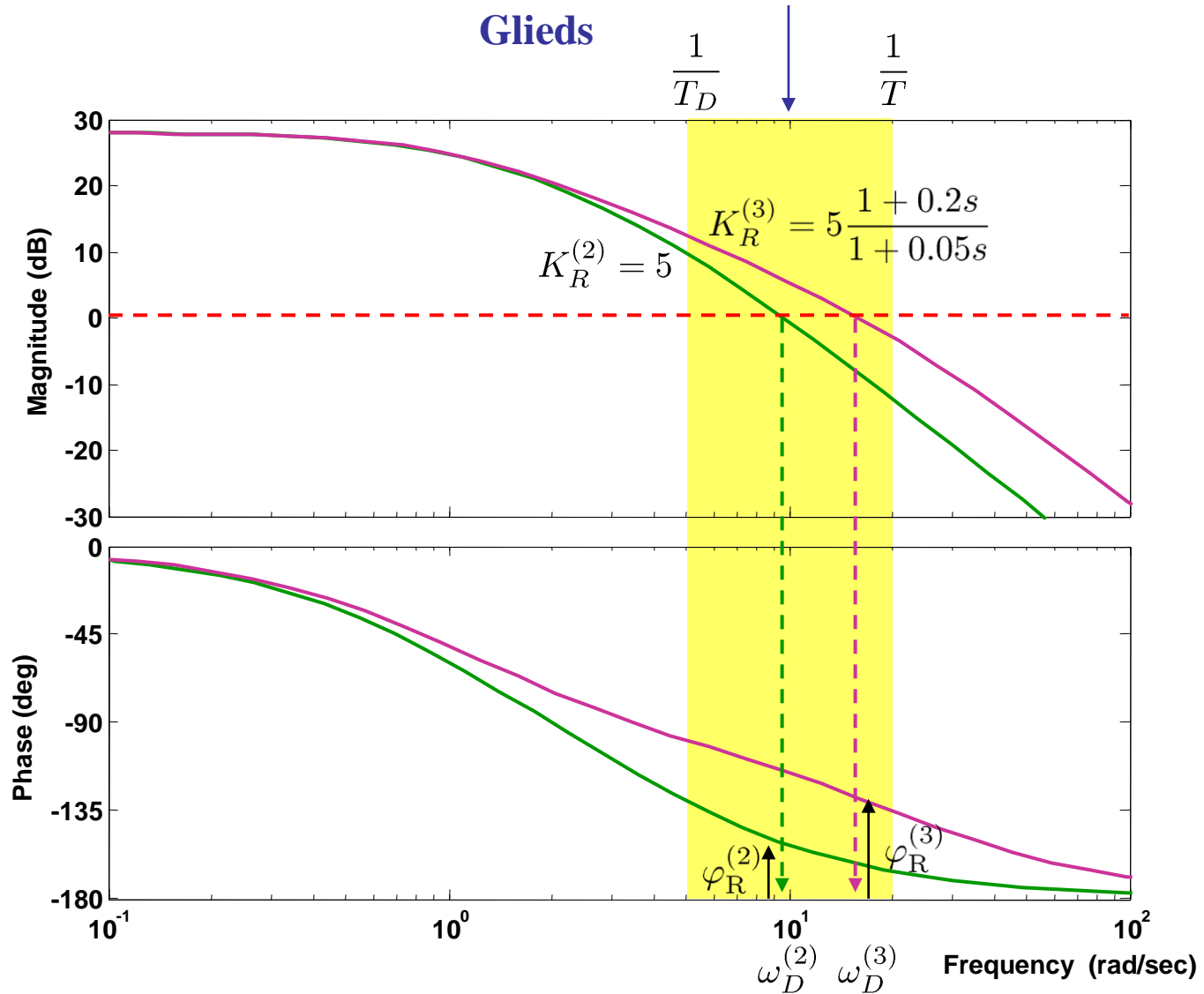
$$G_R(s) = 5 \frac{1 + 0.2s}{1 + 0.05s}$$

5. Wie groß ist nun der Phasenrand und die Amplitudendurchtrittsfrequenz?

Der Phasenrand konnte auf $\varphi_R = 50^\circ$ (Fall 3) erhöht werden. Zusätzlich wurde durch das lokale D-Verhalten des Reglers zwischen T_D und T die Amplitudendurchtrittsfrequenz auf $\omega_D^{(3)} = 17$ rad/sec angehoben und die Bandbreite des Regelkreises erhöht.

14.1 *Lead- und Lag-Glieder

Wirksamer Bereich des Lead-Glieds



14.1 *Lead- und Lag-Glieder

Beispiel für den Einsatz eines Lag-Glieds:

Beginnen wir mit dem Ergebnis des vorherigen Beispiels.

$$\text{Regelstrecke: } G_S(s) = \frac{5}{(s+1)(0.25s+1)} \quad \text{P-Regler mit Lead-Glied: } G_R(s) = 5 \frac{1+0.2s}{1+0.05s}$$

1. Wie groß ist die bleibende Regelabweichung, wenn die Führungsgröße mit einem Einheitssprung beaufschlagt wird?

$$\text{Bleibende Regelabweichung} = 1 - \frac{K}{1+K} = \frac{1}{1+K} = \frac{1}{1+25} = 0.04$$

2. Was kann man tun, um die bleibende Regelabweichung zu reduzieren?
 - a) Einführung eines I-Anteils. → Senkt die Phase um 90° ab und erzeugt Instabilität.
D.h. der Regler müsste komplett neu ausgelegt werden.
 - b) Vergrößerung der Verstärkung. → Reduziert die Phasenreserve.
 - c) Einsatz eines Lag-Glieds. → Vergrößert die Verstärkungen nur für kleine Frequenzen.
Phasenreserve bleibt (nahezu) erhalten!

14.1 *Lead- und Lag-Glieder

3. Ergänze den Regler durch ein Lag-Glied mit den Eckfrequenzen bei 0.05 rad/sec und 0.5 rad/sec, so dass es den Verlauf des Frequenzgangs für höhere Frequenzen kaum beeinflusst.

Das Lag-Glied hat die Übertragungsfunktion:

$$G_{\text{Lag}}(s) = K_{\text{Lag}} \frac{1 + T_D s}{1 + T s} \quad \text{mit } T_D < T$$

Aus den Eckfrequenzen ergibt sich:

$$T_D = 2 \text{ sec}, \quad T = 20 \text{ sec}$$

Die Verstärkung K_{Lag} muss so gewählt werden, dass das Übertragungsverhalten bei hohen Frequenzen nicht verändert wird. Daraus erhält man:

$$K_{\text{Lag}} = 10$$

Der Gesamtregler lautet also:

$$G_R(s) = 50 \frac{1 + 0.2s}{1 + 0.05s} \frac{1 + 2s}{1 + 20s}$$

14.1 *Lead- und Lag-Glieder

4. Wie groß ist die bleibende Regelabweichung mit dem neuen Regler?

$$\text{Bleibende Regelabweichung} = 1 - \frac{K}{1 + K} = \frac{1}{1 + K} = \frac{1}{1 + 50} = 0.02$$

5. Wie hätte man mit dem Lag-Glied eine noch kleinere bleibende Regelabweichung erzwingen können?

Die Übertragungsfunktion des Lag-Glieds ist:

$$G_{\text{Lag}}(s) = K_{\text{Lag}} \frac{1 + T_D s}{1 + T s} \quad \text{mit } T_D < T$$

Die Verstärkung wurde so eingestellt, dass sich für hohe Frequenzen nichts verändert:

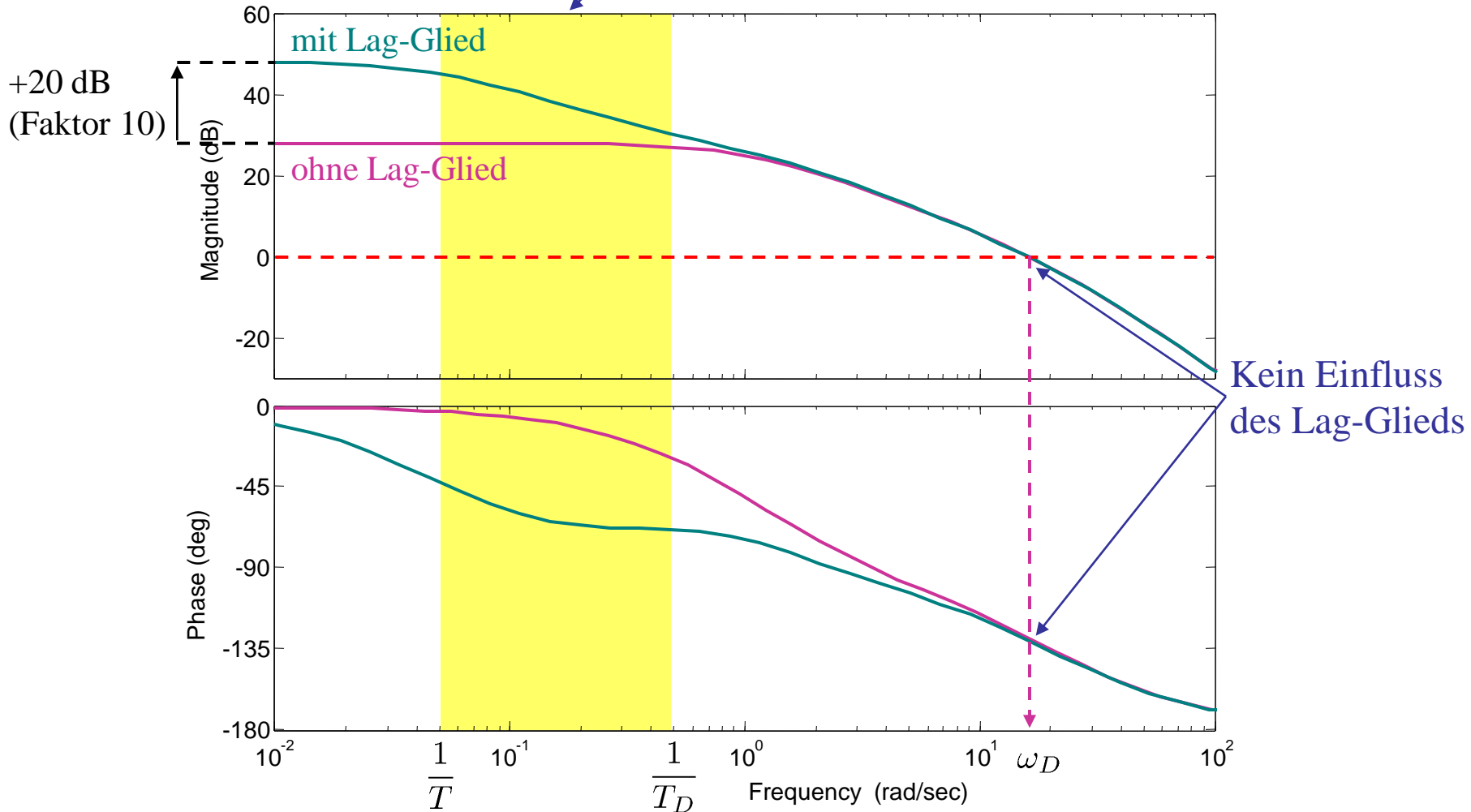
$$K_{\text{Lag}} = \frac{T}{T_D}$$

Je größer der Faktor zwischen T und T_D ist, desto größer wird die Verstärkung K_{Lag} für niedrige Frequenzen.

Wichtig ist, dass $1/T_D$ deutlich kleiner (mindestens eine Zehnerpotenz) ist als die Amplitudendurchtrittsfrequenz ω_D , damit die negative Phasenverschiebung des Lag-Glieds schon abgeklungen ist und nicht mehr (wesentlich) die Phasenreserve reduziert.

14.1 *Lead- und Lag-Glieder

Wirksamer Bereich des Lag-Glieds



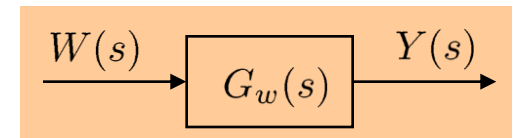
14.2 *Betragsoptimum

Das Betragsoptimum ist eine Einstellregel, also nicht wirklich eine Optimierung, die oft in der Antriebstechnik angewandt wird. Sie hat folgende wichtige Eigenschaften:

- Entwurfsidee kommt aus Überlegungen im Frequenzbereich.
- Führt auf stark gedämpftes (d.h. wenig überschwingendes aber auch langsames) Verhalten des geschlossenen Regelkreises.
- Ist sehr einfach anwendbar.
- Geht von einem PT_2 -Verhalten des geschlossenen Regelkreises aus.

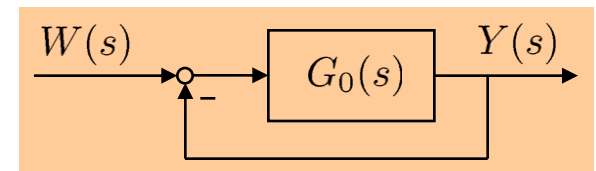
Wir gehen also von folgender Führungsübertragungsfunktion aus:

$$G_w(s) = \frac{\omega_0^2}{s^2 + 2D\omega_0 s + \omega_0^2}$$



Daraus folgt unmittelbar ein IT_1 -Verhalten des *offenen* Regelkreises:

$$G_w(s) = \frac{G_0(s)}{1 + G_0(s)} \quad \rightarrow \quad G_0(s) = \frac{\omega_0^2}{s(s + 2D\omega_0)}$$



14.2 *Betragsoptimum

Ein solches IT_1 -Verhalten ergibt sich in vielen Anwendungen bzw. kann durch geschickte Wahl des Reglers erzeugt werden. Hierzu 4 typische Beispiele:

1. PT_1 -Strecke und I-Regler

$$G_S(s) = \frac{K}{1 + T_1 s}, \quad G_R(s) = \frac{1}{T_I s} \quad \rightarrow \quad G_0(s) = \frac{K}{\underbrace{T_I s}_{\text{freier Reglerparameter}}(1 + T_1 s)}$$

freier Reglerparameter

2. PT_2 -Strecke und PI-Regler mit Kürzung eines Streckenpols

$$G_S(s) = \frac{K}{(1 + T_1 s)(1 + T_2 s)}, \quad G_R(s) = K_P \frac{1 + T_I s}{T_I s} \quad \rightarrow \quad G_0(s) = \frac{\underbrace{K_P K}_{\text{freier Reglerparameter}}}{T_2 s(1 + T_1 s)}$$

mit $T_I = T_2$

3. PT_3 -Strecke und PID-Regler mit Kürzung zweier Streckenpole

$$G_S(s) = \frac{K}{(1 + T_1 s)(1 + T_2 s)(1 + T_3 s)}, \quad G_R(s) = K_P \frac{(1 + T_I s)(1 + T_D s)}{T_I s}$$

$$\rightarrow G_0(s) = \frac{\underbrace{K_P K}_{\text{freier Reglerparameter}}}{T_2 s(1 + T_1 s)} \quad \text{mit } T_I = T_2, T_D = T_3$$

14.2 *Betragsoptimum

4. IT_1 -Strecke und P-Regler

$$G_S(s) = \frac{K}{s(1 + T_1s)}, \quad G_R(s) = K_P \quad \rightarrow \quad G_0(s) = \frac{\overset{\text{freier Reglerparameter}}{\circlearrowleft} K_P K}{s(1 + T_1s)}$$

Wir behandeln nun nur Fall 1; die anderen Fälle lassen sich leicht durch Variablenumbenennung darauf zurück führen.

- Variablenumbenennung: $\frac{1}{T_I} = \frac{K_P}{T_2} \rightarrow K_P = \frac{T_2}{T_I}$ bzw. $K_P = \frac{1}{T_I}$ für Fall 4.
- Im Falle des PI- bzw. PID-Reglers ist es normalerweise sinnvoll, jeweils die langsamsten Pole wegzukürzen, um ein möglichst schnelles Regelverhalten zu ermöglichen.
- In allen Fällen bleibt nur noch 1 freier Reglerparameter zu bestimmen.
- Eine PT_1 -Strecke mit einem I-Regler zu regeln ist natürlich sehr "konservativ"; mit einem PI- oder gar PID-Regler könnte man natürlich viel besseres Regelverhalten erzielen – dann ließe sich aber das Betragsoptimum-Entwurfsverfahren nicht mehr direkt anwenden.

14.2 *Betragsoptimum

Aus Kapitel 10 wissen wir, dass für den *idealen* Regler gilt:

Führungsübertragungsfunktion: $G_w(s) = 1$ Störübertragungsfunktion: $G_d(s) = 0$

Dann würde die Regelgröße $y(t)$ jeder Führungsgröße $w(t)$ perfekt folgen und durch die Störgröße $d(t)$ gar nicht beeinflusst. Wir wissen auch, dass diese Forderungen in Realität nur näherungsweise erfüllt werden können.

Eine sinnvolle Forderung wäre z.B., dass der Amplitudengang möglichst lange (d.h. von $\omega = 0$ bis zu möglichst hohen Frequenzen) möglichst nahe an 1 ist.

Beim Betragsoptimum fordern wir: $|G_w(i\omega)| \approx 1$. Aus dem offenen IT₁-Regelkreis ergibt sich folgendes PT₂-Verhalten des geschlossenen Regelkreises:

$$G_w(s) = \frac{\frac{K}{T_I s(1+T_1 s)}}{1 + \frac{K}{T_I s(1+T_1 s)}} = \frac{K}{T_I T_1 s^2 + T_I s + K}$$

Der Amplitudengang der Führungsübertragungsfunktion ist dann:

$$|G_w(i\omega)| = \frac{K}{\sqrt{(-T_I T_1 \omega^2 + K)^2 + (T_I \omega)^2}} = \frac{K}{\sqrt{T_I^2 T_1^2 \omega^4 + (T_I^2 - 2T_I T_1 K)\omega^2 + K^2}}$$

14.2 *Betragsoptimum

Wir fordern nun $|G_w(i\omega)| \approx 1$ bzw. $|G_w(i\omega)|^2 \approx 1$:

$$\frac{K^2}{T_I^2 T_1^2 \omega^4 + (T_I^2 - 2T_I T_1 K)\omega^2 + K^2} \approx 1$$

Diese Forderung wäre exakt erfüllt, wenn folgende Bedingungen gelten würden:

1. $K^2 = K^2$ (wichtig für niedrige Frequenzen $\omega \ll 1$)
2. $(T_I^2 - 2T_I T_1 K)\omega^2 = 0 \cdot \omega^2$ (wichtig für mittlere Frequenzen $\omega \approx 1$)
3. $(T_I^2 T_1^2)\omega^4 = 0 \cdot \omega^4$ (wichtig für hohe Frequenzen $\omega \gg 1$)

Zu diesen Bedingungen:

1. Durch den I-Anteil gibt es keine bleibende Regelabweichung, d.h. stationär (für $\omega = 0$) ist die Regelung perfekt. Deshalb ist diese Bedingung immer erfüllt.
2. Diese Bedingung ist erfüllt, wenn gilt: $T_I = 2KT_1$
Damit ist der freie Reglerparameter T_I festgelegt.
3. Diese Bedingung lässt sich nie erfüllen. D.h. es wird bei hohen Frequenzen Abweichungen von der idealen Regelung $|G_w(i\omega)| = 1$ geben.

14.2 *Betragsoptimum

Setzen wir nur den Reglerparameter $T_I = 2KT_1$, der nach dem Betragsoptimum ermittelt wurde, in die Führungsübertragungsfunktion ein, erhalten wir:

$$G_w(s) = \frac{K}{T_I T_1 s^2 + T_I s + K} = \frac{K}{2KT_1^2 s^2 + 2KT_1 s + K} = \frac{\frac{1}{2T_1^2}}{s^2 + \frac{1}{T_1} s + \frac{1}{2T_1^2}}$$

Ein Vergleich mit der PT₂-Standardform $G_w(s) = \frac{\omega_0^2}{s^2 + 2D\omega_0 s + \omega_0^2}$ liefert:

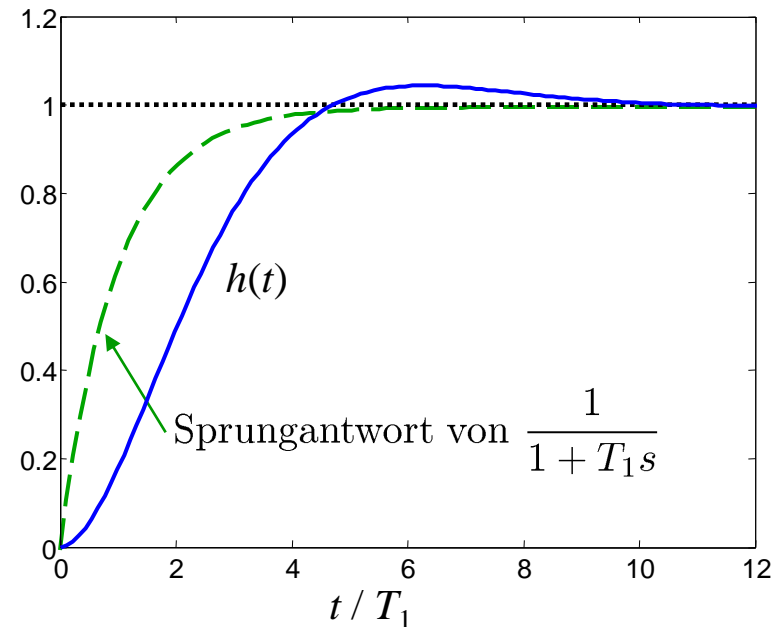
$$\omega_0^2 = \frac{1}{2T_1} \rightarrow \omega_0 = \frac{1}{\sqrt{2}T_1}$$

$$2D\omega_0 = \frac{2D}{\sqrt{2}T_1} = \frac{1}{T_1} \rightarrow D = \frac{1}{\sqrt{2}}$$

Die Sprungantwort $h(t)$ des Regelkreises schwingt also ganz leicht über:

$$\ddot{u} = e^{-\frac{\pi D}{\sqrt{1-D^2}}} = e^{-\frac{\pi \frac{1}{\sqrt{2}}}{\sqrt{1-\frac{1}{2}}}} = 4.3\%$$

Der Vergleich mit der PT₁-Sprungantwort zeigt: Ein I-Regler ist nicht gerade schnell!



14.2 *Betragsoptimum

Vereinfachungen der Regelstrecke, um Betragsoptimum anwenden zu können

Wenn kein offener Regelkreis mit IT_1 -Verhalten vorliegt bzw. sich nicht durch geschickte Wahl der Reglerstruktur erzeugen lässt, dann kann er manchmal durch folgende Vereinfachungen erzeugt werden.

Kleine Totzeiten

$$e^{-T_t s} = \frac{1}{e^{T_t s}} = \frac{1}{1 + \frac{(T_t s)^1}{1!} + \frac{(T_t s)^2}{2!} + \frac{(T_t s)^3}{3!} + \dots} \approx \frac{1}{1 + T_t s}$$

PT_n -Verhalten (Satz von der Summe der kleinen Zeitkonstanten)

Falls man bei PT_2 -Verhalten keinen PI-Regler oder bei PT_3 -Verhalten keinen PID-Regler einsetzen kann/möchte oder $n > 3$ ist:

$$\frac{K}{(1 + T_1)(1 + T_2) \cdot \dots \cdot (1 + T_n)} \approx \frac{K}{1 + T_E s} \quad \text{mit } T_E = T_1 + T_2 + \dots + T_n$$

Diese Vereinfachung ist nur zulässig, wenn in Reihe zu diesem PT_n -Glied ein I-Glied oder ein PT_1 -Glied mit großer Zeitkonstante $T \gg T_E$ geschaltet ist.

14.3 *Symmetrisches Optimum

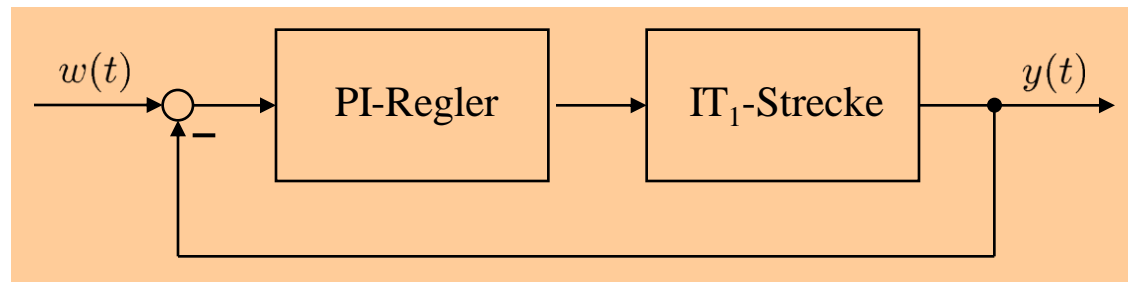
Auch aus der Antriebstechnik kommt eine weitere, einfache Einstellregel: das symmetrische Optimum. Beim ihm geht man von einer IT_1 -Regelstrecke aus, die mit einem PI-Regler geregelt werden soll, nicht mit einem P-Regler: dann wären die Voraussetzung für das Betragsoptimum erfüllt. Im Gegensatz zum Betragsoptimum hat hier der offene Regelkreis *doppelt integrales* Verhalten (einen Integrator in der Strecke, einen weiteren im Regler). Wegen des doppelten I-Anteils darf der PI-Regler *nicht* den Pol der Strecke kürzen:

$$G_R(s) = K_P \frac{T_I s + 1}{T_I s} = K_P \frac{T_1 s + 1}{T_1 s}, \quad G_S(s) = \frac{K}{s(T_1 s + 1)} \quad \rightarrow \quad G_0(s) = \frac{K_P K}{T_1 s^2}$$

$$\rightarrow G_w(s) = \frac{\frac{K_P K}{T_1 s^2}}{1 + \frac{K_P K}{T_1 s^2}} = \frac{K_P K}{T_1 s^2 + K_P K} \quad (\text{oszillatorisch grenzstabil!})$$

Die Nullstelle des Reglers muss also aggressiver gewählt werden, um mehr zur Stabilisierung beizutragen:

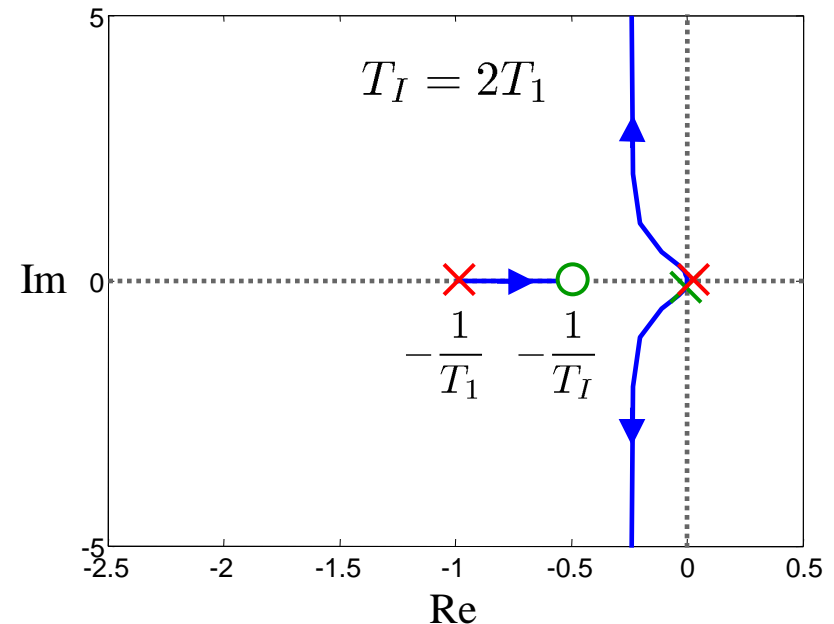
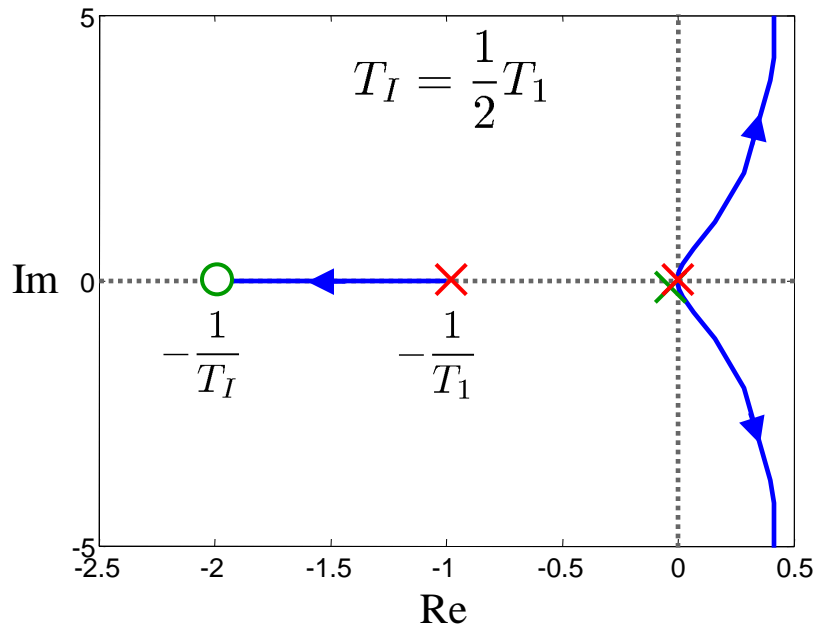
$$\rightarrow T_I > T_1$$



14.3 *Symmetrisches Optimum

Um uns den Zusammenhang zwischen der Reglernullstelle und der Stabilität des geschlossenen Regelkreises klar zu machen, schauen wir uns die WOK an.

- Ohne die Nullstelle des **Reglers** bei $-1/T_I$ und den Pol der **Strecke** bei $-1/T_1$ (bzw. bei einer perfekten Pol/Nullstellen-Kürzung) würde die WOK des Doppelintegrators die Im-Achse hoch- und runterlaufen \rightarrow oszillatorisch grenzstabil!.
- Da Pole die WOK-Äste abstoßen und Nullstellen die WOK-Äste anziehen, gelingt eine Stabilisierung nur, wenn $T_I > T_1$.



14.3 *Symmetrisches Optimum

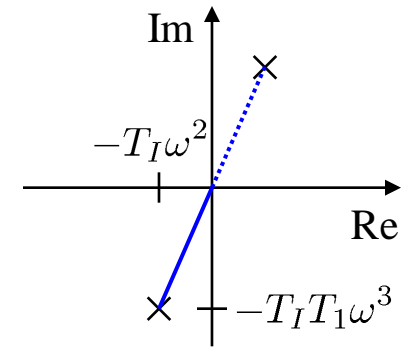
Wir wählen der bequemen Rechnung wegen:

$$T_I = a^2 T_1 \quad \text{mit } a > 1$$

Damit ist die Nachstellzeit T_I des PI-Reglers durch die Wahl von a festgelegt. Die Idee des symmetrischen Optimums ist es, die Reglerverstärkung K_P so zu bestimmen, dass die Phasenreserve (für ein gegebenes a) maximal wird. Berechnen wir also die Phase des *offenen* Regelkreises:

$$G_0(s) = K_P \underbrace{\frac{T_I s + 1}{T_I s}}_{\text{PI-Regler}} \underbrace{\frac{K}{s(T_1 s + 1)}}_{\text{IT}_1\text{-Strecke}} = K_P K \frac{T_I s + 1}{T_I s^2 (T_1 s + 1)}$$

$$G_0(i\omega) = K_p K \frac{1 + iT_I\omega}{-T_I\omega^2 - iT_I T_1 \omega^3}$$



weil im III. Quadranten

$$\varphi(\omega) = \arctan\left(\frac{T_I\omega}{1}\right) - \arctan\left(\frac{-T_I T_1 \omega^3}{-T_I \omega^2}\right) = \arctan(T_I\omega) - \arctan(T_1\omega) - 180^\circ$$

14.3 *Symmetrisches Optimum

$$\varphi(\omega) = \arctan(T_I \omega) - \arctan(T_1 \omega) - 180^\circ$$

Auch an der Gleichung für die Phase erkennt man, dass $T_I > T_1$ sein muss, da sonst die Phase immer unter -180° fallen würde, d.h. der Phasenrand negativ wäre und damit der Regelkreis instabil.

Wo ist die Phase maximal?

$$\frac{d\varphi(\omega)}{d\omega} = \frac{T_I}{1 + T_I^2 \omega^2} - \frac{T_1}{1 + T_1^2 \omega^2} = 0 \rightarrow \omega_{\max} = \frac{1}{\sqrt{T_I T_1}} = \frac{1}{a T_1} \quad \text{geometrisches Mittel von } T_I \text{ und } T_1$$

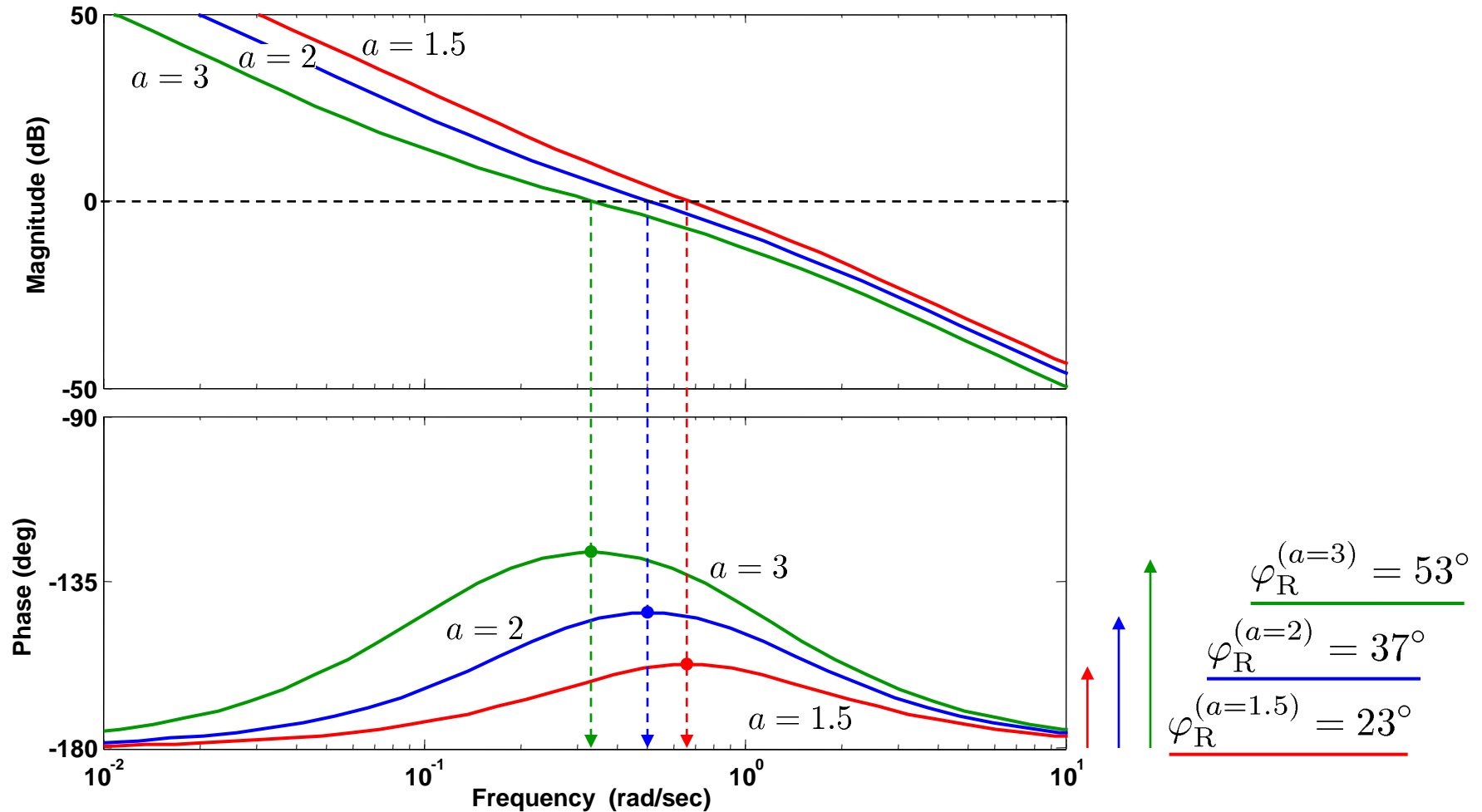
An dieser Frequenz soll der Amplitudengang gleich 1 sein. Dann ist nämlich definitionsgemäß bei dieser Frequenz auch die Phasenreserve abzulesen. Und die Phasenreserve ist dort natürlich maximal, weil ja die Phase dort maximal ist.

$$|G_0(i\omega_{\max})| = \frac{K_p K \sqrt{1 + T_I^2 \omega_{\max}^2}}{T_I \omega_{\max}^2 \sqrt{1 + T_1^2 \omega_{\max}^2}} = 1 \rightarrow K_P = \frac{1}{K \sqrt{T_I T_1}} = \frac{1}{K a T_1}$$

Mit der Wahl des Parameters a kann man nun die Größe der Phasenreserve beeinflussen. Aus a ergibt sich über obige Gleichungen direkt die Reglerverstärkung K_P und die Nachstellzeit T_I des PI-Reglers.

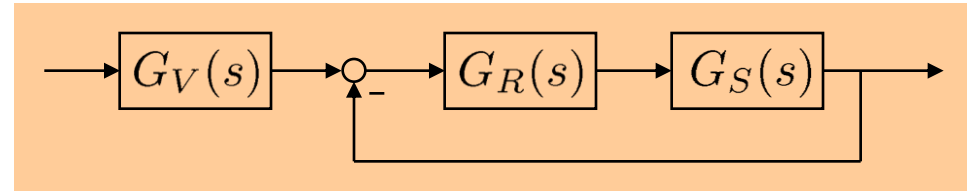
14.3 *Symmetrisches Optimum

Der Name "symmetrisches Optimum" kommt daher, dass ω_{\max} bei logarithmischer Skala genau in der Mitte zwischen $1/T_I$ und $1/T_1$ liegt (geometrisches Mittel). Hier $T_1 = 1$ sec.

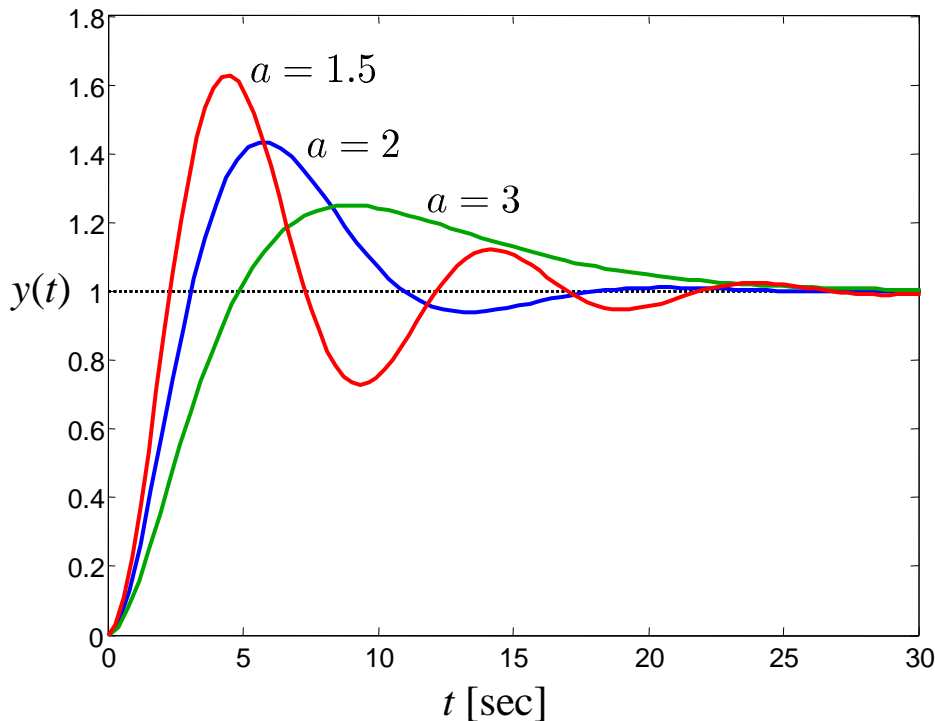


14.3 *Symmetrisches Optimum

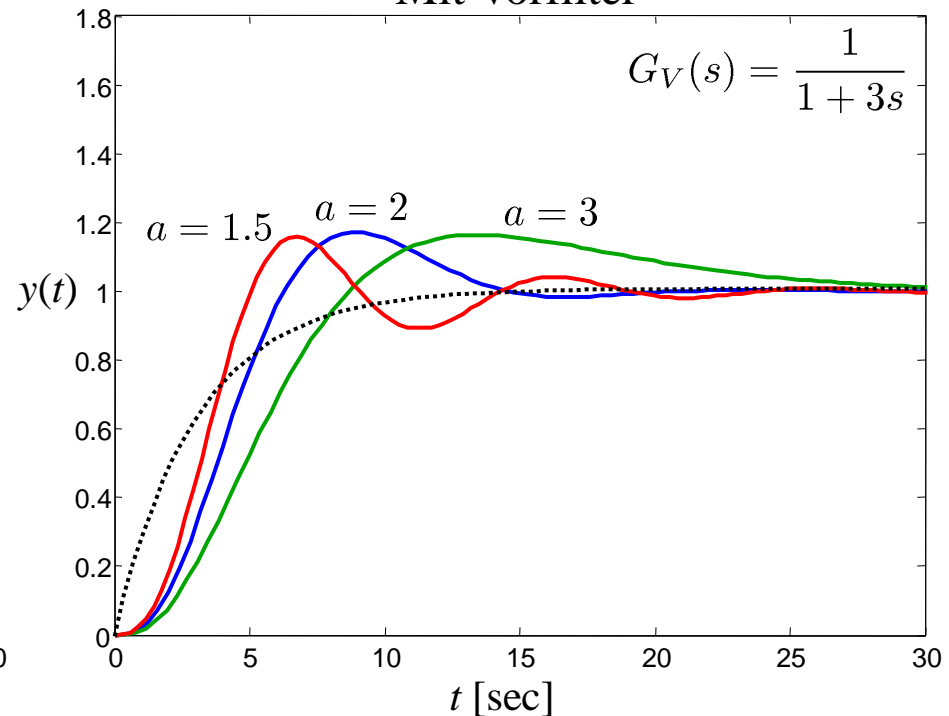
Die Sprungantworten des geschlossenen Regelkreises schwingen wegen des recht starken D-Anteils über, selbst wenn die Pole alle reell sind. Abhilfe schafft ein *Vorfilter* $G_V(s)$, das die Führungsgröße tiefpassfiltert.



Ohne Vorfilter



Mit Vorfilter



14.3 *Symmetrisches Optimum

Die Phasenreserve ergibt sich aus:

$$\varphi_R = 180^\circ - |\varphi(\omega_{\max})|$$

Da $\varphi(\omega_{\max}) < 0$ schreiben wir $|\varphi(\omega_{\max})| = -\varphi(\omega_{\max})$, also:

$$\begin{aligned}\varphi_R &= 180^\circ + \varphi(\omega_{\max}) = 180^\circ + \arctan(T_I \omega_{\max}) - \arctan(T_1 \omega_{\max}) - 180^\circ \\ &= \arctan\left(\frac{T_I}{a T_1}\right) - \arctan\left(\frac{T_1}{a T_1}\right) = \arctan(a) - \arctan\left(\frac{1}{a}\right)\end{aligned}$$

Umformen ergibt folgenden Zusammenhang zwischen a und der Phasenreserve φ_R :

$$\varphi_R = \arctan\left(\frac{a^2 - 1}{2a}\right)$$

$$a = \frac{1 + \sin(\varphi_R)}{\cos(\varphi_R)}$$

Die Vorgehensweise beim Reglerentwurf könnte also mit der Vorgabe einer gewünschten Phasenreserve starten:

$$\varphi_R \rightarrow a \rightarrow K_P, T_I$$

14.3 *Symmetrisches Optimum

Alternativ zur Vorgabe einer gewünschten Phasenreserve könnte man zur Festlegung von a auch nach der Idee des Betragsoptimums vorgehen. Dazu berechnen wir die Führungsübertragungsfunktion, zuvor aber zunächst den offenen Kreis:

$$G_R(s) = K_P \frac{T_I s + 1}{T_I s}, \quad G_S(s) = \frac{K}{s(T_1 s + 1)} \quad \rightarrow \quad G_0(s) = K_P K \frac{T_I s + 1}{T_I s^2 (T_1 s + 1)}$$

Mit den Reglerparametern nach dem symmetrischen Optimum ergibt sich:

$$G_0(s) = \frac{1}{K a T_1} K \frac{a^2 T_1 s + 1}{a^2 T_1 s^2 (T_1 s + 1)} = \frac{a^2 T_1 s + 1}{a^3 T_1^2 s^2 (T_1 s + 1)}$$

Daraus berechnet sich die Führungsübertragungsfunktion zu:

$$G_w(s) = \frac{a^2 T_1 s + 1}{a^3 T_1^2 s^2 (T_1 s + 1) + a^2 T_1 s + 1} = \frac{a^2 T_1 s + 1}{a^3 T_1^3 s^3 + a^3 T_1^2 s^2 + a^2 T_1 s + 1}$$

Zur Vereinfachung wird jetzt für die Forderung $|G_w(i\omega)| \approx 1$ nur noch der Nenner betrachtet. Der Zähler mit recht stark differenzierendem Verhalten führt zu starkem Überschwingen und wird durch einen Tiefpass-Vorfilter, mit dem die Führungsgröße gefiltert wird, entschärft.

14.3 *Symmetrisches Optimum

Das Quadrat des Amplitudengangs (mit Zähler = 1) ergibt sich zu:

$$\begin{aligned} |\tilde{G}_w(i\omega)|^2 &= \frac{1}{(1 - a^3 T_1^2 \omega^2)^2 + (a^2 T_1 \omega - a^3 T_1^3 \omega^3)^2} \\ &= \frac{1}{1 + a^3 T_1^2 \omega^2 (a - 2) + a^5 T_1^4 \omega^4 (a - 2) + a^6 T_1^6 \omega^6} \end{aligned}$$

Man kommt dem idealen Amplitudengang von 1 am nächsten, wenn man wählt:

$$a = 2$$

Dann erhält man aus den Gleichungen nach dem symmetrischen Optimum:

$$K_P = \frac{1}{2KT_1}$$

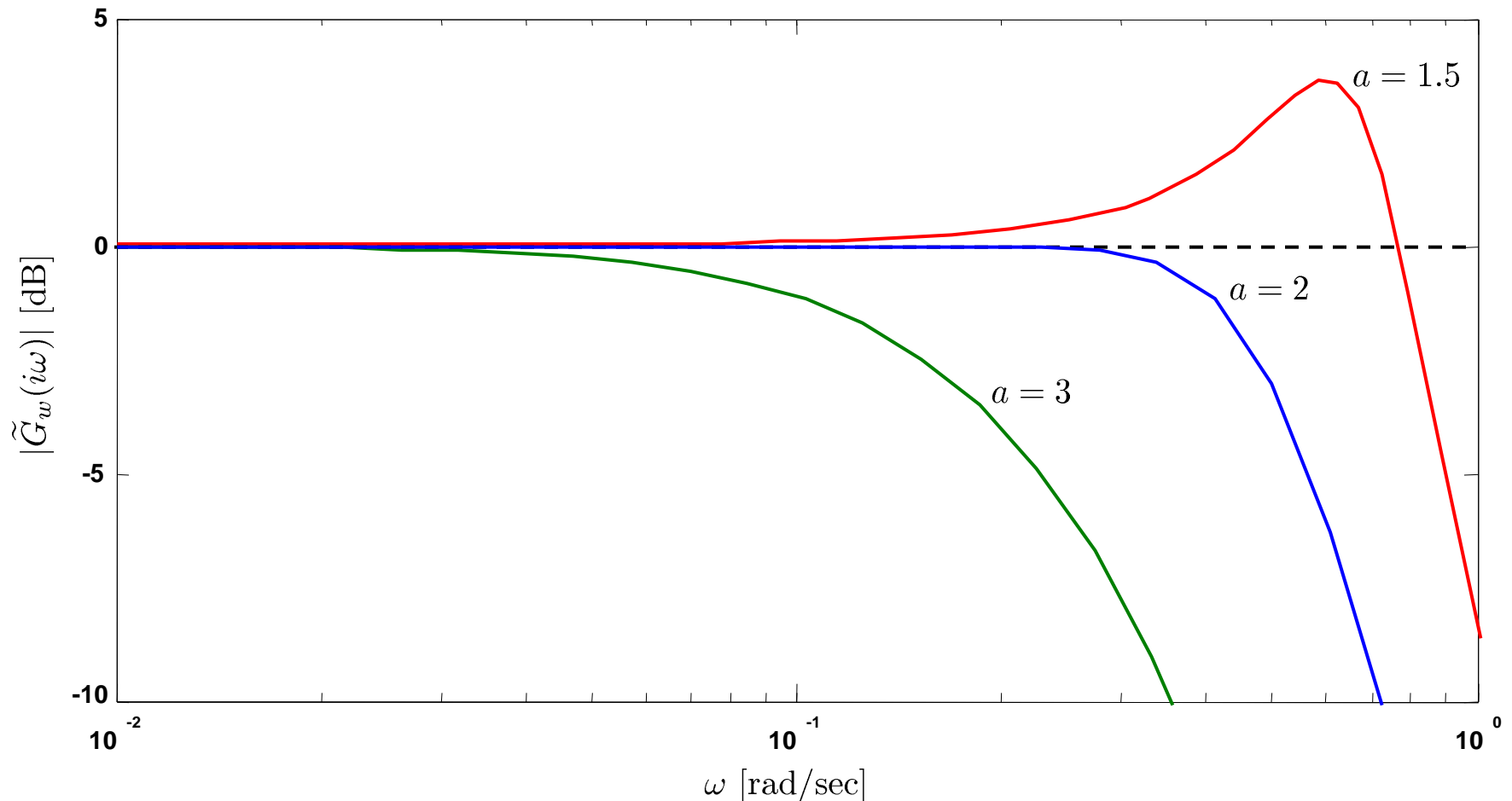
$$T_I = 4T_1$$

Diese Einstellung führt auf eine Phasenreserve von:

$$\varphi_R = 37^\circ$$

14.3 *Symmetrisches Optimum

Vergleich des Betrags der (vereinfachten) Führungsübertragungsfunktion für $T_1 = 1$ sec und für verschiedene Werte von a . Für $a = 2$ ist $|\tilde{G}_w(i\omega)| \approx 1$ am besten erfüllt!



14.3 *Symmetrisches Optimum

Das Verfahren des symmetrischen Optimums lässt sich von der IT_1 -Strecke mit PI-Regler (offener Kreis hat dann I_2T_1 -Verhalten) auf weitere Regelkreisstrukturen erweitern.

IT_n -Strecke mit PID-Regler

Die Strecke hat eine dominante Zeitkonstante T_1 . Die im Vergleich dazu kleinen anderen Zeitkonstanten werden zu T_E zusammengefasst (aufsummiert). Es gilt $T_E \ll T_1$:

$$G_S(s) \approx \frac{K}{s(1 + T_1 s)(1 + T_E s)}, \quad G_R(s) = K_P \frac{(1 + T_I s)(1 + T_D s)}{T_I s}$$

Der PID-Regler kann den dominanten Streckenpol verkürzen. Für $K_P = \frac{1}{K a T_E}$ und $T_I = a^2 T_E$ und $T_D = T_1$ ergibt sich wieder ein offener Kreis mit I_2T_1 -Verhalten:

$$\rightarrow G_0(s) = K_P K \frac{(1 + T_I s)}{T_I s^2 (1 + T_E s)} = \frac{1 + a^2 T_E s}{a^3 T_E^2 s^2 (1 + T_E s)}$$

14.3 *Symmetrisches Optimum

PT_n-Strecke mit PI-Regler

Hat eine PT_n-Strecke einen dominanten Pol $T_1 \gg T_E$, so kann dieser wie ein Integrator behandelt werden. Mit $\tilde{K} = K/T_1$ kann man dann approximieren:

$$G_S(s) = \frac{K}{(1 + T_1 s)(1 + T_E s)} \approx \frac{K}{T_1 s(1 + T_E s)} = \frac{\tilde{K}}{s(1 + T_E s)}$$

Mit einem PI-Regler ergibt sich nun das Standardproblem des symmetrischen Optimums.

PT_n-Strecke mit PID-Regler

Hat eine PT_n-Strecke zwei dominante Pole $T_1 > T_2 \gg T_E$, so kann der dominantere von beiden wie ein Integrator behandelt werden (s.o.) und der zweite kann durch den Zählerfaktor $(1 + T_D s)$ eines PID-Reglers mit $T_D = T_2$ weggekürzt werden.

$$G_S(s) = \frac{K}{(1 + T_1 s)(1 + T_2 s)(1 + T_E s)} \approx \frac{\tilde{K}}{s(1 + T_2 s)(1 + T_E s)}$$

Nach der Pol/Nullstellen-Kürzung ergibt sich wieder das Standardproblem des symmetrischen Optimums aus IT₁-Strecke und PI-Regler(-Rest).

15. Vertiefungen und Erweiterungen des Standardregelkreises

Inhalt Kapitel 15

15. Vertiefungen und Erweiterungen des Standardregelkreises

15.1 Grenzen der Regelung

15.2 Vorsteuerung

15.3 Störgrößenaufschaltung

15.4 Regelkreis mit Hilfsstellgrößen und Hilfsregelgrößen

15.5 Kaskadenregelung

15.6 Mehrgrößenregelung

15.7 Anti-Wind-up-Methoden

15.8 Stoßfreies Umschalten (Bumpless Transfer)

15.9 Smith-Prädiktor

15.10 Internal Model Control

15.11 Prädiktive Regelung

15.1 Grenzen der Regelung

Empfindlichkeit des geschlossenen Regelkreises auf Änderungen der Regelstrecke

Für die Beurteilung der Empfindlichkeit bzw. Robustheit von Regelungen ist entscheidend wie stark sich Änderungen der Regelstrecke auf das Verhalten des geschlossenen Regelkreises auswirken.

Eine Aufgabe der Rückkopplung besteht darin, den Einfluss solcher Streckenänderungen abzuschwächen, den Regelkreis also robust zu machen.

Untersuchen wir, wie sich eine Änderung der Regelstrecke dG_S auf das Verhalten des geschlossenen Regelkreises G_w auswirkt:

$$\frac{dG_w}{dG_S} = \frac{G_R(1 + G_R G_S) - (G_R G_S)G_R}{(1 + G_R G_S)^2} = \frac{G_R}{(1 + G_R G_S)^2}$$

Erinnerung:

$$G_w = \frac{G_R G_S}{1 + G_R G_S}$$

Betrachten wir, wie im Bode-Diagramm, die Logarithmen der Übertragungsfunktionen:

$$\frac{d \log G_w}{d \log G_S} = \frac{\frac{dG_w}{G_w}}{\frac{dG_S}{G_S}} = \frac{dG_w}{dG_S} \frac{G_S}{G_w}$$

denn es gilt:

$$\frac{d \log x}{dx} = \frac{1}{x} \rightarrow d \log x = \frac{dx}{x}$$

15.1 Grenzen der Regelung

Daraus ergibt sich:

$$\frac{d \log G_w}{d \log G_S} = \frac{G_R}{(1 + G_R G_S)^2} \frac{G_S}{G_w} = \frac{G_R}{(1 + G_R G_S)^2} \frac{G_S(1 + G_R G_S)}{G_R G_S} = \frac{1}{1 + G_R G_S} = \frac{1}{1 + G_0}$$

Diese Übertragungsfunktion beschreibt wie stark sich eine Änderung der (logarithmierten) Übertragungsfunktion der Regelstrecke auf die (logarithmierte) Übertragungsfunktion des geschlossenen Regelkreises auswirkt. Man nennt sie daher

Empfindlichkeitsfunktion (*sensitivity function*): $S(s) = \frac{1}{1 + G_0(s)} = G_d(s)$

Als Gegenstück dazu definiert man die

komplementäre Empfindlichkeitsfunktion: $T(s) = \frac{G_0(s)}{1 + G_0(s)} = G_w(s)$

Es gilt daher immer:

$$T(s) + S(s) = 1$$

bzw.

$$G_w(s) + G_d(s) = 1$$

ACHTUNG! Das bedeutet *nicht*: $|T(s)| + |S(s)| = 1$

15.1 Grenzen der Regelung

Identische Forderungen für gutes Führungs- und Störverhalten

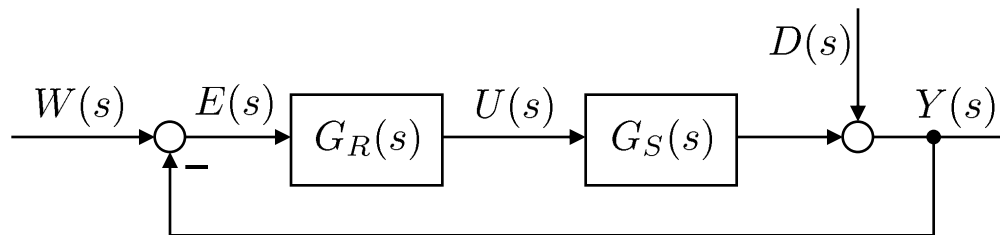
Interpretation der Empfindlichkeitsfunktion

Die Empfindlichkeitsfunktion $S(s)$ beschreibt viele wichtige Zusammenhänge im Regelkreis:

- $S(s)$ ist identisch mit der Störübertragungsfunktion $G_d(s)$: $d \rightarrow y$.
- $S(s)$ ist identisch mit der Übertragungsfunktion von Führungsgröße zu Regelabweichung: $w \rightarrow e$.
- $S(s)$ ist der Faktor, der die Übertragungsfunktion des offenen Kreises $G_0(s)$ von der Übertragungsfunktion des geschlossenen Kreises $G_w(s)$ unterscheidet:

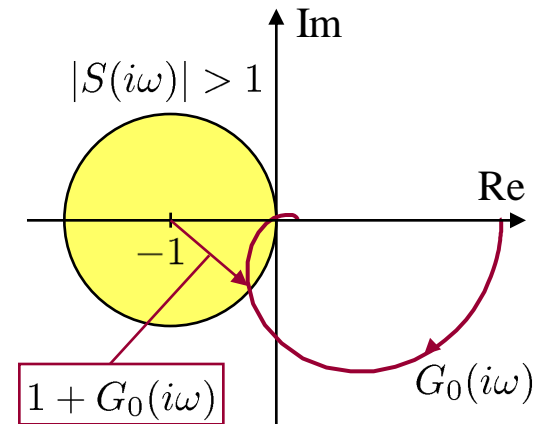
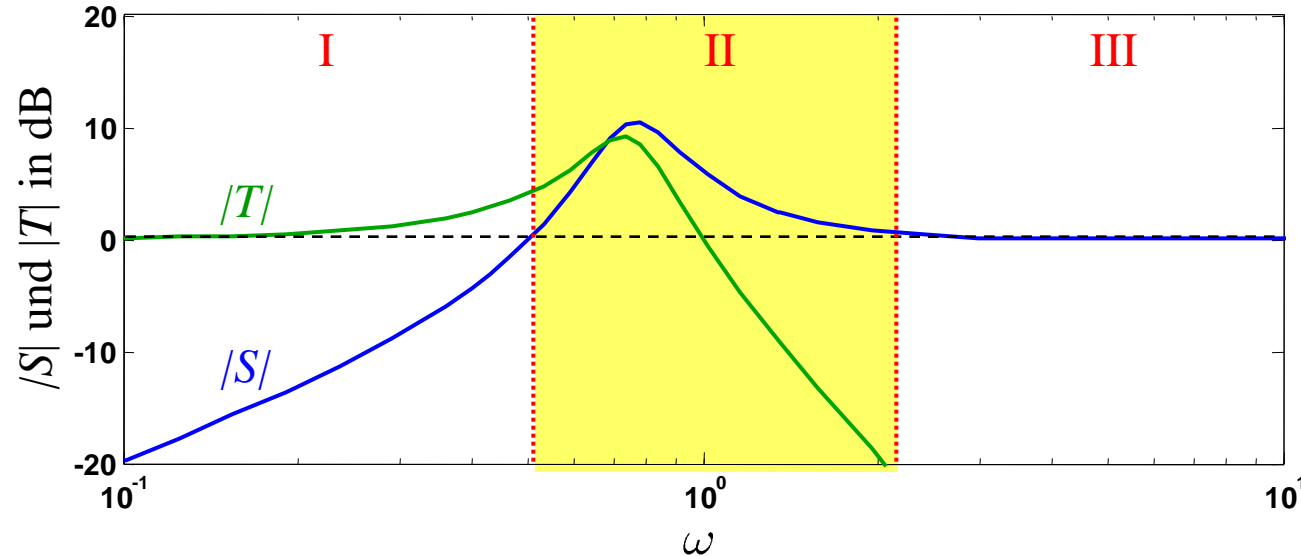
$$G_w(s) = G_0(s)S(s) = \frac{G_0(s)}{1 + G_0(s)}$$

Es entspricht dem "Anteil", der durch die Rückkopplung erzeugt wird, also der Wirkung der Regelung im Vergleich zur Steuerung. Daher wird $S(s)$ auch **dynamischer Regelfaktor** genannt.



15.1 Grenzen der Regelung

Frequenzdarstellung von S und T



Man unterscheidet 3 Frequenzbereiche:

- I. *Gegenkopplungsbereich*: $|S(i\omega)| < 1$. Die Rückkopplung wirkt wie gewünscht. Störungen werden unterdrückt, Führungsgrößen wird gefolgt, Einfluss von Streckenänderungen wird reduziert.
- II. *Mitkopplungsbereich*: $|S(i\omega)| > 1$. Die Rückkopplung wirkt kontraproduktiv.
- III. *Unempfindlichkeitsbereich*: $|S(i\omega)| \approx 1$. Die Rückkopplung wirkt gar nicht.

15.1 Grenzen der Regelung

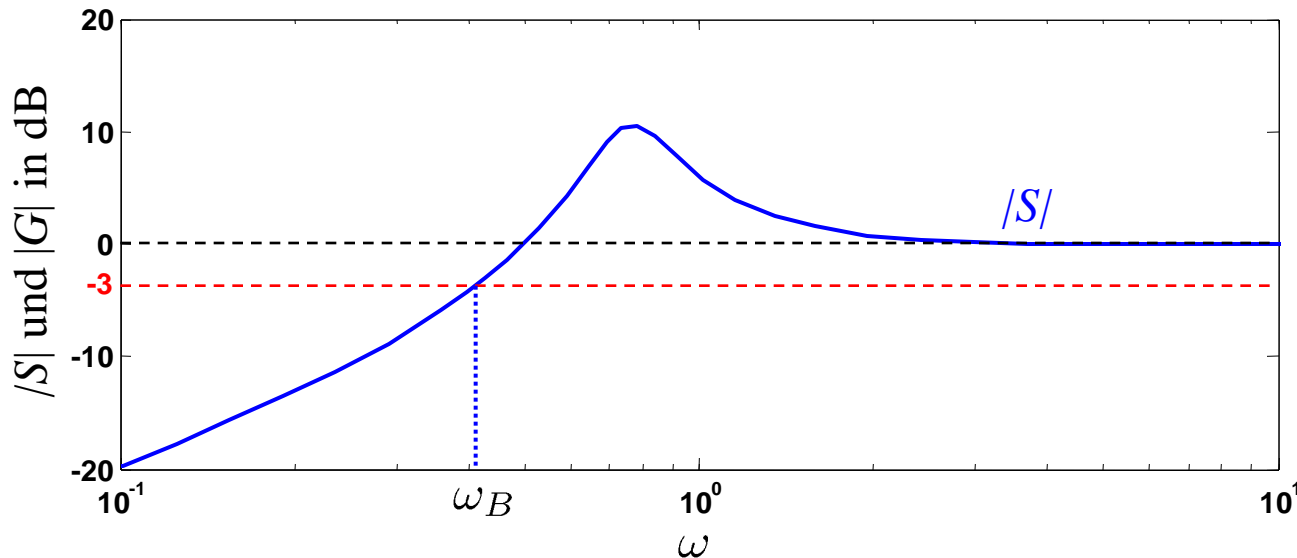
Bandbreite des Regelkreises

Der Regelkreis ist nur im Gegenkopplungsbereich im gewünschten Sinne aktiv. Wir fordern eine Verbesserung der Störunterdrückung im Vergleich zum unregulierten Kreis um mindestens den Faktor $3 \text{ dB} = \sqrt{2}$:

$$|S(i\omega)| < \frac{1}{\sqrt{2}}$$

Die obere Grenze ω_B des Frequenzbereichs, für den diese Forderung erfüllt ist, nennt man **Bandbreite** (des geschlossenen Regelkreises).

Alternativ kann man auch die Phasendurchtrittsfrequenz ω_{-180° als Bandbreite nehmen.



15.1 Grenzen der Regelung

Gleichgewichtstheorem

Man würde sich wünschen, den Bereich I möglichst groß zu machen und den Bereich II möglichst zu vermeiden. Leider ist dies i.A. nicht machbar. Der durch die Rückkopplung erzielte Gewinn (Bereich I) muss durch einen Verlust an anderer Stelle (Bereich II) erkauft werden (no free lunch!).

Das von Bode hergeleitete **Gleichgewichtstheorem** beschreibt diesen Zusammenhang:

$$\int_0^{\infty} \log|S(i\omega)|d\omega = 0$$

Voraussetzungen:

- System stabil
- Polüberschuss ≥ 2

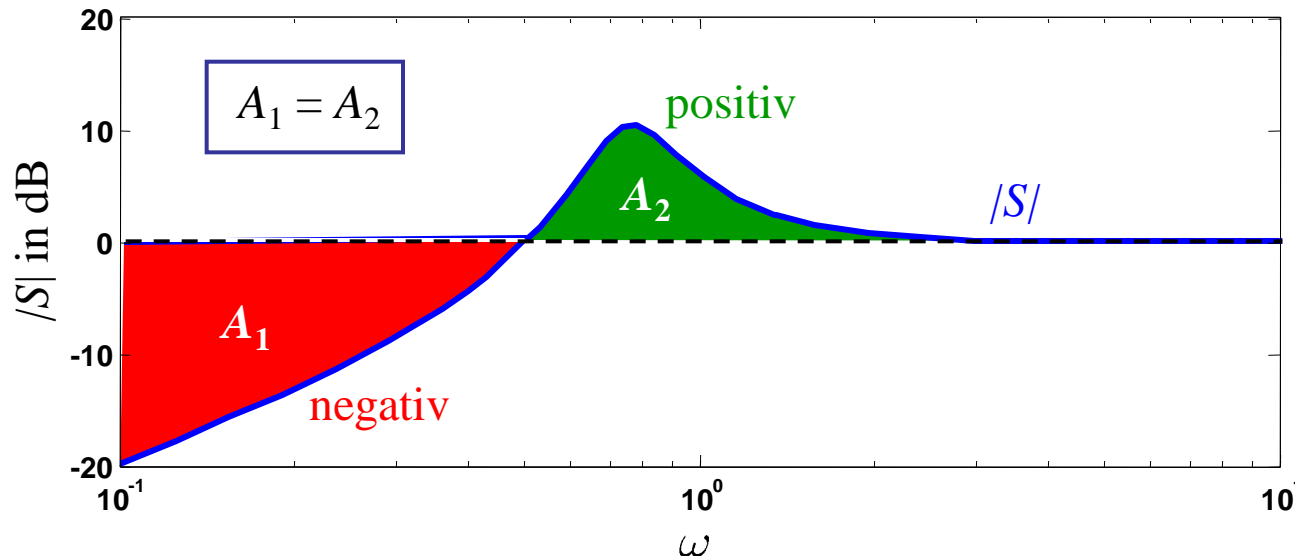
In der Praxis bedeutet die Voraussetzung, dass das System $G_0(s)$ mindestens Polüberschuss 2 (also Zählerordnung \leq Nennerordnung $- 2$) haben muss, keine wirkliche Einschränkung, da $G_0(s)$ aus Regler, Aktor, Strecke und Sensor besteht und jedes der letzten drei Elemente normalerweise einen Polüberschuss von mindestens 1 aufweist, wenn nicht die schnellen Dynamikanteile vernachlässigt werden.

15.1 Grenzen der Regelung

Wasserbett-Effekt

Das Gleichgewichtstheorem sagt aus, dass die Fläche zwischen der "0 dB"-Linie und $|S(i\omega)|$ unterhalb (A_1) der "0 dB"-Linie gleich der Fläche oberhalb (A_2) ist. (Durch die logarithmische Skala der Frequenz ist die Flächengleichheit optisch nicht klar zu erkennen.)

D.h. jede Verbesserung im Gegenkopplungsbereich führt zu einer Verschlechterung im Mitkopplungsbereich. Allerdings kann man versuchen, den Gegenkopplungsbereich auf alle (in der Stör- und Führungsgröße enthaltenen) relevanten Frequenzen auszudehnen; dann liegt der Mitkopplungsbereich sehr hochfrequent und spielt praktisch eine untergeordnete Rolle.



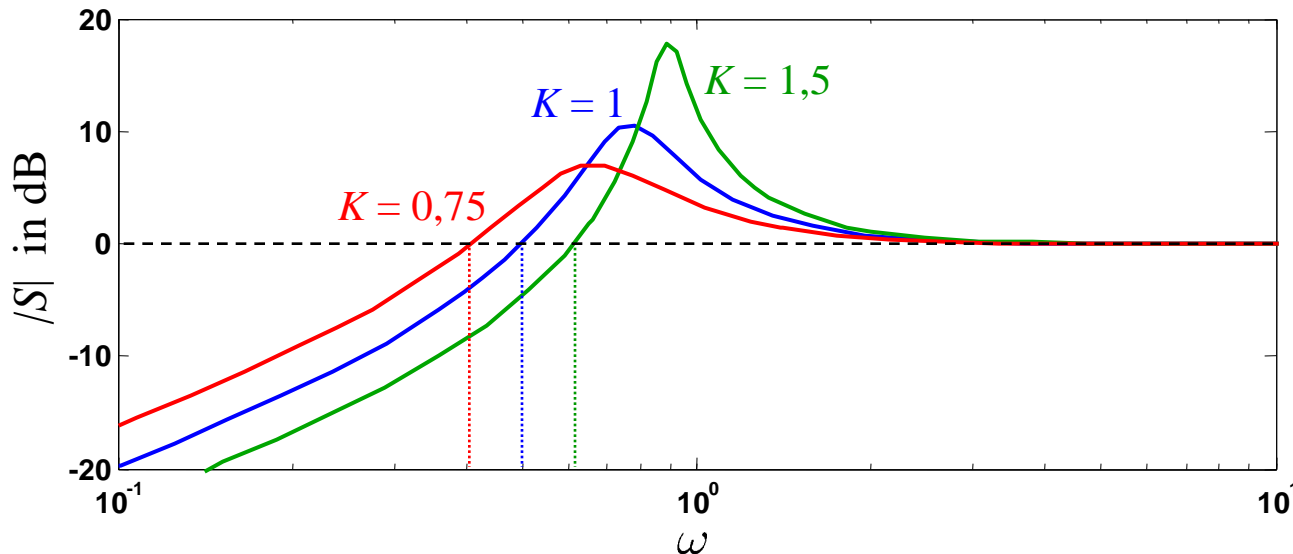
15.1 Grenzen der Regelung

Illustration des Wasserbett-Effekts

Größere Kreisverstärkungen führen zu größeren Bandbreiten, d.h. die Regelung ist bis zu höheren Frequenzen wirksam (Schnittpunkt von $|S(i\omega)|$ mit der "0 dB"-Line) und damit schneller.

Dadurch wird aber die Fläche unter der "0 dB"-Line vergrößert. Das führt wegen des Wasserbett-Effekts zu einer größeren Überhöhung der Empfindlichkeitsfunktion (Fläche über der "0 dB"-Line wird auch größer).

Die maximale Empfindlichkeit steigt mit K an; der Regelkreis verliert Robustheit.



15.1 Grenzen der Regelung

Beispiel für den Wasserbett-Effekt

Betrachten wir einen Prozess 2. Ordnung mit Verstärkung K und einer schnellen und einer langsamen Zeitkonstante T_1 und T_2 :

$$G_S(s) = \frac{K}{(1 + T_1s)(1 + T_2s)}$$

Als Regler nehmen wir einen PI-Regler, der die Prozessverstärkung und die langsame Zeitkonstante des Prozesses wegekürzt, um den Regelkreis zu beschleunigen:

$$G_R(s) = \frac{K_R}{K} \frac{1 + T_2s}{s}$$

Somit ist der offene Regelkreis und auch die Empfindlichkeitsfunktion nur noch von der schnellen Prozesszeitkonstante und der Reglerverstärkung abhängig:

$$G_0(s) = G_R(s)G_S(s) = \frac{K_R}{s(1 + T_1s)}$$

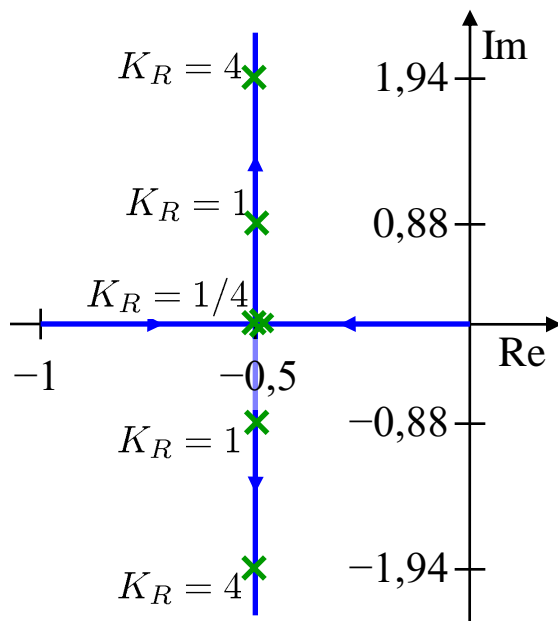
$$S(s) = \frac{1}{1 + G_0(s)} = \frac{1}{1 + \frac{K_R}{s(1 + T_1s)}} = \frac{T_1s^2 + s}{T_1s^2 + s + K_R}$$

$$T(s) = \frac{K_R}{T_1s^2 + s + K_R}$$

15.1 Grenzen der Regelung

Beispiel für den Wasserbett-Effekt: WOK

Als einfaches Zahlenbeispiel wählen wir $T_1 = 1$ sec und vergleichen verschiedene Reglerverstärkungen miteinander. Die tiefsten Einsichten über die Abhängigkeit des Regelkreisverhaltens von der Reglerverstärkung K_R liefert uns die Wurzelortskurve (WOK):



Bis zu einer Verstärkung von $1/4$ ergeben sich zwei reelle Pole, bei $K_R = 1/4$ entsteht ein reeller Doppelpol bei $-0,5$ (aperiodischer Grenzfall). Für größere Reglerverstärkungen fängt der Regelkreis an zu schwingen, und zwar mit immer geringerer Dämpfung (größer werdender Im-Anteil) je weiter K_R erhöht wird.

Theoretisch ist der Regelkreis strukturstabil, da die Verstärkung unendlich groß werden kann und die Pole immer in der linken Halbebene liegen.

In der Praxis wird wegen Modellierungsungenauigkeiten dennoch irgendwann instabiles Verhalten entstehen. Außerdem können wegen stets vorhandenen (aber bei linearen Systemen vernachlässigten) Stellgrößenbeschränkungen, sehr große Reglereingriffe gar nicht mehr praktisch umgesetzt werden.

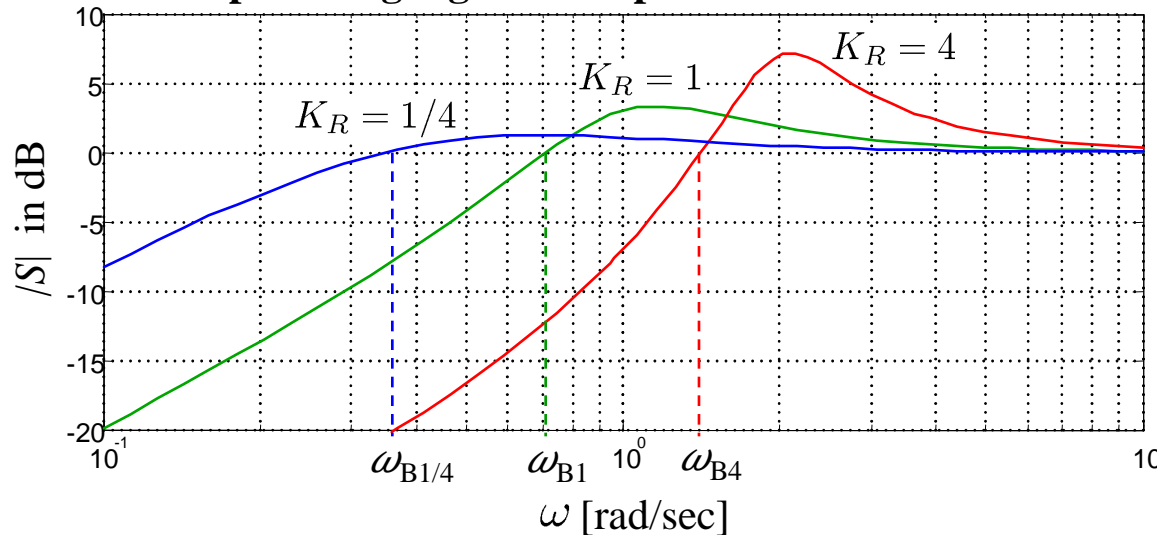
15.1 Grenzen der Regelung

Beispiel für den Wasserbett-Effekt: Empfindlichkeitsfunktionen und Sprungantworten

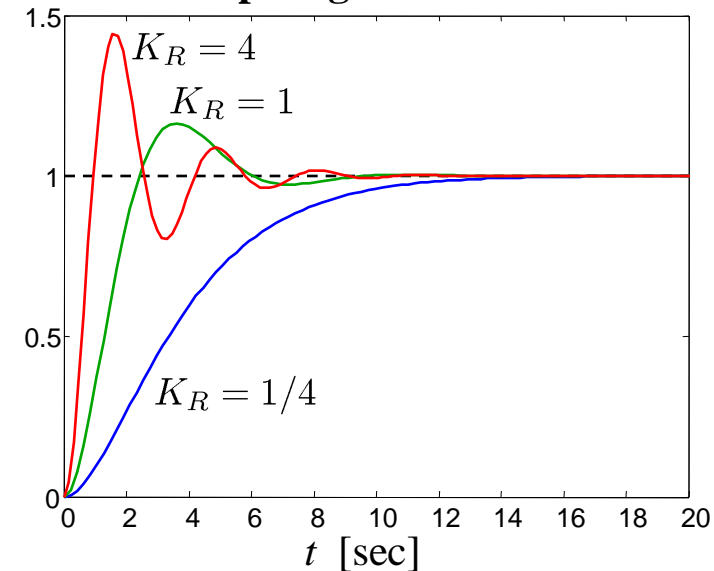
Mit steigender Reglerverstärkung K_R passiert folgendes:

- Die Bandbreite (bis hierhin wird eine Störung abgeschwächt) des Regelkreises steigt an: $\omega_{B1/4} = 0,35 \text{ rad/sec}$, $\omega_{B1} = 0,7 \text{ rad/sec}$, $\omega_{B4} = 1,5 \text{ rad/sec}$.
- Der Regelkreis wird schneller.
- Der Regelkreis schwingt stärker.
- Der Regelkreis wird weniger robust (siehe nächste Seite).

Amplitudengänge der Empfindlichkeitsfunktionen



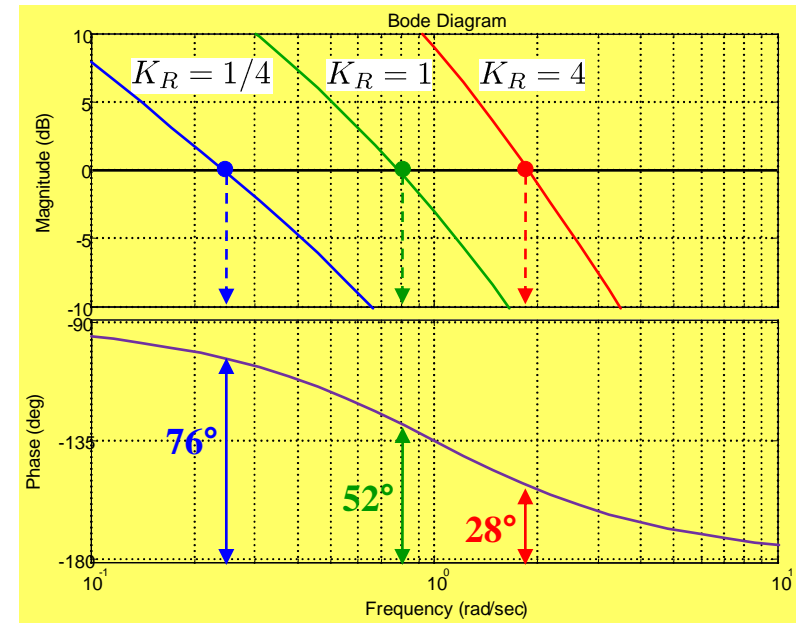
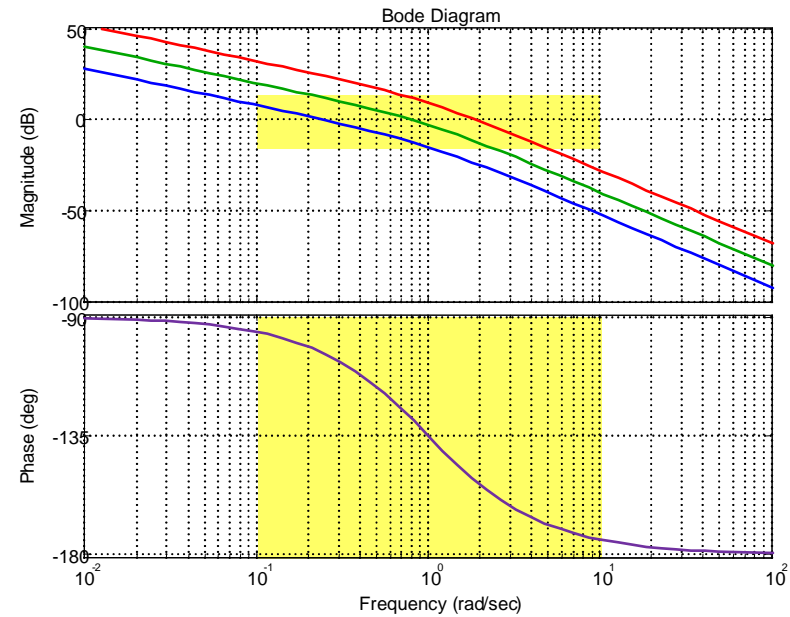
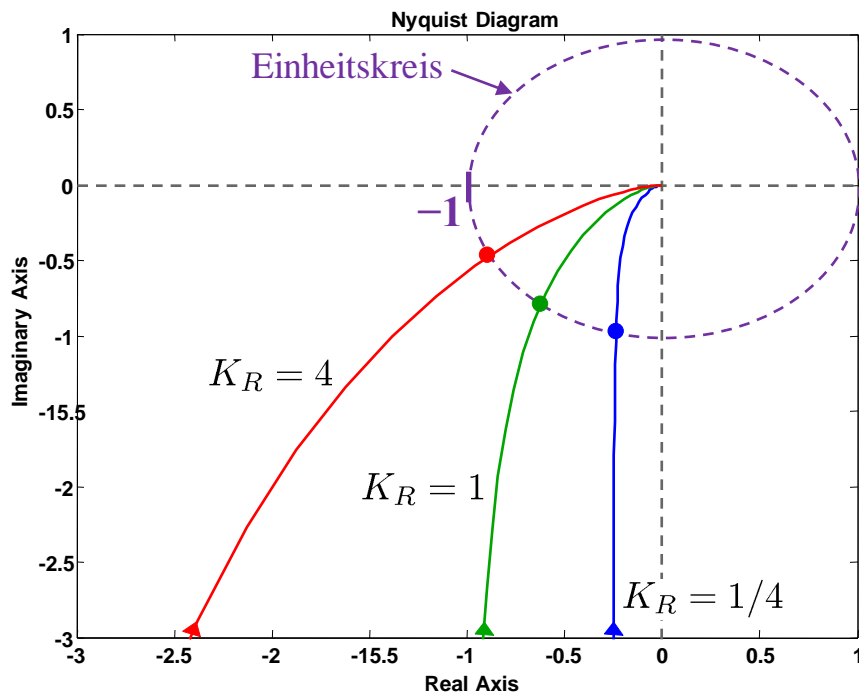
Sprungantworten



15.1 Grenzen der Regelung

Beispiel für den Wasserbett-Effekt: Robustheit

Der Nyquist-Plot, d.h. die Ortskurve des offenen Regelkreises $G_0(i\omega)$, als auch dessen Frequenzgang zeigen wie der Phasenrand und damit die Robustheit des Regelkreises mit steigender Reglererstärke K_R abnimmt.

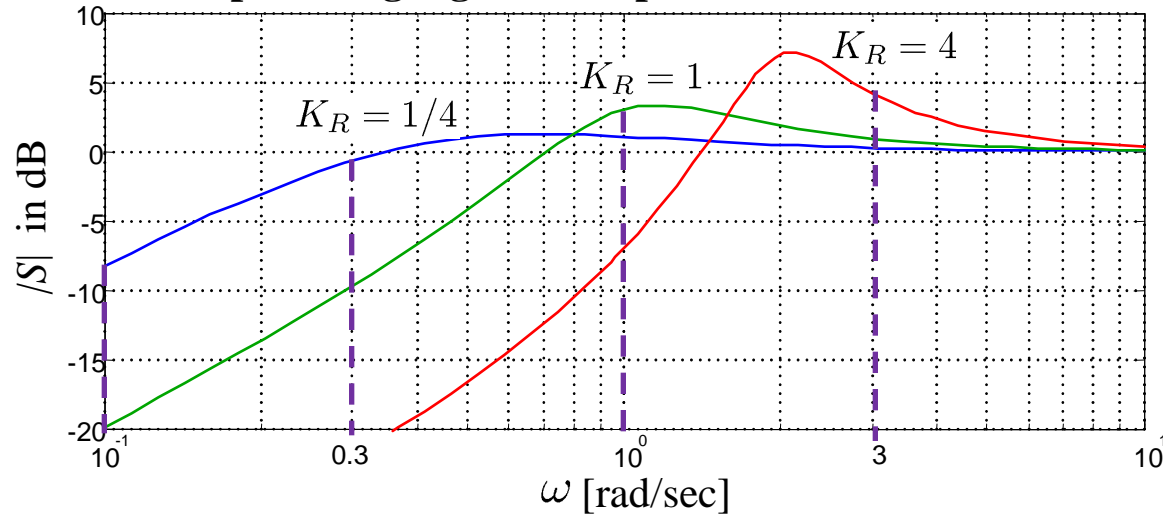


15.1 Grenzen der Regelung

Beispiel: Regelkreisverhalten auf verschiedenen Frequenzen

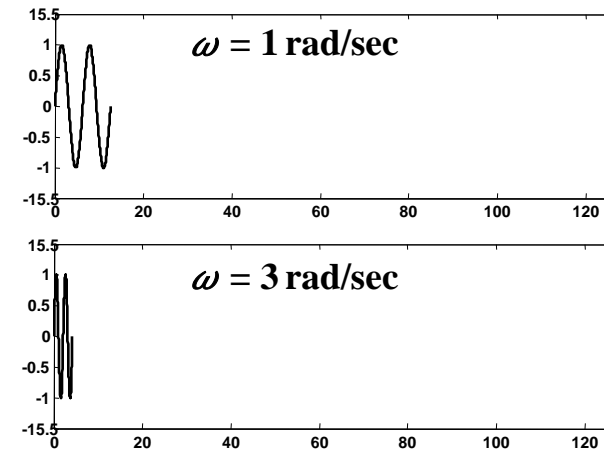
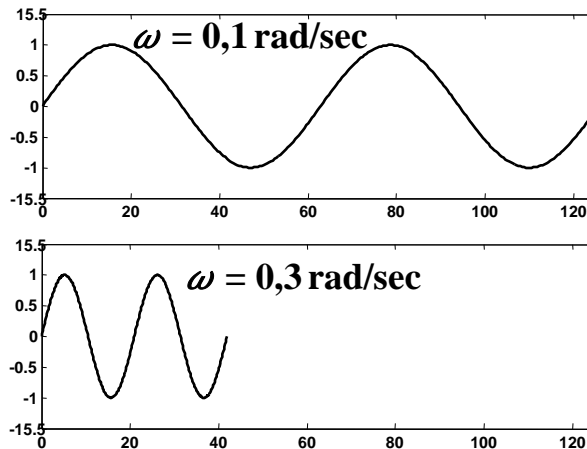
Der Regelkreis wird mit einer sinusförmigen Störung der Amplitude 1 am Ausgang gestört. Wie reagiert der Regelkreis in Abhängigkeit von der Frequenz der Störung? Die Dämpfung der Störung lässt sich direkt aus der Empfindlichkeitsfunktion ablesen!

Amplitudengänge der Empfindlichkeitsfunktionen



Störsignale:

$$u(t) = \sin \omega t$$

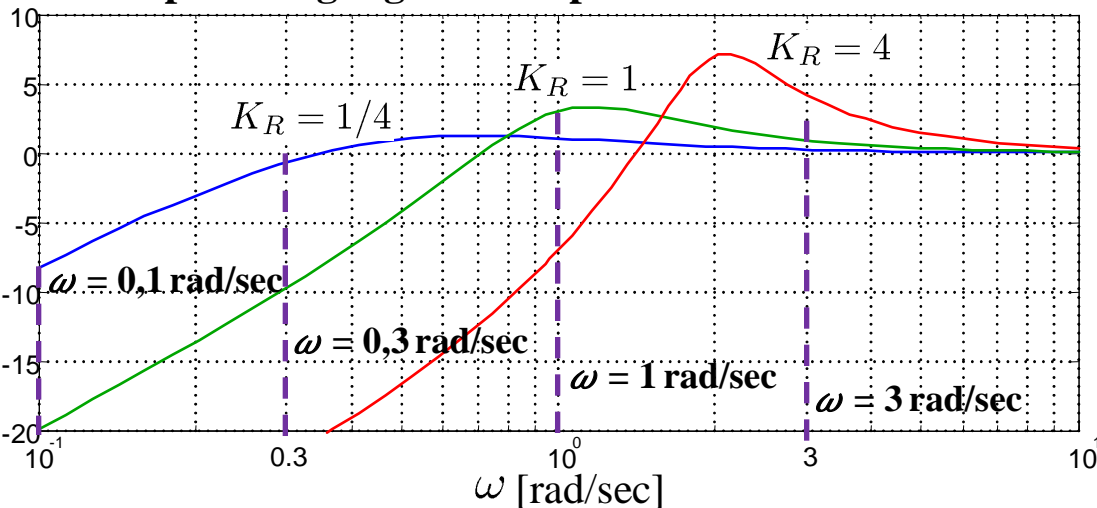


15.1 Grenzen der Regelung

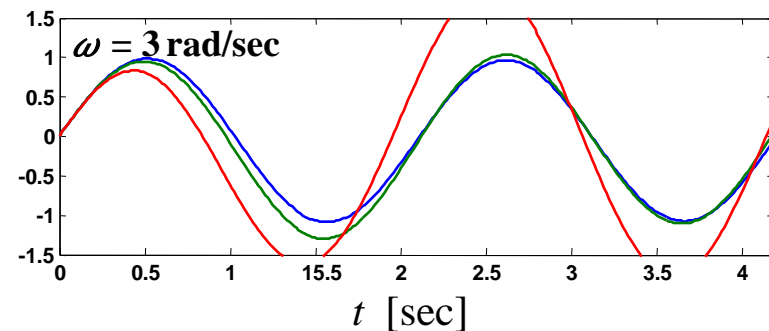
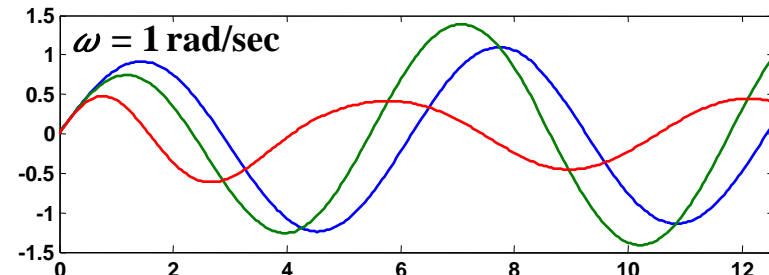
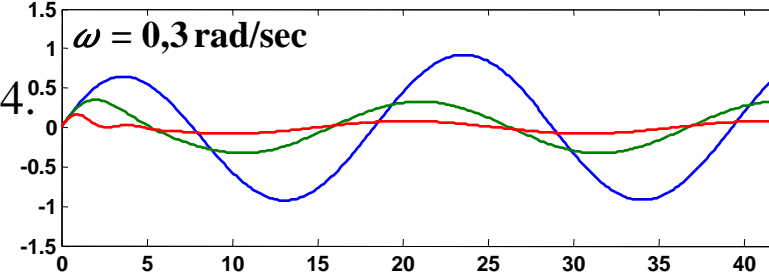
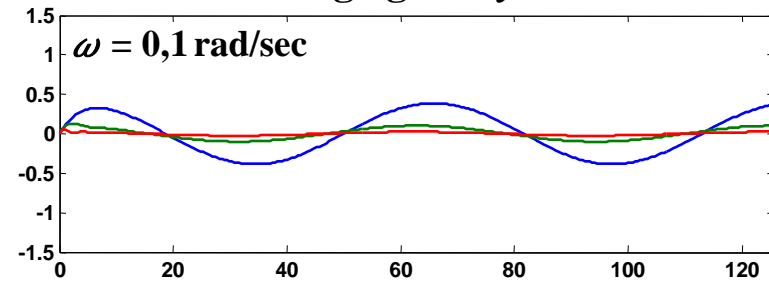
Beispiel: Störunterdrückung bei verschiedenen Frequenzen

- $\omega = 0,1 \text{ rad/sec}$: Regelung funktioniert sehr gut, dämpft selbst bei $K_R = 1/4$ um mehr als Faktor 2.
- $\omega = 0,3 \text{ rad/sec}$: Dämpfung nur noch für $K_R = 1$ u. 4.
- $\omega = 1 \text{ rad/sec}$: Dämpfung nur noch für $K_R = 4$.
- $\omega = 3 \text{ rad/sec}$: Keine Dämpfung für $K_R = 1/4$ u. 1, für $K_R = 4$ kontraproduktiv.

Amplitudengänge der Empfindlichkeitsfunktionen



Regelgröße y



15.1 Grenzen der Regelung

Gleichgewichtstheorem für instabile Systeme

Bodes Gleichgewichtstheorem wurde später auf instabile Systeme erweitert:

$$\int_0^{\infty} \log |S(i\omega)| d\omega = \pi \sum_{i=1}^{n_i} \operatorname{Re}\{p_i\}$$

Voraussetzungen:

- n_i instabile Pole p_i
- Polüberschuss ≥ 2

Für $n_i = 0$ instabile Pole geht diese Gleichung in die Formel für stabile Systeme über. Je mehr instabile Pole vorliegen und desto schneller (größerer Realteil) diese Pole sind, desto weiter verschiebt sich also die Empfindlichkeitsfunktion nach oben. D.h. instabile Pole kosten prinzipiell Regelgüte.

Auch instabile Nullstellen schränken die Möglichkeiten der Regelung ein. Die Formeln sind zu kompliziert, um sie hier zu diskutieren. Interessant ist aber: Eine Strecke ist um zu schwerer zu regeln, je näher sich instabile Pole und instabile Nullstellen kommen.

15.1 Grenzen der Regelung

Instabile Pole im offenen Kreis $G_0(s)$

Die Bandbreite des geschlossenen Regelkreises sollte immer größer sein als der schnellste instabile Pol des offenen Kreises. Ansonsten reagiert die Regelung zu langsam um der instabilen Eigendynamik gegenzuhalten.

Das heißt auch, dass es für instabile Systeme nicht (nur) eine kritische *maximale Verstärkung* gibt, *ab* welcher der Regelkreis instabil wird. Es gibt auch eine kritische *minimale Verstärkung*, *bis* zu welcher der Regelkreis instabil ist.

Instabile Nullstellen im offenen Kreis $G_0(s)$

Nicht phasenminimales Verhalten (Totzeiten und/oder instabile Nullstellen) schränkt die Bandbreite des Regelkreises ein, weil durch die zusätzliche Phasenabsenkung im Vergleich zum phasenminimalen System schon bei kleineren Kreisverstärkungen die Stabilitätsgrenze erreicht wird. Instabile Nullstellen drehen für hohe Frequenzen das Vorzeichen der Verstärkung um. Deshalb auch das typische gegenläufige Verhalten der Sprungantwort. Auch hieran erkennt man, dass die Bandbreite eingeschränkt werden muss.

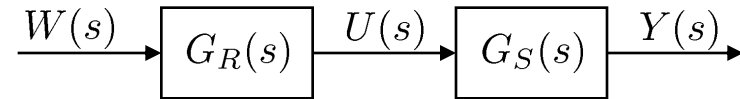
Als Faustregel gilt: **Die Bandbreite sollte kleiner gewählt werden als die langsamste instabile Nullstelle des offenen Kreises.**

15.2 Vorsteuerung

Optimale Steuerung

Wir wissen bereits, dass die Inverse der Strecke die *ideale* optimale Steuerung ist:

$$G_R^{(\text{ideal})}(s) = \frac{1}{G_S(s)}$$



Die Hauptgründe, warum die ideale optimale Steuerung meist nicht realisierbar ist, sind:

1. *Zählergrad ist größer (größer gleich) dem Nennergrad:* Das ergibt sich zwangsläufig, weil für die Strecke $G_S(s)$ gilt, dass ihr Zählergrad kleiner (in Ausnahmefällen gleich) ihrem Nennergrad ist.
2. *Negative Totzeit:* Hat die Strecke eine Totzeit, so ergibt sich für die Inverse eine negative Totzeit, d.h. ein Prädiktor! Ein solcher Prädiktor ist nur realisierbar, wenn der Verlauf der Führungsgröße $w(t)$ um mindestens die Totzeit T_t *im voraus* bekannt ist!
3. *Instabiles Verhalten:* Ist die Strecke nicht phasenminimal (hat also instabile Nullstellen), dann folgt daraus, dass die ideale Steuerung instabile Pole aufweisen müsste. Das kann zu unendlich großen Stellsignalen $u(t)$ führen und bei nicht exakter Pol/Nullstellen-Kürzung zu instabilem Gesamtverhalten. Instabile Strecken lassen sich generell nicht durch eine Steuerung stabilisieren, dafür benötigt man eine Regelung!

15.2 Vorsteuerung

Realisierbare Steuerung

Folgende Strategien bieten sich an, um die ideale optimale Steuerung realisierbar zu machen:

1. *Hinzufügen von Nennerzeitkonstanten bis Zählergrad gleich Nennergrad wird:*

$$G_S(s) = \frac{K}{(1 + T_1s)(1 + T_2s)} \rightarrow G_R^{(\text{real})}(s) = \frac{(1 + T_1s)(1 + T_2s)}{K} \frac{1}{(1 + Ts)^2}$$

2. *Ignorieren der negativen Totzeit:*

$$G_S(s) = \frac{K}{1 + T_1s} e^{-T_t s} \rightarrow G_R^{(\text{real})}(s) = \frac{1 + T_1s}{K} \frac{1}{1 + Ts}$$

$n = n_S - m_S$
Nennergrad von $G_S(s)$
Zählergrad von $G_S(s)$

3. *Ignorieren der instabilen Nullstellen:*

$$G_S(s) = K \frac{(1 - T_n s)}{(1 + T_1s)(1 + T_2s)} \rightarrow G_R^{(\text{real})}(s) = \frac{(1 + T_1s)(1 + T_2s)}{K} \frac{1}{(1 + Ts)^2}$$

Andere Möglichkeiten zur Behandlung instabiler Strecken-Nullstellen wurden bereits in Kapitel 12.2. beschrieben.

Wahl von T : Kompromiss zwischen Rauschempfindlichkeit und Schnelligkeit der Steuerung.

15.2 Vorsteuerung

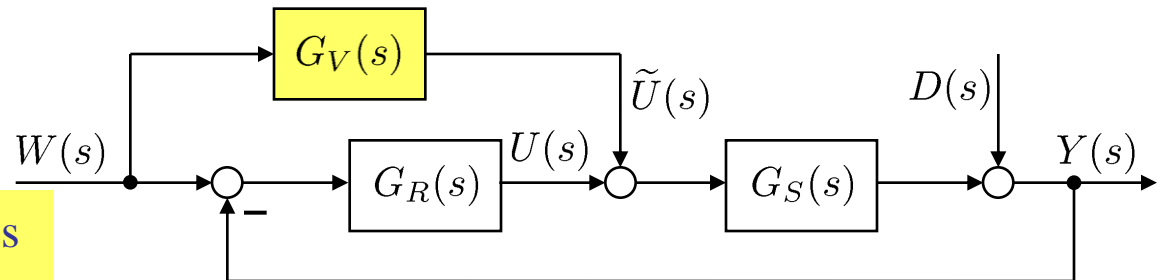
Regler mit 2 Freiheitsgraden (*2 degrees of freedom (2-DOF) control*)

Aus Kapitel 15.1 wissen wir, dass ein Standardregler gleich gut eine Ausgangsstörung $d(t)$ unterdrückt und für die Folge der Führungsgröße $w(t)$ sorgt, solange $d(t)$ und $w(t)$ im gleichen Frequenzbereich liegen, also z.B. beide sprungförmige Signalverläufe aufweisen. Liegen sie in unterschiedlichen Frequenzbereichen, dann muss ein Kompromiss gefunden werden.

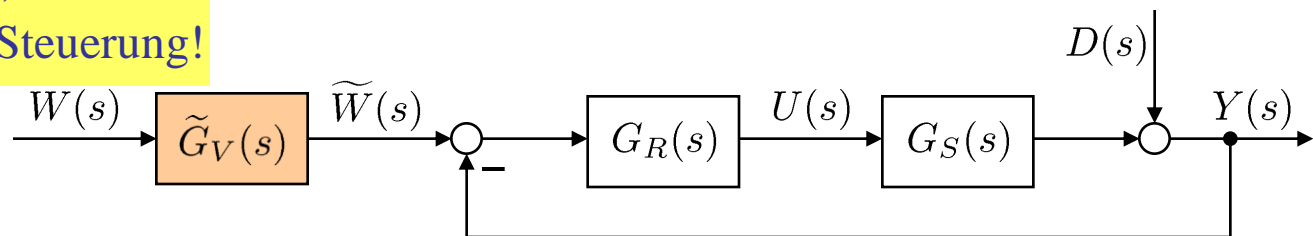
Die Einführung eines 2. Freiheitsgrades in das Regelsystem, nämlich einer **Vorsteuerung** oder eines **Vorfilters**, ermöglicht eine individuelle, optimale Auslegung bezüglich der beiden Anforderungen: *Störgrößenunterdrückung und Führungsgrößenfolge*. Ein Kompromiss ist nicht mehr notwendig.

Vorsteuerung:

Die *Stabilität* des Regelkreises wird *nicht beeinflusst*, denn es handelt sich um eine *Steuerung*!



Vorfilter:



15.2 Vorsteuerung

Vorteile von Regler mit 2 Freiheitsgraden (*2 degrees of freedom (2-DOF) control*)

Die Vor- und Nachteile von Steuerung und Regelung wurden bereits diskutiert. Mit der Erweiterung des Standardregelkreises versucht man die Vorteile beider Ansätze zu vereinen.

Mittels Vorsteuerung kann die **Schnelligkeit** der Steuerung ($w(t)$ wirkt direkt über den Vorwärtszweig auf $y(t)$, ohne auf die Rückkopplung warten zu müssen) kombiniert werden mit der **Robustheit** der Regelung (Störungen und Unbekanntes der Strecke werden ausgeregelt).

Wenn der Mensch als Regler agiert, ist dies meist ebenfalls eine Kombination aus beidem. Z.B. beim Auto fahren um eine Kurve wird das Bekannte **gesteuert** und könnte auch mit verbundenen Augen passieren (kein Sensor). Dieser Anteil nimmt mit Kenntnis der Strecke und Erfahrung des Fahrers zu und geschieht sehr schnell.

Auf das Unbekannte (Unsicherheit oder Veränderung der Strecke, Baustellen, Menschen etc.) wird **geregelt**. Dies benötigt die Augen (Sensor) und geschieht recht langsam.

15.2 Vorsteuerung

Grundidee der Vorsteuerung

1. Zuerst wird die Vorsteuerung als (näherungsweise) Inverse der Strecke entworfen.
2. Danach wird der Regler so entworfen, dass er möglichst gut die Störung $d(t)$ unterdrückt. Dazu ist eine möglichst genaue Kenntnis des typischen Spektrums vom $D(s)$ nützlich.

Grundidee des Vorfilters

1. Zuerst wird der Regler so entworfen, dass er möglichst gut die Störung $d(t)$ unterdrückt. Dazu ist eine möglichst genaue Kenntnis des typischen Spektrums von $D(s)$ nützlich.
2. Danach wird der Vorfilter so entworfen, dass er die Dynamik des geschlossenen Regelkreises verbessert, d.h. die Regelgröße besser der Führungsgröße folgt.

Die Interpretation und Vorgehensweise beider Ansätze ist zwar unterschiedlich, sie können aber leicht ineinander umgerechnet werden.

Dafür fordern wir identische Führungsübertragungsfunktionen $w \rightarrow y$ für beide Ansätze...

15.2 Vorsteuerung

Äquivalenz von Vorsteuerung und Vorfilter

Die Führungsübertragungsfunktion beider Ansätze sind:

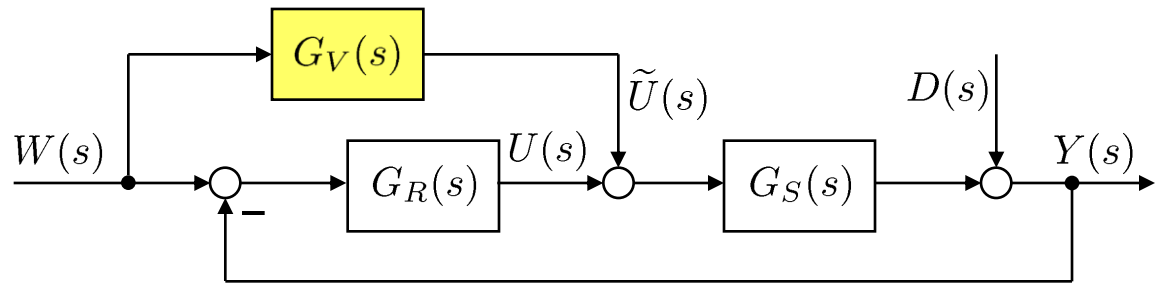
$$\text{Vorsteuerung: } G_w = \frac{G_R G_S + G_V G_S}{1 + G_R G_S}$$

$$\text{Vorfilter: } G_w = \frac{\tilde{G}_V G_R G_S}{1 + G_R G_S}$$

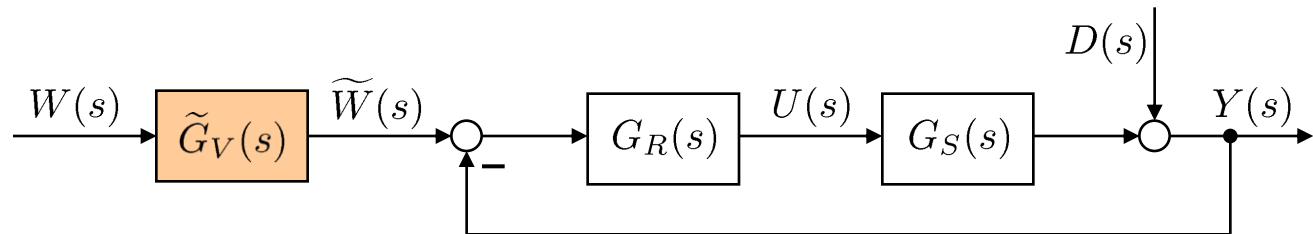
Gleichsetzen liefert:

$$G_R G_S + G_V G_S = \tilde{G}_V G_R G_S \quad \rightarrow \quad G_R + G_V = \tilde{G}_V G_R$$

$$G_V = G_R(\tilde{G}_V - 1)$$



$$\tilde{G}_V = 1 + \frac{G_V}{G_R}$$



15.2 Vorsteuerung

Idealer Fall

Ohne Beachtung der Realisierbarkeit würden wir die Vorsteuerung als Inverse der Strecke wählen:

$$G_V^{(\text{ideal})}(s) = \frac{1}{G_S(s)}$$

Das ist äquivalent mit folgender Wahl des Vorfilters:

$$\tilde{G}_V^{(\text{ideal})}(s) = 1 + \frac{1}{G_R(s)G_S(s)} = \frac{1 + G_R(s)G_S(s)}{G_R(s)G_S(s)}$$

Das ist, wie zu erwarten, genau die Inverse der Übertragungsfunktion des geschlossenen Regelkreises (ohne Vorfilter).

Die Vorsteuerung *hängt* im Gegensatz zum Vorfilter *nicht vom Regler* ab! Dies bietet den Vorteil, dass sie nicht geändert werden muss, wenn der Regler z.B. durch nachträgliches Adaptieren oder Optimieren verändert wird.

15.2 Vorsteuerung

Alternative Entwurfsidee für Vorfilter

Annahme: Die Führungsübertragungsfunktion $G_{\tilde{w}} : \tilde{w} \rightarrow y$ weist „störende“ D-Anteile (Nullstellen) auf. Diese wirken oft störend, da sie zu evtl. starkem Überschwingen der Sprungantwort führen können (siehe Kapitel 7). Solche D-Anteile können auf zwei Arten entstehen:

- Regelstrecke weist D-Anteile auf.
- Regler weist D-Anteile auf. Diese sind zur Stabilisierung und Verbesserung der dynamischen Eigenschaften des geschlossenen Regelkreises aber gewollt. Ihr Einfluss auf den Nenner (die Pole) von $G_{\tilde{w}} : \tilde{w} \rightarrow y$ ist erwünscht. Ihr Einfluss auf den Zähler (die Nullstellen) von $G_{\tilde{w}} : \tilde{w} \rightarrow y$ ist meist unerwünscht und muss zwangsläufig in Kauf genommen werden.

Das Vorfilter kann in beiden Fällen dazu genutzt werden, die störenden Nullstellen in $G_{\tilde{w}} : \tilde{w} \rightarrow y$ wegzukürzen. Zunächst kann man sich also beim Entwurf des Reglers auf die Veränderung der Pole des geschlossenen Regelkreises konzentrieren. Die evtl. negativen Nebenwirkungen eines starken D-Anteils auf das Führungsverhalten werden danach mit dem Vorfilter kompensiert.

15.2 Vorsteuerung

Beispiel: Vorfilter kompensiert Nullstellen der Führungsübertragungsfunktion

Regelstrecke: $G_S(s) = \frac{5}{s(s+5)^2}$ (grenzstabil)

Regler: PD-Regler zur Stabilisierung. Kein I-Anteil nötig, da in Strecke schon enthalten.

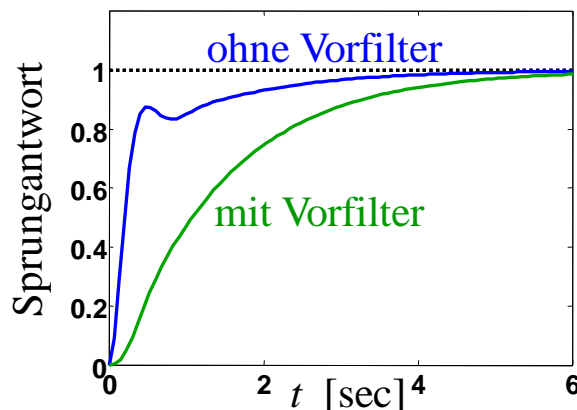
$G_R(s) = K_R(s+1)$

Führungsverhalten: $G_{\tilde{w}}(s) = \frac{\frac{5K_R(s+1)}{s(s+5)^2}}{1 + \frac{5K_R(s+1)}{s(s+5)^2}} = \frac{5K_R(s+1)}{s^3 + 10s^2 + (5K_R + 25)s + 5K_R}$

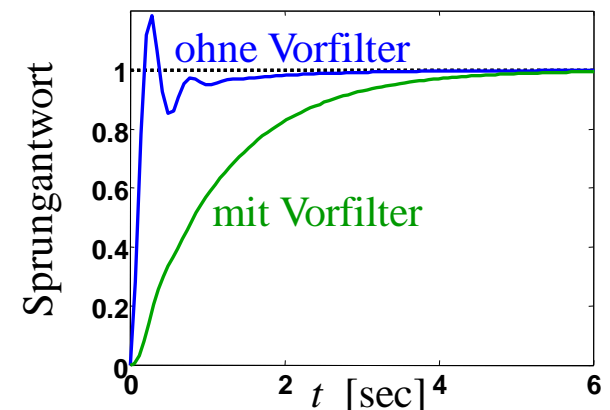
← störende Nullstelle

Vorfilter: $\tilde{G}_V(s) = \frac{1}{s+1}$ ($\tilde{G}_V(0) = 1$, denn Gesamtverstärkung soll unverändert bleiben!)

$K_R = 10$:



$K_R = 30$:



15.3 Störgrößenaufschaltung

Grundidee der Störgrößenaufschaltung

Wenn man eine *Störgröße messen* kann, dann kann sie ähnlich wie die Führungsgröße behandelt werden. Allerdings mit dem Unterschied, dass der Einfluss der Störgröße möglichst klein sein soll.

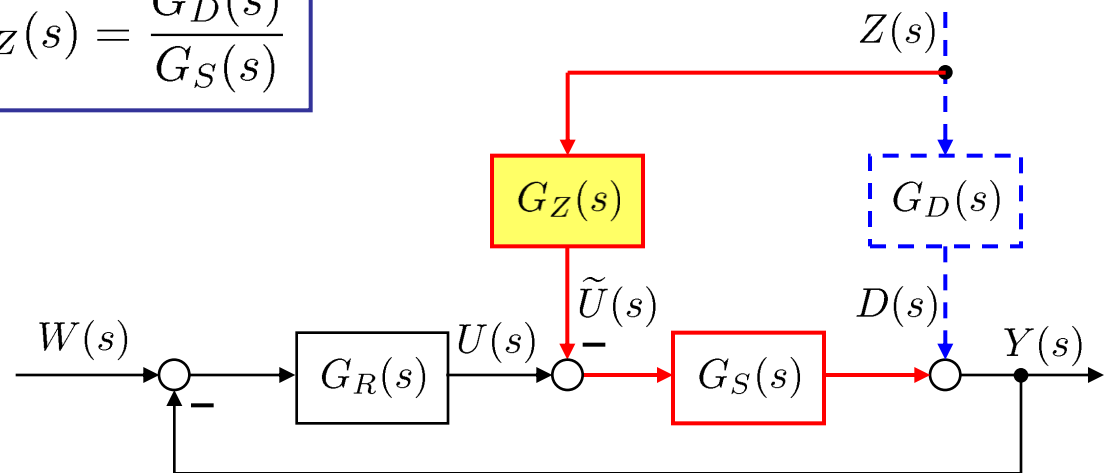
Überträgt man das Konzept der Vorsteuerung auf messbare Störgrößen erhält man die sog. Störgrößenaufschaltung. Eine messbare Störung $z(t)$ wirkt über ein dynamisches System $G_D(s)$ additiv auf die Regelgröße $y(t)$. Eine solche Störung lässt sich (im Idealfall) perfekt durch die Störgrößenaufschaltung $G_Z(s)$ kompensieren, falls gilt:

$$G_Z(s)G_S(s) = G_D(s) \rightarrow G_Z(s) = \frac{G_D(s)}{G_S(s)}$$

Gleiche Formel wie bei Vorsteuerung für $G_D(s) = 1$.

Die *Stabilität* des Regelkreises wird *nicht beeinflusst*, denn es handelt sich um eine Steuerung!

roter Pfad = blauer gestrichelter Pfad



15.3 Störgrößenaufschaltung

Realisierbarkeit der Störgrößenaufschaltung

Im Unterschied zur Vorsteuerung ist die Störgrößenaufschaltung meist leichter realisierbar, weil das Problem durch das typischerweise verzögernde Verhalten von $G_D(s)$ entschärft wird. Je stärker $G_D(s)$ verzögernd wirkt (größere Totzeit, stärkeres Tiefpassverhalten), um so leichter lässt sich $G_Z(s)$ realisieren.

Beispiel:
$$G_S(s) = \frac{K}{(1 + T_1s)(1 + T_2s)} e^{-2s}$$

Fall 1: $G_D(s) = 1$ (vergleichbar mit der Vorsteuerungsaufgabe)

$$G_Z^{(\text{ideal})}(s) = \frac{(1 + T_1s)(1 + T_2s)}{K} e^{2s} \xrightarrow[\text{Abweichungen}]{\text{große}} G_Z^{(\text{real})}(s) = \frac{(1 + T_1s)(1 + T_2s)}{K(1 + Ts)^2}$$

Fall 2: $G_D(s) = \frac{1}{1 + T_Ds} e^{-3s}$ ← größere Totzeit als $G_S(s)$

$$G_Z^{(\text{ideal})}(s) = \frac{(1 + T_1s)(1 + T_2s)}{K(1 + T_Ds)} e^{-s} \xrightarrow[\text{Abweichungen}]{\text{kleine}} G_Z^{(\text{real})}(s) = \frac{(1 + T_1s)(1 + T_2s)}{K(1 + T_Ds)(1 + Ts)} e^{-s}$$

15.3 Störgrößenaufschaltung

Simulationsergebnisse für Fall 2

$$G_S(s) = \frac{5}{(1 + 3s)(1 + 10s)} e^{-2s}$$

$$G_D(s) = \frac{1}{1 + 20s} e^{-3s}$$

$$G_R(s) = 0.02 \frac{1 + 10s}{s}$$

Störgrößenaufschaltung:

- keine:

$$G_Z(s) = 0$$

- statisch:

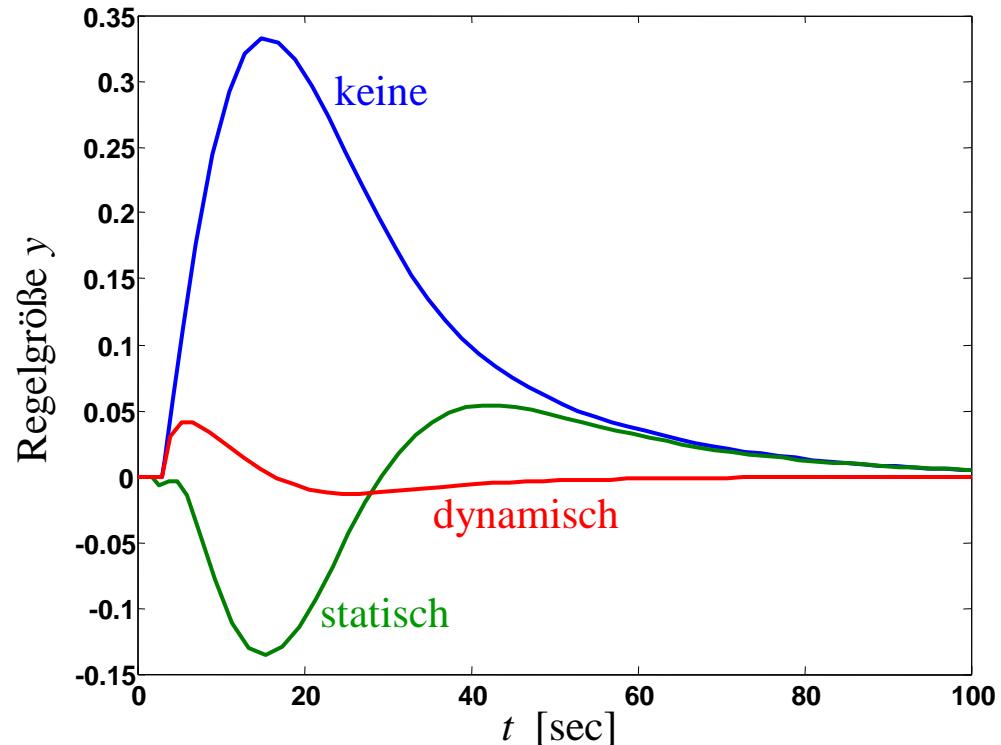
$$G_Z(s) = \frac{1}{5}$$

- dynamisch:

$$G_Z(s) = \frac{(1 + 3s)(1 + 10s)}{5(1 + 20s)(1 + Ts)} e^{-s}$$

mit $T = 1$ sec.

Einheitssprung der Störung bei $t = 0$ sec

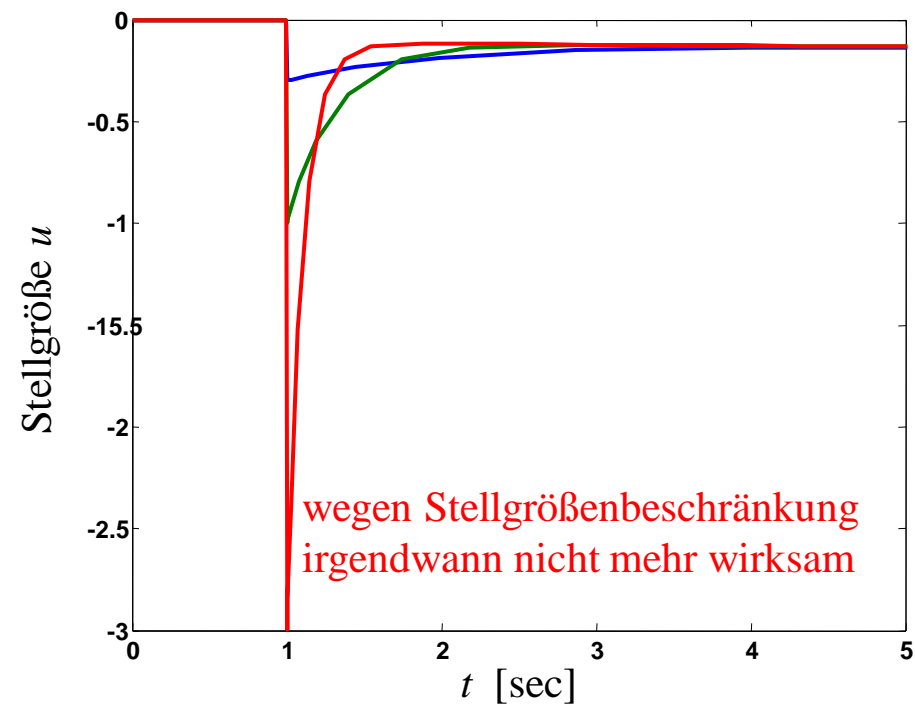
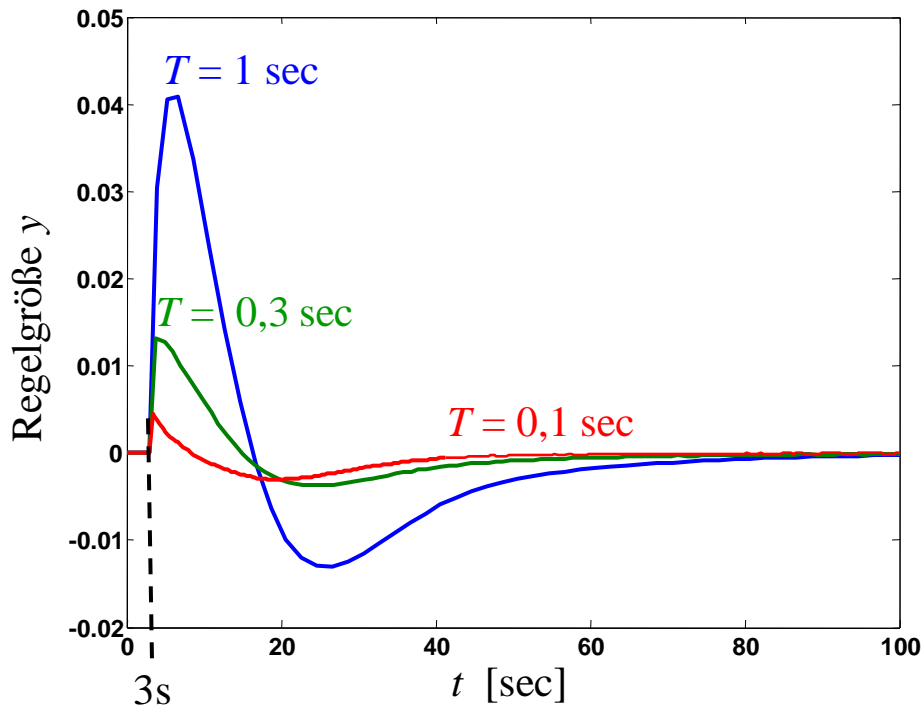


15.3 Störgrößenaufschaltungung

Dynamische Störgrößenaufschaltungung: Wahl der Zeitkonstante T

- $T \rightarrow 0$ ergibt die ideale Inverse und liefert damit eine perfekte Störunterdrückung.
- Der differenzierende Charakter nimmt mit kleiner werdendem T stark zu.
- Die (betragsmäßige) Stellgrößenauslenkung nimmt mit kleiner werdendem T stark zu.

Einheitssprung der Störung bei $t = 0$ sec



15.3 Störgrößenaufschaltung

Realisierbarkeit der Störgrößenaufschaltung

An den 2 Fällen des letzten Beispiels erkennen wir:

- Wenn die Störung mit einer Totzeit größer gleich der Streckentotzeit auf die Regelgröße einwirkt, dann kann die Kompensation ohne Zeitverzug wirken.
- Je stärker das Verzögerungsverhalten ist, mit dem die Störung auf die Regelgröße einwirkt, desto einfacher ist es, die Störung zu kompensieren. Das macht sich dadurch bemerkbar, dass $G_Z(s)$ weniger extremes Hochpassverhalten (also schwächer differenzierenden Charakter) hat und damit unempfindlicher auf der Störung überlagertes Messrauschen reagiert.
- Um für $G_Z(s)$ die Bedingung Zählergrad \leq Nennergrad einhalten zu können, muss ein PT_n -Glied der folgenden Ordnung ergänzt werden:

$$n = n_S - m_S - (n_D - m_D)$$

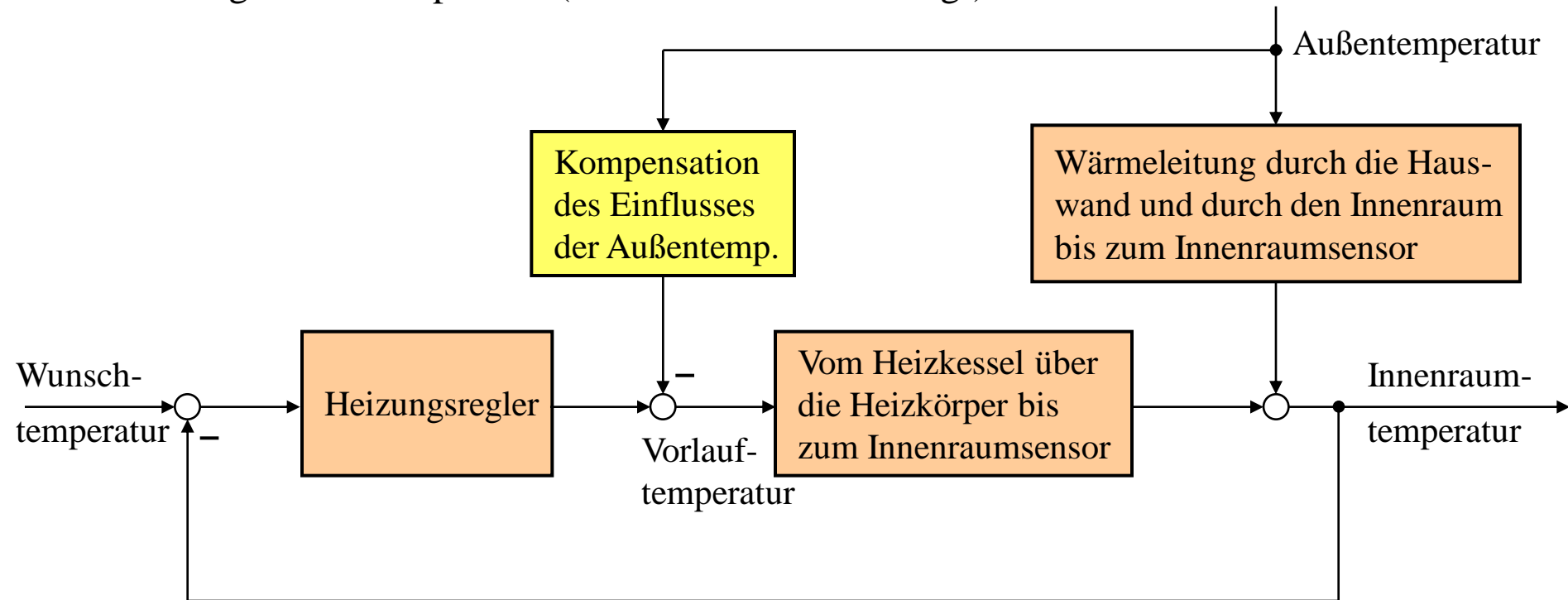
Nennergrad von $G_S(s)$ Zählergrad von $G_S(s)$ Nennergrad von $G_D(s)$ Zählergrad von $G_D(s)$

Ist der Polüberschuss von $G_D(s)$ gleich dem von $G_S(s)$, ist $n = 0$.

15.3 Störgrößenaufschaltung

Beispiel: Innenraumtemperaturregelung mit Außentempersensor

- Regelgröße: gemessene Innenraumtemperatur (Sensor im Zimmer notwendig!)
- Führungsgröße: gewünschte Innenraumtemperatur
- Stellgröße: Vorlauftemperatur der Heizung
- Störung: Außentemperatur (Sensor außen notwendig!)



15.4 Regelkreis mit Hilfsstellgrößen und Hilfsregelgrößen

Was sind Hilfsstell- und Hilfsregelgrößen?

In einem Standardregelkreis steht *eine Stellgröße* zur Verfügung, mit welcher die Regelstrecke im gewünschten Sinne beeinflusst werden kann. Und es steht *eine Regelgröße* zur Verfügung, die zur Rückkopplung genutzt werden kann.

Falls es die zusätzliche Möglichkeit gibt, die Regelstrecke über einen *weiteres Stellglied* zu beeinflussen, kann eine solche **Hilfsstellgröße** genutzt werden, um das dynamische Verhalten des Regelkreises zu verbessern.

Falls es die zusätzliche Möglichkeit gibt, relevante Signale aus dem Inneren der Regelstrecke über einen *weiteren Sensor* zu messen, kann eine solche Messgröße als **Hilfsregelgröße** genutzt werden, um das dynamische Verhalten des Regelkreises zu verbessern.

Beide Erweiterungen des Standardregelkreises weisen u.a. folgende Eigenschaften auf:

- Verbesserung des dynamischen Verhaltens, insbesondere bei Strecken hoher Ordnung.
- Zusätzlicher Hardwareaufwand (Stellglied bzw. Sensor).
- Eingriff in den geschlossenen Regelkreis. → Stabilität wird beeinflusst.

15.4 Regelkreis mit Hilfsstellgrößen und Hilfsregelgrößen

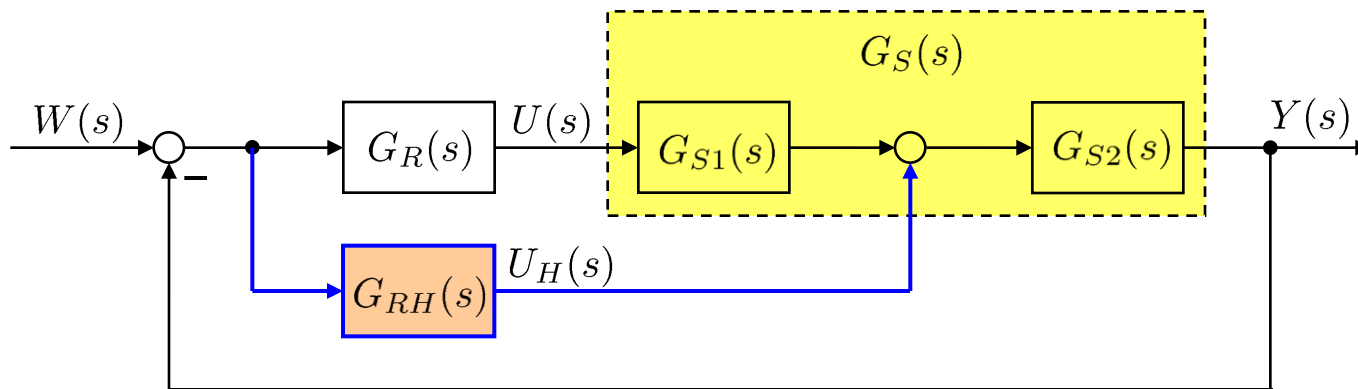
Hilfsstellgrößen

Aus der Regelabweichung wird über einen **Hilfsregler** $G_{RH}(s)$ eine Hilfsstellgröße $u_H(t)$ erzeugt, die im Innern der Strecke eingreift. Da damit ein Teil der Strecke, nämlich $G_{S1}(s)$, überbrückt wird, kann die Regelung über den Hilfsregler *schneller agieren*.

Typischerweise stellt man mit $u(t)$ das stationäre Verhalten ein, weil dies z.B. aus verfahrenstechnischen oder energetischen Gründen günstiger ist, und $u_H(t)$ verbessert die Dynamik.

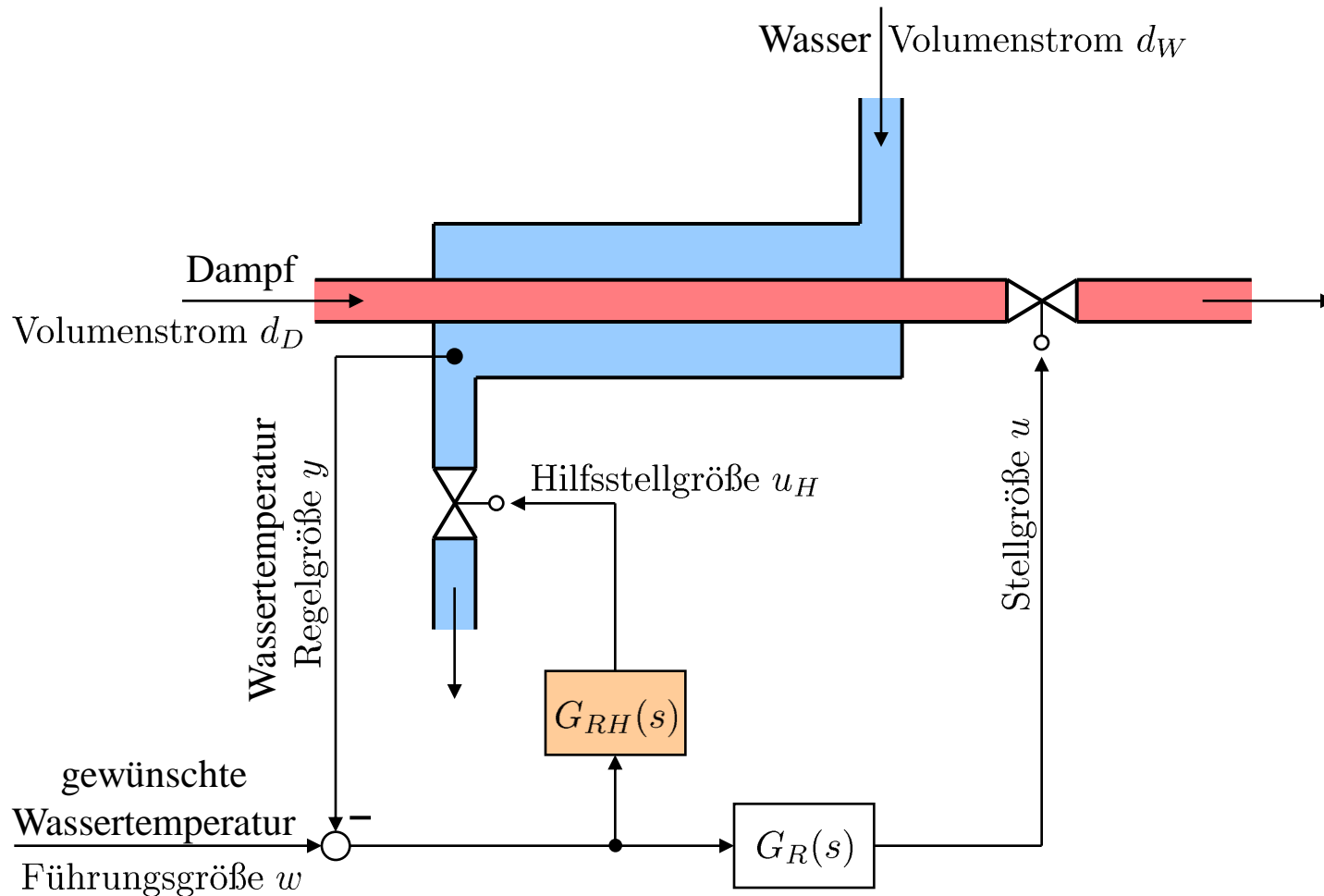
Durch die Hilfsstellgröße verändert sich die Übertragungsfunktion des offenen und des geschlossenen Regelkreises:

$$G_0(s) = [G_R(s)G_{S1}(s) + G_{RH}(s)]G_{S2}(s) \quad \text{und} \quad G_w(s) = \frac{G_0(s)}{1 + G_0(s)}$$



15.4 Regelkreis mit Hilfsstellgrößen und Hilfsregelgrößen

Beispiel: Hilfsstellgröße bei einem Gegenstromwärmetauscher



15.4 Regelkreis mit Hilfsstellgrößen und Hilfsregelgrößen

Funktionsweise des Gegenstromwärmetauschers

Strecke $u \rightarrow y$ (langsam):

Die Wassertemperatur wird mittels der Stellgröße Dampfventil u beeinflusst. Je weiter das Ventil geöffnet wird, desto mehr Dampf strömt durch das Rohr und desto mehr Wärmeenergie kann auf das Wasser übertragen werden, d.h. die Wassertemperatur steigt.

Strecke $u_H \rightarrow y$ (schnell):

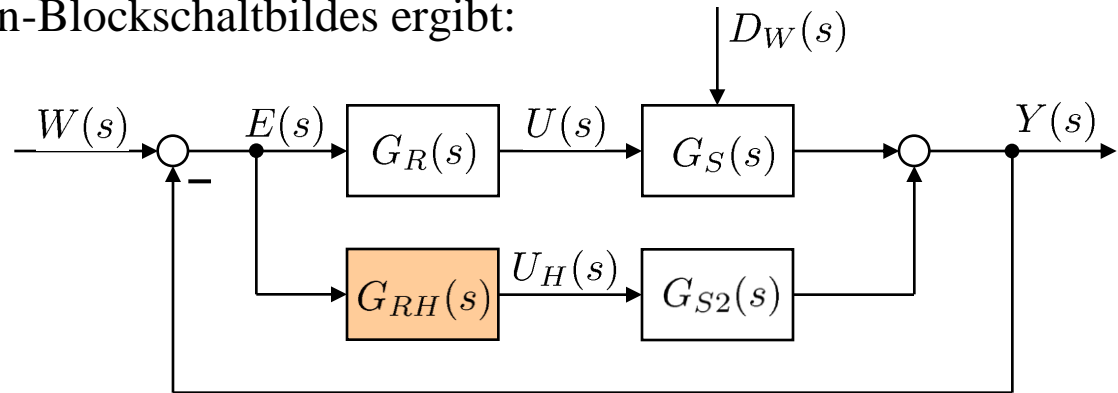
Eine zweite Möglichkeit zur Beeinflussung der Wassertemperatur liefert die Hilfsstellgröße Wasserventil u_H . Je weiter dieses Ventil geschlossen (negative Verstärkung!) wird, desto weniger Wasser fließt durch den Wärmetauscher und desto länger kann Energie vom Dampf auf das Wasser übertragen werden, d.h. die Wassertemperatur steigt.

Es wird normalerweise ein bestimmter Wasservolumenstrom d_w und eine bestimmte Wassertemperatur y gewünscht. Deshalb sollen Eingriffe in den Wasservolumenstrom über die Hilfsstellgröße u_H nur vorübergehend erfolgen, um möglichst schnell die gewünschte Wassertemperatur zu erreichen bzw. zu stabilisieren. Langfristig soll der ursprüngliche Wasservolumenstrom wieder hergestellt werden.

15.4 Regelkreis mit Hilfsstellgrößen und Hilfsregelgrößen

Blockschaltbild der Regelung

Umzeichnen des Hilfsstellgrößen-Blockschaltbildes ergibt:



Aus den bisherigen Überlegung folgt:

- $G_S(s)$ ist langsam, $G_{S2}(s)$ ist schnell.
 - Stellgröße u soll stationären Zustand einstellen. \rightarrow I-Anteil im Regler $G_R(s)$, damit $u = \text{konst.}$ für Regelabweichung $e = 0$.
 - Hilfsstellgröße u_H soll vorübergehenden Regelabweichungen schnell entgegenwirken, weil sich die Regelgröße y schneller über $G_{S2}(s)$ als über $G_S(s)$ beeinflussen lässt. Stationär ($e = 0$) soll die Hilfsstellgröße aber den Wasservolumenstrom nicht ändern.
- \rightarrow **Wähle z.B.: $G_R(s)$ als PI-Regler und $G_{RH}(s)$ als (P)DT₁-Regler.**

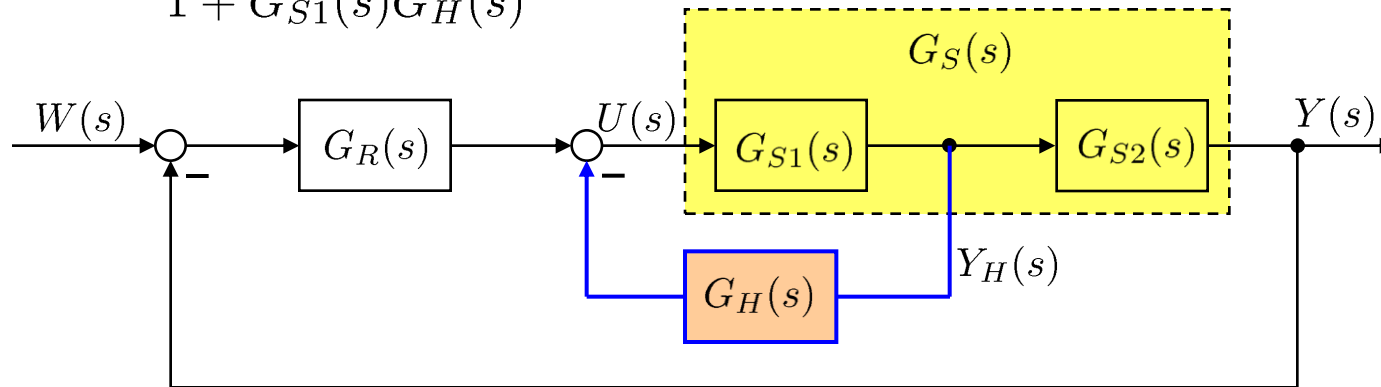
15.4 Regelkreis mit Hilfsstellgrößen und Hilfsregelgrößen

Hilfsregelgrößen

Neben der Regelgröße $y(t)$ wird im Inneren der Regelstrecke eine weitere Größe $y_H(t)$ gemessen und als Hilfsregelgröße verwendet. Sie wird über die Hilfsrückführung $G_H(s)$ auf die Stellgröße rückgekoppelt. Da hiermit Informationen aus dem Inneren der Strecke zur Regelung herangezogen werden, kann die Regelung auf sich abzeichnende Abweichungen *schneller reagieren*. Beispielsweise wird eine Störung, die zwischen $G_{S1}(s)$ und $G_{S2}(s)$ angreift sofort "erkannt" und nicht erst, wenn sie durch $G_{S2}(s)$ hindurch gelaufen ist.

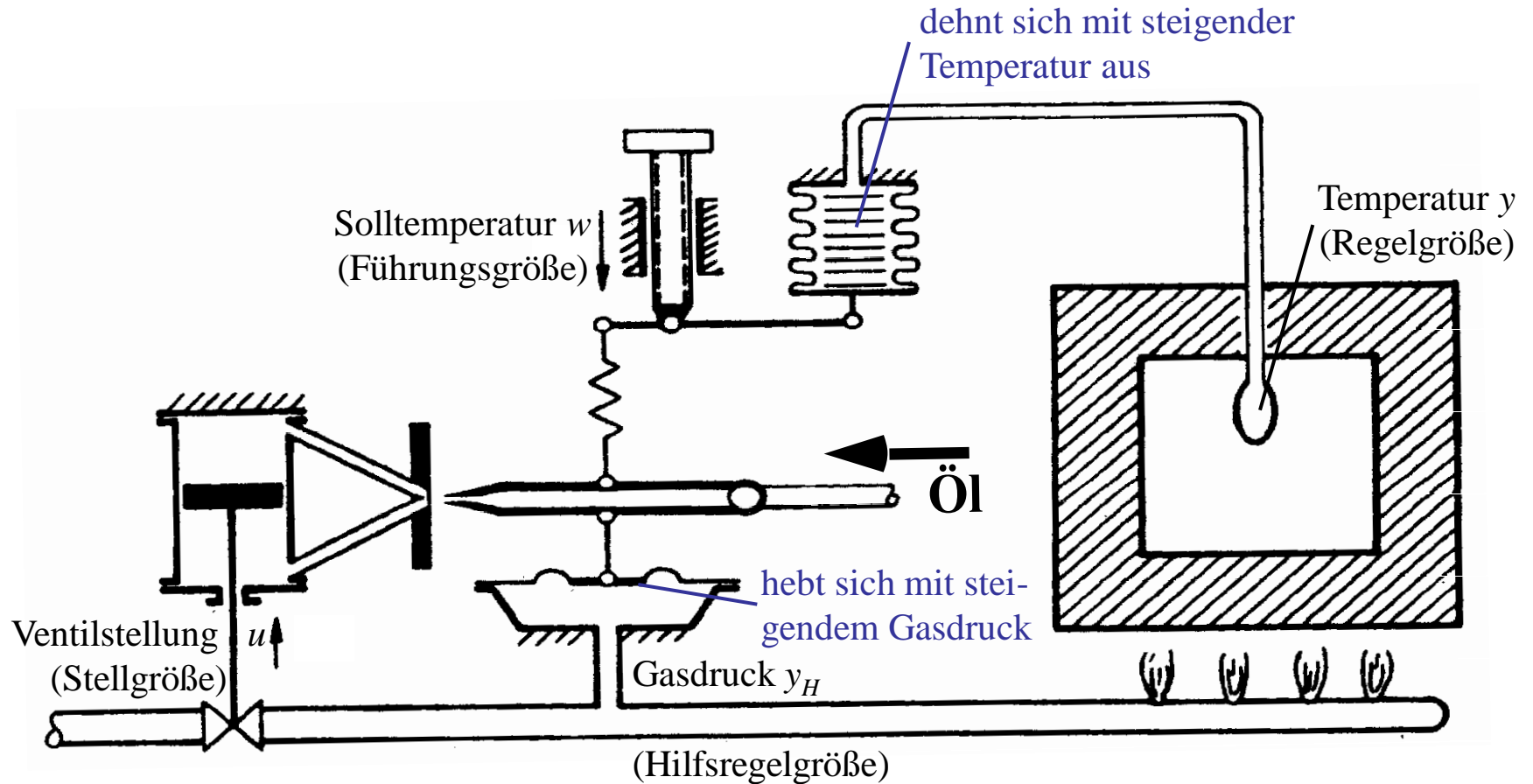
Durch die Hilfsregelgröße verändert sich die Übertragungsfunktion des offenen und des geschlossenen Regelkreises:

$$G_0(s) = G_R(s) \frac{G_{S1}(s)}{1 + G_{S1}(s)G_H(s)} G_{S2}(s) \quad \text{und} \quad G_w(s) = \frac{G_0(s)}{1 + G_0(s)}$$



15.4 Regelkreis mit Hilfsstellgrößen und Hilfsregelgrößen

Beispiel: Hilfsregelgröße bei einer Ofentemperaturregelung



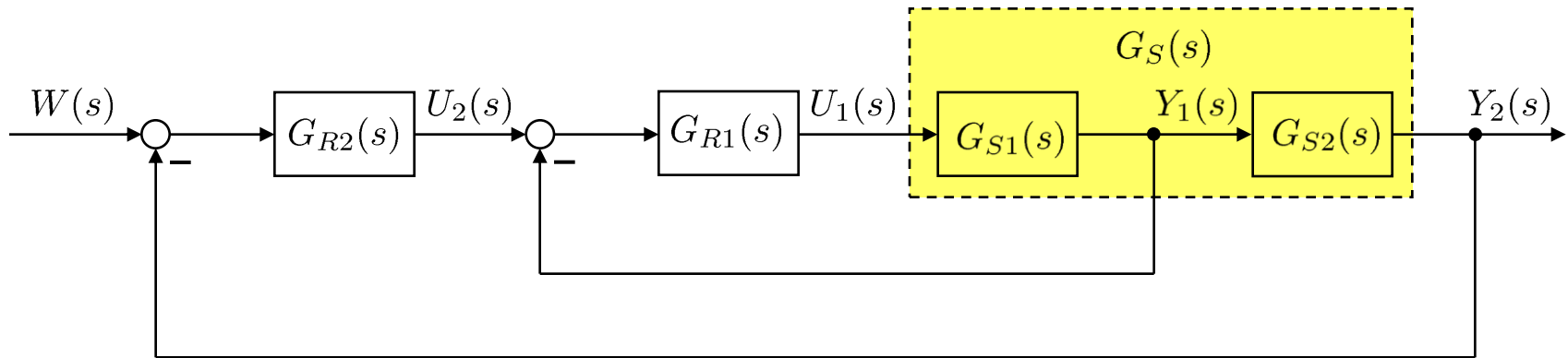
15.5 Kaskadenregelung

Prinzip der Kaskadenregelung

Die **Kaskadenregelung** besteht aus hierarchisch ineinander geschachtelten Standardregelkreisen. Sie kann als Kombination von Hilfsstellgrößen- und Hilfsregelgrößenregelung gesehen werden. Dafür müssen 2 (oder mehr) Regelgrößen und 2 (oder mehr) dazu passende Stellgrößen zur Verfügung stehen.

Es gibt einen inneren Regelkreis, der Teil einer äußeren Strecke ist. Die Führungsgröße des inneren Regelkreises ist die Stellgröße des äußeren Reglers. Diese Hierarchie lässt sich auch auf mehr als 2 ineinander geschachtelte Regelkreise ausdehnen.

Die Kaskadenregelung vereint die Geschwindigkeits- und Robustheitsvorteile der Hilfsstellgrößen- und Hilfsregelgrößenregelung und deren Nachteil des erhöhten Hardwareaufwands.

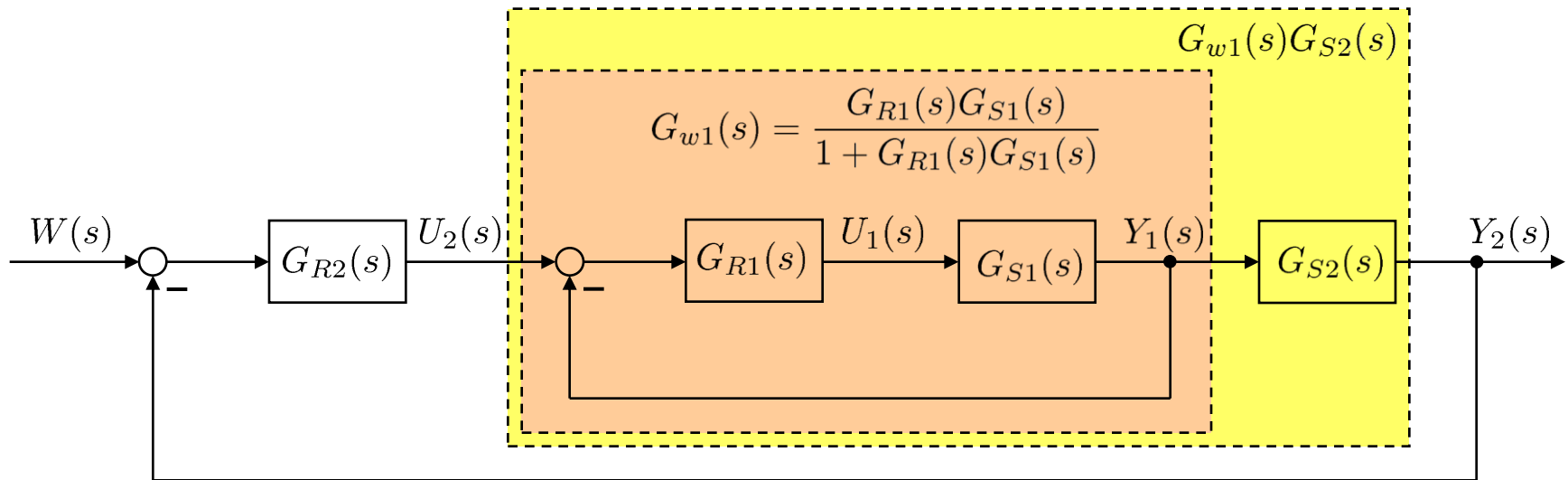


15.5 Kaskadenregelung

Entwurf einer Kaskadenregelung

Wie üblich bei hierarchischen Problemen beginnt man zweckmäßigerweise mit der innersten Struktur, also hier dem inneren Regelkreis 1.

1. Der innere Regler $G_{R1}(s)$ lässt sich *im Prinzip* so auslegen als gäbe es den äußeren Regelkreis nicht. Es handelt sich als um ein Standardproblem.
2. Ist $G_{R1}(s)$ entworfen, so kann man das Führungsverhalten $G_{w1}(s)$ berechnen (rosa Kasten).
3. Nun lässt sich der äußere Regler $G_{R2}(s)$ für die Strecke $G_{w1}(s) \cdot G_{S2}(s)$ entwerfen. Das ist wieder ein Standardproblem...



15.5 Kaskadenregelung

Besondere Aspekte beim Entwurf von Kaskadenregelungen

Obwohl, wie auf der vorherigen Folie erklärt, alle Entwurfsmethoden für Standardregelkreise auch für die Kaskadenregelung verwendet werden können, gibt es wichtige *Unterschiede* bei den *Zielen* für den Reglerentwurf:

- Die Wahl eines inneren Reglers (z.B. $G_{R1}(s)$) beeinflusst entscheidend die Struktur und Parameter der inneren Führungsübertragungsfunktion (z.B. $G_{w1}(s)$).
- Diese Führungsübertragungsfunktion ist aber ein Teil der Regelstrecke für den nächst äußeren Regelkreis. Somit bestimmt der Entwurf des inneren Reglers darüber, wie schwierig die Regelungsaufgabe für den nächst äußeren Regler wird. Wenn insgesamt schnelles Regelverhalten gewünscht wird, so muss schon der innere Regler entsprechend aggressiv ausgelegt sein.
- Oft interessiert nur die Regelgüte des äußersten Regelkreises; alle inneren Regelkreise sind nur Mittel zum Zweck. Dann spielen Entwurfsziele wie "keine bleibende Regelabweichung" oder "nur geringes Überschwingen" für die inneren Regelkreise keine entscheidende Rolle. Z.B. kann eine bleibende Regelabweichung mit einem I-Anteil im äußersten Regler vermieden werden. Die inneren Regelkreise dürfen ruhig bleibende Regelabweichungen produzieren. Die inneren Regler können dann vom Typ P- oder PDT_1 sein und damit schneller als PI o.ä.

15.5 Kaskadenregelung

Impliziter D-Anteil

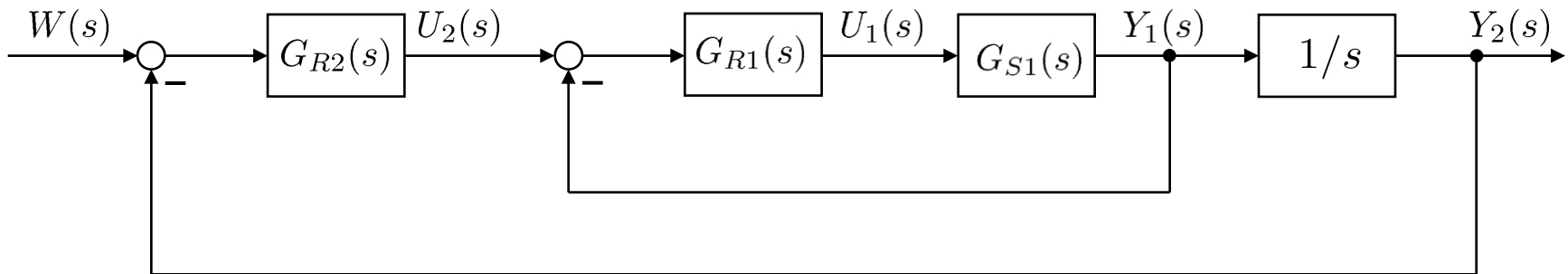
Streckenanteile, z.B. $G_{s2}(s)$, haben meist verzögernden Charakter. Entspricht $G_{s2}(s)$ beispielsweise einem PT_1 mit Verstärkung 1 und Zeitkonstanten T , dann gilt:

$$Y_1(s) = (1 + Ts)Y_2(s)$$

Beim einem Integrator (siehe Bild unten) gilt:

$$Y_1(s) = sY_2(s)$$

Eine Rückführung von $Y_1(s)$ erzeugt somit die Ableitung von $Y_2(s)$, OHNE differenzieren zu müssen. Also werden alle damit verbundenen Nachteile (Rauschempfindlichkeit, Zählerordnung > Nennerordnung) vermieden und der D-Anteil hat nur Vorteile! Allerdings muss man dafür $Y_1(s)$ messen.

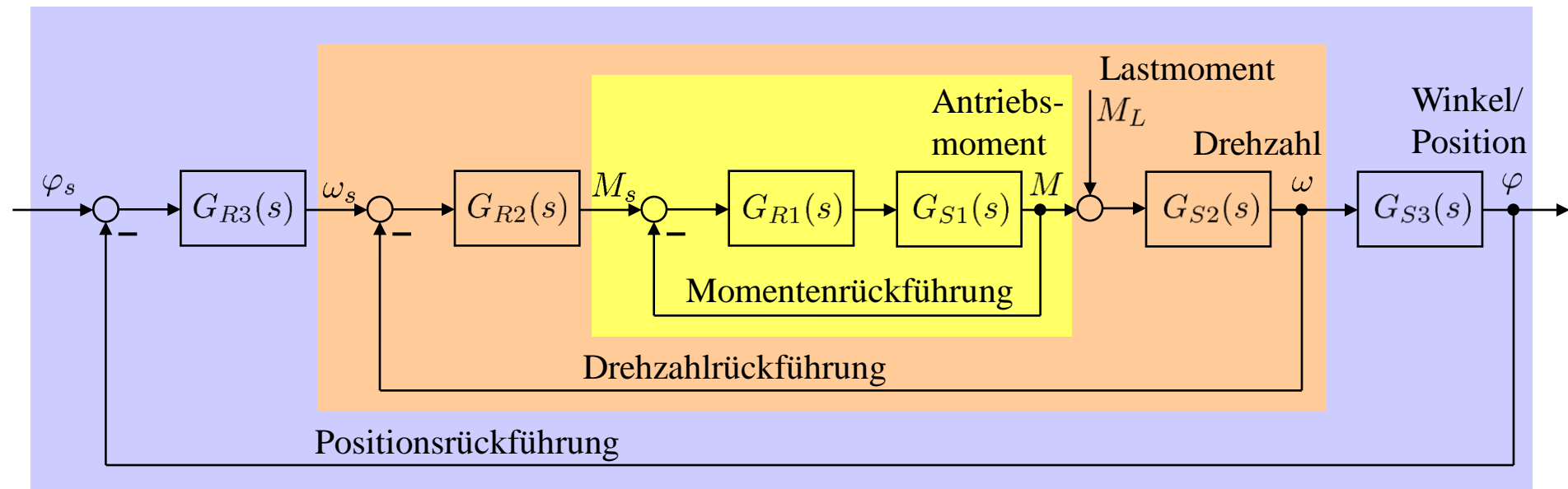


15.5 Kaskadenregelung

Beispiel: Lage-/Positionsregelung

Die Kaskadenstruktur ist insbesondere bei der Regelung *elektrischer Antriebe* besonders weit verbreitet. Typischerweise liegt bei Lage- bzw. Positionsregelungsaufgaben eine 3-stufige Kaskadenregelung vor:

1. Innerster Regelkreis: Strom/Drehmomentregelung.
2. Mittlerer Regelkreis: Drehzahlregelung.
3. Äußerer Regelkreis: Lage-/Positionsregelung.



15.6 Mehrgrößenregelung

SISO = Single Input, Single Output

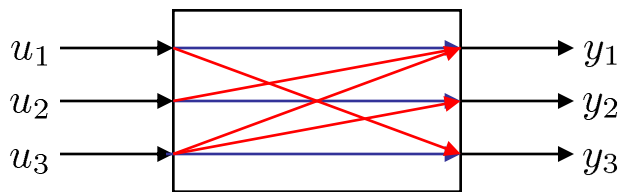
MIMO = Multiple Input, Multiple Output

Wann braucht man Mehrgrößenregelungen?

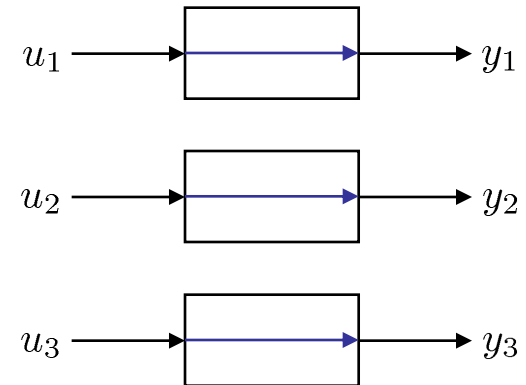
Ein Mehrgrößenregelung ist immer dann erforderlich, wenn

- mehrere Stell- und Regelgrößen vorliegen
- und diese stark miteinander **gekoppelt** sind.

Ist die *Kopplung* nur *schwach*, kann man die Mehrgrößenregelung in separate Eingrößenregelungen zerlegen, bei denen die Kopplung vernachlässigt bzw. als externe Störgröße aufgefasst wird. Dann kann man jedes Entwurfsverfahren für Eingrößenregelkreise einsetzen.



wenn **Kopplungen**
schwach



Wenn die *Kopplungen stark* sind, muss das Mehrgrößensystem als Ganzes betrachtet werden. Mit ausschließlich solchen Fällen wollen wir uns im Folgenden befassen.

15.6 Mehrgrößenregelung

Beispiele für Mehrgrößenregelungen

Starke Wechselwirkung der Regelgrößen:

- *Dampferzeuger:* Regelgrößen Druck und Temperatur.
- *Klimaanlage:* Regelgrößen Temperatur und Luftfeuchtigkeit.
- *Brenner:* Regelgrößen Temperatur, Flammenhöhe und Rauchgaszusammensetzung.
- *Roboterarm:* Regelgrößen Position, Geschwindigkeit und Kraft.

Koordinierter Eingriff der Stellgrößen:

- *Robotergreifen:* Mehrere Antriebe müssen gleichzeitig positionsgeregelt werden.
- *Kurvenflug eines Flugzeugs:* Roll- und Gierbewegung müssen gleichzeitig ausgeführt werden.
- *Stabilisierung eines Bioreaktors:* Temperatur, pH-Wert und Biomasseverteilung müssen gleichzeitig in bestimmten Grenzen gehalten werden, um den Arbeitspunkt zu stabilisieren.

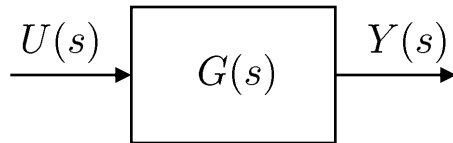
15.6 Mehrgrößenregelung

Matrix/Vektor-Darstellung

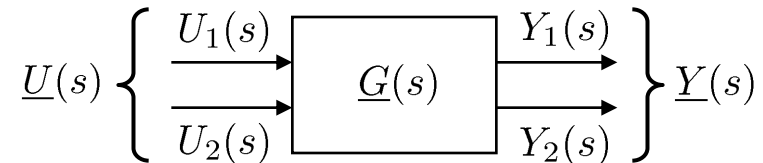
Um das Verhalten von Mehrgrößensystemen möglichst effizient beschreiben zu können, greifen wir auf die Matrix/Vektor-Darstellung zurück. Um die Notation so einfach wie möglich zu halten, gehen wir im Folgenden immer von Systemen mit 2 Eingängen und 2 Ausgängen aus. Eine Erweiterung auf Systeme mit beliebig vielen Ein- und Ausgängen ist aber immer leicht möglich.

Das System kann sowohl eine Regelstrecke als auch einen Regler oder anderes beschreiben.

Eingrößensystem



Mehrgrößensystem



Übertragungsfunktion:

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

Übertragungsmatrix:

$$\underline{Y}(s) = \underline{G}(s)\underline{U}(s)$$

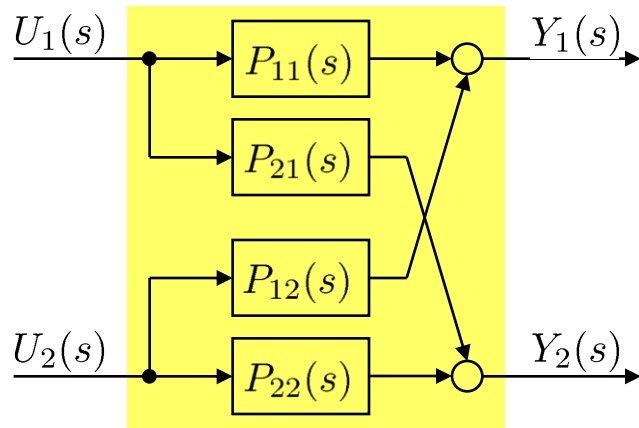
$$\begin{pmatrix} Y_1(s) \\ Y_2(s) \end{pmatrix} = \begin{pmatrix} G_{11}(s) & G_{12}(s) \\ G_{21}(s) & G_{22}(s) \end{pmatrix} \begin{pmatrix} U_1(s) \\ U_2(s) \end{pmatrix}$$

15.6 Mehrgrößenregelung

Systeme in P- und V-kanonischer Struktur

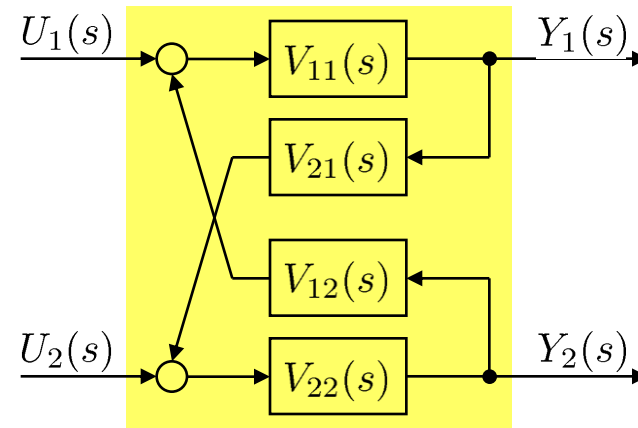
Die beiden folgenden Strukturen treten bei Mehrgrößensystemen besonders häufig auf.

P-kanonische Struktur



- Entspricht direkt der Übertragungsmatrix.
- Leichter regelungstechnisch zu behandeln.
- $P_{ij}(s)$ haben oft keine 1:1-Entsprechung zum realen System.
- "Weiter weg" von der Physik.

V-kanonische Struktur



- Ergibt sich häufig aus der Modellbildung.
- Schwieriger regelungstechnisch zu behandeln.
- $V_{ij}(s)$ haben oft eine 1:1-Entsprechung zum realen System.
- Näher an der Physik.

15.6 Mehrgrößenregelung

Umrechnung von der V-kanonischen in die P-kanonische Struktur

Moderne regelungstechnische Methoden basieren fast alle auf der P-kanonischen Struktur. Liegt das System nach der Modellbildung zunächst in V-kanonischer Struktur vor, kann es wie folgt umgerechnet werden:

$$\begin{pmatrix} Y_1(s) \\ Y_2(s) \end{pmatrix} = \begin{pmatrix} V_{11}(s) & 0 \\ 0 & V_{22}(s) \end{pmatrix} \left[\begin{pmatrix} U_1(s) \\ U_2(s) \end{pmatrix} + \begin{pmatrix} 0 & V_{12}(s) \\ V_{21}(s) & 0 \end{pmatrix} \begin{pmatrix} Y_1(s) \\ Y_2(s) \end{pmatrix} \right] \\ = \begin{pmatrix} V_{11}(s) & 0 \\ 0 & V_{22}(s) \end{pmatrix} \begin{pmatrix} U_1(s) \\ U_2(s) \end{pmatrix} + \begin{pmatrix} 0 & V_{11}(s)V_{12}(s) \\ V_{22}(s)V_{21}(s) & 0 \end{pmatrix} \begin{pmatrix} Y_1(s) \\ Y_2(s) \end{pmatrix}$$

$$\begin{pmatrix} 1 & -V_{11}(s)V_{12}(s) \\ -V_{22}(s)V_{21}(s) & 1 \end{pmatrix} \begin{pmatrix} Y_1(s) \\ Y_2(s) \end{pmatrix} = \begin{pmatrix} V_{11}(s) & 0 \\ 0 & V_{22}(s) \end{pmatrix} \begin{pmatrix} U_1(s) \\ U_2(s) \end{pmatrix}$$

$$\begin{pmatrix} Y_1(s) \\ Y_2(s) \end{pmatrix} = \begin{pmatrix} 1 & -V_{11}(s)V_{12}(s) \\ -V_{22}(s)V_{21}(s) & 1 \end{pmatrix}^{-1} \begin{pmatrix} V_{11}(s) & 0 \\ 0 & V_{22}(s) \end{pmatrix} \begin{pmatrix} U_1(s) \\ U_2(s) \end{pmatrix}$$

15.6 Mehrgrößenregelung

Grob kann man 3 Konzepte zur Mehrgrößenregelung unterscheiden:

- *Dezentrale Regelung:* Es gibt für jedes Stell-/Regelgrößenpaar nur einen Regler. Die auf diesen Regelkreis durch die Kopplungen einwirkenden "Störungen" werden beim Entwurf berücksichtigt. Die Nachteile dieses Verfahrens werden anhand des nachfolgenden Beispiels erläutert.
- *Regelung mit Entkopplung:* Es gibt für jedes Stell-/Regelgrößenpaar je einen **Hauptregler** und für alle anderen Kopplungen zwischen den Stell- und Regelgrößen je einen **Entkopplungsregler**. Die Aufgabe der Entkopplungsregler ist es, den Einfluss der anderen Stellgrößen auf die jeweilige Regelgröße zu eliminieren oder zumindest soweit zu reduzieren, dass man auch die ursprünglich stark verkoppelten Regelkreise separat voneinander betrachten kann und den Hauptregler mit Methoden der Eingrößenregelung entwerfen kann. Dieses Verfahren wird im Folgenden ausführlich vorgestellt.
- *Echte Mehrgrößenregelung:* Der Mehrgrößenregler hat so viele Eingänge wie Regelgrößen (und Regelabweichungen) und so viele Ausgänge wie Stellgrößen. Die Kopplungen zwischen den Größen wird voll in einem einheitlichen Entwurf berücksichtigt und wo es möglich ist, zum Vorteil genutzt. Diese Verfahren gehen über den Stoffumfang dieser Vorlesung hinaus.

15.6 Mehrgrößenregelung

Dezentrale Regelung

Gegeben sei folgende Regelstrecke: $\underline{G}(s) = \begin{pmatrix} \frac{2}{s^2+3s+2} & \frac{k_{12}}{s+1} \\ \frac{k_{21}}{s^2+2s+1} & \frac{6}{s^2+5s+6} \end{pmatrix}$

Die Kopplung von Eingang u_2 auf Ausgang y_1 wird durch die Verstärkung k_{12} bestimmt; die Kopplung von Eingang u_1 auf Ausgang y_2 durch die Verstärkung k_{21} .

Bei der **dezentralen Regelung** wird jeder Regler ohne direkte Berücksichtigung der Kopplung entworfen. Die Kopplungseinflüsse werden als Störungen aufgefasst, die der jeweilige Regler auszuregulieren hat. Diese Sicht der Dinge ist natürlich falsch, weil Störungen immer unabhängige Größen sind, während die Kopplungen von den Größen des Regelkreises beeinflusst werden. Daher kann es bei der dezentralen Regelung zu Stabilitätsproblemen kommen, insbesondere für starke Kopplungen. Für schwache Kopplungen ist dies aber i.a. eine akzeptable Vorgehensweise.

Die Regler werden als Kompensationsregler so gewählt, dass für beide Regelkreise folgende Führungsübertragungsfunktion entsteht:

$$G_{w1}(s) = G_{w2}(s) = \frac{9}{s^2 + 4s + 9}$$

15.6 Mehrgrößenregelung

Entwurf der dezentralen Regler

Dieses Führungsverhalten wird unter Ignorierung der Kopplungen ($k_{12} = k_{21} = 0$) mit folgenden Reglern $R_1(s)$ und $R_2(s)$ für die Strecken $G_{11}(s)$ und $G_{22}(s)$ erreicht:

$$R_1(s) = \frac{4.5(s^2 + 3s + 2)}{s(s + 4)} \quad R_2(s) = \frac{1.5(s^2 + 5s + 6)}{s(s + 4)}$$

Gilt nun wirklich $k_{12} = k_{21} = 0$, dann wird das gewünschte Regelverhalten erreicht.

Regelverhalten bei schwacher Kopplung ($k_{12} = k_{21} = 0.1$)

Einflüsse der Kopplungen werden schnell ausgeregelt.

Regelverhalten bei starker Kopplung ($k_{12} = -1, k_{21} = 0.5$)

Die Güte der Regelungen leidet spürbar, aber die Regelkreise sind noch stabil.

Regelverhalten bei sehr starker Kopplung ($k_{12} = -2, k_{21} = -1$)

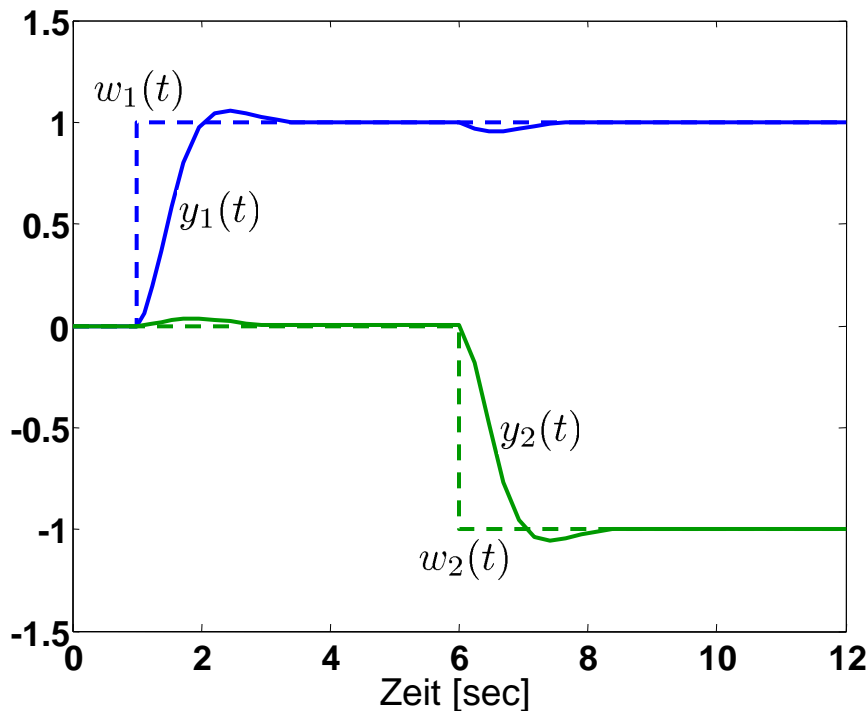
Regelkreise werden instabil. Spätestens an diesem Punkt, müssen die Kopplung mit in den Reglerentwurf integriert werden!

15.6 Mehrgrößenregelung

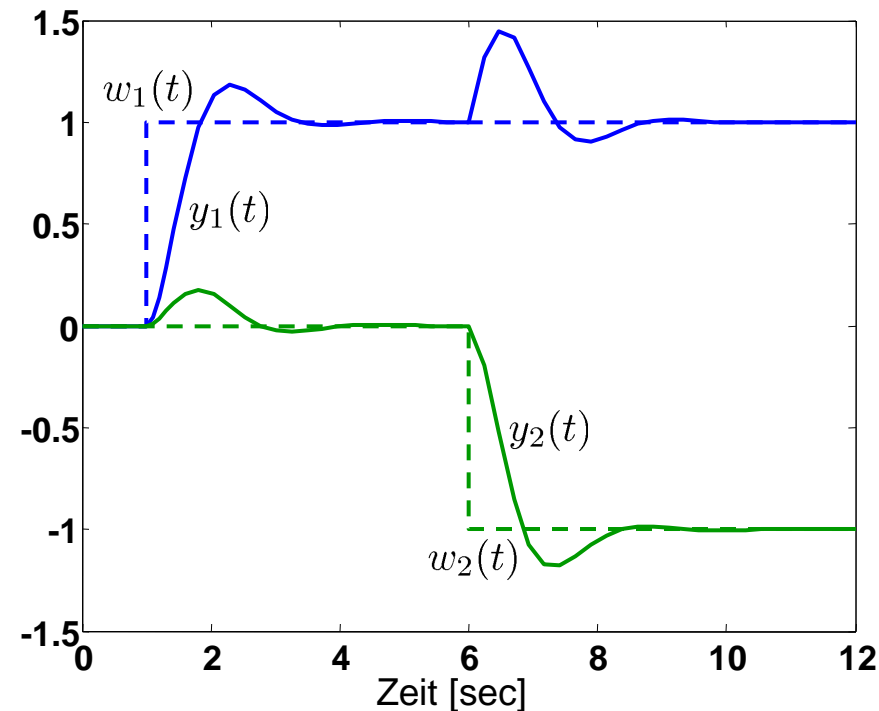
Simulationsergebnisse

Die Simulationsergebnisse zeigen, dass das einfache Konzept der dezentralen Regelung für schwach gekoppelte Systeme durchaus erfolgreich sein kann. Für starke Kopplungen wird hingegen ein Ansatz benötigt, der diese Kopplungen explizit mit in den Entwurf einbezieht.

Schwache Kopplung ($k_{12} = k_{21} = 0.1$)



Starke Kopplung ($k_{12} = -1, k_{21} = 0.5$)

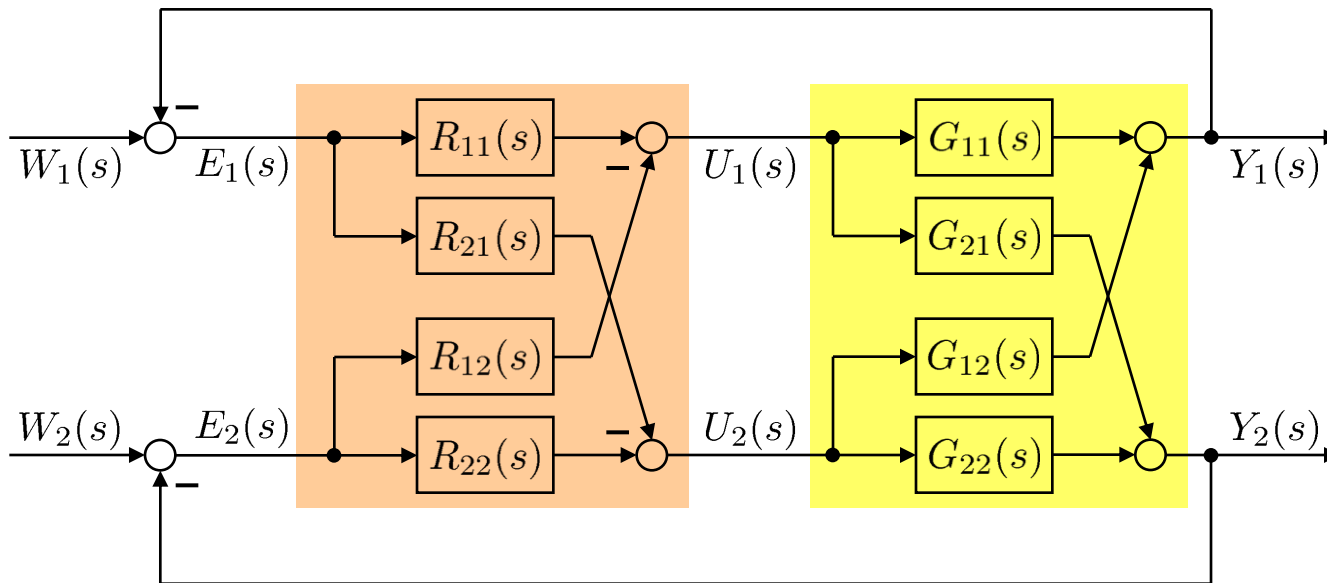


15.6 Mehrgrößenregelung

Grundidee der Entkopplung

Der Reiz der Entkopplung besteht darin, mittels der gleichen Idee wie bei der Störgrößen-aufschaltung, die Kopplungen zwischen den Regelkreisen zu eliminieren oder zumindest so weit abzuschwächen, dass die Mehrgrößenregelung zu mehreren Eingrößenregelungen vereinfacht werden kann. Diese können dann separat mit Hilfe der bekannten Methoden behandelt werden.

Mehrgrößenregelkreis



15.6 Mehrgrößenregelung

Aufbau des Reglers

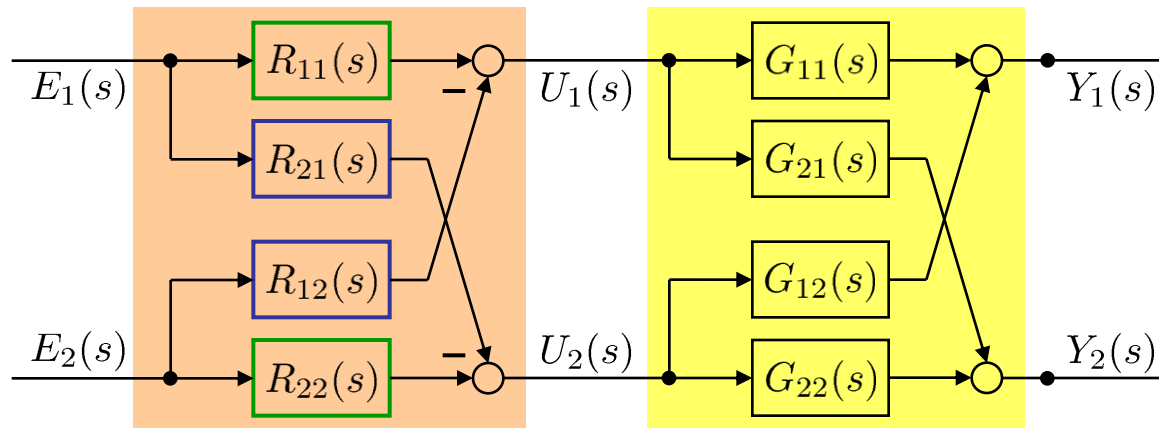
Der Regler enthält 2 **Entkopplungsregler** $R_{12}(s)$ und $R_{21}(s)$:

- $R_{12}(s)$ soll den Einfluss von $U_2(s)$ auf $Y_1(s)$ eliminieren.
- $R_{21}(s)$ soll den Einfluss von $U_1(s)$ auf $Y_2(s)$ eliminieren.

Des Weiteren enthält der Regler 2 **Hauptregler** $R_{11}(s)$ und $R_{22}(s)$:

- $R_{11}(s)$ soll mit der Stellgröße $U_1(s)$ die Regelgröße $Y_1(s)$ regeln.
- $R_{22}(s)$ soll mit der Stellgröße $U_2(s)$ die Regelgröße $Y_2(s)$ regeln.

Mehrgrößenregler und Strecke



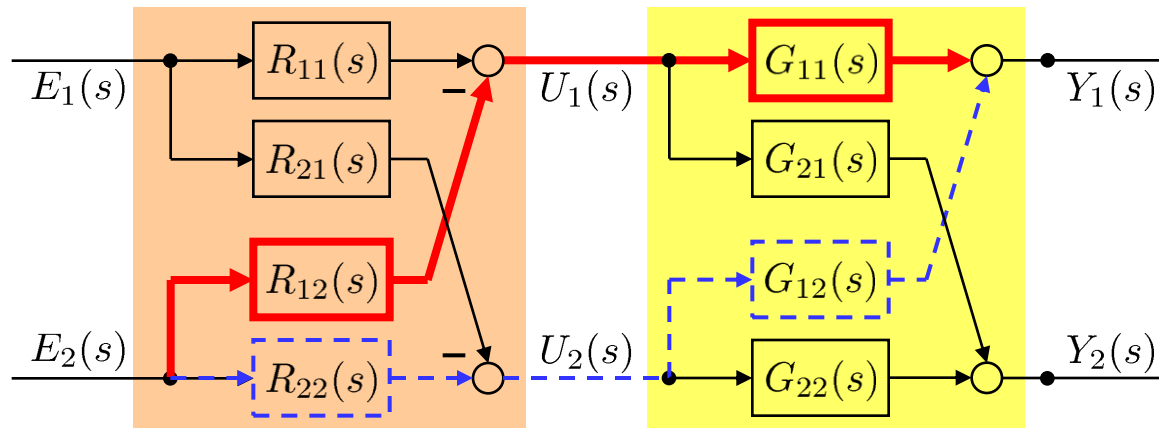
15.6 Mehrgrößenregelung

Entkopplung des 1. Regelkreises

Um den Einfluss des 2. Regelkreises auf den 1. zu eliminieren, muss die Übertragungsfunktion von $E_2(s) \rightarrow Y_1(s)$ entlang des blauen gestrichelten und roten Pfades identisch sein. Weil der rote Pfad mit negativem Vorzeichen wirkt löschen sich die beiden Einflüssen dann aus:

$$\underbrace{R_{12}(s)G_{11}(s)}_{\text{red arrow}} = \underbrace{R_{22}(s)G_{12}(s)}_{\text{blue dashed arrow}} \rightarrow R_{12}(s) = R_{22}(s) \frac{G_{12}(s)}{G_{11}(s)}$$

Mehrgrößenregler und Strecke



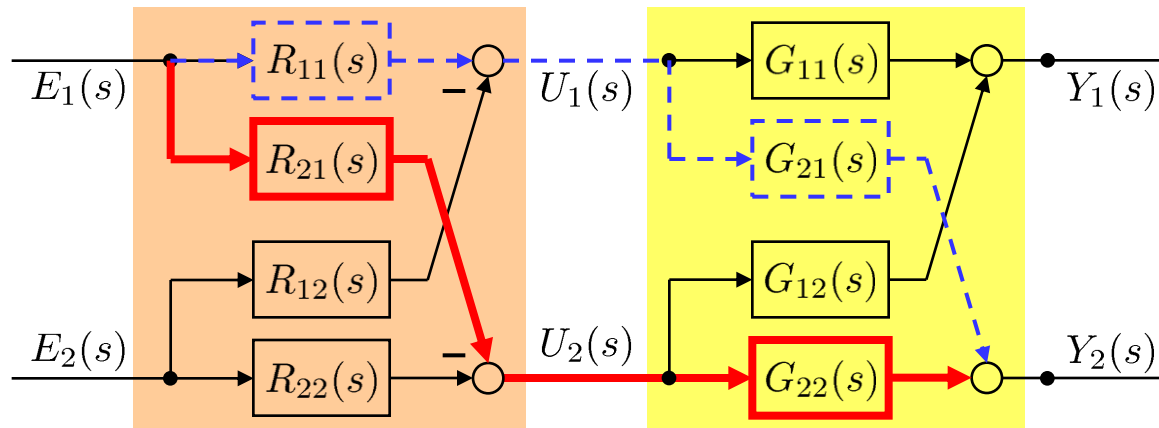
15.6 Mehrgrößenregelung

Entkopplung des 2. Regelkreises

Um den Einfluss des 1. Regelkreises auf den 2. zu eliminieren, muss die Übertragungsfunktion von $E_1(s) \rightarrow Y_2(s)$ entlang des blauen gestrichelten und roten Pfades identisch sein. Weil der rote Pfad mit negativem Vorzeichen wirkt löschen sich die beiden Einflüssen dann aus:

$$R_{21}(s)G_{22}(s) = R_{11}(s)G_{21}(s) \quad \rightarrow \quad R_{21}(s) = R_{11}(s) \frac{G_{21}(s)}{G_{22}(s)}$$

Mehrgrößenregler und Strecke



15.6 Mehrgrößenregelung

Realisierbarkeit der Entkopplung

Bezüglich der Realisierbarkeit der Entkopplungsregler gelten im Wesentlichen die gleichen Überlegungen wie bei der Störgrößenaufschaltung. Lässt sich eine perfekte Entkopplung nicht erreichen, dann sollte man eine näherungsweise Entkopplung versuchen. Einfachstes Mittel ist die rein statische Entkopplung, bei der die Übertragungsfunktionen auf ihre Verstärkungen reduziert werden.

Im Allgemeinen ist eine gute Entkopplung umso leichter zu realisieren, je langsamer die Kopplung zwischen den Regelkreisen ist. Die ist physikalisch unmittelbar einsichtig, ergibt sich aber auch aus der Entkopplungsgleichung:

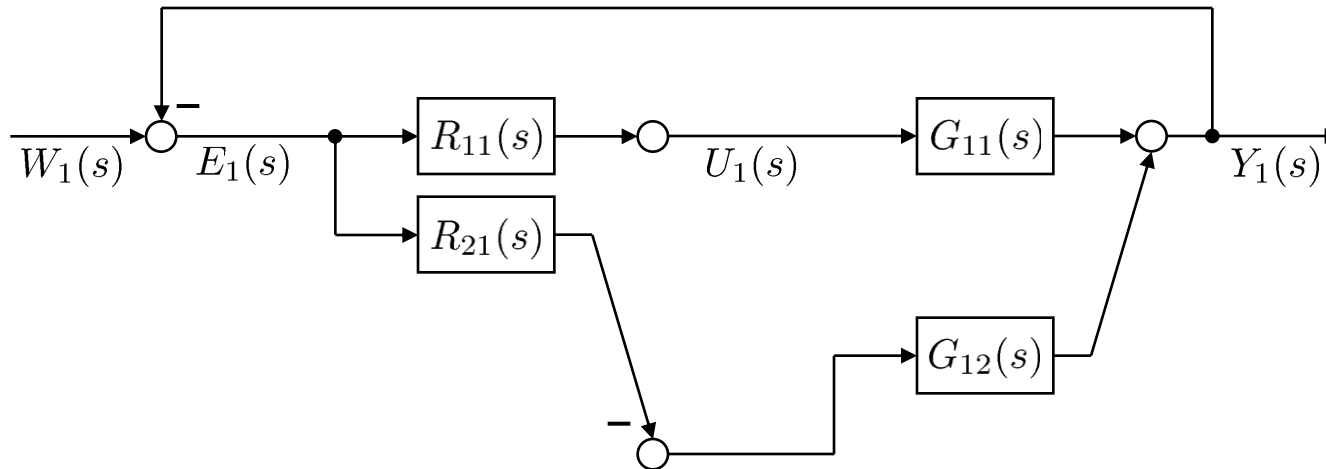
$$\text{Entkopplungsregler} = \text{Hauptregler} \frac{\text{Kopplung}}{\text{Hauptstrecke}}$$

Da der Hauptregler (wie jeder Regler) typischerweise "Zählergrad = Nennergrad" aufweist, muss der Polüberschuss der Kopplungsübertragungsfunktion größer gleich dem Polüberschuss der Hauptstrecke sein, damit der Entkopplungsregler realisierbar ist. Daneben gelten die bekannten Einschränkungen für Totzeiten und instabile Pol- und Nullstellen.

15.6 Mehrgrößenregelung

Entkoppelter 1. Regelkreis

Für den Fall der perfekten Entkopplung bleibt für den 1. Regelkreis folgender Teil der Mehrgrößenregelung übrig:



Damit ergibt sich die Übertragungsfunktion des 1. *offenen* Regelkreises:

$$G_{01}(s) = R_{11}(s)G_{11}(s) - R_{21}(s)G_{12}(s)$$

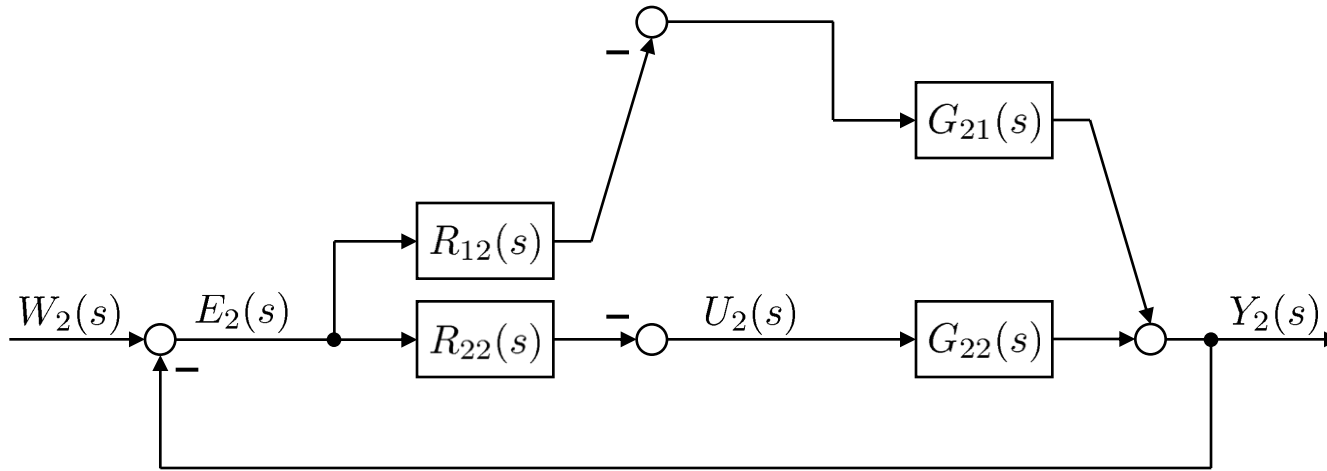
Wenn man die Gleichung für den Entkopplungsregler einsetzt, erhält man:

$$G_{01}(s) = R_{11}(s)G_{11}(s) - R_{11}(s) \frac{G_{21}(s)}{G_{22}(s)} G_{12}(s) = R_{11}(s) \underbrace{\frac{G_{11}(s)G_{22}(s) - G_{12}(s)G_{21}(s)}{G_{22}(s)}}_{G_{S1}(s)}$$

15.6 Mehrgrößenregelung

Entkoppelter 2. Regelkreis

Für den Fall der perfekten Entkopplung bleibt für den 2. Regelkreis folgender Teil der Mehrgrößenregelung übrig:



Damit ergibt sich die Übertragungsfunktion des 2. *offenen* Regelkreises:

$$G_{02}(s) = R_{22}(s)G_{22}(s) - R_{12}(s)G_{21}(s)$$

Wenn man die Gleichung für den Entkopplungsregler einsetzt, erhält man:

$$G_{02}(s) = R_{22}(s)G_{22}(s) - R_{22}(s) \frac{G_{12}(s)}{G_{11}(s)} G_{21}(s) = R_{22}(s) \frac{G_{11}(s)G_{22}(s) - G_{12}(s)G_{21}(s)}{G_{11}(s)}$$

15.6 Mehrgrößenregelung

Entkopplung mittels Matrix/Vektorrechnung

Der Zusammenhang zwischen Regelfehler $E(s)$ und Stellgröße $U(s)$ und zwischen Stellgröße $U(s)$ und Regelgröße $Y(s)$ lässt sich wie folgt beschreiben:

$$\begin{pmatrix} U_1(s) \\ U_2(s) \end{pmatrix} = \begin{pmatrix} R_{11}(s) & -R_{12}(s) \\ -R_{21}(s) & R_{22}(s) \end{pmatrix} \begin{pmatrix} E_1(s) \\ E_2(s) \end{pmatrix}$$

$$\begin{pmatrix} Y_1(s) \\ Y_2(s) \end{pmatrix} = \begin{pmatrix} G_{11}(s) & G_{12}(s) \\ G_{21}(s) & G_{22}(s) \end{pmatrix} \begin{pmatrix} U_1(s) \\ U_2(s) \end{pmatrix}$$

Damit ergibt sich für den offenen Mehrgrößenregelkreis:

$$\begin{aligned} \begin{pmatrix} Y_1(s) \\ Y_2(s) \end{pmatrix} &= \begin{pmatrix} G_{11}(s) & G_{12}(s) \\ G_{21}(s) & G_{22}(s) \end{pmatrix} \begin{pmatrix} R_{11}(s) & -R_{12}(s) \\ -R_{21}(s) & R_{22}(s) \end{pmatrix} \begin{pmatrix} E_1(s) \\ E_2(s) \end{pmatrix} \\ &= \begin{pmatrix} R_{11}(s)G_{11}(s) - R_{21}(s)G_{12}(s) & -R_{12}(s)G_{11}(s) + R_{22}(s)G_{12}(s) \\ R_{11}(s)G_{21}(s) - R_{21}(s)G_{22}(s) & -R_{12}(s)G_{21}(s) + R_{22}(s)G_{22}(s) \end{pmatrix} \begin{pmatrix} E_1(s) \\ E_2(s) \end{pmatrix} \end{aligned}$$

Eine Entkopplung bedeutet, dass wir eine Matrix in Diagonalgestalt fordern, d.h. $E_1(s)$ wirkt nur auf $Y_1(s)$ und $E_2(s)$ wirkt nur auf $Y_2(s)$. Alle anderen Matrixelemente müssen = 0 sein.

15.6 Mehrgrößenregelung

Äquivalenz der beiden Vorgehensweisen

Wir haben also die Forderungen:

$$\left(\begin{array}{cc} \underbrace{R_{11}(s)G_{11}(s) - R_{21}(s)G_{12}(s)}_{=0} & \underbrace{-R_{12}(s)G_{11}(s) + R_{22}(s)G_{12}(s)}_{=0} \\ R_{11}(s)G_{21}(s) - R_{21}(s)G_{22}(s) & -R_{12}(s)G_{21}(s) + R_{22}(s)G_{22}(s) \end{array} \right)$$

Das sind genau die selben Entkopplungsgleichungen, die wir zuvor über die Blockschaltbildbetrachtungen erhalten haben. Wenn die Blockschaltbilder komplizierter werden, ist die Matrix/Vektorrechnung vorzuziehen, da dabei weniger Flüchtigkeitsfehler unterlaufen.

Problem bei dieser Art der Entkopplung:

Der Entwurf des Haupt- und des Entkopplungsreglers hängen voneinander ab. Es wäre wünschenswert, wenn der Entkopplungsregler unabhängig vom Hauptregler entworfen werden könnte. Bei einer späteren Änderung (Adaption, Optimierung, etc.) des Hauptreglers wäre dann der Entkopplungsregler gar nicht tangiert.

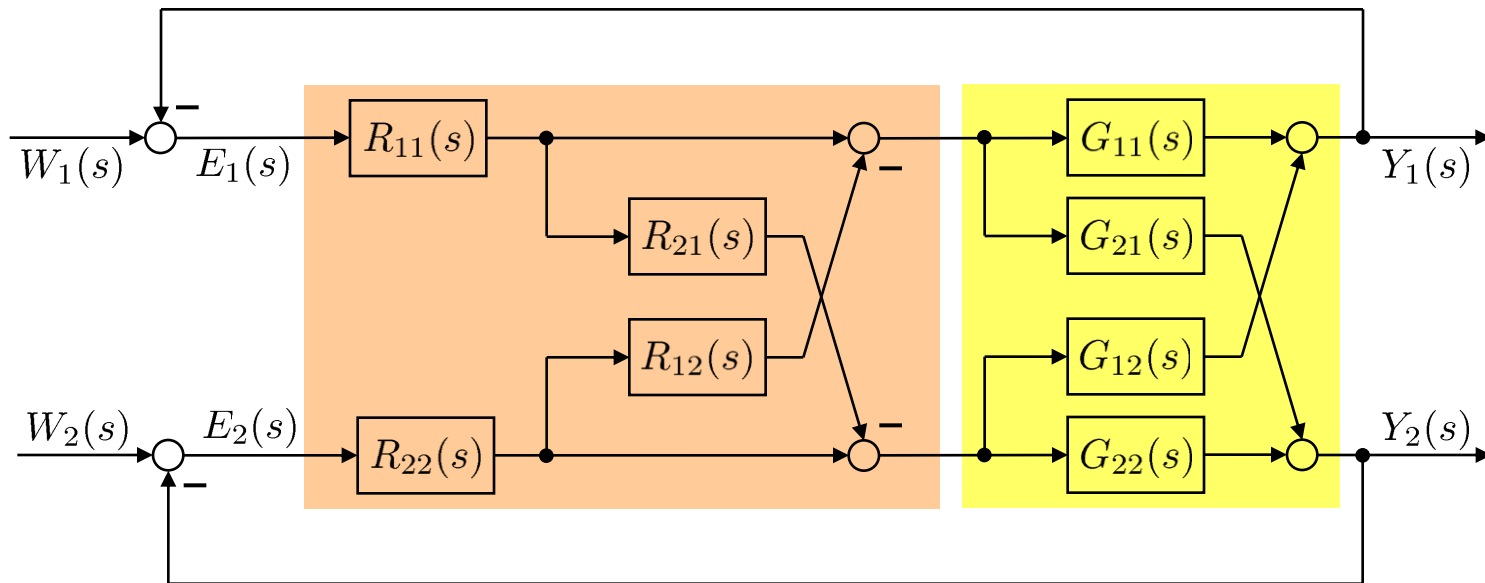
15.6 Mehrgrößenregelung

Entkopplung: Alternative Vorgehensweise

Die im Blockschaltbild dargestellte Art der Entkopplung ermöglicht den Entwurf des Entkopplungsreglers unabhängig von der Wahl des Hauptreglers. Von einer nachträglichen Änderung (Adaption, Optimierung, etc.) des Hauptreglers ist der Entkopplungsregler nicht betroffen.

Die Entkopplungsregler haben nun die jeweiligen Stellgrößen als Eingangssignale:

Mehrgrößenregelkreis



15.6 Mehrgrößenregelung

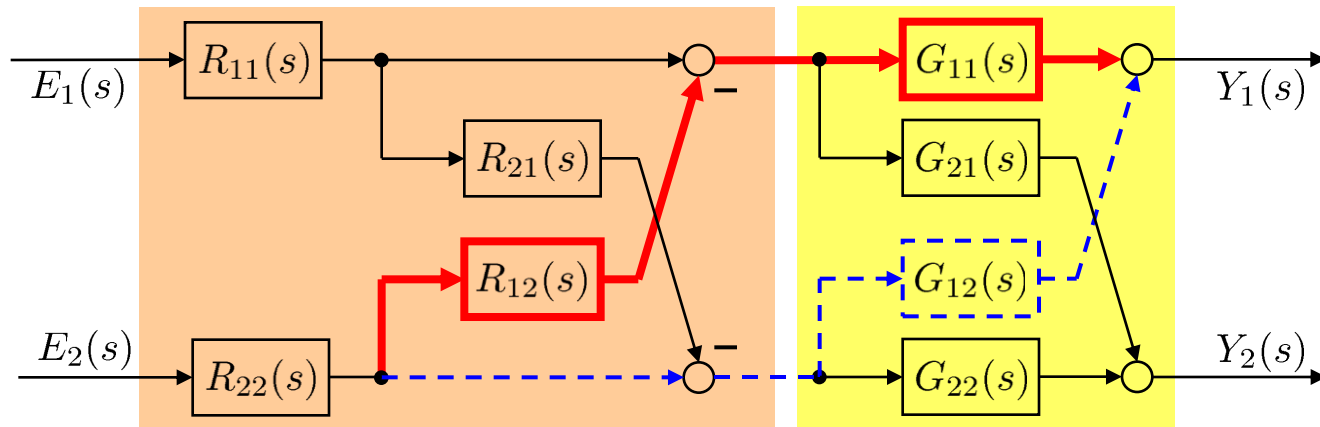
Alternative Entkopplung des 1. Regelkreises

Der blaue gestrichelte und der rote Pfad müssen sich kompensieren:

$$\underbrace{R_{12}(s)G_{11}(s)}_{\text{red arrow}} = \underbrace{G_{12}(s)}_{\text{blue dashed arrow}} \rightarrow \boxed{R_{12}(s) = \frac{G_{12}(s)}{G_{11}(s)}}$$

Dieser Entkopplungsregler hängt nur von der Regelstrecke ab! Er wird zuerst entworfen und ist dann Basis für den Entwurf des Hauptreglers.

Mehrgrößenregler und Strecke



15.6 Mehrgrößenregelung

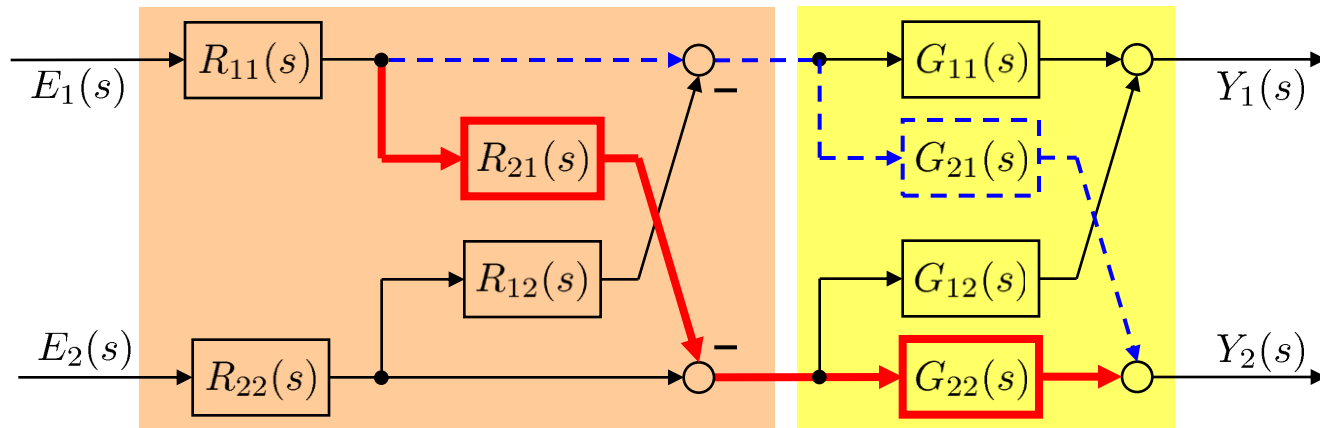
Alternative Entkopplung des 2. Regelkreises

Der blaue gestrichelte und der rote Pfad müssen sich kompensieren:

$$\underbrace{R_{21}(s)}_{\text{red arrow}} G_{22}(s) = \underbrace{G_{21}(s)}_{\text{blue dashed arrow}} \rightarrow \boxed{R_{21}(s) = \frac{G_{21}(s)}{G_{22}(s)}}$$

Dieser Entkopplungsregler hängt nur von der Regelstrecke ab! Er wird zuerst entworfen und ist dann Basis für den Entwurf des Hauptreglers.

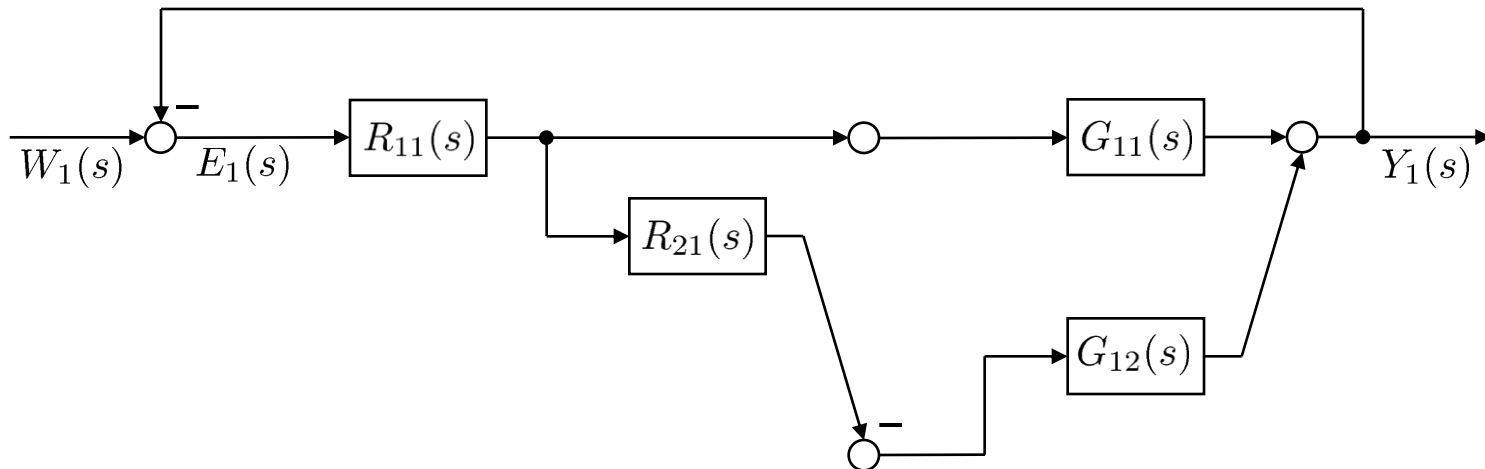
Mehrgrößenregler und Strecke



15.6 Mehrgrößenregelung

Alternativ entkoppelter 1. Regelkreis

Für den Fall der perfekten Entkopplung bleibt für den 1. Regelkreis folgender Teil der Mehrgrößenregelung übrig:



Damit ergibt sich die Übertragungsfunktion des 1. *offenen* Regelkreises:

$$G_{01}(s) = R_{11}(s)[G_{11}(s) - R_{21}(s)G_{12}(s)]$$

Wenn man die Gleichung für den Entkopplungsregler einsetzt erhält man:

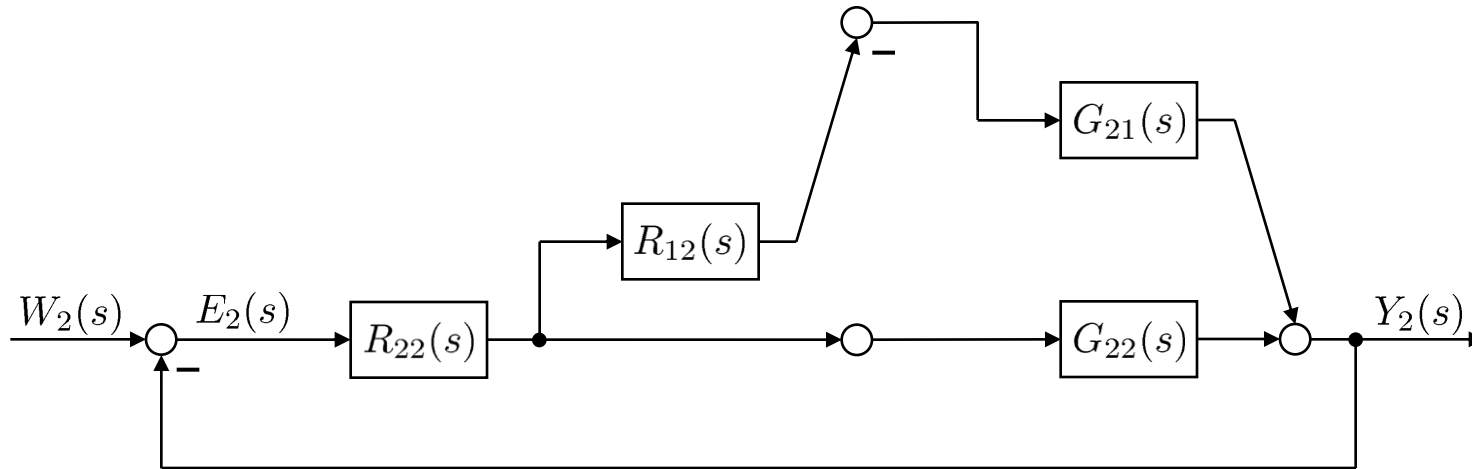
$$G_{01}(s) = R_{11}(s) \left[G_{11}(s) - \frac{G_{21}(s)}{G_{22}(s)} G_{12}(s) \right] = R_{11}(s) \frac{G_{11}(s)G_{22}(s) - G_{12}(s)G_{21}(s)}{G_{22}(s)}$$

$G_{S1}(s)$ (with an arrow pointing to the fraction in the equation above)

15.6 Mehrgrößenregelung

Alternativ entkoppelter 2. Regelkreis

Für den Fall der perfekten Entkopplung bleibt für den 2. Regelkreis folgender Teil der Mehrgrößenregelung übrig:



Damit ergibt sich die Übertragungsfunktion des 2. *offenen* Regelkreises:

$$G_{02}(s) = R_{22}(s)[G_{22}(s) - R_{12}(s)G_{21}(s)]$$

Wenn man die Gleichung für den Entkopplungsregler einsetzt erhält man:

$$G_{02}(s) = R_{22}(s) \left[G_{22}(s) - \frac{G_{12}(s)}{G_{11}(s)} G_{21}(s) \right] = R_{22}(s) \frac{G_{11}(s)G_{22}(s) - G_{12}(s)G_{21}(s)}{G_{11}(s)}$$

$G_{S2}(s)$

15.6 Mehrgrößenregelung

Entwurf des Hauptreglers

Ist der Mehrgrößenregelkreis *entkoppelt*, egal nach welcher Methode(!), dann können also die einzelnen Hauptregler für folgende Strecken separat entworfen werden:

$$R_{11}(s) : G_{S1}(s) = \frac{G_{11}(s)G_{22}(s) - G_{12}(s)G_{21}(s)}{G_{22}(s)} = G_{11}(s) \left[1 - \frac{G_{12}(s)G_{21}(s)}{G_{11}(s)G_{22}(s)} \right]$$

$$R_{22}(s) : G_{S2}(s) = \frac{G_{11}(s)G_{22}(s) - G_{12}(s)G_{21}(s)}{G_{11}(s)} = G_{22}(s) \left[1 - \frac{G_{12}(s)G_{21}(s)}{G_{11}(s)G_{22}(s)} \right]$$

Den folgenden Ausdruck nennt man

Koppelfaktor: $\kappa(s) = \frac{G_{12}(s)G_{21}(s)}{G_{11}(s)G_{22}(s)}$

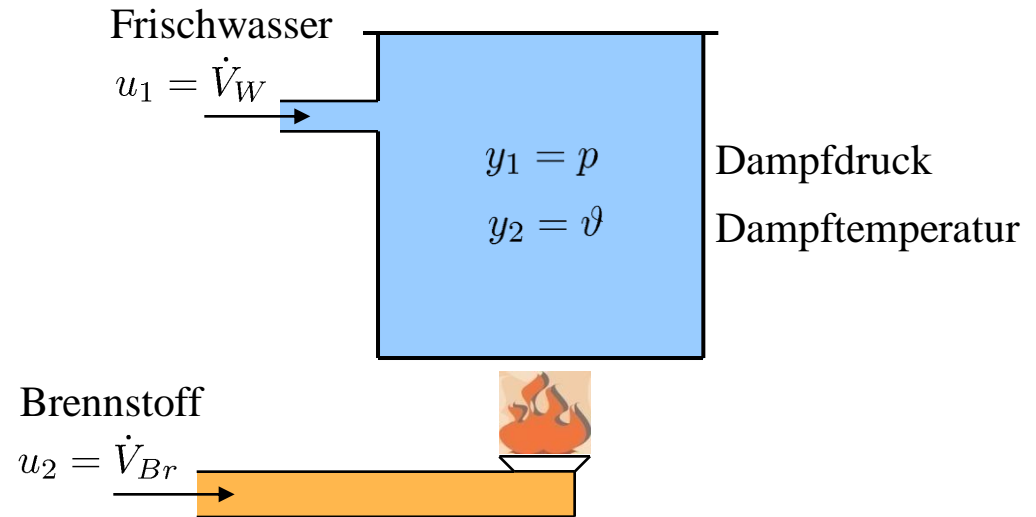
Ist der Koppelfaktor gleich Null, so können die Hauptregler ohne Berücksichtigung der Kopplungen entworfen werden, d.h. $R_{11}(s)$ für $G_{11}(s)$ und $R_{22}(s)$ für $G_{22}(s)$. Das ist dann der Fall, wenn $G_{12}(s)$ oder $G_{21}(s)$ gleich Null sind.

15.6 Mehrgrößenregelung

Beispiel: Dampferzeuger

Die beiden Volumenströme für Frischwasser und Brennstoff sind die Eingangsgrößen.

Der Druck und die Temperatur des Dampfes sind die Ausgangsgrößen.



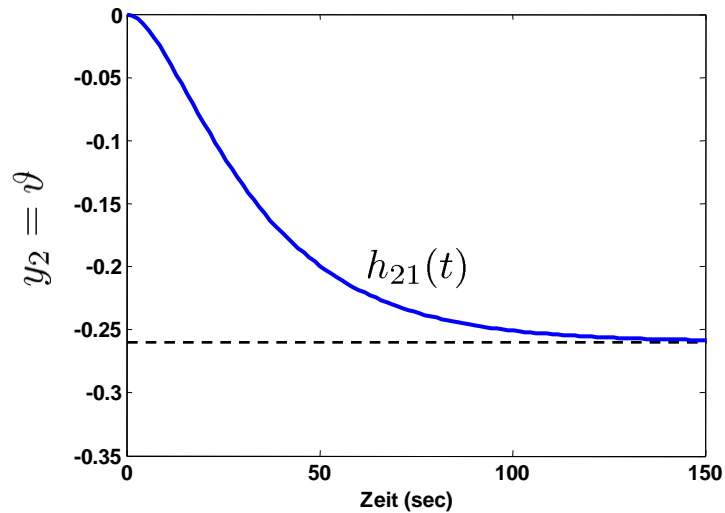
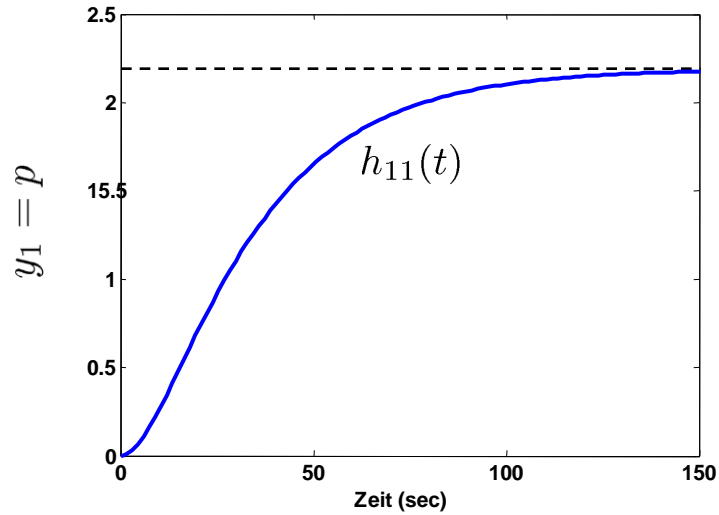
In [Lunze 2] wird für den Dampferzeuger folgendes Modell angegeben:

$$\begin{pmatrix} Y_1(s) \\ Y_2(s) \end{pmatrix} = \begin{pmatrix} \frac{0.0081}{(s+0.037)(s+0.1)} & \frac{0.00006353}{(s+0.0385)(s+0.05)} \\ \frac{-0.001}{(s+0.0385)(s+0.1)} & \frac{0.0001125}{(s+0.025)(s+0.05)} \end{pmatrix} \begin{pmatrix} U_1(s) \\ U_2(s) \end{pmatrix}$$

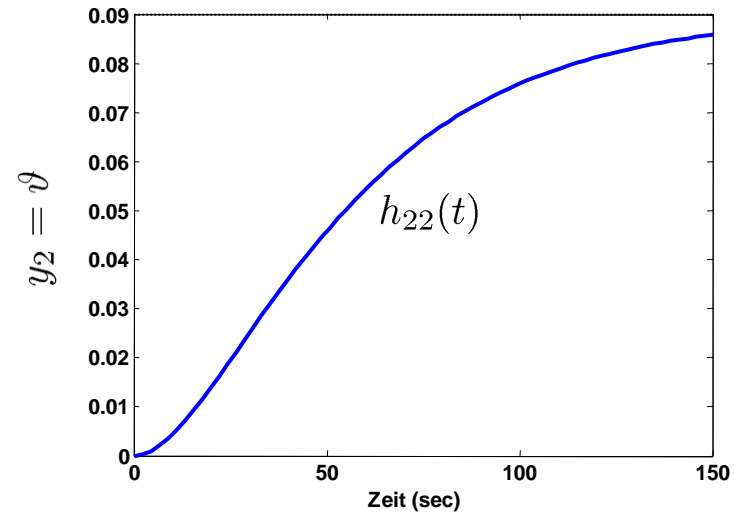
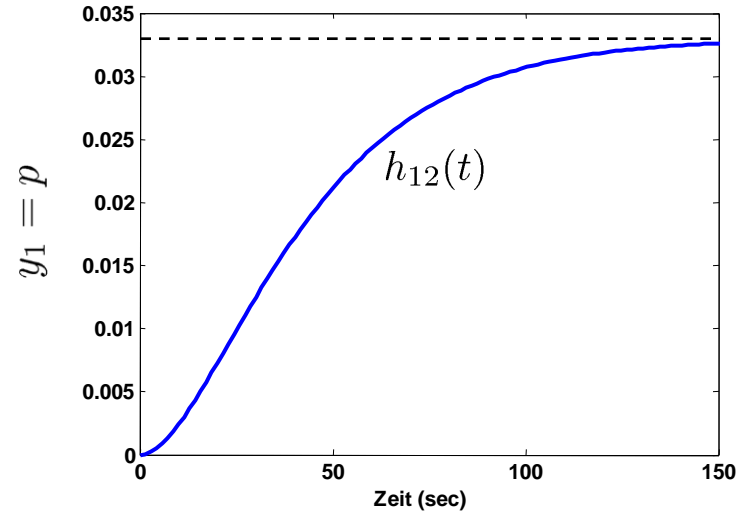
Die negative Verstärkung der Übertragungsfunktion $G_{21}(s)$ vom Frischwasservolumenstrom u_1 zur Dampftemperatur y_2 lässt sich so erklären: Bei einer erhöhten Menge an Frischwasser muss die zur Verfügung stehende Heizenergie auf eine größere Dampfmenge verteilt werden, was ein Absinken der Dampftemperatur zur Folge hat.

15.6 Mehrgrößenregelung

Sprung in $u_1 = \dot{V}_W$



Sprung in $u_2 = \dot{V}_{Br}$



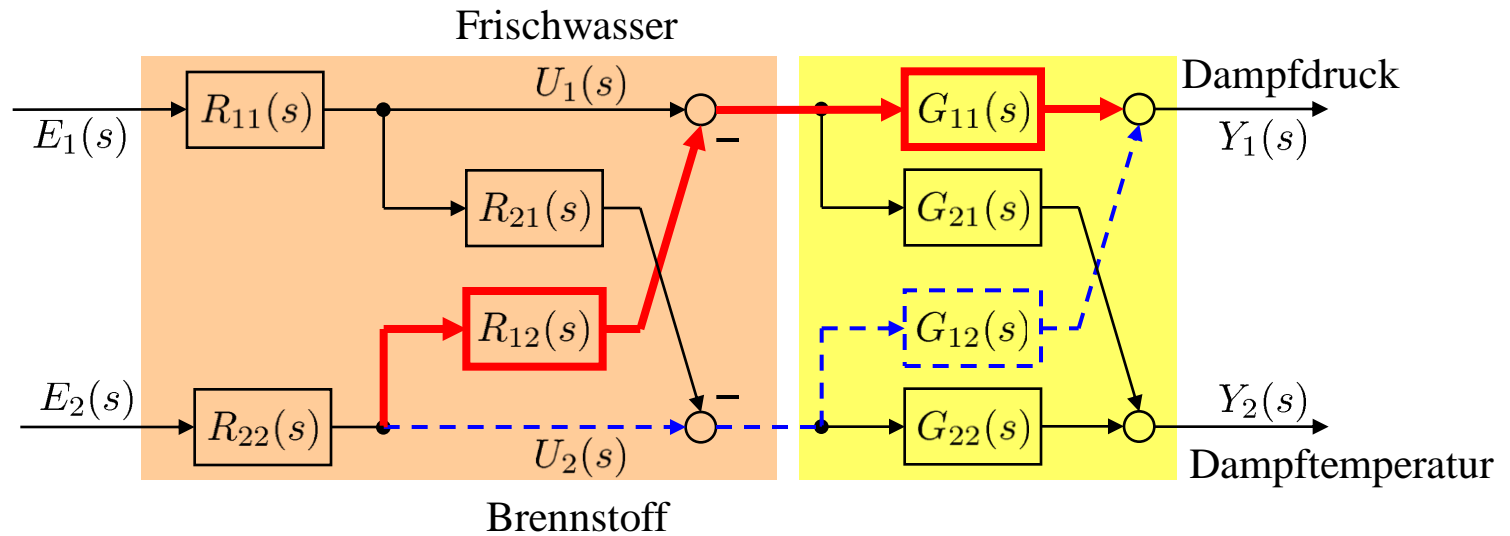
15.6 Mehrgrößenregelung

Entkopplung des Dampfdruckregelkreises

$$R_{12}(s) = \frac{G_{12}(s)}{G_{11}(s)} = 0.00784 \frac{(s + 0.037)(s + 0.1)}{(s + 0.0385)(s + 0.05)} \rightarrow \text{realisierbar!}$$

Stationäre Entkopplung:

$$R_{12}(0) = 0.0151$$



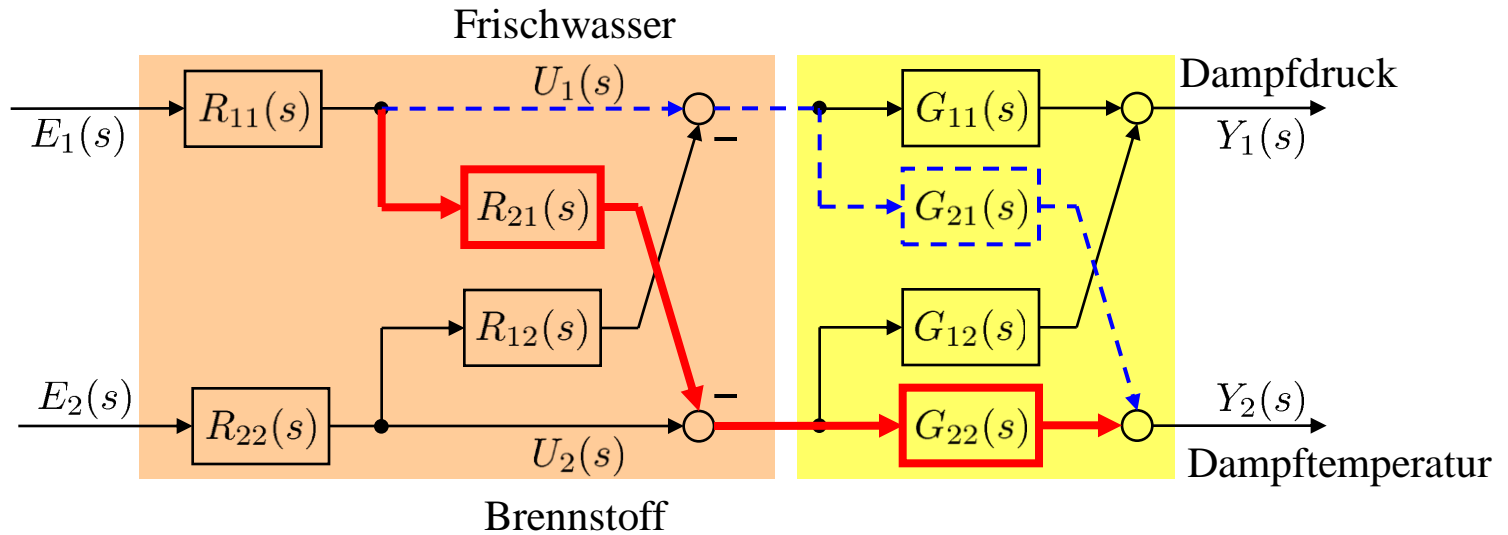
15.6 Mehrgrößenregelung

Entkopplung des Dampftemperaturregelkreises

$$R_{21}(s) = \frac{G_{21}(s)}{G_{22}(s)} = -8.89 \frac{(s + 0.025)(s + 0.05)}{(s + 0.0385)(s + 0.1)} \rightarrow \text{realisierbar!}$$

Stationäre Entkopplung:

$$R_{21}(0) = -2.886$$



15.6 Mehrgrößenregelung

Entkoppelte Regelkreise

$$\kappa(s) = \frac{G_{12}(s)G_{21}(s)}{G_{11}(s)G_{22}(s)} = -0.0697 \frac{(s + 0.037)\cancel{(s + 0.1)}(s + 0.025)\cancel{(s + 0.05)}}{(s + 0.0385)^2\cancel{(s + 0.1)}\cancel{(s + 0.05)}}$$

Dampfdruckregelkreis

$$\begin{aligned} G_{S1}(s) &= G_{11}(s) \left[1 - \frac{G_{12}(s)G_{21}(s)}{G_{11}(s)G_{22}(s)} \right] \\ &= \frac{0.0081}{(s + 0.037)(s + 0.1)} \left[1 + 0.0697 \frac{(s + 0.037)(s + 0.025)}{(s + 0.0385)^2} \right] \end{aligned}$$

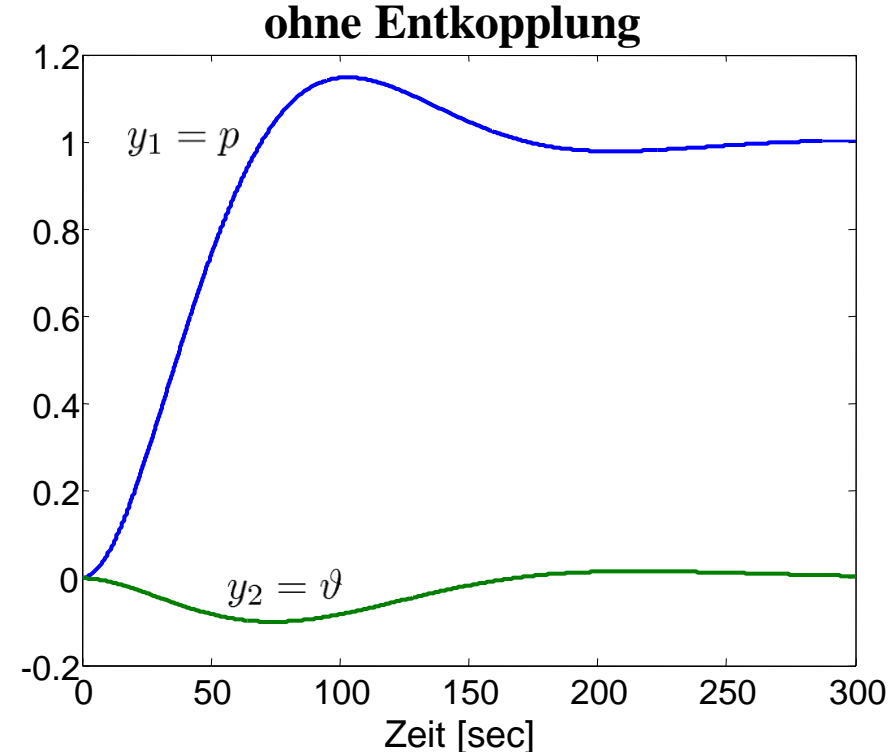
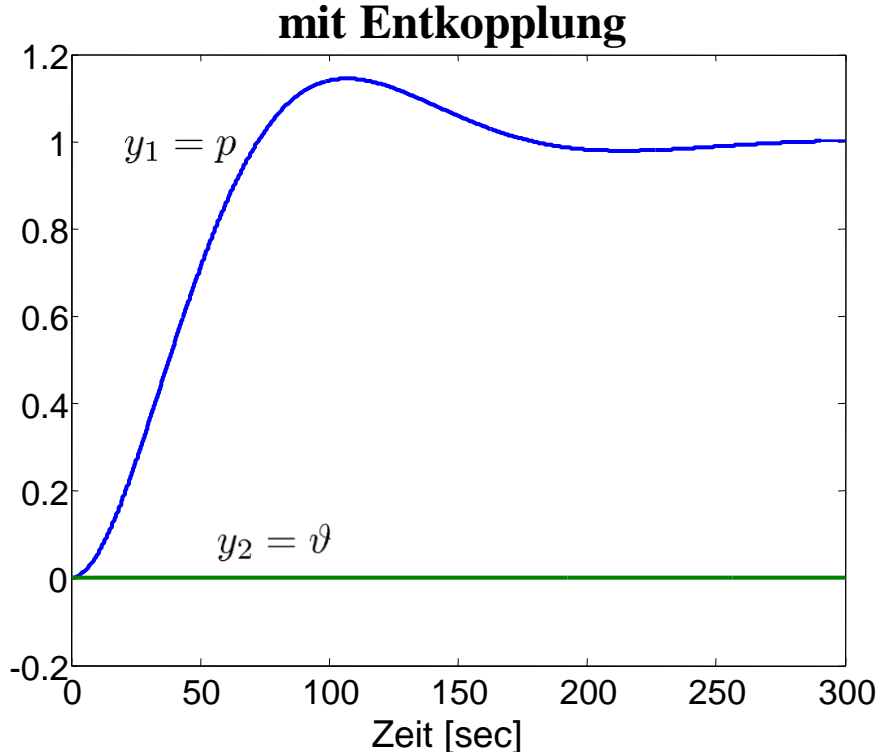
Dampftemperaturregelkreis

$$\begin{aligned} G_{S2}(s) &= G_{22}(s) \left[1 - \frac{G_{12}(s)G_{21}(s)}{G_{11}(s)G_{22}(s)} \right] \\ &= \frac{0.0001125}{(s + 0.025)(s + 0.05)} \left[1 + 0.0697 \frac{(s + 0.037)(s + 0.025)}{(s + 0.0385)^2} \right] \end{aligned}$$

15.6 Mehrgrößenregelung

Simulationsergebnisse

Mit den oben berechneten Entkopplungsreglern und PI-Hauptreglern, die jeweils den langsameren Streckenpol kürzen, erhält man folgendes Einregelverhalten der Regelgrößen für einen Einheitssprung im Sollwert des Dampfdruckes w_1 (links). Ohne Entkopplung würde sich der Dampfdruckregelkreis auf die Dampftemperatur auswirken (rechts).

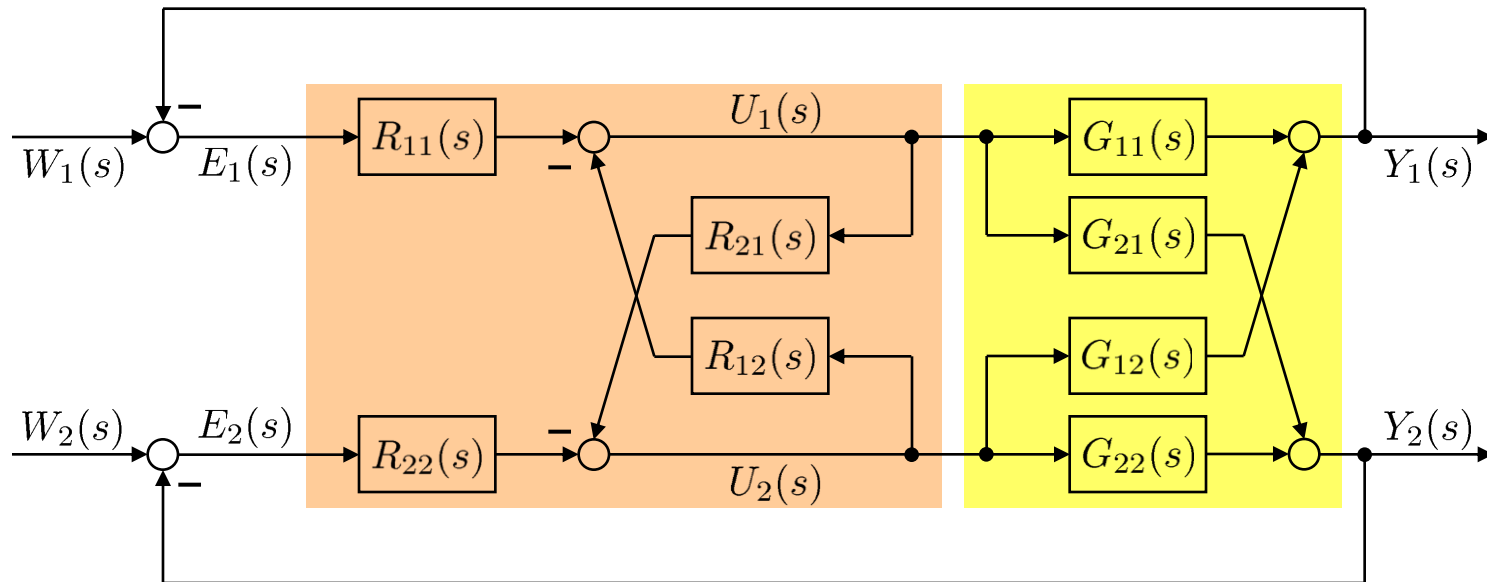


15.6 Mehrgrößenregelung

Entkopplung: Weitere alternative Vorgehensweise mit V-kanonischem Regler

Die im Blockschaltbild dargestellte Art der Entkopplung mit Hilfe einer V-kanonischen Reglerstruktur ermöglicht wie zuvor einen Entwurf des Entkopplungsreglers unabhängig von der Wahl des Hauptreglers. Außerdem wird der Koppelfaktor $\kappa(s) = 0$, d.h. im entkoppelten Regelkreis 1 muss *kein* Anteil des 2. Regelkreises berücksichtigt werden und umgekehrt.

Mehrgrößenregelkreis



15.6 Mehrgrößenregelung

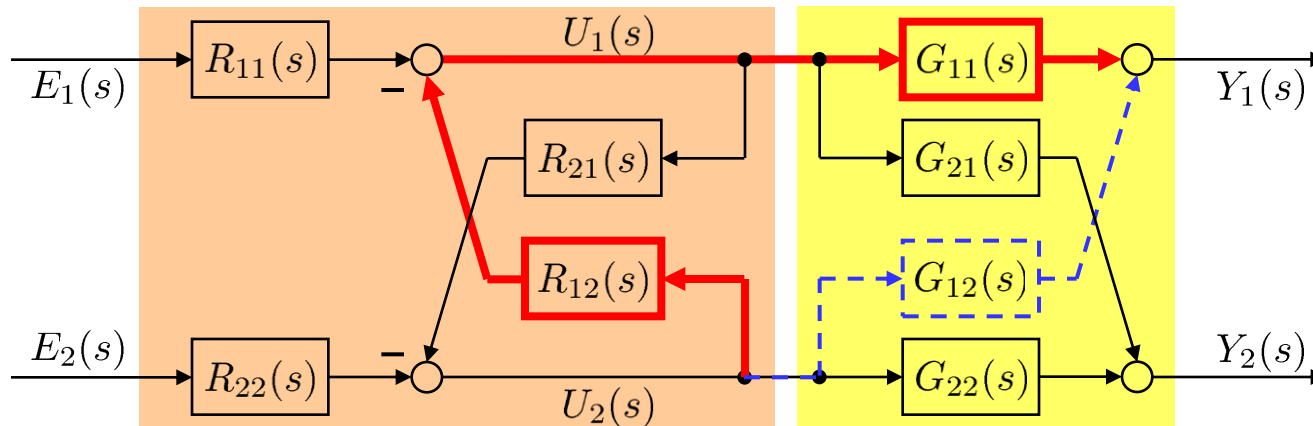
Entkopplung des 1. Regelkreises mit V-kanonischem Regler

Der blaue gestrichelte und der rote Pfad müssen sich kompensieren:

$$\begin{array}{c}
 \xrightarrow{\text{rot}} R_{12}(s)G_{11}(s) = G_{12}(s) \xrightarrow{\text{blau gestrichelt}} \\
 \rightarrow \boxed{R_{12}(s) = \frac{G_{12}(s)}{G_{11}(s)}}
 \end{array}$$

Die Entkopplungsbedingung ist identisch mit dem zuvor behandelten P-kanonischen Regler!

Mehrgrößenregler und Strecke



15.6 Mehrgrößenregelung

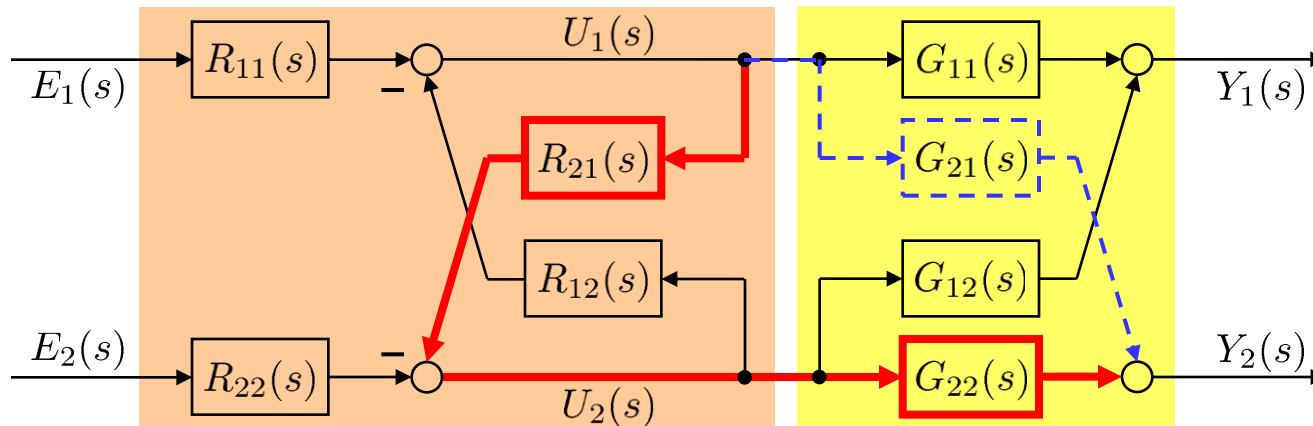
Alternative Entkopplung des 2. Regelkreises mit V-kanonischem Regler

Der blaue gestrichelte und der rote Pfad müssen sich kompensieren:

$$\begin{array}{c}
 \xrightarrow{\text{rot}} \\
 R_{21}(s)G_{22}(s) = G_{21}(s) \quad \rightarrow \quad \boxed{R_{21}(s) = \frac{G_{21}(s)}{G_{22}(s)}} \\
 \xrightarrow{\text{blau gestrichelt}}
 \end{array}$$

Die Entkopplungsbedingung ist identisch mit dem zuvor behandelten P-kanonischen Regler!

Mehrgrößenregler und Strecke

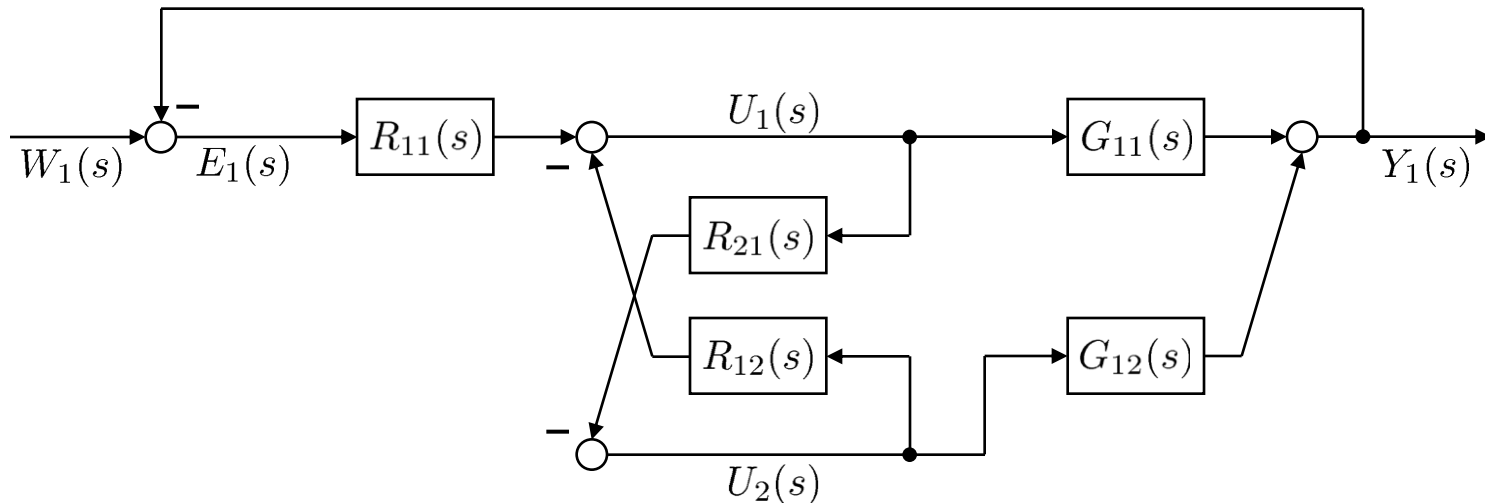


15.6 Mehrgrößenregelung

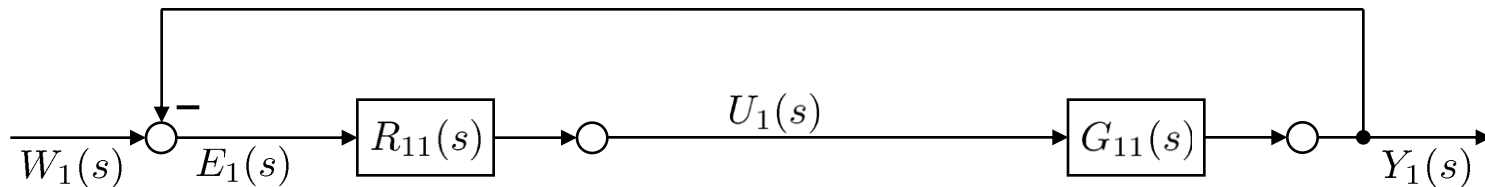
Entkoppelter 1. Regelkreis mit V-kanonischem Regler

Für den Fall der perfekten Entkopplung bleibt für den 1. Regelkreis folgender Teil der Mehrgrößenregelung übrig:

Mehrgrößenregelkreis



Da sich die Wirkung von $U_2(s)$ nach $Y_1(s)$ aufhebt, bleibt nur noch der 1. Regelkreis übrig:



15.6 Mehrgrößenregelung

Entkoppelter 1. Regelkreis mit V-kanonischem Regler

D.h. der entkoppelte 1. Regelkreis entspricht exakt einem Eliminieren des 2. Regelkreises, also:

$$G_{S1}(s) = G_{11}(s) \quad \rightarrow \quad G_{01}(s) = R_{11}(s)G_{11}(s)$$

Entsprechendes gilt für den entkoppelten 2. Regelkreis:

$$G_{S2}(s) = G_{22}(s) \quad \rightarrow \quad G_{02}(s) = R_{22}(s)G_{22}(s)$$

Damit ist der Koppelfaktor bei dieser Anordnung gleich Null:

$$\kappa(s) = 0$$

Die Kombination aus P-kanonischer Strecke und V-kanonischem Regler liefert also ein besonders einfaches entkoppeltes Ergebnis, bei dem die Ersatzregelstrecken $G_{S1}(s)$ und $G_{S2}(s)$ identisch mit den ursprünglichen Regelstrecken $G_{11}(s)$ und $G_{22}(s)$ sind!

15.6 Mehrgrößenregelung

Entkopplung mit V-kanonischem Regler

Eine Rechnung ist möglicherweise leichter verständlich als obige Blockschaltbildumformung:

Reglergleichungen:

$$E_1(s)R_{11}(s) - R_{12}(s)U_2(s) = U_1(s)$$

$$E_2(s)R_{22}(s) - R_{21}(s)U_1(s) = U_2(s)$$

In Matrixform umgerechnet:

$$\underline{U}(s) = \frac{1}{1 - R_{12}(s)R_{21}(s)} \begin{pmatrix} R_{11}(s) & -R_{12}(s)R_{22}(s) \\ -R_{11}(s)R_{21}(s) & R_{22}(s) \end{pmatrix} \underline{E}(s)$$

Mit der Regelstrecke

$$\underline{G}(s) = \begin{pmatrix} G_{11}(s) & G_{12}(s) \\ G_{21}(s) & G_{22}(s) \end{pmatrix}$$

ergibt sich der offene Regelkreis („s“ entfällt aus Platzgründen):

$$\underline{Y}(s) = \frac{1}{1 - R_{12}R_{21}} \begin{pmatrix} R_{11}[G_{11} - G_{12}R_{21}] & R_{22}[G_{12} - G_{11}R_{12}] \\ R_{11}[G_{21} - G_{22}R_{21}] & R_{22}[G_{22} - G_{21}R_{12}] \end{pmatrix} \underline{E}(s)$$

Damit erhält man die Entkopplungsbedingungen (Nebendiagonalelemente = 0!):

15.6 Mehrgrößenregelung

Entkopplung mit V-kanonischem Regler

$$R_{22}(s)[G_{12}(s) - G_{11}(s)R_{12}(s)] = 0$$

$$R_{11}(s)[G_{21}(s) - G_{22}(s)R_{21}(s)] = 0$$

$$R_{12}(s) = \frac{G_{12}(s)}{G_{11}(s)}$$

$$R_{21}(s) = \frac{G_{21}(s)}{G_{22}(s)}$$

Diese Entkopplungsbedingungen sind identisch mit den zuvor aus Blockschaltbildumformungen berechneten! Bei größeren Mehrgrößensystemen ist i.a. die Matrix/Vektorrechnung effektiver und weniger fehleranfällig und damit vorzuziehen.

Einsetzen der Entkopplungsbedingungen führt auf:

$$\underline{Y}(s) = \frac{1}{1 - \frac{G_{12} G_{21}}{G_{11} G_{22}}} \begin{pmatrix} R_{11} [G_{11} - G_{12} \frac{G_{21}}{G_{22}}] & R_{22} [G_{12} - G_{11} \frac{G_{12}}{G_{11}}] \\ R_{11} [G_{21} - G_{22} \frac{G_{21}}{G_{22}}] & R_{22} [G_{22} - G_{21} \frac{G_{12}}{G_{11}}] \end{pmatrix} \underline{E}(s)$$

$$= \frac{1}{1 - \frac{G_{12} G_{21}}{G_{11} G_{22}}} \begin{pmatrix} R_{11} G_{11} [1 - \frac{G_{12} G_{21}}{G_{11} G_{22}}] & 0 \\ 0 & R_{22} G_{22} [1 - \frac{G_{21} G_{12}}{G_{22} G_{11}}] \end{pmatrix} \underline{E}(s)$$

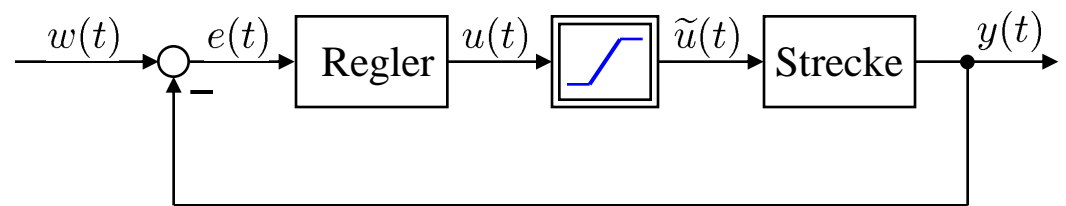
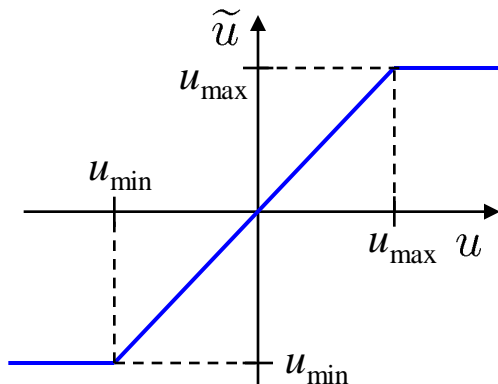
$$Y(s) = \begin{pmatrix} R_{11} G_{11} & 0 \\ 0 & R_{22} G_{22} \end{pmatrix} \underline{E}(s)$$

15.7 Anti-Wind-up-Methoden

Stellgrößenbeschränkungen

In der Realität ist kein Regelkreis wirklich linear. Die vorhanden Nichtlinearitäten können aber in vielen Fällen vernachlässigt werden, zumal ein Regelkreis ja robust ausgelegt ist, so dass kleine Abweichungen zwischen Modell und Realität verkraftet werden können.

Eine wesentliche Nichtlinearität sollte man allerdings bei der Analyse und Synthese von Regelkreisen immer im Hinterkopf haben: die **Beschränkungen des Stellglieds**. Jedes Stellglied hat eine untere und obere Begrenzung (Anschlag) über den hinaus die Stellgröße nicht verändert werden kann: Ventil (offen / geschlossen), Drosselklappe (offen / geschlossen), Motor oder Pumpe (Stillstand / Maximalleistung), etc. D.h. die Stellgröße bewegt sich immer in einem vorgegebenen Intervall $[u_{\min}, u_{\max}]$. Im einfachsten Fall ist die Charakteristik dazwischen linear, so dass sich eine Kennlinie folgenden Typs ergibt:



15.7 Anti-Wind-up-Methoden

Behandlung von Stellgrößenbeschränkungen beim Reglerentwurf

Wenn der Stellgrößenwunsch $u(t)$, den der Regler ausgibt, außerhalb des Intervalls $[u_{\min}, u_{\max}]$ liegt, entsteht eine Abweichung zwischen dem Stellgrößenwunsch und der tatsächlich vom Stellglied ausgegebenen Stellgröße $\tilde{u}(t)$. Die Wirkung des Reglers ist dann nicht so stark (aggressiv) wie ursprünglich gedacht und beim Reglerentwurf geplant. Es gibt drei Strategien, dieser Tatsache beim Reglerentwurf Rechnung zu tragen:

1. Man entwirft den Regler so vorsichtig, dass die Wunschstellgröße $u(t)$ niemals das Intervall $[u_{\min}, u_{\max}]$ verlässt. Dies führt zu sehr langsamen Regelungen mit geringen Kreisverstärkungen und ist daher oft nicht akzeptabel.
2. Man entwirft den Regler aggressiver als nach 1. und überprüft die Auswirkungen der Stellgrößenbeschränkung per Simulation. Im Reglerentwurfsverfahren (Parameter—optimierung) selbst bleibt die Stellgrößenbeschränkung aber unberücksichtigt. Ihre Auswirkungen werden erst im Nachhinein überprüft. Das ist die Standardstrategie. Hier kann es zu **Wind-up-Effekten** kommen, siehe nächste Folie.
3. Man berücksichtigt die Stellgrößenbeschränkungen schon während des Reglerentwurfs (prädiktive Regelung). Das erfordert den Einsatz von *Optimierungsverfahren mit Nebenbedingungen*. Solche Ansätze sind komplizierter und werden hier nicht behandelt.

15.7 Anti-Wind-up-Methoden

Erklärung des Wind-up-Effekts

Ein mögliches Beispielszenario wäre folgendes:

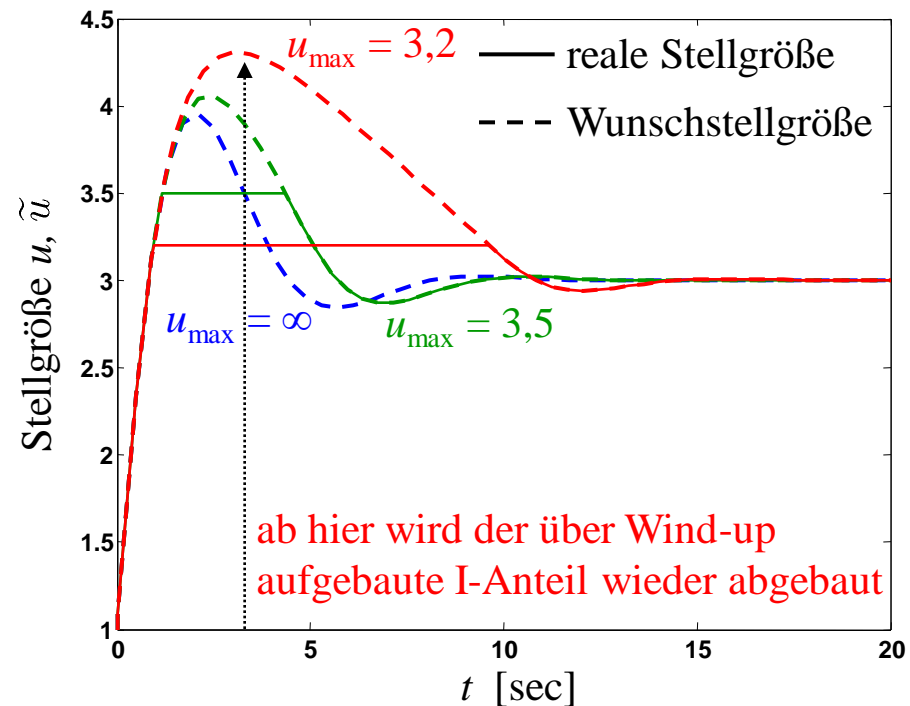
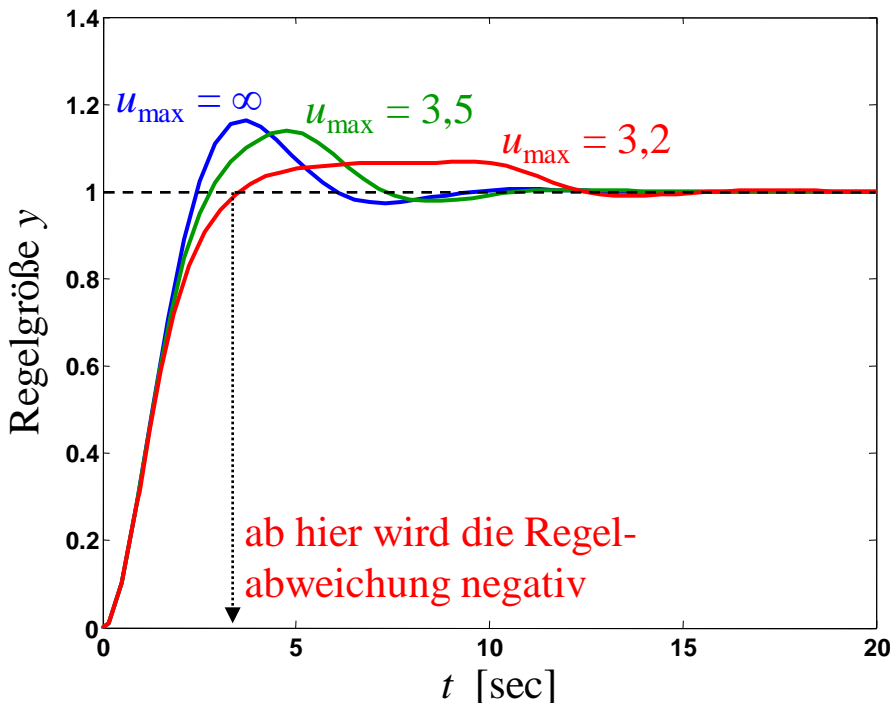
1. Ein PI-Regler gibt steigende Stellgrößen aus. Der I-Anteil wächst, solange die Regelabweichung positiv ist (Regelgröße $<$ Führungsgröße).
2. Die Wunschstellgröße überschreitet u_{\max} , d.h. es gibt eine Abweichung zwischen der Wunschstellgröße $u(t)$ und tatsächlich vom Stellglied ausgegebenen Größe u_{\max} .
3. Weil ein kleineres Stellsignal auf die Regelstrecke gegeben wird als eigentlich vom Regler gewünscht, baut sich die Regelabweichung langsamer ab, als es für ein lineares Modell (ohne Beschränkungen) zu erwarten wäre.
4. Der I-Anteil, und damit die Wunschstellgröße, steigt weiter an, weil die Regelabweichung aufintegriert wird. Erst beim Verschwinden der Regelabweichung bleibt der I-Anteil stehen.
5. Während der Zeit, die die Stellgröße an ihrer oberen Beschränkung hing, hat sich durch den steigenden I-Anteil die Wunschstellgröße immer weiter nach oben entfernt. Wenn nun die Regelabweichung negativ wird, braucht der I-Anteil lange Zeit, bis er wieder auf "Normalmaß" (nämlich entsprechend einer Wunschstellgröße $u(t) \leq u_{\max}$) gesunken ist.

15.7 Anti-Wind-up-Methoden

Illustration des Wind-up-Effekts

Je schärfer die Begrenzung wirkt (d.h. je kleiner u_{\max} ist), desto schlechter wird die Regelgüte. Das liegt an zwei Faktoren:

- Die Stellgröße kann nicht mehr so stark eingreifen.
- Durch den Windup-Effekt hängt die Wunschstellgröße lange über u_{\max} fest.

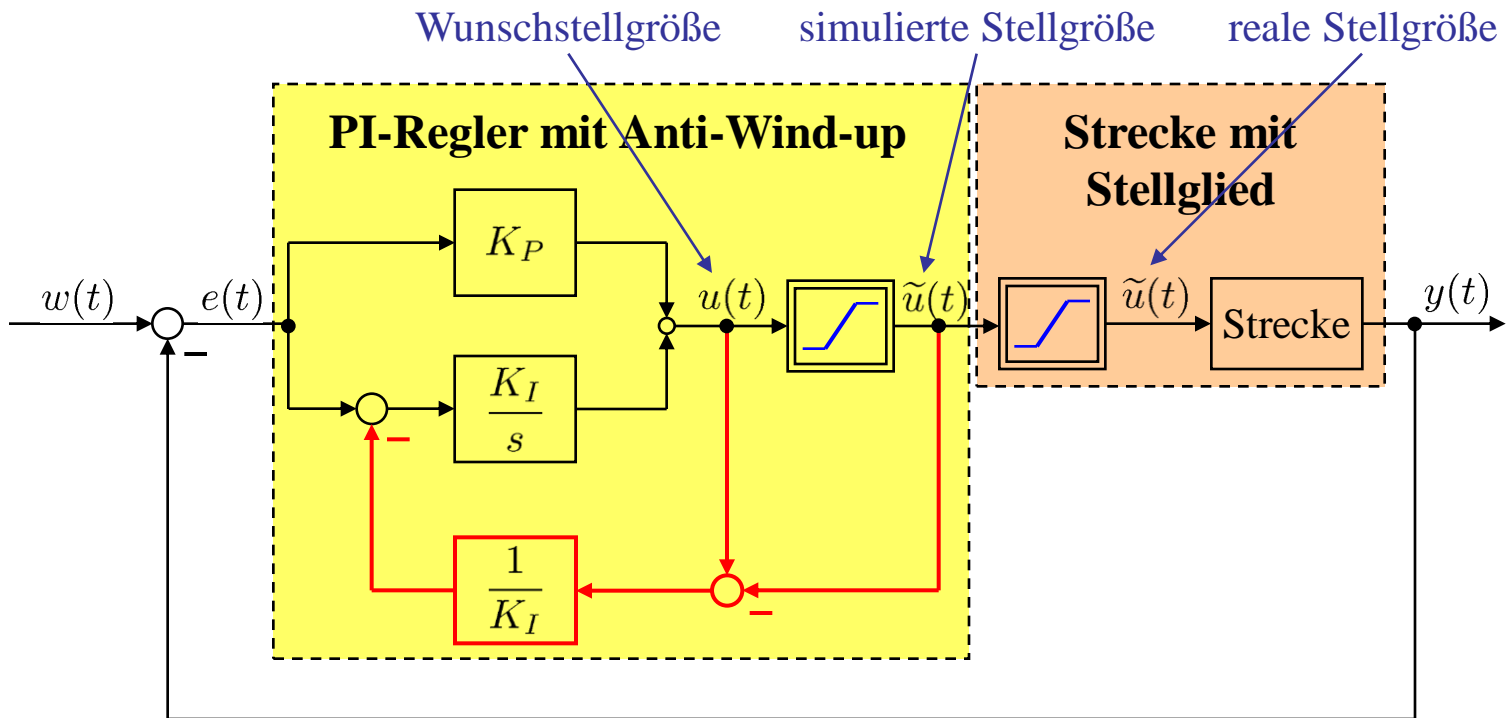


15.7 Anti-Wind-up-Methoden

Anti-Wind-up-Methode

Der P-Anteil bleibt mit K_P unverändert. Der I-Anteil K_I kann durch geeignete Anti-Wind-up-Methoden abgeschwächt werden:

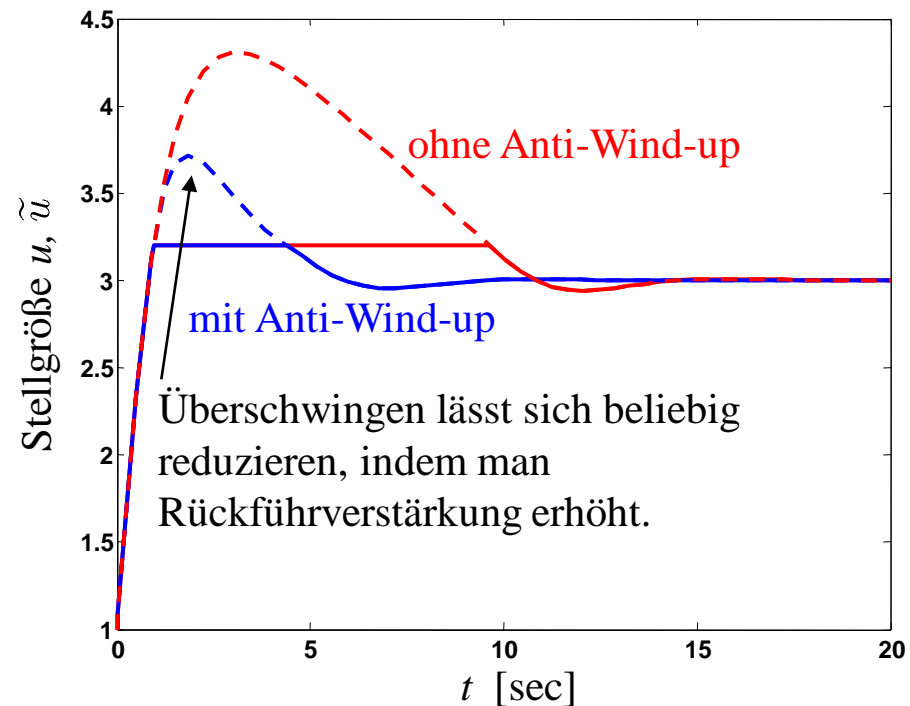
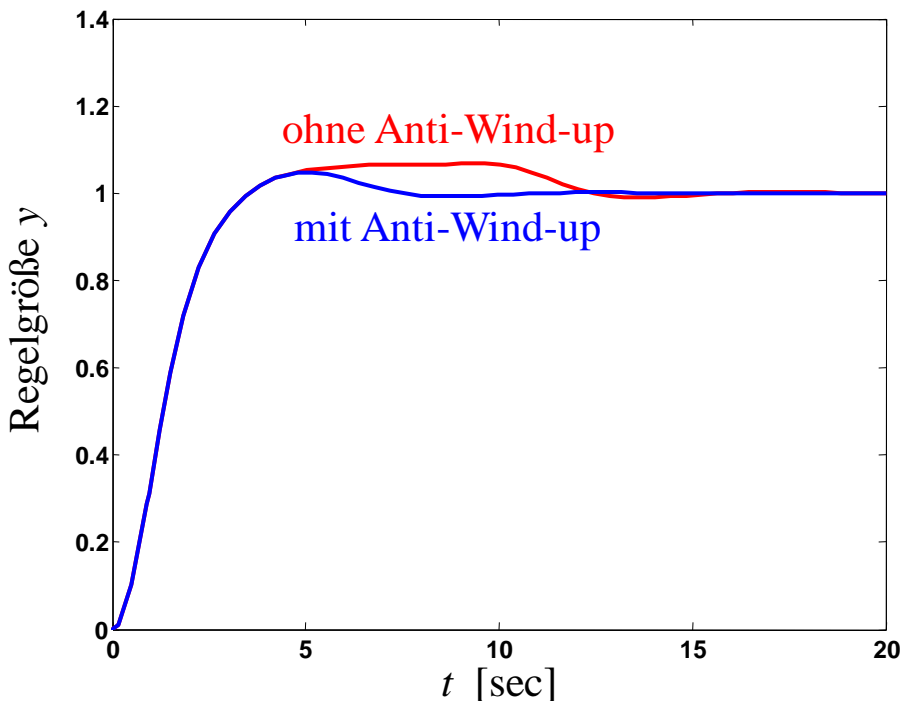
Eine Rückkopplung der Abweichung zwischen Wunschstellgröße $u(t)$ und realer Stellgröße $\tilde{u}(t)$ auf den Integrator wirkt dem Wind-up Effekt entgegen.



15.7 Anti-Wind-up-Methoden

Illustration der Anti-Wind-up-Methode

- Zur Rückführung wird das simulierte $\tilde{u}(t)$, nicht das gemessene reale Signal verwendet, um Verzögerungen, Verzerrungen und Rauschen des Sensors zu vermeiden.
- Verstärkung der Anti-Wind-up-Rückführung zu $1/K_I$ gewählt (grobes Ausprobieren).
- Je größer die Verstärkung gewählt wird, desto mehr wird der Wind-up-Effekt unterdrückt.

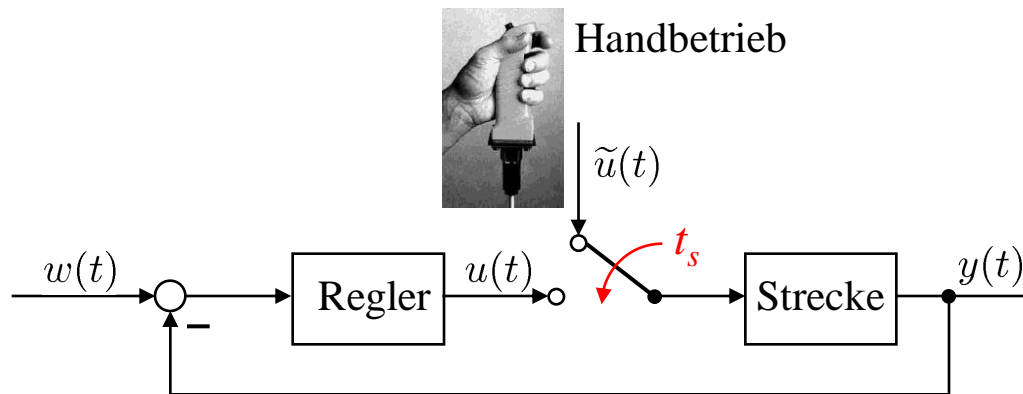


15.8 Stoßfreies Umschalten (Bumpless Transfer)

In der Praxis kommt es häufig vor, dass ein Regler auf Handbetrieb umgeschaltet werden muss, bzw. vom Handbetrieb wieder zurück auf den automatischen Regler. Auch kann es sinnvoll sein, z.B. während der Weitentwicklungsphase des Reglers, zwischen zwei Reglern hin- und herzuschalten oder ein komplexer adaptiver Regler schaltet automatisch auf einen einfacheren Backup-Regler um, wenn die Überwachungsebene erkennt, dass z.B. bei der Adaption etwas schief gelaufen ist. Evtl. stehen auch verschiedene Regler zur Verfügung, die für unterschiedliche Betriebsbereiche optimiert worden sind und zwischen denen umgeschaltet werden soll.

In all diesen Situationen stellt sich die Frage: **Wie schaltet man stoßfrei um?**

D.h. zum Umschaltzeitpunkt t_s wird gefordert: $u(t_s) = \tilde{u}(t_s)$



15.8 Stoßfreies Umschalten (Bumpless Transfer)

Schauen wir uns dieses Problem für die Umschaltung von Handbetrieb auf einen PID-Regler genauer an. Der PID-Regler liefert zum Zeitpunkt t_s die Stellgröße:

$$u(t_s) = K_P e(t_s) + K_D \dot{e}(t_s) + K_I \int_0^{t_s} e(\tau) d\tau,$$

wenn sein Integrator zum Zeitpunkt $t = 0$ mit der Anfangsbedingung 0 initialisiert wurde. Wenn zuvor, d.h. für $t < t_s$, der Regler jedoch gar nicht aktiv war, man nicht einmal sicher sein kann, dass der Prozess überhaupt geregelt wurde, dann ist die Integration des Fehlers über diese Zeit total sinnlos. Der Integrator muss also zum Zeitpunkt $t = t_s$ neu initialisiert werden, und zwar mit der Anfangsbedingung, die ein stoßfreies Umschalten sicherstellt:

$$K_P e(t_s) + K_D \dot{e}(t_s) + K_I u_{\text{init}} = \tilde{u}(t_s).$$

Aus dieser Bedingung lässt sich die Anfangsbedingung u_{init} berechnen. Für $t \geq t_s$ berechnet sich die Stellgröße dann zu:

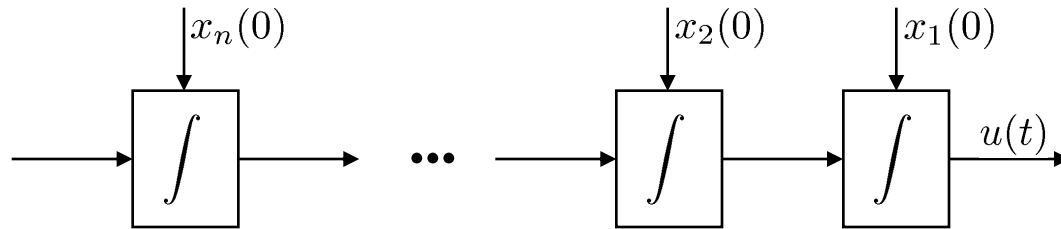
$$u(t) = K_P e(t) + K_D \dot{e}(t) + K_I \int_{t_s}^t e(\tau) d\tau + u_{\text{init}}.$$

$\underbrace{\hspace{10em}}_{= 0 \text{ für } t = t_s}$

15.8 Stoßfreies Umschalten (Bumpless Transfer)

Stoßfreies Umschalten bei Reglern höherer Ordnung

Enthält der Regler mehr als einen Integrator, wie dies bei Reglern höherer Ordnung der Fall ist, so müssen mehrerer Anfangswerte bestimmt werden. Dann kann man wie folgt vorgehen: Zuerst zeichnet man das Blockschaltbild des Reglers auf. Darin muss eine Integratorkette vorkommen (der Rest des Blockschaltbildes ist hier unerheblich):



Man sieht, dass sich der Anfangswert des Integrators ganz rechts sofort auf $u(t)$ auswirkt. Deshalb muss dieser Anfangswert so gewählt werden, dass die Stellgröße nicht springt:

$$u(t_s) = \tilde{u}(t_s)$$

Der Anfangswert des Integrators links davon beeinflusst die Ableitung von $u(t)$. Er muss deshalb so gewählt werden, dass die Ableitung der Stellgröße nicht springt:

$$\dot{u}(t_s) = \dot{\tilde{u}}(t_s)$$

usw.... So bekommt man für jeden Anfangswert eine Gleichung zu dessen Bestimmung.

15.9 Smith-Prädiktor

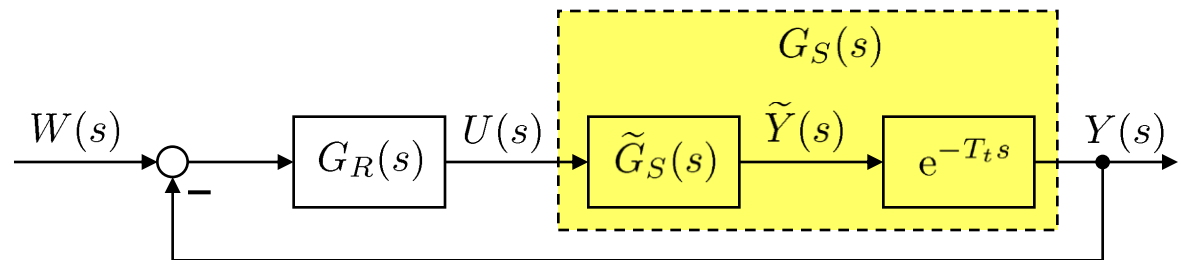
Regelung von Strecken mit Totzeit

Totzeiten erschweren die Regelung. Einerseits senken sie im Frequenzgang die Phase ab (Betrag bleibt unverändert), was zur Destabilisierung beiträgt und damit die maximal mögliche Kreisverstärkung und Bandbreite beschränkt. Andererseits basieren viele regelungstechnischen Methoden auf gebrochen rationalen Übertragungsfunktionen und lassen sich auf Systeme mit Totzeit nicht, jedenfalls nicht direkt, anwenden: z.B. das Hurwitz-Kriterium oder der Polvorgabe-Regler.

Die erstgenannte Einschränkung ist prinzipieller Art und kann durch nichts aufgehoben werden. Der zweitgenannte Nachteil soll durch den Smith-Prädiktor umgangen werden.

Zunächst wird die Regelstrecke in ihren Totzeit-Anteil und den Rest zerlegt:

$$G_S(s) = \tilde{G}_S(s)e^{-T_t s}$$



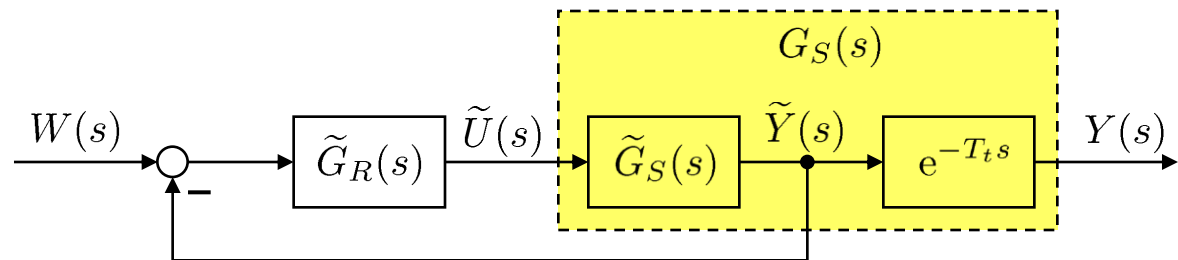
15.9 Smith-Prädiktor

Grundidee des Smith-Prädiktors

Die Totzeit behindert den Reglerentwurf, weil jede Regelabweichung ja schon um die Zeit T_t früher feststeht, als sie sich am Streckenausgang $y(t)$ messen lässt. Wenn sich $\tilde{y}(t)$ messen ließe, könnte diese schädliche Verzögerung vermieden werden, und der Regler könnte sofort auf den Ausgang des ersten Streckenteils $\tilde{G}_S(s)$ reagieren.

Wenn man nun $\tilde{y}(t)$ schon nicht messen kann, dann sollte es doch möglich sein, es auszurechnen und dann anstelle der wirklichen, gemessenen Regelgröße $y(t)$ die virtuelle, berechnete Regelgröße $\tilde{y}(t)$ rückzuführen. Das ist die Grundidee des Smith-Prädiktors. Dadurch lässt sich die Totzeit zwar nicht vermeiden (das geht nicht), aber sie stört den Reglerentwurf nicht mehr.

Welcher Zusammenhang besteht nun zwischen dem Smith-Prädiktor-Regler $\tilde{G}_R(s)$ und dem Standardregler $G_R(s)$?



15.9 Smith-Prädiktor

Berechnung des Smith-Prädiktor-Reglers

Dazu setzen wir die Führungsübertragungsfunktionen beider Regelkreis gleich:

$$\frac{G_R(s)\tilde{G}_S(s)e^{-T_t s}}{1 + G_R(s)\tilde{G}_S(s)e^{-T_t s}} = \frac{\tilde{G}_R(s)\tilde{G}_S(s)}{1 + \tilde{G}_R(s)\tilde{G}_S(s)} e^{-T_t s}$$

$$[1 + \tilde{G}_R(s)\tilde{G}_S(s)]G_R(s) = [1 + G_R(s)\tilde{G}_S(s)e^{-T_t s}]\tilde{G}_R(s)$$

$$[1 + \tilde{G}_R(s)\tilde{G}_S(s) - \tilde{G}_R(s)\tilde{G}_S(s)e^{-T_t s}]G_R(s) = \tilde{G}_R(s)$$

$$G_R(s) = \frac{\tilde{G}_R(s)}{1 + \tilde{G}_R(s)\tilde{G}_S(s)(1 - e^{-T_t s})}$$

Es wird zuerst der Regler $\tilde{G}_R(s)$ für die Strecke $\tilde{G}_S(s)$ (also ohne Totzeit) mit irgendeinem Standardverfahren entworfen. Dann verwendet man obige Formel, um den äquivalenten Regler im Standardregelkreis für die Strecke $G_S(s)$ (also mit Totzeit) auszurechnen.

15.9 Smith-Prädiktor

Blockschaltbild des Smith-Prädiktor-Reglers

Die Übertragungsfunktion des Smith-Prädiktor-Reglers hat die Form eines geschlossenen "Regelkreises". Man kann den Regler selbst also als eine Art Regelkreis interpretieren.

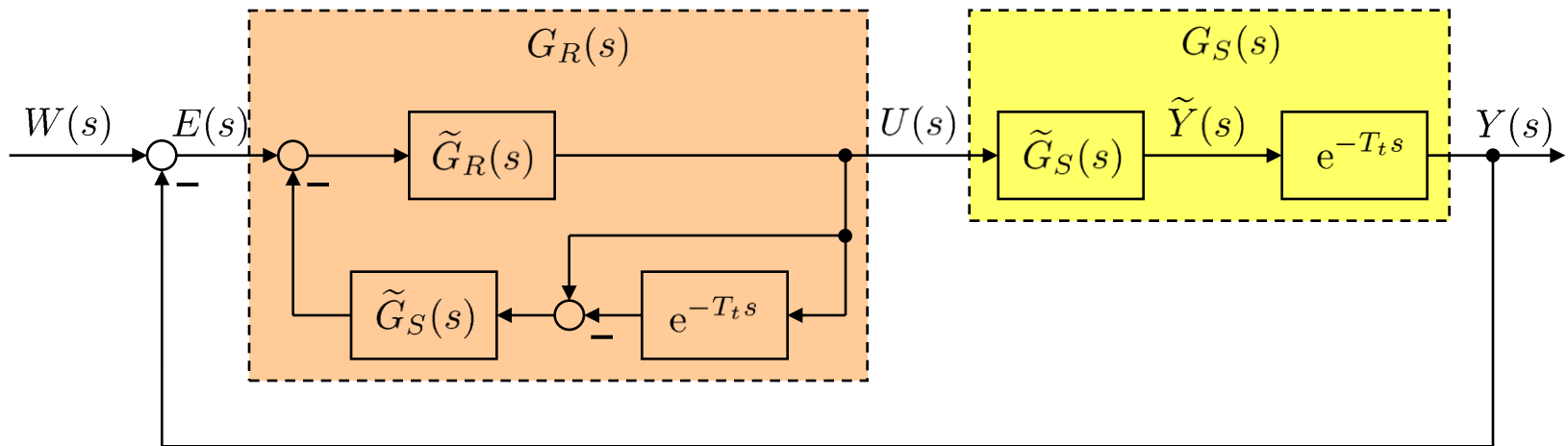
Dabei liegt im Vorwärtszweig

$$\tilde{G}_R(s)$$

und der offene Kreis $G_0(s)$ ist

$$\tilde{G}_R(s)\tilde{G}_S(s)(1 - e^{-T_t s})$$

Damit ergibt sich folgendes Blockschaltbild:



15.9 Smith-Prädiktor

Beispiel: Regelung eines Totzeitprozesses mit dem Smith-Prädiktor

Als Beispielprozess nehmen wir ein PT_1 -System mit Totzeit mit $K = 1$, $T = 1$ sec, $T_t = 3$ sec:

$$G_S(s) = \frac{K}{1 + Ts} e^{-T_t s}$$

Wir wollen dafür einen PI-Regler entwerfen, der die Prozessverstärkung und den Prozesspol wegekürzt. Ignorieren bzw. vernachlässigen wir die Totzeit des Prozesses ergibt sich damit folgender Regler:

$$G_R(s) = \frac{K_R}{K} \frac{1 + Ts}{s}$$

Berücksichtigen wir die Totzeit und entwerfen nach dem Smith-Prädiktor-Prinzip ergibt sich:

$$\tilde{G}_R(s) = \frac{K_R}{K} \frac{1 + Ts}{s}$$

Dies entspricht nach der Smith-Prädiktor-Formel dem komplizierten konventionellen Regler:

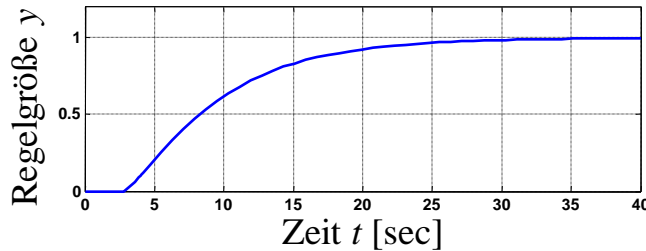
$$G_R(s) = \frac{\frac{K_R}{K} \frac{1+Ts}{s}}{1 + \frac{K_R}{K} \frac{1+Ts}{s} \frac{K}{1+Ts} (1 - e^{-T_t s})} = \frac{K_R(1 + Ts)}{K(s + K_R(1 - e^{-T_t s}))}$$

15.9 Smith-Prädiktor

Beispiel: Regelung mit Standard-PI-Regler

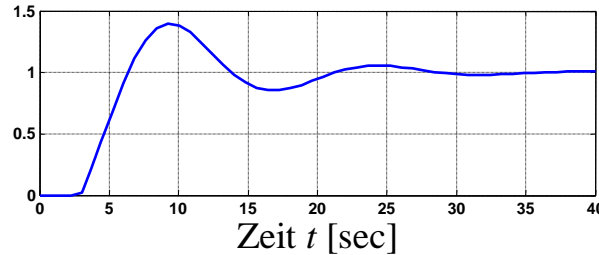
Als einziger freier Parameter des Reglers bleibt nun noch K_R einzustellen bzw. zu tunen. Die relative große Prozesstotzeit von $T_t = 3$ sec, macht den Vorteil des Smith-Prädiktors besonders deutlich. Die Führungsgröße w ist ein Einheitssprung.

$K_R = 0,1$



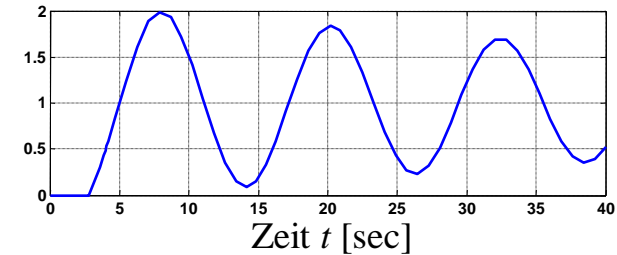
super langsam

$K_R = 0,3$

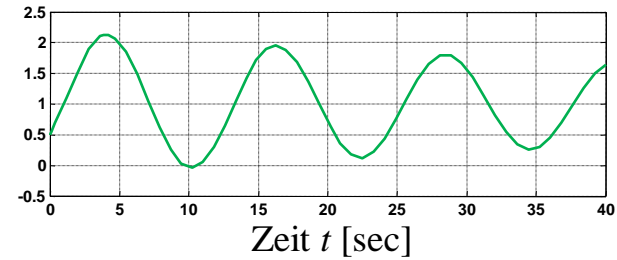
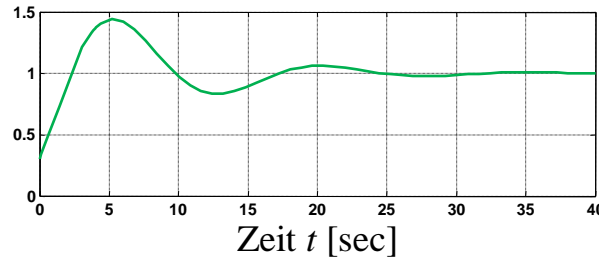
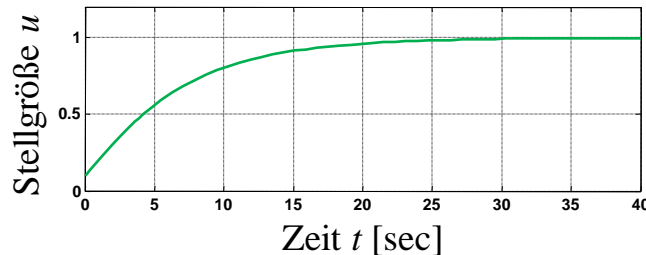


Kompromiss

$K_R = 0,5$



kurz vor Stabilitätsgrenze



Sehr schlechte Performance! Die Einschwingzeit des geschlossenen Regelkreises ist viel größer als die des offenen, da die Reglerverstärkung sehr klein gewählt werden muss.

15.9 Smith-Prädiktor

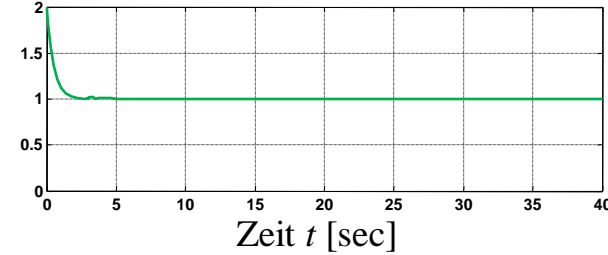
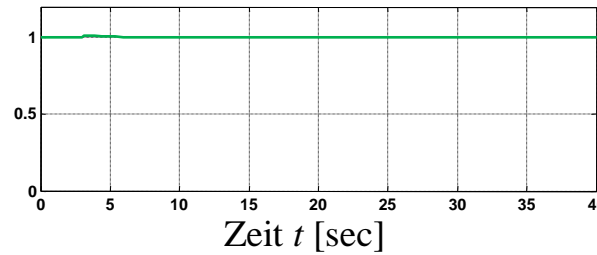
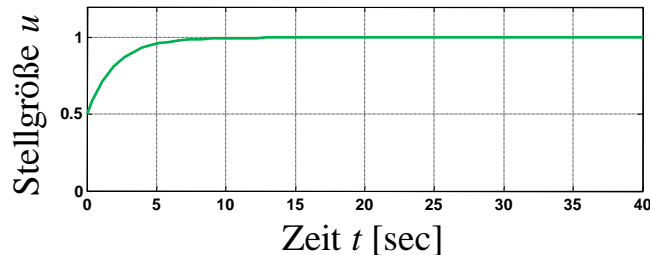
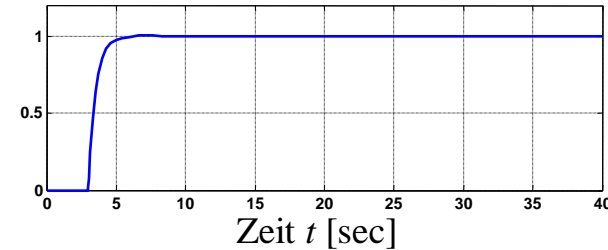
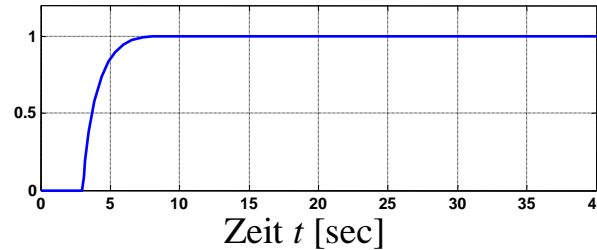
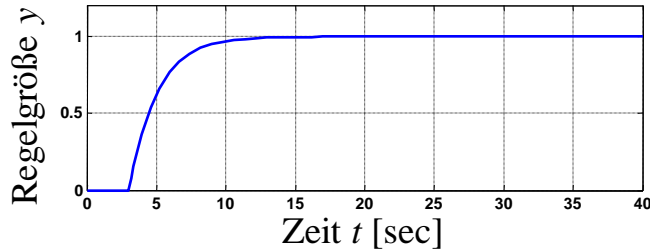
Beispiel: Regelung mit Smith-Prädiktor-PI-Regler

Da mit dem Smith-Prädiktor die Totzeit explizit berücksichtigt wird, kann die Reglerverstärkung K_R wesentlich größer gewählt werden. Der Regelkreis ist um den Faktor 10 schneller. In der Praxis ist dies aber wegen Stellgrößenbeschränkungen und Modellfehlern begrenzt.

$K_R = 0,5$

$K_R = 1,0$

$K_R = 2,0$



schnell

sehr schnell

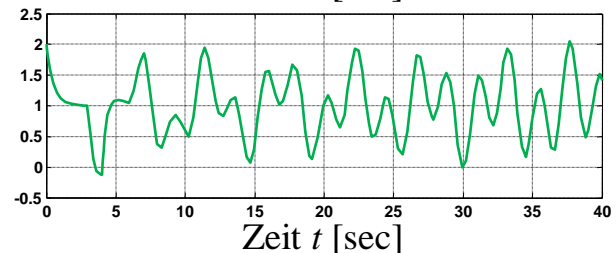
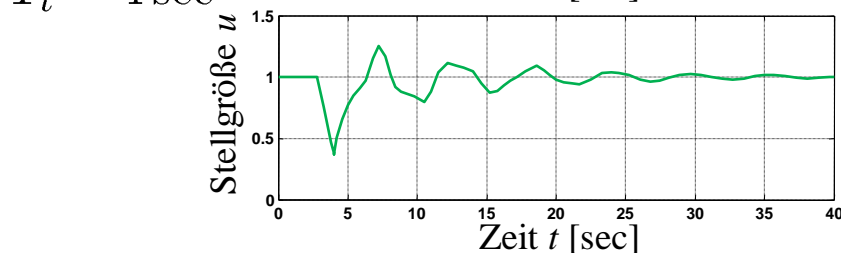
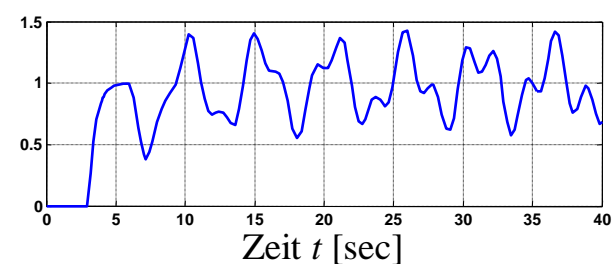
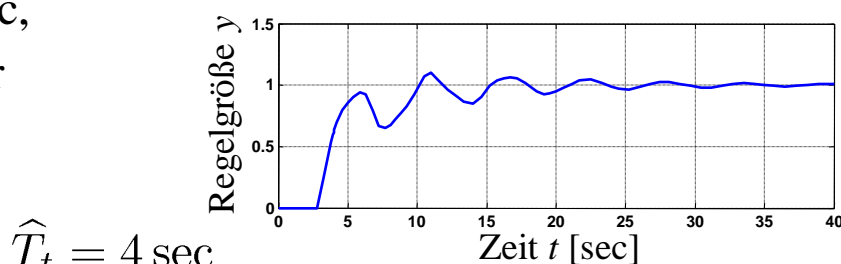
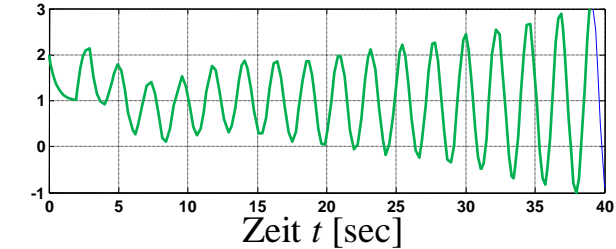
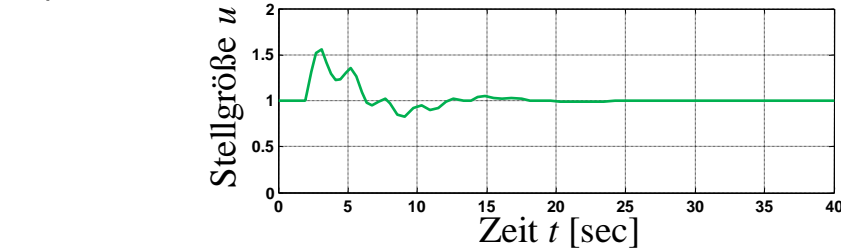
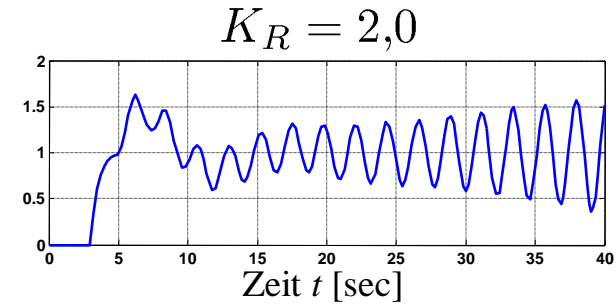
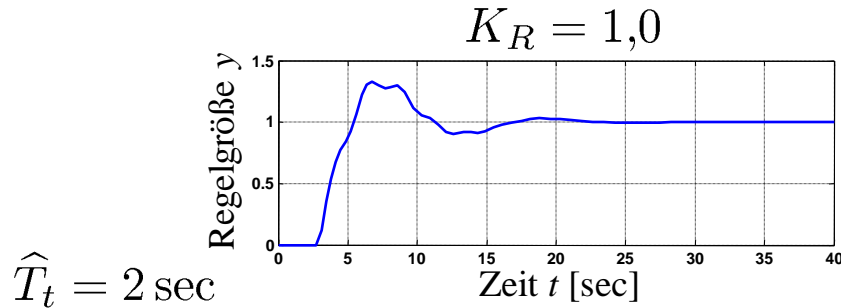
super schnell

Im perfekten Fall, d.h. keine Stellgrößenbeschränkungen und Modellfehler, wird die Performance mit wachsendem K_R immer besser. Aber die Regler werden auch weniger robust!

15.9 Smith-Prädiktor

Beispiel: Regelung mit Smith-Prädiktor-PI-Regler bei Modellfehlern

Ist die modellierte Totzeit 2 sec bzw. 4 sec lang statt der korrekten $T_t = 3$ sec, reagiert der Regler sehr empfindlich!

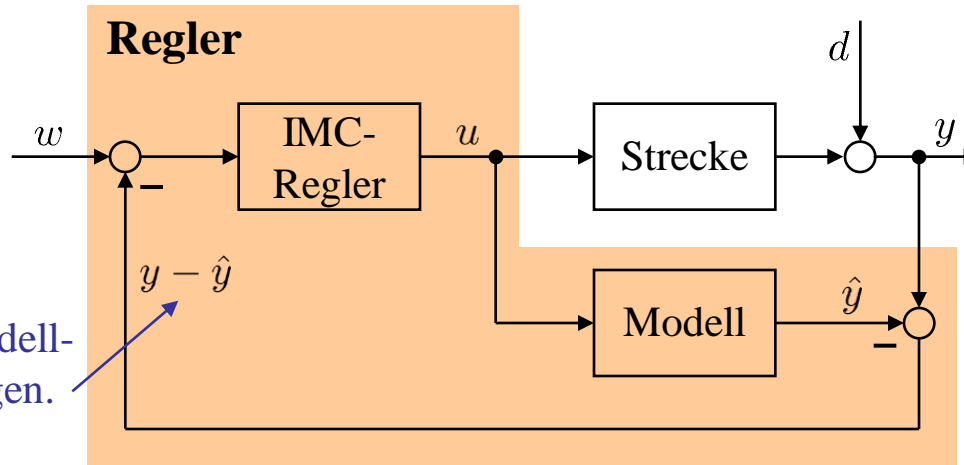


15.10 Internal Model Control

Grundidee

Internal Model Control (IMC) ist eine moderne Regelstrategie, die ein Modell der Regelstrecke voraussetzt. Der Standardregelkreis wird dabei wie folgt erweitert:

- Parallel zur Strecke wird ein Modell der Strecke geschaltet.
- Statt der Regelgröße wird die Abweichung zwischen Streckenausgang und Modellausgang rückgekoppelt.
- Im Idealfall, d.h. ohne Störungen ($d = 0$) und mit perfektem Modell, ist $y - \hat{y} = 0$. Dann ist die Rückkopplung inaktiv und der Regelkreis wird zu einer **Steuerung!**



Bedingt durch Modellfehler und Störungen.
Im Idealfall = 0.

15.10 Internal Model Control

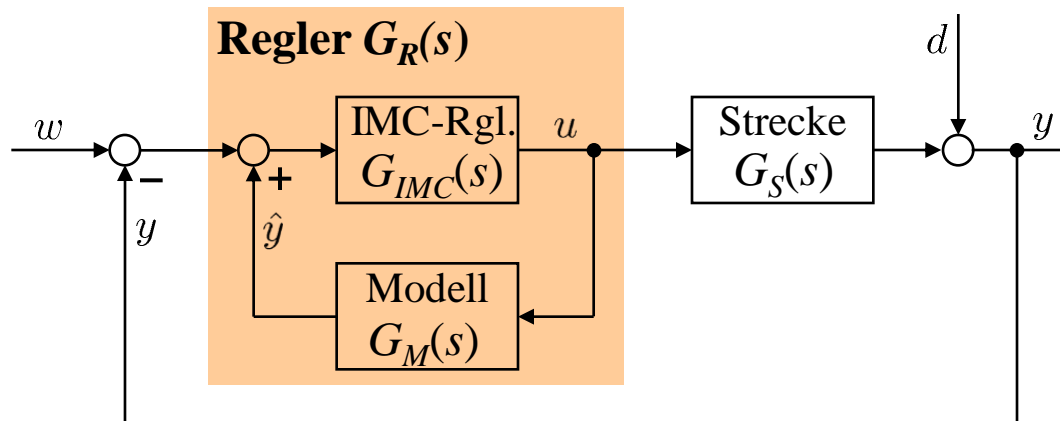
IMC-Regler

Der IMC-Regler lässt sich wie unten gezeigt in die Standardregelkreisform bringen. Mit Strecke $G_S(s)$, Modell $G_M(s)$, Regler $G_R(s)$ und IMC-Regler $G_{IMC}(s)$ gilt für den Regler:

$$G_R(s) = \frac{G_{IMC}(s)}{1 - G_{IMC}(s)G_M(s)}$$

Und für den geschlossenen Regelkreis:

$$G_w(s) = \frac{\frac{G_{IMC}(s)G_S(s)}{1 - G_{IMC}(s)G_M(s)}}{1 + \frac{G_{IMC}(s)G_S(s)}{1 - G_{IMC}(s)G_M(s)}} \rightarrow G_w(s) = \frac{G_{IMC}(s)G_S(s)}{1 + G_{IMC}[G_S(s) - G_M(s)]}$$



15.10 Internal Model Control

Idealfall: Modell = Strecke

Beim Reglerentwurf geht man zunächst davon aus, dass die Regelstrecke und ihr Modell gleich sind. Dann gilt für den geschlossenen Regelkreis:

$$G_w(s) = G_{IMC}(s)G_S(s)$$

$G_{IMC}(s)$ ist eine *Steuerung*! Das hat bemerkenswerte Konsequenzen:

- Für den Reglerentwurf kann im 1. Schritt eine optimale (oder näherungsweise optimale) Steuerung entworfen werden, siehe Kapitel 15.2 und Kapitel 12. Im 2. Schritt wird dann dieser IMC-Regler $G_{IMC}(s)$ in den Standardregler $G_R(s)$ umgerechnet.
- Jedes stabile $G_{IMC}(s)$ führt bei stabiler Strecke auch auf einen stabilen Regelkreis $G_w(s)$! Man kann zeigen, dass mit einer entsprechenden Wahl von (stabilem) $G_{IMC}(s)$ jeder beliebige Regler $G_R(s)$ erzeugt werden kann, der den Regelkreis stabilisiert. Deshalb nennt man die Formel $G_R(s) = f(G_{IMC}(s))$ die **Parametrierung aller stabilisierenden Regler** oder eine **Q-Parametrierung** (weil $G_{IMC}(s)$ oft auch mit $Q(s)$ benannt wird) bzw. **Youla-Parametrierung** (nach dem Entdecker dieses Zusammenhangs für Mehrgrößensysteme).

15.10 Internal Model Control

Youla-Parametrierung

Also noch mal: Jeder stabile IMC-Regler $G_{IMC}(s)$ führt bei stabiler Strecke auf einen stabilen geschlossenen Regelkreis $G_w(s)$.

Das ist eine sehr bemerkenswerte Eigenschaft. Machen wir uns zunächst klar, dass dies für den Standardregler *nicht* gilt:

$$G_w(s) = \frac{G_R(s)G_S(s)}{1 + G_R(s)G_S(s)}$$

Die Frage, für welche Regler $G_R(s)$ der geschlossene Regelkreis $G_w(s)$ stabil ist, kann man nicht einfach beantworten. Wir haben viele Stabilitätstests kennen gelernt, um diese Frage für konkrete Beispielfälle möglichst einfach zu beantworten. Aber es gibt dafür keine einfache allgemeine Lösung.

Beim IMC-Regler hingegen müssen wir nur die stabilen $G_{IMC}(s)$ untersuchen und wissen für stabile Strecken damit automatisch, dass auch $G_w(s)$ stabil ist. Das erleichtert den Reglerentwurf. Auch die Tatsache, dass die Führungs- und Störübertragungsfunktion (im Gegensatz zum Standardregler) **linear** von $G_{IMC}(s)$ abhängen, erleichtert die **Optimierung** des Reglers sehr. Viele moderne Reglerentwurfsverfahren basieren daher auf der Youla-Parametrierung.

15.10 Internal Model Control

IMC-Reglerentwurfverfahren

Beim Entwurf eines IMC-Reglers für eine stabile Strecke geht man wie folgt vor:

1. Entwurf einer idealen optimalen Steuerung $G_{IMC}^{(opt,ideal)}(s)$. Das ist einfach die Inverse der Strecke $G_S(s)$. Falls die Strecke nicht phasenminimal ist, wird zur Behandlung der instabilen Nullstellen nach Methode 2 in Kapitel 12.2 vorgegangen, d.h. es wird der Amplitudeneinfluss der instabilen Nullstellen bei der Inversion durch Spiegelung dieser Nullstellen kompensiert.
2. Falls die optimale Steuerung nicht realisierbar ist, weil der Zählergrad höher ist als der Nennergrad, wird ein PT_n -Filter der Form

$$F(s) = \frac{1}{(1 + Ts)^l} \quad \rightarrow \quad G_{IMC}^{(opt,real)}(s) = G_{IMC}^{(opt,ideal)}(s)F(s)$$

ergänzt, wobei dessen Ordnung l so gewählt wird, dass für die Steuerung gilt:
Zählergrad = Nennergrad. Damit wird die Realisierbarkeit sichergestellt.

4. Der Parameter T ist der einzige Tuningfaktor des Reglers. Mit ihm wird zwischen Bandbreite (Geschwindigkeit) und Robustheit des geschlossenen Regelkreises abgewogen.
5. Der entsprechende Standardregler $G_R(s)$ kann aus $G_{IMC}^{(opt,real)}(s)$ berechnet werden.

15.10 Internal Model Control

Erklärungen zum IMC-Reglerentwurfsverfahren

Optimale Steuerung:

Als Optimierungskriterium wird üblicherweise die H_2 -Norm, d.h. das Integral über die quadratische Regelabweichung $e(t)$, herangezogen:

$$J = \int_0^{\infty} e^2(t) dt = \frac{1}{2\pi} \int_0^{\infty} |E(i\omega)|^2 d\omega \rightarrow \min.$$

Bei Anregung mit einem Impuls $w(t) = \delta(t)$, d.h. $W(s) = 1$, werden alle Frequenzen gleich gewichtet. Im Frequenzbereich erhalten wir wegen $E(s) = W(s) - Y(s)$ dann die Beziehung:

$$J = \frac{1}{2\pi} \int_0^{\infty} \left| 1 - G_{IMC}^{(\text{opt,ideal})}(i\omega) G_S(i\omega) \right|^2 d\omega \rightarrow \min.$$

Für *phasenminimale* Strecken gelingt die (ideale) Inversion perfekt, d.h.

$$G_{IMC}^{(\text{opt,ideal})}(s) = \frac{1}{G_S(s)}$$

Dann ist der Fehler exakt = 0 und die Verlustfunktion $J = 0$.

15.10 Internal Model Control

Erklärungen zum IMC-Reglerentwurfsverfahren

Optimale Steuerung:

Für *nicht phasenminimale* Strecken wird die Strecke in eine phasenminimalen Anteil und einen Allpass aufgeteilt (siehe Kapitel 7.5):

$$G_s(s) = G_s^{(\text{PM})}(s)G_s^{(\text{AP})}(s)$$

Für die (ideale) optimale Steuerung wird dann nur der phasenminimale Teil invertiert:

$$G_{IMC}^{(\text{opt,ideal})}(s) = \frac{1}{G_s^{(\text{PM})}(s)}$$

Das ist genau die in Kapitel 12.2 beschriebene Methode 2:

- Totzeiten werden für die Inversion ignoriert (sie wandern in den Allpass-Teil).
- Für instabile Nullstellen muss im Allpass-Teil eine gespiegelte Polstelle erzeugt werden. Deshalb entsteht dieser *gespiegelte* Term im Nenner von $G_s^{(\text{AP})}(s)$ und im Zähler von $G_s^{(\text{PM})}(s)$ und damit durch die Inversion wieder im Nenner von $G_{IMC}^{(\text{opt,ideal})}(s)$.

Diese optimale Steuerung wird im Folgeschritt durch das Hinzufügen eines PT_n -Filters suboptimal. Der endgültige Regler hat also keine Optimalitätseigenschaften.

15.10 Internal Model Control

Erklärungen zum IMC-Reglerentwurfsverfahren

Wahl der Zeitkonstanten T im PT_n -Filter:

- Diese Zeitkonstante T ist der einzige freie Parameter im Reglerentwurf. Das gilt völlig unabhängig von der Komplexität (Ordnung) der Strecke und des Reglers. Im Gegensatz dazu müssen schon für simple PID-Regler 3 Parameter gleichzeitig eingestellt werden und für Regler höherer Ordnung noch mehr. Das Tuning von T ist also eine sehr einfache und übersichtliche Aufgabe; außerdem lässt sich T unmittelbar interpretieren.
- Weil der IMC-Regler den phasenminimalen Teil der Strecke wegekürzt, berechnet sich die Übertragungsfunktion des geschlossenen Regelkreises zu:

$$G_w(s) = G_s^{(\text{AP})}(s)F(s)$$

Für phasenminimale Strecken gilt vereinfachend $G_s^{(\text{AP})}(s) = 1$. D.h. mit der Zeitkonstante T des Filters $F(s)$ wird direkt das dynamische Verhalten des Regelkreises eingestellt. Hier muss man, wie üblich, einen guten Kompromiss zwischen Bandbreite und Robustheit finden.

- T entspricht damit in etwa der Stellgrößenbestrafung in einer Verlustfunktion, ist also ein Meta-Parameter ($T \rightarrow 0$: max. Performance, $T \rightarrow \infty$: max. Robustheit).

15.10 Internal Model Control

IMC-Reglerstruktur

Das IMC-Reglerentwurfsverfahren liefert einen Regler, dessen **Struktur und Parameter** sich automatisch aus der Regelstrecke ergeben. Das ist ein grundsätzlicher Unterschied zu vielen anderen Ansätzen, die eine bestimmte Reglerstruktur vorgeben (z.B. PID-Regler) und dann bloß noch die zugehörigen Parameter einstellen.

Ein IMC-Regler ist deswegen viel mächtiger als solche strukturfixen Ansätze und kann daher auch ohne Erweiterung direkt auf Prozesse mit **Totzeit** (ergibt automatisch(!) eine Smith-Prädiktor-ähnliche Struktur) und auf **nichtlineare** Prozesse angewandt werden.

Ordnung des IMC-Reglers

Wir gehen von einer stabilen Streckenübertragungsfunktion mit Nennergrad n aus. Da der Nennergrad immer größer gleich dem Zählergrad ist, hat die Strecke auch die Ordnung n . Da alle Pole stabil sind, werden sie durch die Inversion im Zähler des IMC-Reglers auftauchen. D.h. der IMC-Regler hat einen Zählergrad n . Der Nennergrad wird durch das Filter $F(s)$ ebenfalls auf Grad n gebracht. D.h. die *Ordnung des Reglers* beträgt auch n und ist damit identisch zur Streckenordnung. Der Regler hat also die gleiche Komplexität wie die Strecke.

15.10 Internal Model Control

Beispiel: IMC-Regler für eine PT_1 -Strecke

$$G_S(s) = \frac{K}{1 + T_1 s}$$

Die Strecke ist phasenminimal. Der ideale optimale IMC-Regler ist deshalb die Inverse:

$$G_{IMC}^{(\text{opt,ideal})}(s) = \frac{1 + T_1 s}{K}$$

Um diesen Regler realisierbar (und robuster) zu machen, wird ein PT_1 -Filter mit einstellbarer Zeitkonstante T ergänzt, um für den Regler "Zählergrad = Nennergrad" zu erreichen:

$$G_{IMC}^{(\text{opt,real})}(s) = \frac{1 + T_1 s}{K(1 + T s)}$$

Daraus ergibt sich für den geschlossenen Regelkreis:

$$G_w(s) = G_{IMC}^{(\text{opt,real})}(s)G_S(s) = \frac{1 + T_1 s}{K(1 + T s)} \frac{K}{1 + T_1 s} = \frac{1}{1 + T s}$$

15.10 Internal Model Control

Beispiel: IMC-Regler für eine PT_1 -Strecke

Mit dem Streckenmodell $G_M(s) = G_S(s)$ ergibt sich der äquivalente Standardregler:

$$G_R(s) = \frac{G_{IMC}(s)}{1 - G_{IMC}(s)G_M(s)} = \frac{\frac{1+T_1s}{K(1+Ts)}}{1 - \frac{1+T_1s}{K(1+Ts)} \frac{K}{1+T_1s}} = \frac{\frac{1+T_1s}{K(1+Ts)}}{1 - \frac{1}{1+Ts}} = \frac{1 + T_1s}{KTs}$$

D.h. es ergibt sich folgender PI-Standardregler:

$$G_R(s) = \frac{T_1}{KT} + \frac{1}{KTs} = \frac{T_1}{KT} \left(1 + \frac{1}{T_1s} \right)$$

Aus dem IMC-Entwurf resultiert ein PI-Regler, dessen Nullstelle den Pol der Strecke kürzt; ein Verfahren das wir intuitiv schon öfters zuvor gewählt haben wird hier also automatisch erzeugt! Mit der einstellbaren Zeitkonstante T wird antiproportional nur noch die Regler- und damit die Kreisverstärkung bestimmt.

15.10 Internal Model Control

Beispiel: IMC-Regler für eine PT_2 -Strecke

$$G_S(s) = \frac{K}{(1 + T_1s)(1 + T_2s)}$$

Die Strecke ist phasenminimal. Der ideale optimale IMC-Regler ist deshalb die Inverse:

$$G_{IMC}^{(\text{opt,ideal})}(s) = \frac{(1 + T_1s)(1 + T_2s)}{K}$$

Um diesen Regler realisierbar (und robuster) zu machen, wird ein PT_2 -Filter mit einstellbarer Zeitkonstante T ergänzt, um für den Regler "Zählergrad = Nennergrad" zu erreichen:

$$G_{IMC}^{(\text{opt,real})}(s) = \frac{(1 + T_1s)(1 + T_2s)}{K(1 + Ts)^2}$$

Daraus ergibt sich für den geschlossenen Regelkreis:

$$G_w(s) = G_{IMC}^{(\text{opt,real})}(s)G_S(s) = \frac{(1 + T_1s)(1 + T_2s)}{K(1 + Ts)^2} \frac{K}{(1 + T_1s)(1 + T_2s)} = \frac{1}{(1 + Ts)^2}$$

15.10 Internal Model Control

Beispiel: IMC-Regler für eine PT₂-Strecke

Mit dem Streckenmodell $G_M(s) = G_S(s)$ ergibt sich der äquivalente Standardregler:

$$\begin{aligned} G_R(s) &= \frac{G_{IMC}(s)}{1 - G_{IMC}(s)G_M(s)} = \frac{\frac{(1+T_1s)(1+T_2s)}{K(1+Ts)^2}}{1 - \frac{(1+T_1s)(1+T_2s)}{K(1+Ts)^2} \frac{K}{(1+T_1s)(1+T_2s)}} = \frac{\frac{(1+T_1s)(1+T_2s)}{K(1+Ts)^2}}{1 - \frac{1}{(1+Ts)^2}} \\ &= \frac{\frac{(1+T_1s)(1+T_2s)}{K}}{T^2s^2 + 2Ts} = \frac{(1 + T_1s)(1 + T_2s)}{KTs(Ts + 2)} = \frac{(T_1 + T_2) + \frac{1}{s} + T_1T_2s}{KT(Ts + 2)} \end{aligned}$$

Dies ist ein mit einem PT₁-Glieder gefilterter PID-Regler. Dessen beide Nullstellen werden wieder zur Kürzung der Streckenpole verwendet. Mit der einstellbaren Zeitkonstante T wird antiproportional die Regler- und damit die Kreisverstärkung bestimmt und proportional die Zeitkonstante des PT₁-Filters.

Allgemein ergibt sich für eine PT _{n} -Strecke ein Standardregler mit n Nullstellen, welche die n Streckenpole kürzen. Außerdem hat der Regler einen I-Anteil und wird durch ein PT _{$n-1$} -Filter ergänzt.

15.10 Internal Model Control

Beispiel: IMC-Regler für eine nicht phasenminimale PDT₂-Strecke mit Totzeit

$$G_S(s) = \frac{K(1 - T_D s)}{(1 + T_1 s)(1 + T_2 s)} e^{-T_t s} \quad \text{mit } T_1, T_2, T_D, T_t > 0$$

Zunächst muss die nicht phasenminimale Strecke in einen phasenminimalen Anteil und einen Allpass zerlegt werden.

$$G_S(s) = \underbrace{\frac{K(1 + T_D s)}{(1 + T_1 s)(1 + T_2 s)}}_{\text{phasenminimaler Anteil}} \underbrace{\frac{1 - T_D s}{1 + T_D s}}_{\text{Allpass}} e^{-T_t s}$$

Der ideale optimale IMC-Regler ist die Inverse des phasenminimalen Anteils. D.h. Totzeiten der Originalstrecke werden ignoriert und Nullstellen werden gespiegelt invertiert:

$$G_{IMC}^{(\text{opt,ideal})}(s) = \frac{(1 + T_1 s)(1 + T_2 s)}{K(1 + T_D s)}$$

↙
gespiegelte
instabile Nullstelle

15.10 Internal Model Control

Beispiel: IMC-Regler für eine nicht phasenminimale PDT₂-Strecke mit Totzeit

Um diesen Regler realisierbar (und robuster) zu machen, wird ein PT₁-Filter mit einstellbarer Zeitkonstante T ergänzt, um für den Regler "Zählergrad = Nennergrad" zu erreichen:

$$G_{IMC}^{(\text{opt,real})}(s) = \frac{(1 + T_1 s)(1 + T_2 s)}{K(1 + T_D s)(1 + T s)}$$

Daraus ergibt sich für den geschlossenen Regelkreis:

$$\begin{aligned} G_w(s) &= G_{IMC}^{(\text{opt,real})}(s)G_s(s) = \frac{(1 + T_1 s)(1 + T_2 s)}{K(1 + T_D s)(1 + T s)} \frac{K(1 - T_D s)}{(1 + T_1 s)(1 + T_2 s)} e^{-T_t s} \\ &= \frac{1 - T_D s}{(1 + T_D s)(1 + T s)} e^{-T_t s} \end{aligned}$$

D.h. der geschlossene Regelkreis zeigt das Allpassverhalten mit der Totzeit ergänzt um den PT₁-Filter zur Realisierbarkeit des Reglers.

Auch hier sehen wir wieder, dass ein Regler –egal welcher Struktur– nichts an instabilen Nullstellen und an Totzeiten verändern kann.

15.10 Internal Model Control

Beispiel: IMC-Regler für eine nicht phasenminimale PDT_2 -Strecke mit Totzeit

Mit dem Streckenmodell $G_M(s) = G_S(s)$ ergibt sich der äquivalente Standardregler:

$$\begin{aligned} G_R(s) &= \frac{G_{IMC}(s)}{1 - G_{IMC}(s)G_M(s)} = \frac{\frac{(1+T_1s)(1+T_2s)}{K(1+T_Ds)(1+Ts)}}{1 - \frac{(1+T_1s)(1+T_2s)}{K(1+T_Ds)(1+Ts)} \frac{K(1-T_Ds)}{(1+T_1s)(1+T_2s)} e^{-T_t s}} \\ &= \frac{\frac{(1+T_1s)(1+T_2s)}{K(1+T_Ds)(1+Ts)}}{1 - \frac{1-T_Ds}{(1+T_Ds)(1+Ts)} e^{-T_t s}} \end{aligned}$$

Dieser Ausdruck lässt sich nicht mehr sinnvoll vereinfachen. Für den *Spezialfall* ohne Totzeit ergibt sich:

$$G_R(s) = \frac{(1 + T_1s)(1 + T_2s)}{Ks(TT_Ds + T + 2T_D)}$$

Das ist ein ähnliches Ergebnis wie für die PT_2 -Strecke: Ein PID-Regler ergänzt durch einen PT_1 -Filter.

15.10 Internal Model Control

Eigenschaften des Internal Model Control

- Sehr allgemeines Verfahren. Lässt sich auf alle Arten von Prozessen anwenden, sofern ein geeignetes Streckenmodell zur Verfügung steht.
- Die Anwendung auf instabile Prozesse ist möglich, jedoch aufwändiger und wurde deshalb hier nicht behandelt.
- Es wird keine Reglerstruktur vorgegeben, sondern die Reglerstruktur ergibt sich aus der Regelstrecke und den Realisierbarkeitsbedingungen.
- Es ist nur ein Parameter T einzustellen, der einen Kompromiss zwischen Bandbreite des geschlossenen Regelkreises und dessen Robustheit ermöglicht.
- Im Gegensatz zur prädiktiven Regelung (siehe nächstes Kapitel), die ähnliche Vorteile aufweist, ist der Entwurfs- und Rechenaufwand für IMC gering!
- IMC kann als Entwurfsverfahren gesehen werden, um Standardregler zu entwerfen. Alternativ kann man aber auch den IMC-Regelkreis implementieren. Letzteres bietet gewisse Vorteile, wenn Nichtlinearitäten in der Strecke und im Modell vorhanden sind, denn solange diese sowohl in der Strecke als auch im Modell enthalten sind, entstehen keinerlei Stabilitätsprobleme, da (im idealen Fall) das rückgekoppelte Signal = 0 ist.

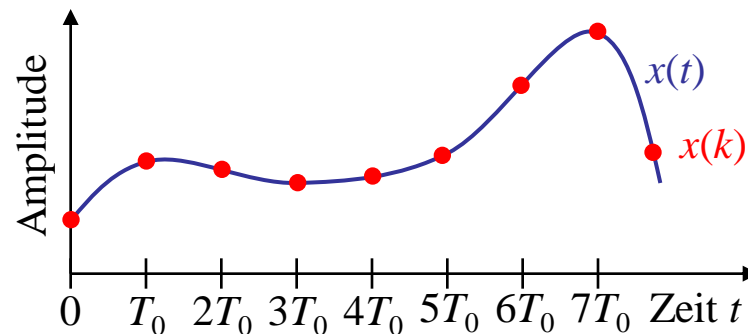
15.11 Prädiktive Regelung

Einführung

Die prädiktive Regelung verallgemeinert das Internal Model Control Konzept. Sie basiert daher auch auf einem **Modell der Regelstrecke**. Auch wird zunächst das Problem der **optimalen Steuerung** gelöst und anschließend durch die Einbeziehung der gemessenen Regelgröße zur Regelung erweitert.

Am Einfachsten lässt sich die prädiktive Regelung für zeitdiskrete (d.h. digitale bzw. abgetastete) Systeme erklären. Wir wollen dies hier tun, ohne auf die Theorie der digitalen Regelung im Allgemeinen näher einzugehen. Siehe dazu Kapitel 17.

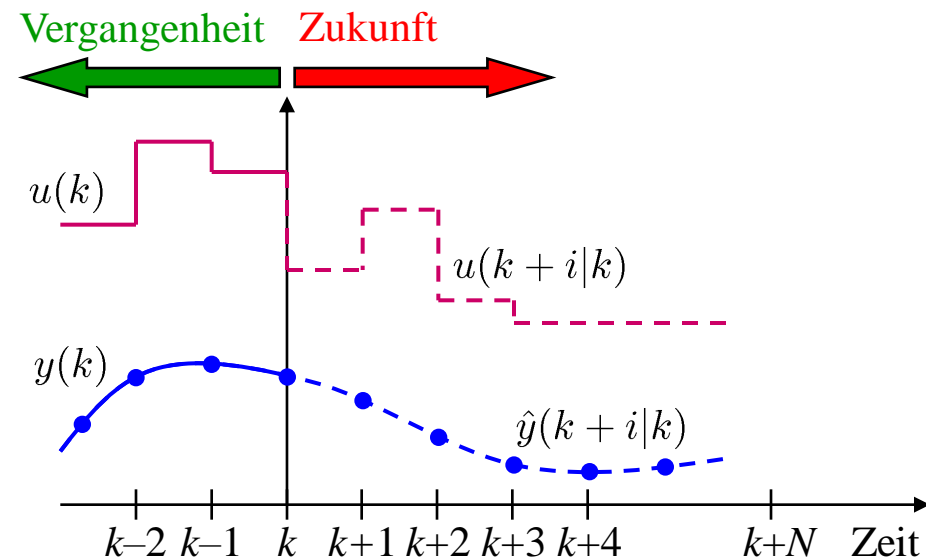
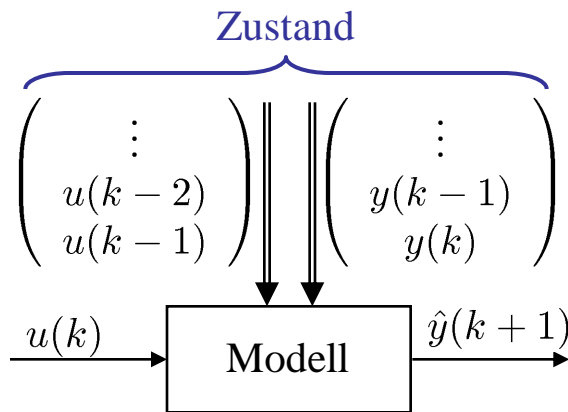
Wir gehen also davon aus, dass alle Signale zu synchronen Zeitpunkten $t = kT_0$ abgetastet, eingelesen, ausgegeben und verarbeitet werden. Dabei ist $k = 0, 1, 2, \dots$ ein Laufindex und T_0 die Abtastzeit. Abkürzend schreiben wir $x(k)$ statt $x(kT_0)$. Die Signalamplitude wird als kontinuierlich angenommen.



15.11 Prädiktive Regelung

Prädiktion

Gehen wir von einer nicht sprungfähigen Strecke und daher einem nicht sprungfähigen Modell aus, dann bestimmt $u(k)$ und die alten Ein-/Ausgangswerte den nächsten Ausgangswert $\hat{y}(k + 1|k)$. Dies ist eine Vorhersage oder **Prädiktion** um einen Zeitschritt in die Zukunft (Zeitpunkt $k+1$) auf Basis der Informationen zum Zeitpunkt k (deshalb die Schreibweise " $|k$ ". Benutzt man diese Ein-Schritt-Prädiktion zusammen mit dem nächsten Eingangswert $u(k+1)$ kann man die Zwei-Schritt-Prädiktion $\hat{y}(k + 2|k)$ berechnen usw. So kann für jeden möglichen Verlauf der Eingangsgröße der Verlauf der Ausgangsgröße (Regelgröße) prädiziert werden.



15.11 Prädiktive Regelung

Grundidee der prädiktiven Regelung

Eine prädiktive Regelung läuft folgendermaßen ab:

1. Zum Zeitpunkt k wird basierend auf den vergangenen Ein- und Ausgangssignalen der Strecke mit Hilfe des Streckenmodells für einen zukünftigen Stellgrößenverlauf $u(k+i|k)$, $i = 0, 1, \dots, N-1$, der zugehörige Regelgrößenverlauf $\hat{y}(k+i|k)$, $i = 0, 1, \dots, N$, berechnet.
2. Von allen möglichen Stellsignalverläufen $u(k+i|k)$, $i = 0, 1, \dots, N-1$, wird durch ein Optimierungsverfahren derjenige bestimmt, der eine gegebene Verlustfunktion minimiert. D.h. die Stellgrößen sind für das Optimierungsverfahren die zu optimierenden Parameter.
3. Der optimalen Stellsignalverlauf, repräsentiert die optimale Steuerungsfolge. Hiervon wird allerdings nur der erste Wert $u(k|k)$ an das Stellglied ausgegeben. Dann wird ein Zeitschritt gewartet und die daraus entstandene Regelgröße $y(k+1)$ gemessen. Mit Hilfe der neu gemessenen Regelgröße wird erneut die Optimierung aus 2. durchgeführt. Von dieser wird wieder der erste Wert $u(k+1|k+1)$ an das Stellglied ausgegeben. Danach wird wieder ein Zeitschritt gewartet und die daraus entstandene Regelgröße $y(k+2)$ gemessen usw.

15.11 Prädiktive Regelung

Erklärungen zum Ablauf der prädiktiven Regelung

- In 2. wird mit den Informationen zum Zeitpunkt k die optimale Stellgrößenfolge oder Steuerungsfolge berechnet. Diese N Schritte in die Zukunft basieren auf den Vorhersagen des Modells über den Verlauf der Regelgröße. Es handelt sich um eine reine **Steuerung**. Die zukünftige Regelgröße kann ja nicht gemessen werden.
- Um aus Steuerungsfolge eine Regelung zu machen, wird nur die erste Stellgröße der Folge $u(k|k)$ an das Stellglied ausgegeben. Die nächste Stellgröße wird dann nicht aus dieser Folge genommen, also *nicht* $u(k+1|k)$. Statt dessen wird auf der Basis der neu gemessenen Regelgröße $y(k+1)$ eine neue optimale Stellgrößenfolge berechnet und deren erste Stellgröße $u(k+1|k+1)$ an das Stellglied ausgegeben.

Der Unterschied zwischen $u(k+1|k)$ und $u(k+1|k+1)$ besteht darin, dass erstere mit der *prädizierten* Regelgröße $\hat{y}(k+1|k)$ optimiert wurde, während letztere mit der *gemessenen* Regelgröße $y(k+1)$ optimiert wurde. Und $u(k+1|k+1)$ ist also neue Information enthalten. Auf diese Weise wird die Steuerung zur Regelung, denn mit der gemessenen Regelgröße finden Modellfehler und Störungen in der Optimierung Berücksichtigung.

15.11 Prädiktive Regelung

Erklärungen zum Ablauf der prädiktiven Regelung

- Eine solche Strategie, nämlich eine optimale Folge zu berechnen, den ersten Wert davon zu verwenden und dann einen Schritt später, basierend auf neuer Information, wieder eine optimale Folge zu berechnen usw. nennt man "*receding horizon*"-Strategie.
- Als Verlustfunktion in 2. wir üblicherweise auf uns schon bekannte quadratische Formen zurückgegriffen, nur jetzt in zeitdiskreter Schreibweise, d.h. mit Summen statt Integralen:

$$J = \sum_{i=1}^N [w(k+i) - \hat{y}(k+i|k)]^2 + \alpha \Delta u^2(k+i-1) \rightarrow \min_{u(k), u(k+1), \dots, u(k+N-1)}$$

wobei die quadratische Stellgrößenänderung $\Delta u(k) = u(k) - u(k-1)$ mit dem Faktor α gewichtet wird. Ist die Regelstrecke nicht nur nicht sprungfähig sondern hat zusätzlich eine Totzeit von dT_0 , dann bestimmt $u(k-1)$ ja nicht $y(k)$ sondern erst den Ausgang $y(k+d)$. Dann ändert sich die Verlustfunktion zu:

$$J = \sum_{i=1}^N [w(k+i+d) - \hat{y}(k+i+d|k)]^2 + \alpha \Delta u^2(k+i-1) \rightarrow \min_{u(k), u(k+1), \dots, u(k+N-1)}$$

15.11 Prädiktive Regelung

Erklärungen zum Ablauf der prädiktiven Regelung

- Damit die prädiktive Regelung gut funktionieren kann, muss N hinreichend groß gewählt werden. Es sollte sich an der langsamsten Zeitkonstante der Strecke orientieren. Wenn z.B. bei nicht phasenminimalen Strecken ein zu kleines N gewählt wird, kann der Regler instabil werden, weil nur das gegenphasige Unterschwingen einer Sprungantwort in die Verlustfunktion eingeht, was dann fälschlicherweise als negative Verstärkung interpretiert werden würde.
- Da N also typischerweise relativ groß gewählt wird, ist es sinnvoll die Anzahl an freien Stellgrößen (also die Anzahl der Parameter der Optimierung) einzuschränken, um die Optimierung schneller und robuster zu machen. Normalerweise schränkt man die Anzahl der Freiheitsgrade auf $N_u = 1, 2$ oder 3 ein. Dies kann beispielsweise dadurch geschehen, dass man den Stellgrößenverlauf ab $u(k+N_u|k)$ konstant hält. D.h. statt der N freien Stellgrößen:

$$u(k), u(k+1|k), u(k+2|k), \dots, u(k+N-1|k)$$

gibt es nur noch N_u Freiheitsgrade:

$$u(k), u(k+1|k), \dots, u(k+N_u-1|k) = u(k+N_u|k) = u(k+N_u+1|k) = \dots = u(k+N-1|k)$$

15.11 Prädiktive Regelung

Erklärungen zum Ablauf der prädiktiven Regelung

- Damit ergibt sich dann auch eine leicht abgewandelte Verlustfunktion:

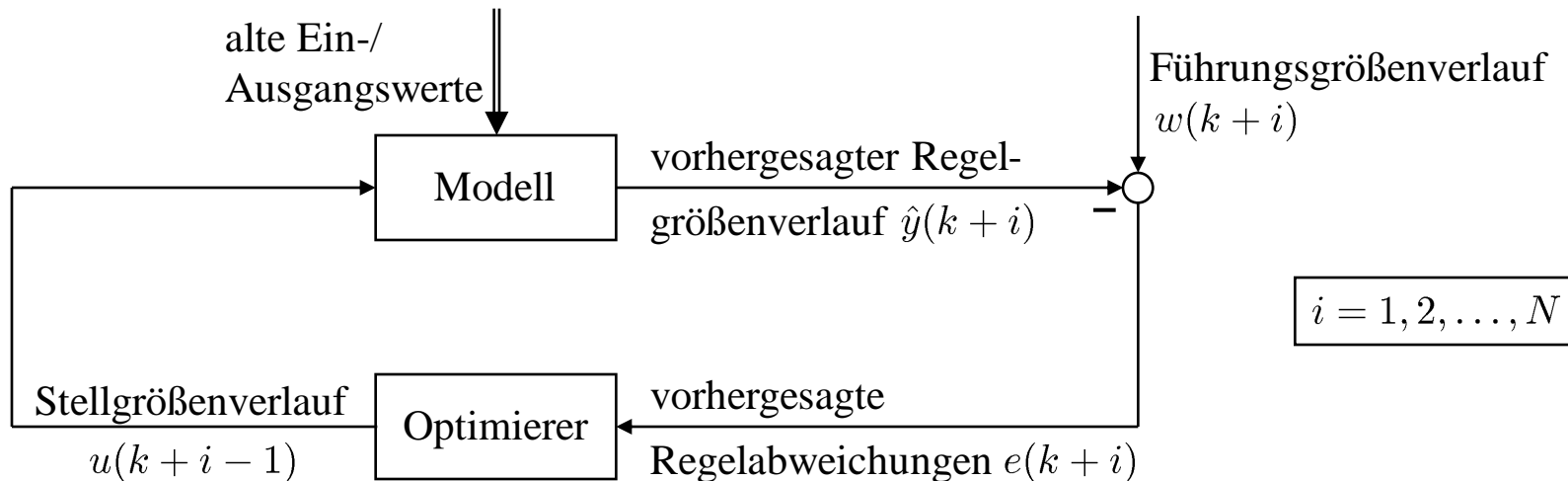
$$J = \sum_{i=1}^N [w(k+i) - \hat{y}(k+i|k)]^2 + \sum_{i=1}^{N_u} \alpha \Delta u^2(k+i-1) \rightarrow \min_{u(k), u(k+1), \dots, u(k+N_u-1)}$$

- Anstelle des Konstanthaltens der Stellgröße ab Zeitpunkt $k = N_u$ während der Optimierung gibt es auch andere Möglichkeiten die Freiheitsgrade des Stellgrößenverlaufs einzuschränken. Z.B. kann man den Verlauf als Polynom mit Zeit t als Eingangsgröße beschreiben. Auch andere Basisfunktionen sind denkbar.
- Der hier beschriebene prädiktive Regler existiert gar nicht in Form einer Übertragungsfunktion o.ä. Vielmehr wird in jedem Abtastschritt eine Optimierung durchgeführt und daraus die ausgegebene Stellgröße abgeleitet. Für lineare Systeme und Verlustfunktionen vom oben angegebenen quadratischen Typ lässt sich allerdings eine geschlossene Lösung berechnen. Für nichtlineare Systeme oder wenn Nebenbedingungen wie z.B. Stellgrößenbeschränkungen in die Optimierung einbezogen werden sollen, dann muss man iterative numerische Optimierungsverfahren einsetzen.

15.11 Prädiktive Regelung

Optimierungsverfahren bei der prädiktiven Regelung

Das Bild veranschaulicht den entscheidenden Schritt eines prädiktiven Regelungsverfahrens: die Optimierung des Stellgrößenverlaufs. Der Optimierer variiert die N_u freien Stellgrößen so lange, bis der optimale Stellgrößenverlauf gefunden ist. Der optimale Stellgrößenverlauf ist durch den minimalen Verlustfunktionswert gekennzeichnet. In diese Verlustfunktion geht wesentlich die vorhergesagte Regelabweichung ein, die sich aus dem mittels des Modells vorhergesagten Regelgrößenverlaufs und dem vorgegebenen Führungsgrößenverlauf ergibt. Kennt man den zukünftigen Führungsgrößenverlauf nicht, so werden darüber realistische Annahmen getroffen, z.B. $w(k+i) = \text{konstant}$ für alle $i \geq 0$.



15.11 Prädiktive Regelung

Optimierungsverfahren bei der prädiktiven Regelung

Abhängig vom Prozess sind folgende drei Klassen von prädiktiven Reglern typisch:

- *Linearer Prozess ohne Nebenbedingungen:* Der einfachste Fall. Es existiert ein einziges Optimum. Kann geschlossen mit der **Methode der kleinsten Quadrate (*Least Squares, LS*)** gelöst werden. Erfordert eine Matrixinversion der Größe $N_u \times N_u$. Bekannte prädiktive Regelungsverfahren dieser Klasse sind Generalized Predictive Control (GPC), Dynamic Matrix Control (DMC), Model Algorithmic Control (MAC) und Predictive Functional Control (PFC).
- *Linearer Prozess mit Nebenbedingungen:* Lässt sich auch relativ schnell lösen. Es existiert ebenfalls ein einziges Optimum. Die Optimierung erfordert aber ein iteratives Vorgehen mit dem Verfahren der **Quadratischen Programmierung (QP)**. Da das globale Optimum in vertretbarer Zeit garantiert gefunden wird, sind auch diese Verfahren recht weit verbreitet.
- *Nichtlinearer Prozess mit oder ohne Nebenbedingungen:* Der schwierigste Fall. Das Auffinden des globalen Optimums kann nicht garantiert werden. Hoher Rechenaufwand ist erforderlich. Kann daher nur bei langsamen Prozessen eingesetzt werden. Ein Backup-Regler sollte implementiert sein, für den Fall, dass bei der Optimierung etwas schief läuft.

15.11 Prädiktive Regelung

Eigenschaften der prädiktiven Regelung

- Sehr allgemeines Verfahren. Lässt sich auf alle Arten von Prozessen anwenden, sofern ein geeignetes Streckenmodell zur Verfügung steht.
- Es wird keine Reglerstruktur vorgegeben, sondern die optimale Reglerstruktur ergibt sich aus der Regelstrecke und den Freiheitsgraden für den Stellgrößenverlauf N_u .
- Es ergibt sich aus der Optimierung automatisch eine Kombination aus Vorsteuerung und Regelung.
- Nebenbedingungen wie z.B. Stellgrößenbeschränkungen aber auch Grenzen für andere modellierte Größen, lassen sich leicht in die Optimierung direkt einbauen. Damit ist die prädiktive Regelung das wichtigste Verfahren für Anwendungen, die durch Nebenbedingungen wesentlich geprägt sind.
- Die Optimierung ist rechenaufwendig. Für nichtlineare Optimierungen ist es schwierig Konvergenz und Auffinden eines guten lokalen Optimums sicherzustellen. Daher wird die prädiktive Regelung häufig bei langsamen Prozessen in der Verfahrenstechnik eingesetzt.
- Die Einstellung der Meta-Parameter (N, N_u, α, \dots) ist schwierig und kann langwierig sein.

16. Nichtlineare Regelung

Inhalt Kapitel 16

16. Nichtlineare Regelung

- 16.1 Kennlinien und Kennfelder
- 16.2 Nichtlineare Effekte
- 16.3 Stabilität nichtlinearer Systeme
- 16.4 Kennlinienregler
- 16.5 Übersicht wichtiger nichtlinearer Regelungsansätze
 - Beschreibungsfunktionen (Harmonische Balance)
 - Phasenebene und zeitoptimale Regelung
 - Direkte Methode von Ljapunow
 - Neuronale Netze und Fuzzy-Systeme

16.1 Kennlinien und Kennfelder

Eigenschaften nichtlinearer Systeme

- Für nichtlineare Systeme gilt das Superpositionsprinzip i.a. *nicht!* D.h.

$$f(x_1) + f(x_2) \neq f(x_1 + x_2)$$

- Für nichtlineare Systeme gilt das Verstärkungsprinzip i.a. *nicht!* D.h.

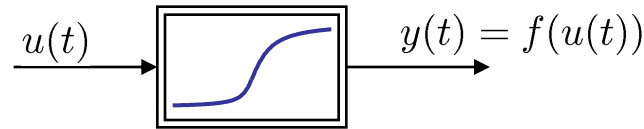
$$K \cdot f(x) \neq f(K \cdot x)$$

- Es ist sehr schwer alle nichtlinearen Systeme in bestimmte allgemein gültige Klassen einzuteilen. Jedes nichtlineare System ist in gewisser Weise einzigartig. Deshalb gibt es keine geschlossene Theorie nichtlinearer Systeme, so wie wir sie für lineare Systeme kennen.
- Nichtlineare Differentialgleichungen lassen sich i.a. nur numerisch lösen. Das ist einerseits aufwändig und vermittelt andererseits meist keine tieferen Einsichten. (So wie z.B. die numerische Berechnung des Hurwitz-Kriteriums zwar eine Ja/Nein-Aussage über die Stabilität liefert, aber nichts Konstruktives zum Reglerentwurf beisteuert).
- Deshalb sollte man wenn möglich versuchen, durch Vereinfachung, Transformation, Kompensation, o.ä. das nichtlineare System auf ein lineares abzubilden.
- Kennlinien spielen für die Betrachtung nichtlineare Systeme eine große Rolle.

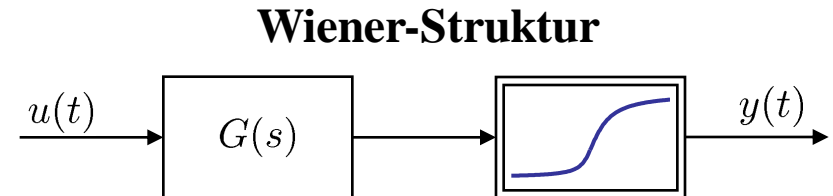
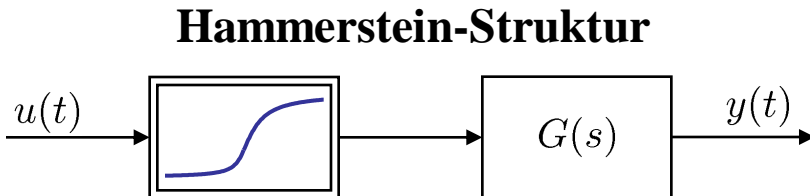
16.1 Kennlinien und Kennfelder

Kennlinien

Kennlinien beschreiben einen **nichtlinearen statischen** Zusammenhang zwischen einer Eingangsgröße und einer Ausgangsgröße.



Oft lassen sich nichtlineare dynamische Systeme in Teilblöcke zerlegen, die entweder *linear dynamisch* oder *nichtlinear statisch* sind. Gelingt es, ein nichtlineares dynamisches System in genau einen linear dynamischen (symbolisiert durch eine Übertragungsfunktion) und einen nichtlinear statischen Block zu zerlegen, so sind die folgende beiden Strukturen typisch:



Beide Strukturen können in ihrem statischen Verhalten und im dynamischen Verhalten mit kleinen Auslenkungen um einen Arbeitspunkt herum durchaus identisch sein; sie weisen aber anderes Großsignalverhalten auf: **Im Nichtlinearen gilt das Kommutativgesetz nicht!**

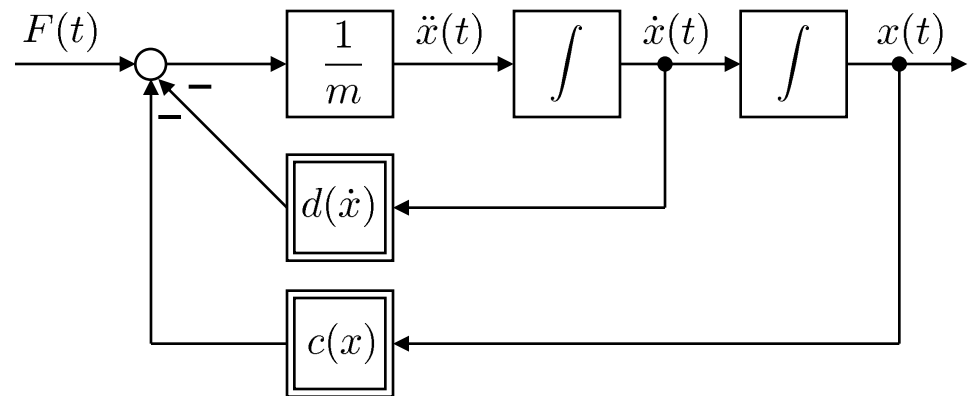
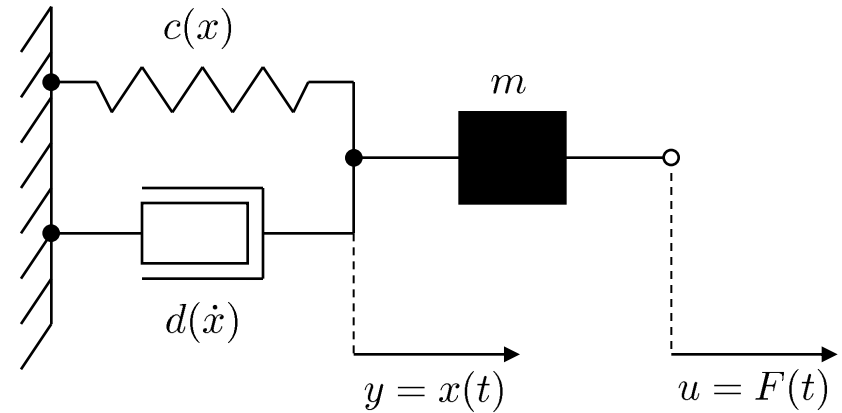
16.1 Kennlinien und Kennfelder

Beispiele für Kennlinien

Masse-Feder-Dämpfer-System mit nichtlinearer Feder- und Dämpfercharakteristik

$$F(t) = m\ddot{x}(t) + d(\dot{x})\dot{x}(t) + c(x)x$$

Nichtlinearitäten hängen von der Art der Feder und des Dämpfers ab.



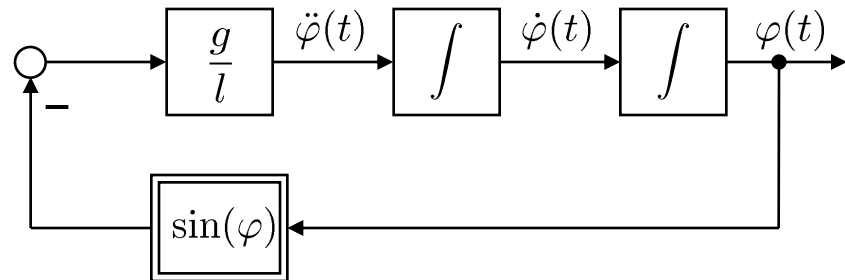
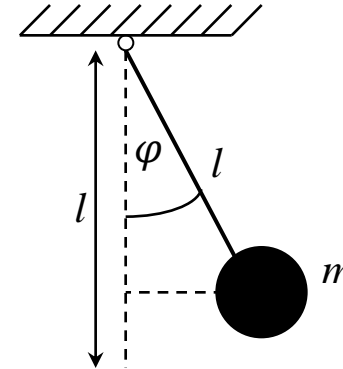
16.1 Kennlinien und Kennfelder

Beispiele für Kennlinien

Freie Schwingung eines ungedämpften Pendels

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

Nichtlinearität ist durch die Physik des Problems definiert.



16.1 Kennlinien und Kennfelder

Beispiele für Kennlinien

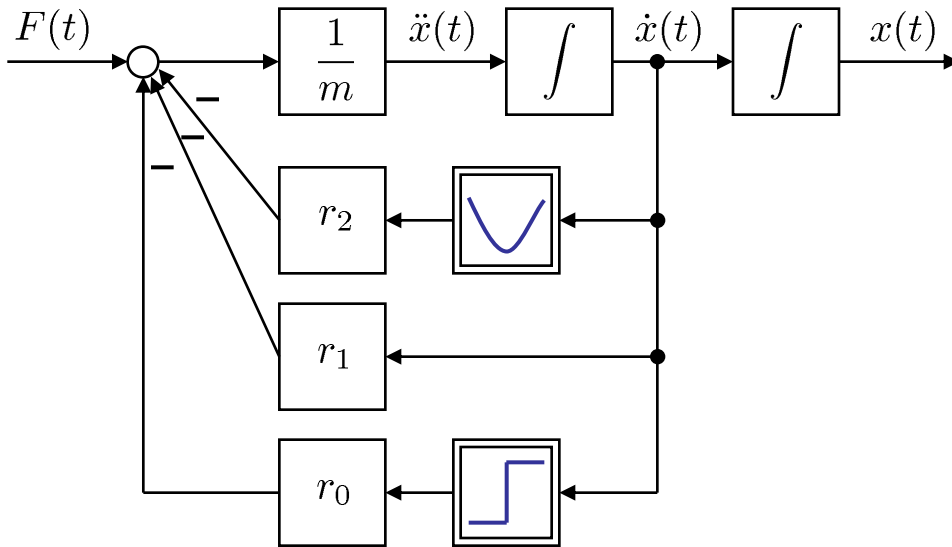
Längsdynamik eines Fahrzeuges

$$F(t) = m\ddot{x}(t) + r_2\dot{x}^2(t) + r_1\dot{x}(t) + r_0\text{sign}(\dot{x}(t))$$

Luftwiderstand
(quadratisch)

viskose Reibung
(linear)

trockene Reibung
(nichtlinear)



16.1 Kennlinien und Kennfelder

Beispiele für Kennlinien

Fremderregter Gleichstrommotor

Feldkreis: $u_F = R_F i_F + \dot{\Phi}_F$

Ankerkreis: $u_A = R_A i_A + L_A \dot{i}_A + e_M$

Gegen-EMK: $e_M = c \omega \Phi_F$

Antriebsmoment: $M_A = c i_A \Phi_F$

Antriebsgleichung: $\Theta \dot{\omega} = M_A - M_L$

Magnetisierungs-

Kennlinie: $\phi_F = F(i_F)$

inverse

Kennlinie: $i_F = \tilde{F}(\phi_F) = F^{-1}(\phi_F)$

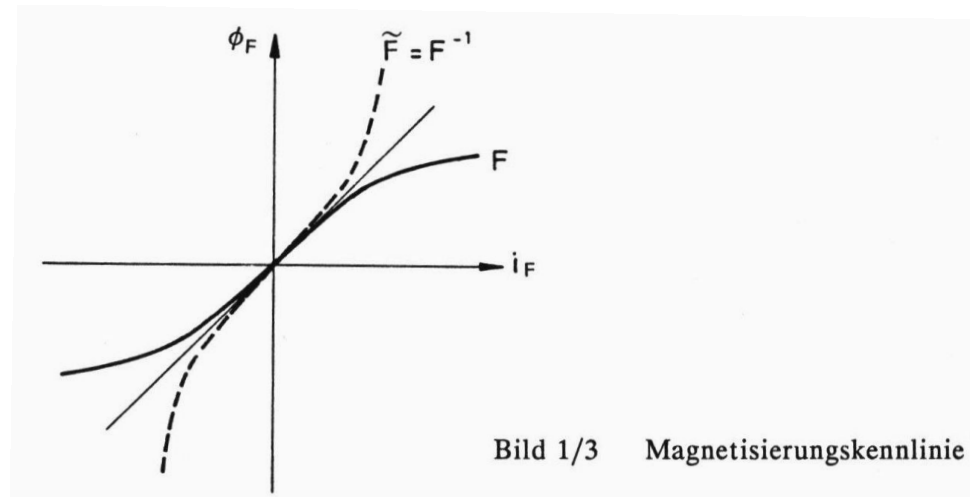
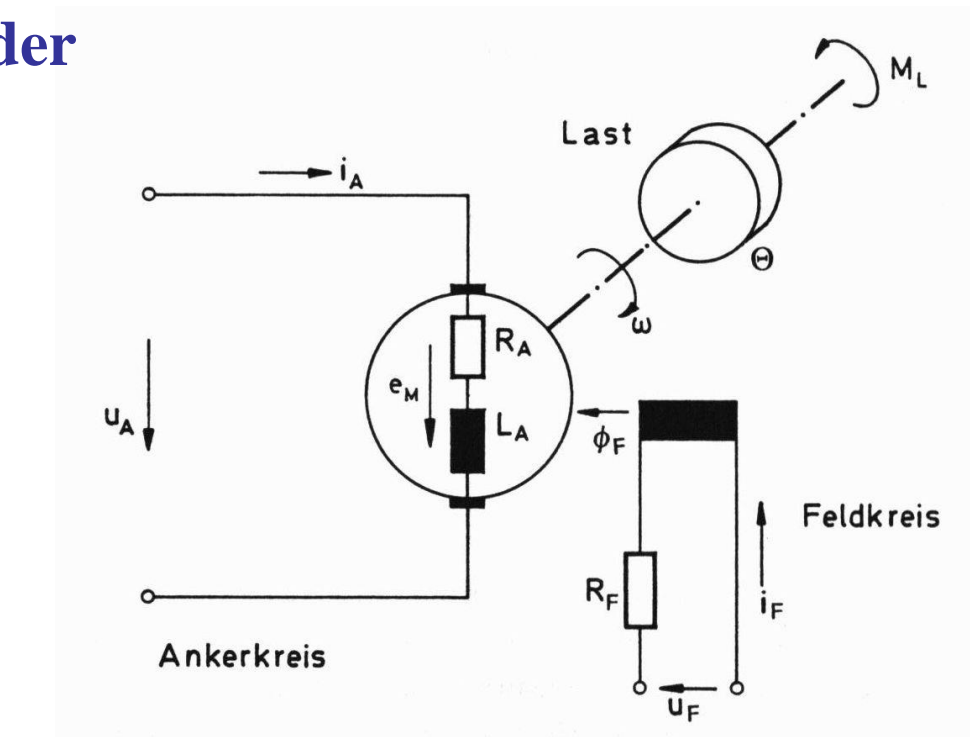
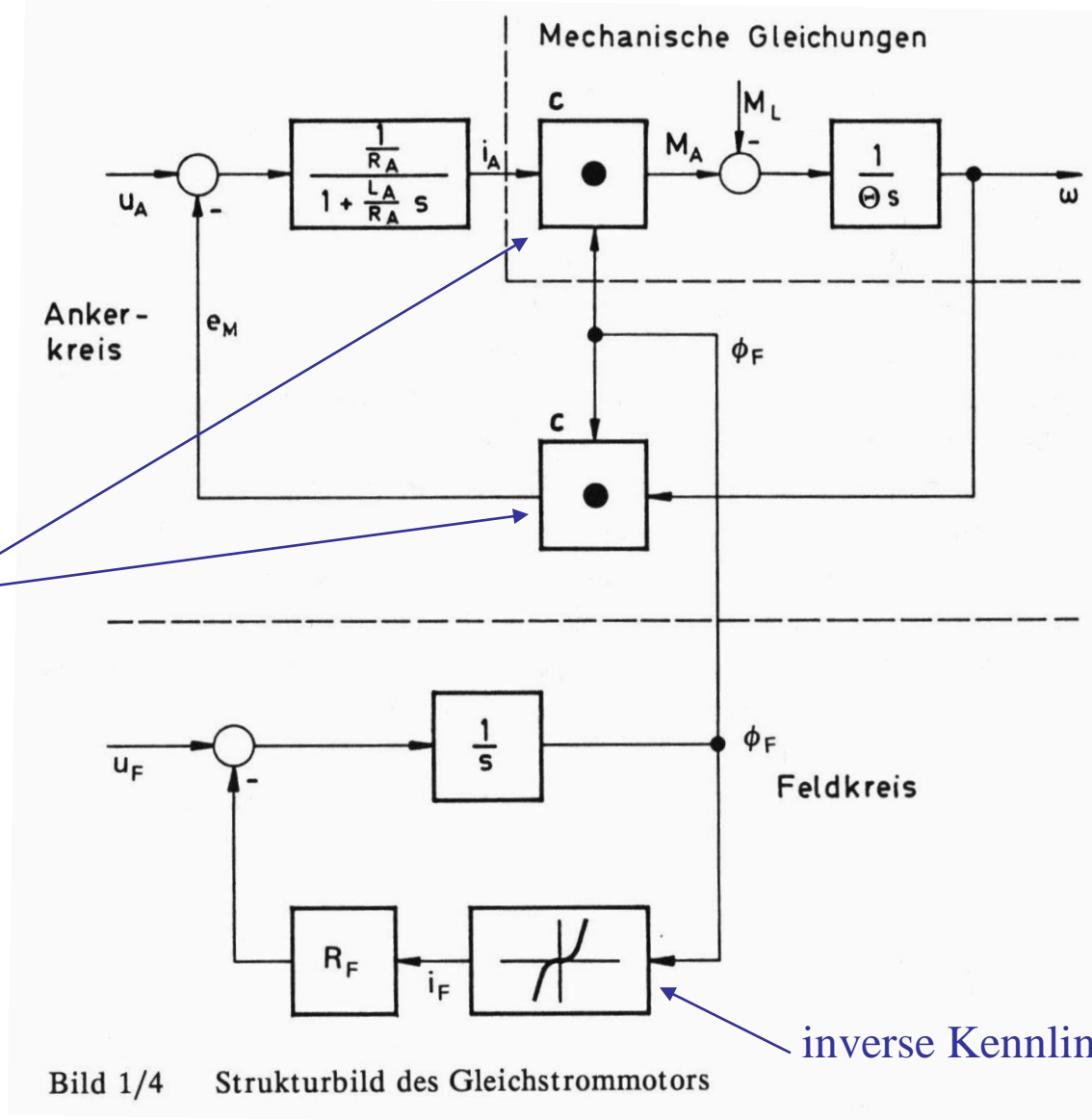


Bild 1/3 Magnetisierungskennlinie

16.1 Kennlinien und Kennfelder

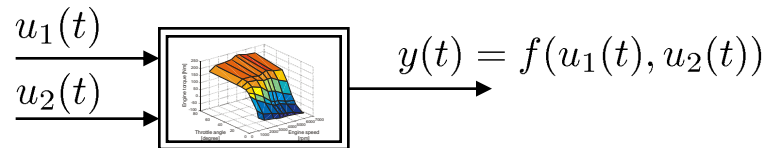


Keine Kennlinien!
 Produkte sind auch
 nichtlineare Elemente!

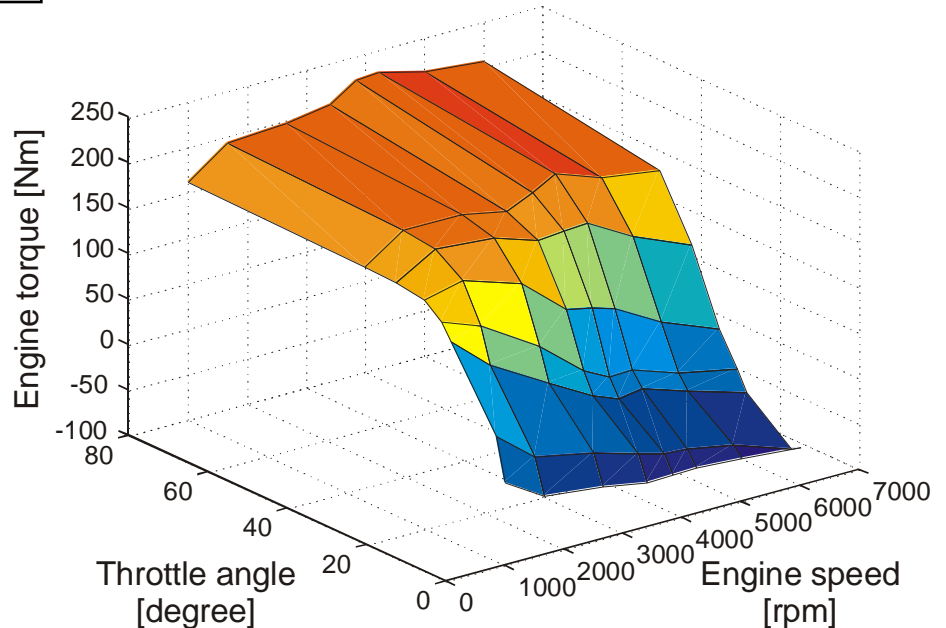
16.1 Kennlinien und Kennfelder

Kennfelder

Wenn eine Größe auf nichtlineare statische Weise von mehr als einer anderen Größen abhängt, muss die Kennlinie zum **Kennfeld** erweitert werden. Für mehr als 2 Eingangsgrößen lassen sich allerdings nur noch Schnitte graphisch darstellen.



Ein typisches Kennfeld aus dem Bereich der elektronischen Steuerung von Verbrennungsmotoren stellt die Abhängigkeit des Motordrehmoments von den Eingangsgrößen Motordrehzahl und Drosselklappenstellung (beim Ottomotor) bzw. Einspritzmenge (beim Dieselmotor) dar.



16.1 Kennlinien und Kennfelder

Rasterkennfelder

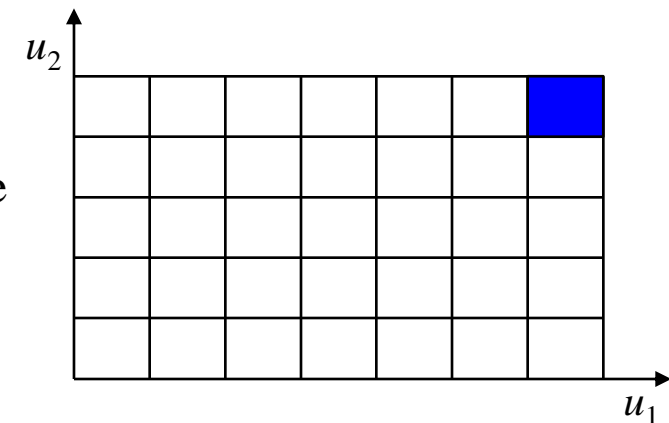
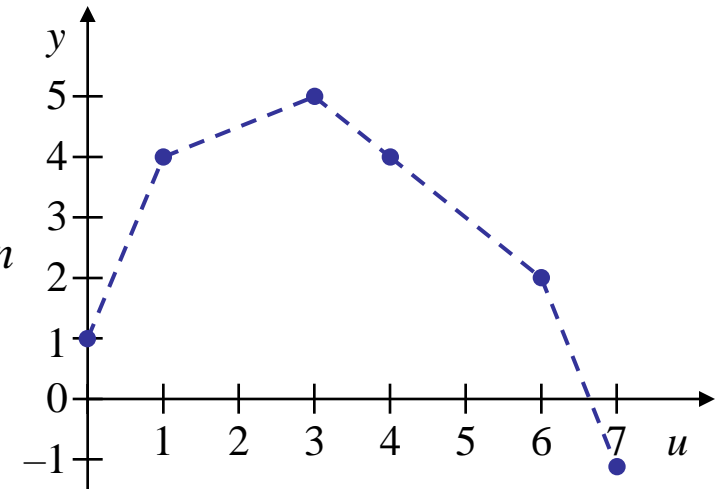
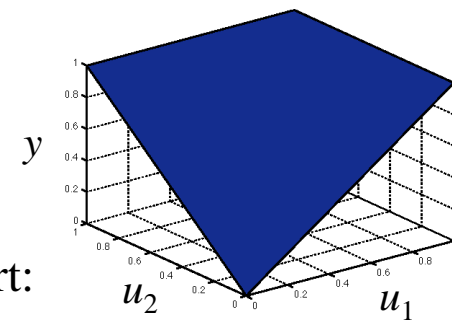
Klassischerweise werden Kennlinien/Kennfelder als Tabelle gespeichert:

u	0	1	3	4	6	7
y	1	4	5	4	2	-1

Ausgangswerte für Eingangswerte *zwischen* den tabellierten Eingangswerten werden durch *Interpolation* berechnet, typischerweise durch lineare Interpolation, d.h. indem man die Punkte mit Geraden (gestrichelt) verbindet.

Ausgangswerte für Eingangswerte außerhalb des tabellierten Bereichs (hier $[0, 7]$) müssen entweder geraten oder durch Vorwissen über die Zusammenhänge sinnvoll bestimmt werden. Typische Extrapolationsverläufe sind: auf Null abklingend, konstant, linear.

Für Kennfelder müssen die Punkte rasterförmig auf einem Gitter im u_1 - u_2 -Eingangsraum verteilt werden.



16.1 Kennlinien und Kennfelder

Andere Realisierungen von Kennfeldern

Die Anzahl zu tabellierender Punkte steigt bei Rasterkennfeldern exponentiell mit der Anzahl der Eingänge an. Außerdem ist der Ausgang nur stetig, aber nicht differenzierbar (Ableitung springt an den tabellierten Eingangswerten). Wichtige Alternativen zur Repräsentation von Kennfeldern sind:

- Polynome
- Splines (lokal gültige Polynome)
- Fuzzy-Logik (es existieren viele verschiedenen Typen)
- Neuronale Netze (es existieren viele verschiedenen Typen)
- Assoziativspeicher

Diese nichtlinearen Modelle haben Parameter, welche mit einem Optimierungsverfahren an Messdaten angepasst werden. Bei Rasterkennfeldern geschieht diese Anpassung meist durch das einfache Abspeichern der Messwerte in einer Tabelle. Das ist aber auch nur dann so einfach möglich, wenn die Daten rasterförmig vermessen werden. Ansonsten ist für die Bestimmung der Kennfeldpunkthöhen auch eine Optimierung notwendig.

16.2 Nichtlineare Effekte

Besondere Eigenschaften nichtlinearer dynamischer Systeme

Das Superpositions- und Verstärkungsprinzip gilt nicht:

- Blöcke in einem Blockschaltbild dürfen i.a. nicht vertauscht werden.
- Betrachtungen im Frequenzbereich nicht mehr (so einfach) möglich.
- Laplace-Transformation nicht mehr anwendbar.
- Analyse wird wesentlich erschwert.
- Komplexität der Betrachtungen wird wesentlich erhöht.

Nichtlinearitäten verändern die Frequenzen:

- Ein lineares System antwortet auf eine Schwingung mit einer Frequenz ω_1 , immer auch mit einer Schwingung der selben Frequenz (im eingeschwungenen Zustand). Lediglich die Amplitude und die Phase der Schwingung wird verändert. Um wie viel, lässt sich am Frequenzgang des linearen Systems ablesen.
- Für nichtlineare Systeme gilt das Alles nicht! Sie können das Spektrum des Eingangssignals beliebig verändern.

16.2 Nichtlineare Effekte

Besondere Eigenschaften nichtlinearer dynamischer Systeme

Nichtlinearitäten verändern die Frequenzen:

- Beispiel: Quadratische Kennlinie**

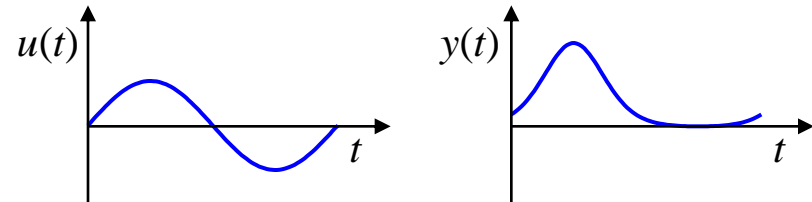
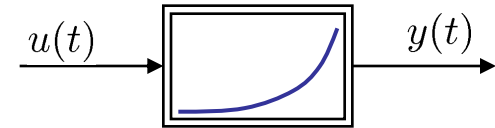
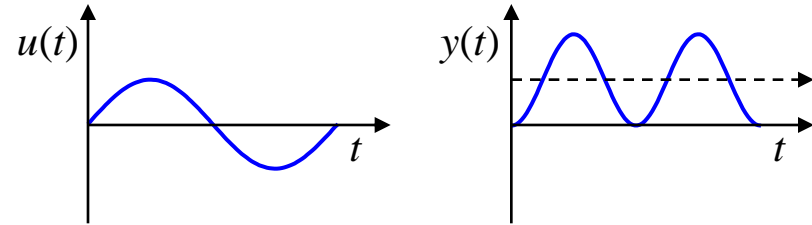
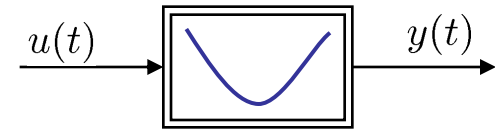
$$u(t) = A \sin(\omega_1 t)$$

$$y(t) = A^2 \sin^2(\omega_1 t) = \frac{A^2}{2} (1 - \cos(2\omega_1 t))$$

D.h. aus einem Eingangssignal mit der Frequenz ω_1 ist ein Ausgangssignal mit einem Gleichanteil (Frequenz 0) und einem periodischen Anteil doppelter Frequenz $2\omega_1$ geworden.

- Beispiel: Exponentielle Kennlinie**

Der sinusförmige Verlauf des Eingangssignals ist im Ausgangssignal überhaupt nicht mehr zu erkennen. Es enthält ein sehr breites Spektrum an Frequenzen.



16.2 Nichtlineare Effekte

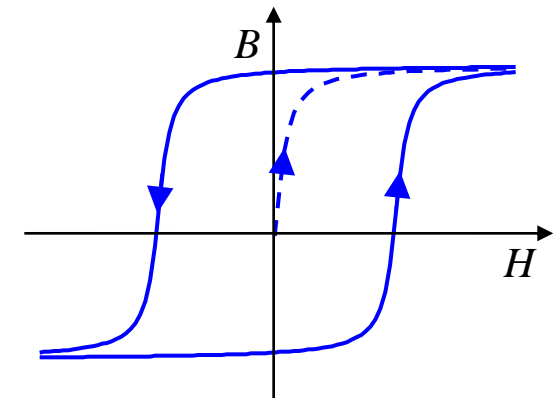
Besondere Eigenschaften nichtlinearer dynamischer Systeme

Hysterese

Hysterese (gr. hysteros: hinterher) ist das Fortdauern einer Wirkung nach Wegfall der Ursache. Sie bezeichnet eine Erscheinung, die durch eine Hysteresekurve gekennzeichnet wird (siehe Bild). Typisch für die Hysterese ist das Auftreten von bistabilem Verhalten. Bei gleichen Umgebungsbedingungen ist der Zustand von der Vergangenheit abhängig.

Man kann zwischen zwei Typen von Hysterese unterscheiden:

- Hysterese, bei welcher der durchlaufene Ast neben der Eingangsgröße nur noch von der Änderungsrichtung (also dem Vorzeichen der Ableitung) der Eingangsgröße abhängt. Bei einem Richtungswechsel springt der Arbeitspunkt von einem Ast auf den anderen. Beispiel: Komperatorschaltungen mit Operationsverstärkern.
- Hysterese, bei welcher der durchlaufene Pfad von der gesamten Vergangenheit der Eingangsgröße abhängt. Der Pfad verläuft auch im Inneren der Hysteresefläche und springt nicht. Beispiel: Magnetisierung ferromagnetischer Stoffe (Eisen).



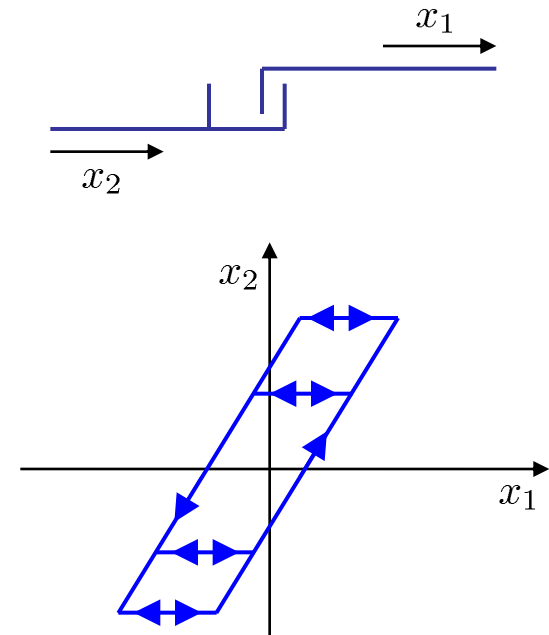
16.2 Nichtlineare Effekte

Besondere Eigenschaften nichtlinearer dynamischer Systeme

Lose oder "Spiel" (backlash)

Lose oder "Spiel" kann bei zwei bewegten Teilen auftreten, zwischen denen eine Kraft oder ein Drehmoment übertragen werden soll. Zwischen ineinander greifenden Zahnrädern tritt z.B. typischerweise Lose auf. Bei einem Getriebe addieren sich dann die Loseffekte der im Kraftfluss liegenden Zahnräder auf und die Gesamtlose kann evtl. nicht mehr vernachlässigt werden.

Die Schwierigkeit bei der Modellierung von Lose besteht darin, dass die beiden Auslenkungen x_1 und x_2 innerhalb des Losebereichs unabhängig voneinander sind.



16.2 Nichtlineare Effekte

Besondere Eigenschaften nichtlinearer dynamischer Systeme

Stabile Dauerschwingungen:

- Ein *lineares* System, das Dauerschwingungen ausführt ist *grenzstabil*; es weist ein (oder mehrere) konjugiert komplexe Polpaar(e) auf der imaginären Achse auf.
- Ein *lineares* System, das Dauerschwingungen ausführt, schwingt immer rein *sinusförmig*. Die Frequenz hängt von dem Imaginärteil des konjugiert komplexen Polpaars ab.
- Bei einem nichtlinearen System ist das nicht so! Dauerschwingungen treten typischerweise bei schaltenden Systemen auf, z.B. Systeme mit Zweipunktreglern. Aber auch nicht schaltende Systeme können Dauerschwingungen ausführen. Dauerschwingungen nichtlinearer Systeme sind normalerweise auch nicht sinusförmig.
- **Beispiel: Räuber/Beute-Modell (Volterra-Lotka-Modell)**

Beute: $\dot{x}_1 \sim x_1$

Räuber: $\dot{x}_2 \sim x_2$

D.h. Beute und Räuber vermehren sich proportional zu ihrem Bestand, also je mehr Tiere es gibt, desto mehr werden neu geboren.

16.2 Nichtlineare Effekte

Besondere Eigenschaften nichtlinearer dynamischer Systeme

Stabile Dauerschwingungen:

Gäbe es keine Räuber, würde sich die Beute ungehemmt exponentiell vermehren. Gäbe es genug Beute, so würden sich die Räuber ungehemmt exponentiell vermehren. Eine große Anzahl Räuber x_2 *reduziert* die Beuteanzahl, weil sie gefressen wird. Aber eine große Anzahl Beute x_1 *erhöht* die Räuberanzahl, weil sie mehr zu fressen hat.

Beute: $\dot{x}_1 = f(x_2) x_1$

Räuber: $\dot{x}_2 = g(x_1) x_2$

Wobei $f(x_2)$ monoton fallend und $g(x_1)$ monoton steigend ist. Im einfachsten Fall nehmen wir lineare Funktionen an:

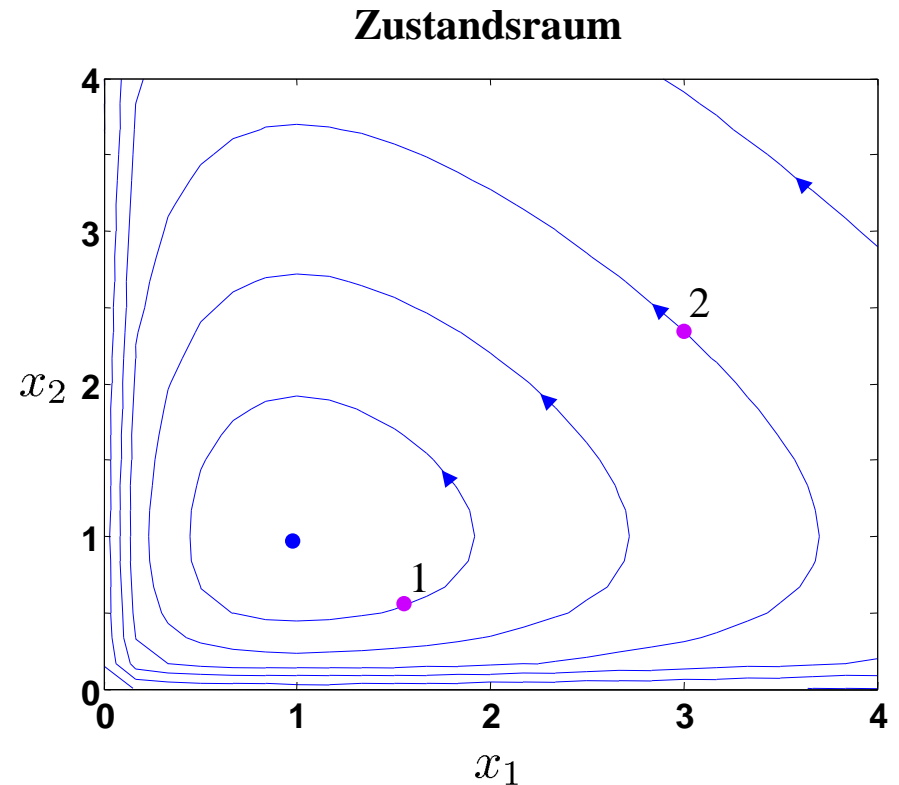
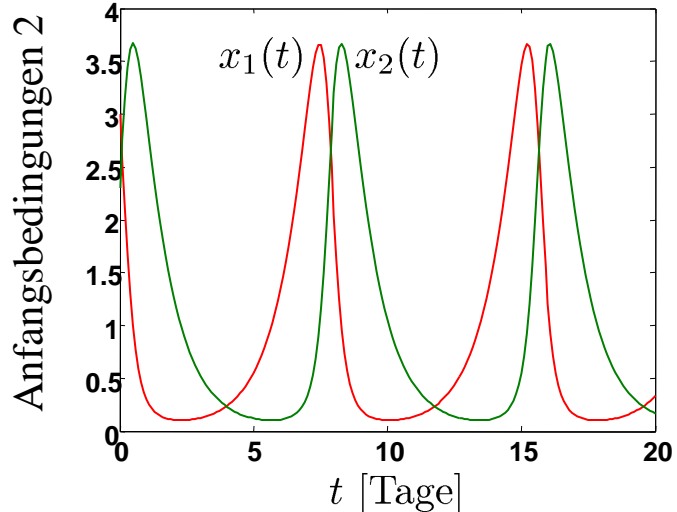
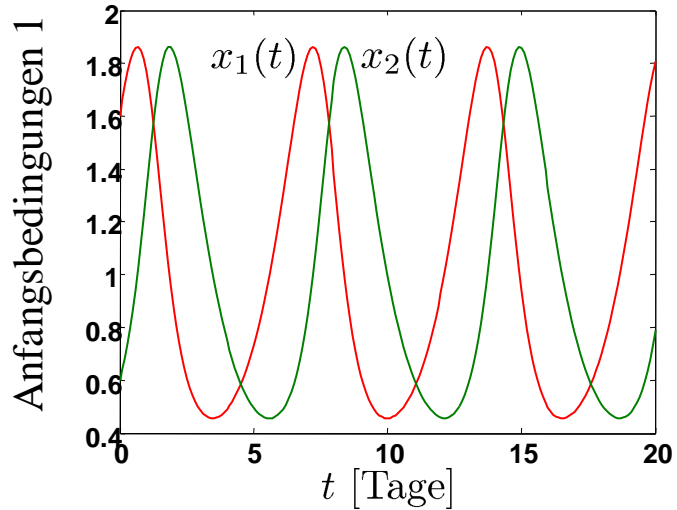
Beute: $\dot{x}_1 = (a - bx_2) x_1$

Räuber: $\dot{x}_2 = (cx_1 - d) x_2$

Für jede Anfangsbedingung $x_1 > 0$, $x_2 > 0$ ergeben diese nichtlinearen Differentialgleichungen eine stabile Dauerschwingung.

16.2 Nichtlineare Effekte

Besondere Eigenschaften nichtlinearer dynamischer Systeme



16.2 Nichtlineare Effekte

SIR-Modell zur Virus-Ausbreitung

Bevölkerung wird (prozentual) aufgeteilt in 3 Gruppen:

S : Susceptible (anfällig) I : Infected (infiziert) R : Recovered (erholt und immun)

SIR-Modell (aus <https://www.youtube.com/watch?v=k6nLfCbAzgo>):

1. $\dot{S} = -\text{transmit} \cdot S \cdot I$
2. $\dot{I} = \text{transmit} \cdot S \cdot I - \text{recover} \cdot I$
3. $\dot{R} = \text{recover} \cdot I$

zu 2: Die Anzahl an Menschen, die sich neu infizieren (Änderungsrate \dot{I}) ist proportional zur Übertragungs/Ansteckungsrate „transmit“ und zur Anzahl der bereits Infizierten I und zur Anzahl der noch Infizierbaren/Anfälligen S anzüglich der Änderungsrate der Immunen, d.h. Infizierten, die sich erholt haben; Erholungsrate ist „recover“.

zu 1: Die Anzahl der Anfälligen ändert sich entsprechend der Rate der Neuinfizierten.

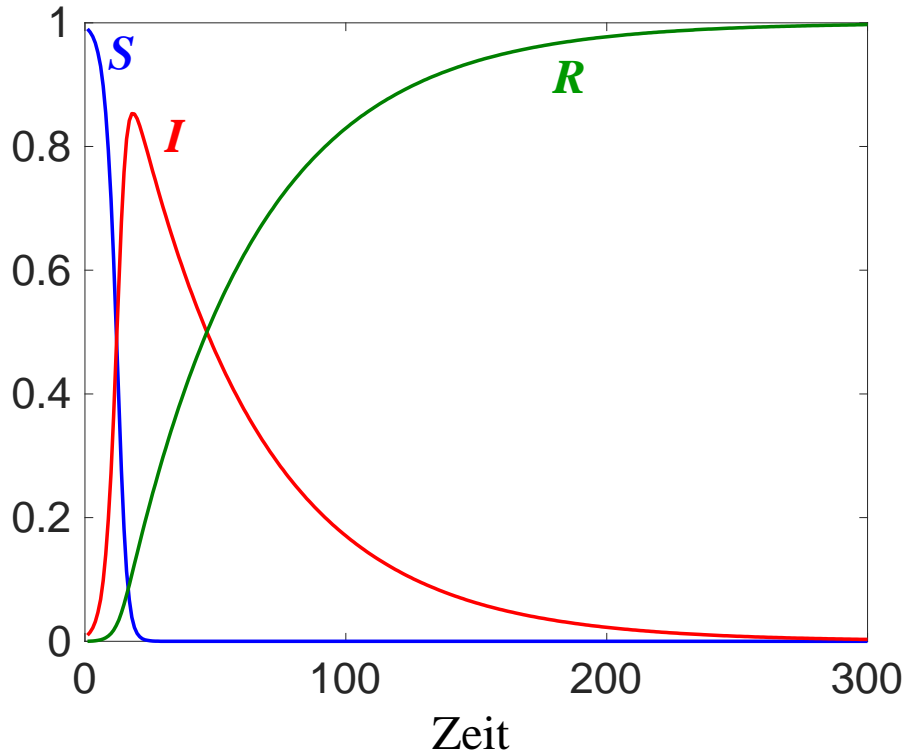
zu 3: Die Anzahl der Erholten ändert sich mit der Erholungsrate „recover“ mal der Anzahl an Infizierten I .

16.2 Nichtlineare Effekte

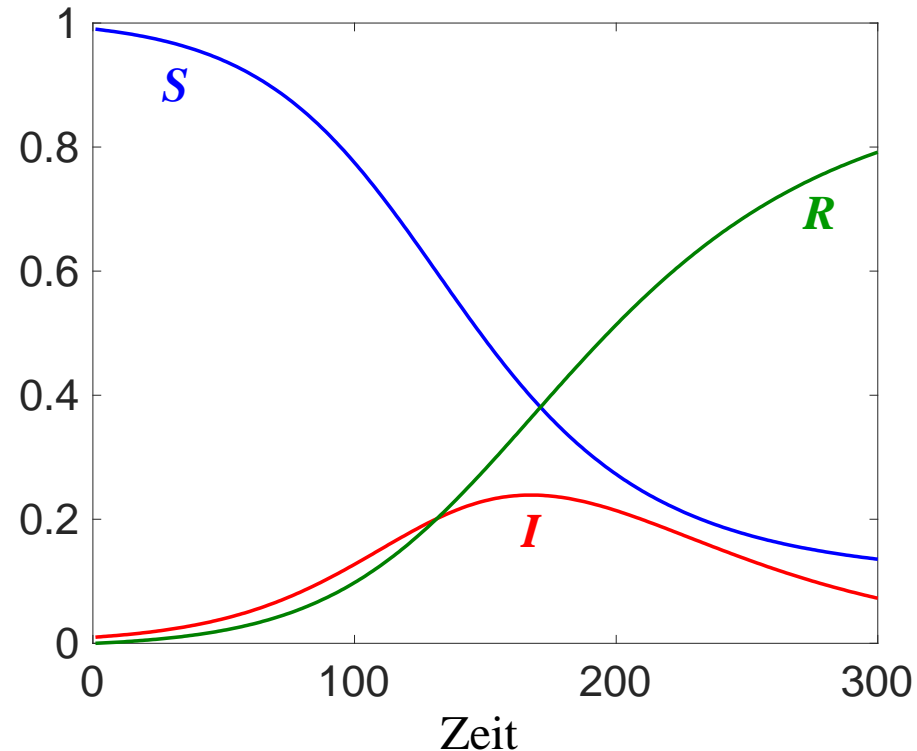
SIR-Modell zur Virus-Ausbreitung

Um das System aus 3 nichtlinearen DGLs zu simulieren, benötigt man Anfangsbedingung, hier sind das die Initialwerte $I = 0.01$, $S = 0.99$, $R = 0$. Und man braucht Annahmen für die Übertragungs/Ansteckungsrate „transmit“ und die Erholungsrate „recover“.

transmit = 0.5, recover = 0.02



transmit = 0.05, recover = 0.02



16.2 Nichtlineare Effekte

Besondere Eigenschaften nichtlinearer dynamischer Systeme

Chaos:

Nichtlineare dynamische Systeme können sich chaotisch verhalten, d.h. sie zeigen *scheinbar* zufälliges Verhalten, obwohl sie theoretisch deterministisch sind. Es gelingt nicht, chaotische Systeme über einen längeren Zeitraum zu simulieren, da sich kleinste Fehler (z.B. in den Anfangsbedingungen) drastisch auswirken.

Beispielsweise wiederholt sich die Bewegung eines chaotischen Systems wie eines doppelten Pendels unter gewissen Voraussetzungen (Anfangsbedingungen) niemals, sondern verändert sich stetig, so dass sein Verhalten zufällig und ungeordnet erscheint. Man spricht von chaotischen Systemen, wenn man nicht vorhersagen kann, wie sich ein System entwickeln wird. Formal ist ein deterministisches System dann chaotisch, wenn unter nur minimalen Veränderungen der Anfangsbedingungen die zeitliche Entwicklung des Systems nicht berechenbar ist und zu deutlich unterschiedlichen Endzuständen führt. Andernfalls (kleine Ursache, kleine Wirkung oder große Ursache, große Wirkung) spricht man von stetigem Verhalten. Chaotische Systeme, wie z.B. das Wetter, sind nicht weit in die Zukunft vorhersagbar, kleine Änderungen können enorm (insbesondere exponentiell) verstärkt werden, große Änderungen können gedämpft werden. Ein Prozess wird als chaotisch bezeichnet, wenn Zustände des Prozesses zu einem beliebigen Zeitpunkt mit exponentieller Geschwindigkeit divergieren.

16.3 Stabilität nichtlinearer Systeme

Unterschiede zwischen der Stabilität linearer und nichtlinearer Systeme

Die Stabilität nichtlinearer Systeme ist eine wesentlich kompliziertere Eigenschaft als bei linearen Systemen. Bei linearen Systemen unterscheiden wir zwischen:

1. *Stabile Systeme*: Die Zustände klingen von jeder Anfangsbedingung exponentiell schnell auf den Ruhezustand ab.
2. *Grenzstabile Systeme*: Die Zustände nähern sich weder dem Ruhezustand noch streben sie gegen unendlich.
3. *Instabile Systeme*: Mindestens ein Zustand strebt betragsmäßig gegen unendlich.

Die Untersuchung der Gewichtsfunktion $g(t)$ oder der Pole der Übertragungsfunktion des Systems gibt Aufschluss darüber, welcher Fall vorliegt. Folgende Fakten sind dabei besonders bemerkenswert und gelten *nicht* für nichtlineare Systeme:

- Stabilität ist ausschließlich eine Eigenschaft des Systems; sie hängt von nichts anderem ab.
- Insbesondere gilt die Stabilitätseigenschaft im gesamten Zustandsraum.
- Es gibt nur einen einzigen Ruhezustand, typischerweise der Ursprung des Zustandsraums.
- Stabilität garantiert exponentiell schnelles Abklingen.

16.3 Stabilität nichtlinearer Systeme

Zur Stabilität nichtlinearer Systeme

- Stabilität ist keine globale Eigenschaft. Es kann stabile und instabile Bereiche des Zustandsraums geben.
- Da es verschiedene Stabilitätsbereiche geben kann, ist es offensichtlich, dass es auch verschiedene Ruhelagen geben kann. Zu jedem Stabilitätsbereich gehört eine Ruhelage.
- Ruhelagen können stabil, grenzstabil oder instabil sein. Das kann man durch Linearisierung der nichtlinearen Zustandsgleichungen herausfinden:

Nichtlineare Zustandsgleichungen für $u(t) = 0$: $\dot{\underline{x}}(t) = f(\underline{x}(t))$

Linearisierung um \underline{x}_0 : $\Delta\dot{\underline{x}}(t) = \underline{A}\Delta\underline{x}(t)$ mit $\underline{A} = \left. \frac{df(\underline{x})}{d\underline{x}} \right|_{\underline{x}=\underline{x}_0}$

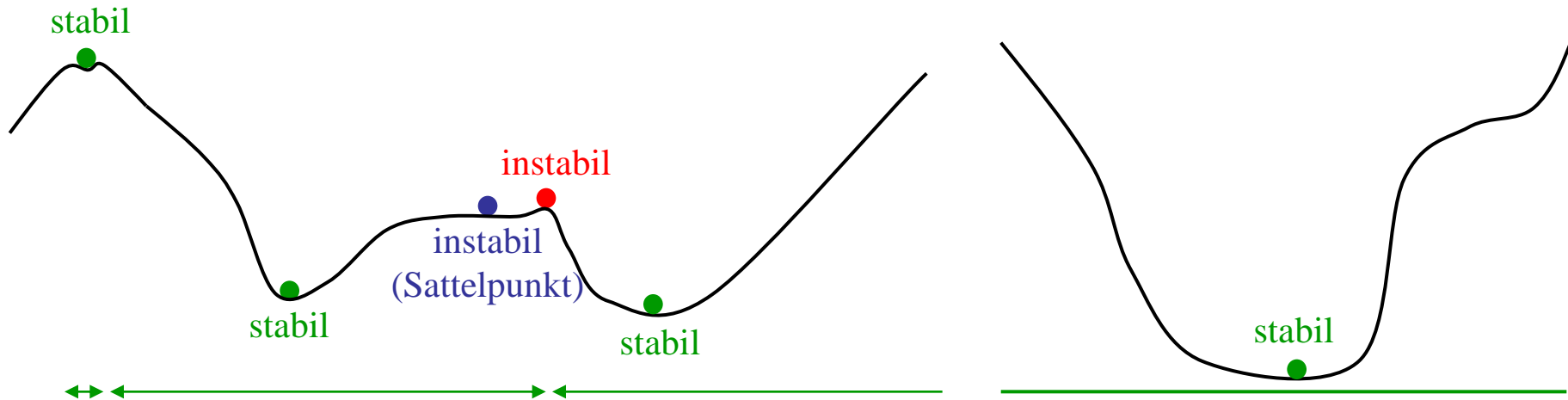
mit $\Delta\underline{x} = \underline{x} - \underline{x}_0$. Ist die durch Linearisierung gefundene Systemmatrix stabil, so gilt dies auch für die zugehörige Ruhelage $\underline{x} = \underline{x}_0$. Diese Aussage gilt aber nur für den Punkt \underline{x}_0 selbst. Über die Größe des Stabilitätsgebiets rund um diesen Punkt ist damit noch nichts ausgesagt. Daher ist das Ergebnis aus der Linearisierung für die Praxis nicht besonders hilfreich.

16.3 Stabilität nichtlinearer Systeme

Ruhelagen für nichtlineare dynamische Systeme

Im Gegensatz zu linearen Systemen können nichtlineare Systeme mehrere Ruhelage haben. Mittels Linearisierung lässt sich zwar die Stabilität einer Ruhelage bestimmen, aber sie erlaubt keine Aussage über die **Größe des Einzugsgebiets** (*region/domain/basin of attraction*). Ist das Einzugsgebiet winzig (stabile Ruhelage ganz links), dann können schon kleine Störungen das System destabilisieren und eine solche stabile Ruhelage ist praktisch instabil! Im Idealfall ist das *Einzugsgebiet unendlich groß* (Bild rechts) und damit das System **global stabil**.

Die Bestimmung der Größe des Einzugsgebiets ist eine schwierige Aufgabe. Oft wird man nur ein Teilgebiet des wahren Einzugsgebiets finden.



16.3 Stabilität nichtlinearer Systeme

Beispiel: Ruhelagen eines Pendels

Ein Pendel hat 2 Ruhelagen:

- Stabile Ruhelage bei $\varphi = 0^\circ$ mit dem Einzugsbereich $(-180^\circ, 180^\circ)$
- Instabile Ruhelage bei $\varphi = 180^\circ$

offenes Intervall

stabile
Ruhelage



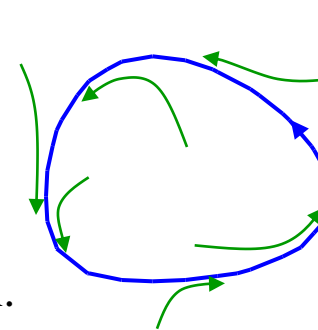
instabile
Ruhelage



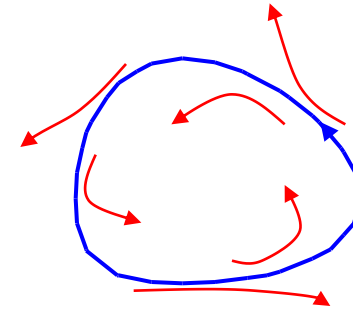
Stabile und instabile Dauerschwingungen

Wie das Räuber/Beute-Beispiel zeigt, sind Dauerschwingungen auch ein typisches Phänomen nichtlinearer Systeme. Eine Dauerschwingung kann genau wie Ruhelagen stabil oder instabil sein. Eine **stabile Dauerschwingung** ist dadurch gekennzeichnet, dass sie sich durch eine kleine Störung nicht aus der "Bahn werfen" lässt. Das System strebt nach Störungen immer wieder dieser Dauerschwingung zu. **Instabile Dauerschwingungen** brechen bei einer Störung zusammen und das System strebt einer anderen Dauerschwingung oder einer Ruhelage zu. Sie können daher in der Realität nicht beobachtet werden.

stabile
Dauerschwingung



instabile
Dauerschwingung



16.4 Kennlinienregler

Kennlinien als Regler

Neben Nichtlinearitäten im Prozess, die oft unerwünscht oder zumindest für den Reglerentwurf unbequem sind, werden Nichtlinearitäten auch in Reglern verwendet. Breite Anwendung finden dabei einfache schaltende Regler z.B. realisiert durch Zweipunktschalter.

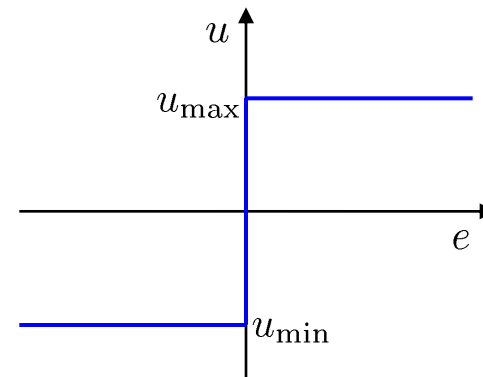
Anwendungsbeispiele für Kennlinienregler

- Bügeleisen, Elektroherd.
- Kühlschrank, Gefrierschrank.
- Thermostat an der Heizung.

Eigenschaften von Kennlinienreglern

- Einfach und leicht zu implementieren.
- Zweipunktregler nutzt meist die maximal mögliche Stellenergie aus (oberer und unterer Anschlag $[u_{\min}, u_{\max}]$), d.h. ist bei geschickter Schaltstrategie **zeitoptimal**.
- Starke Beanspruchung des Stellglieds (Hin- und Herschalten zwischen den Anschlägen).
- Dauerschwingungen.

Zweipunktschalter



Oft gilt:
 $u_{\min} = -u_{\max}$
Oder:
 $u_{\min} = 0$

16.4 Kennlinienregler

Zweipunktregler ohne und mit Hysterese

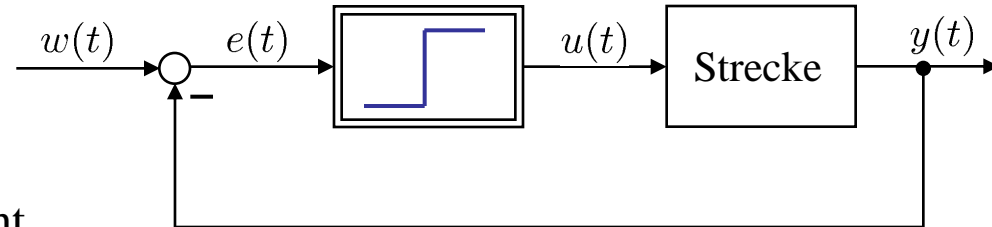
Betrachten wir zunächst einen Zweipunktregler ohne Hysterese zur Regelung einer I-Strecke. Andere (phasenminimale)

Strecken zeigen ein kompliziertes aber nicht grundsätzlich unterschiedliches Verhalten.

Deshalb wollen wir uns hier mit einer I-Strecke begnügen.

Ist der Regelkreis eingeschwungen, also $y = w$, dann führt die Stellgröße $u = u_{\max}$ zu einem sofortigen kleinen Anstieg der Regelgröße y . Dadurch wird die Regelabweichung e leicht negativ und der Zweipunktregler schaltet um auf $u = u_{\min}$. Das führt wiederum zu einer Reduktion von y und damit zu einer positiven Regelabweichung e und es folgt wieder ein Umschalten des Reglers auf $u = u_{\max}$.

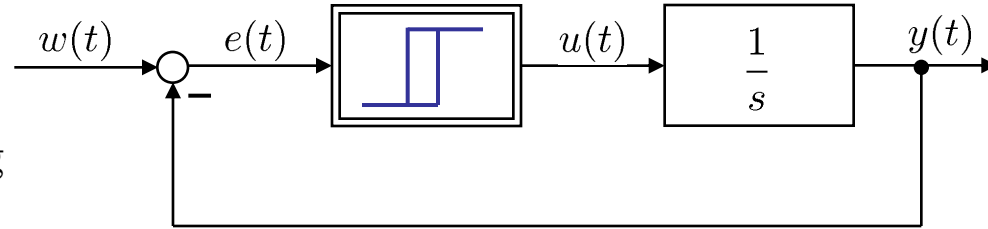
Ein reiner Zweipunktregler führt also zu einem (theoretisch unendlich schnellen) ständigen Hin- und Herschalten des Reglers (**Rattern**, *chattering*). Die Folge ist eine enorme Beanspruchung des Stellglieds. Um die zeitlichen Abstände zwischen dem Hin- und Herschalten zu vergrößern, verwendet man deshalb in der Regel einen Zweipunktregler mit Hysterese.



16.4 Kennlinienregler

Zweipunktregler mit Hysterese für I-Strecke

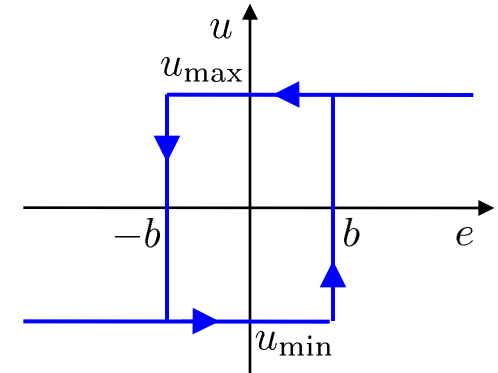
Eine Hysteresebreite von b bedeutet, dass der Zweipunktregler erst dann schaltet, wenn die Regelabweichung e betragsmäßig größer als b geworden ist.



Amplitude und Frequenz der Dauerschwingung

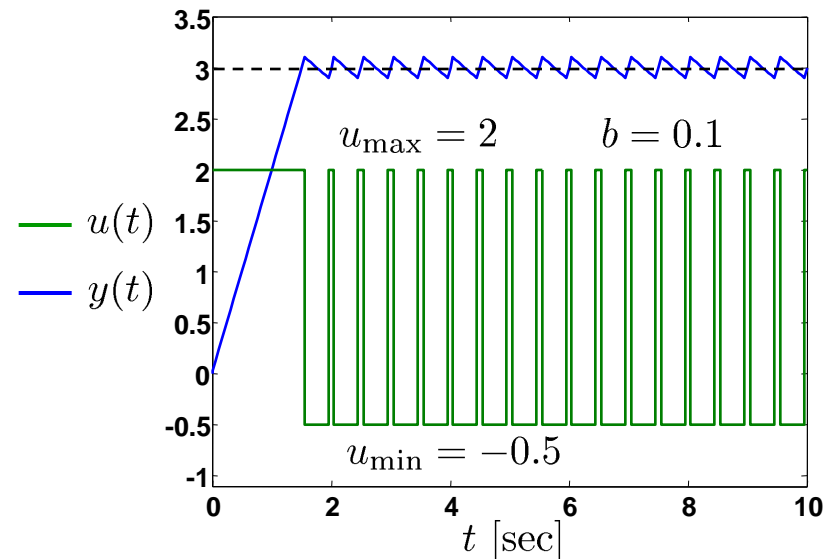
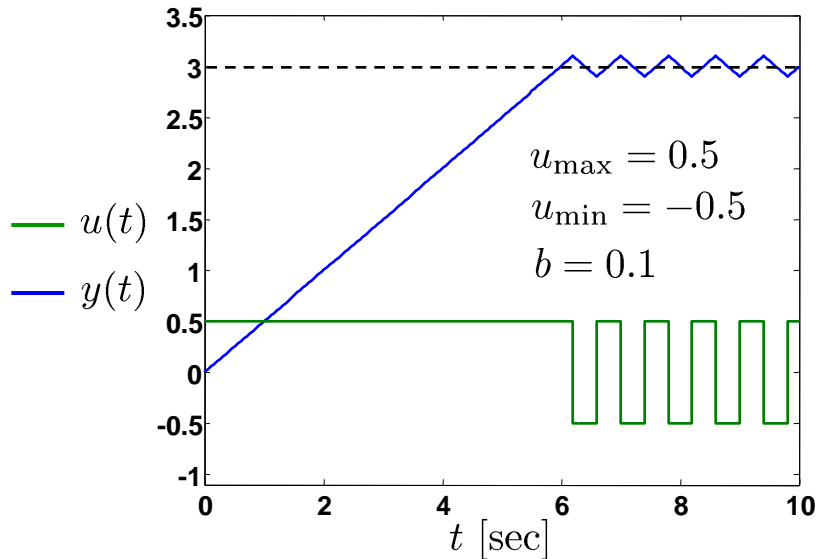
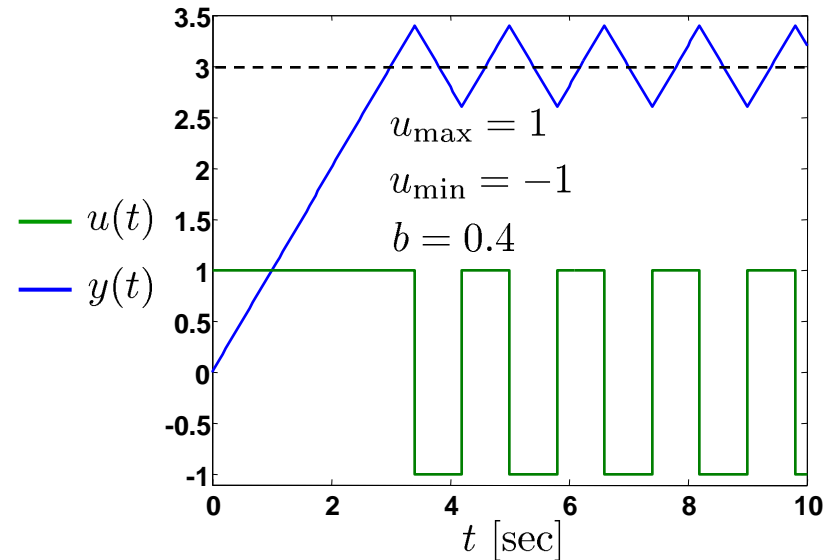
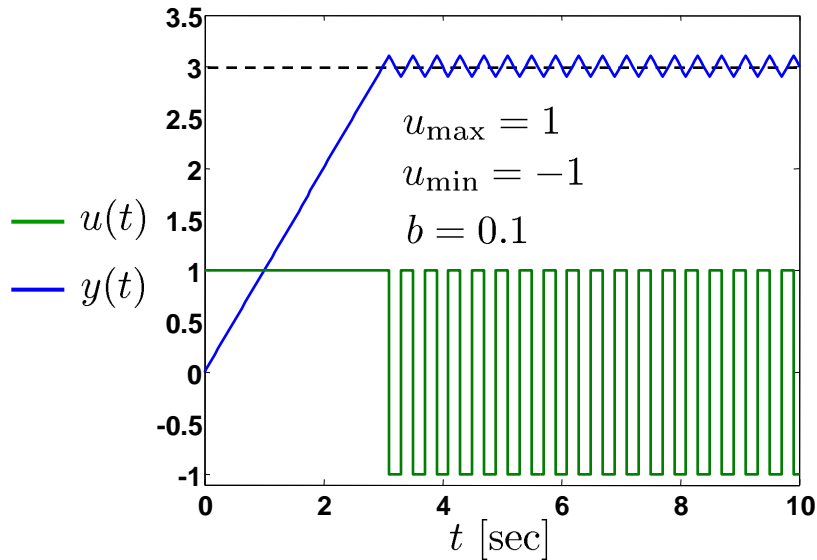
Auch mit dem Zweipunktregler mit Hysterese stellt sich nach dem Erreichen des Sollwertes eine Dauerschwingung ein. Die Amplitude der Dauerschwingung ist durch die Hysteresebreite gegeben, da die Regelgröße y bei $e > b$ bzw. $e < -b$ jeweils die Richtung wechselt. Die Frequenz der Dauerschwingung ist proportional zu u_{\max} , wenn $u_{\min} = -u_{\max}$, weil u_{\max} die Steigung des Anstiegs von y bestimmt. Und die Frequenz der Dauerschwingung ist antiproportional zu b , weil b die Zeit des Anstiegs und Abfalls von y und damit die Periodendauer bestimmt.

Zweipunktschalter mit Hysterese



Dauerschwingung: Amplitude $\sim b$ Frequenz $\sim \frac{u_{\max}}{b}$

16.4 Kennlinienregler

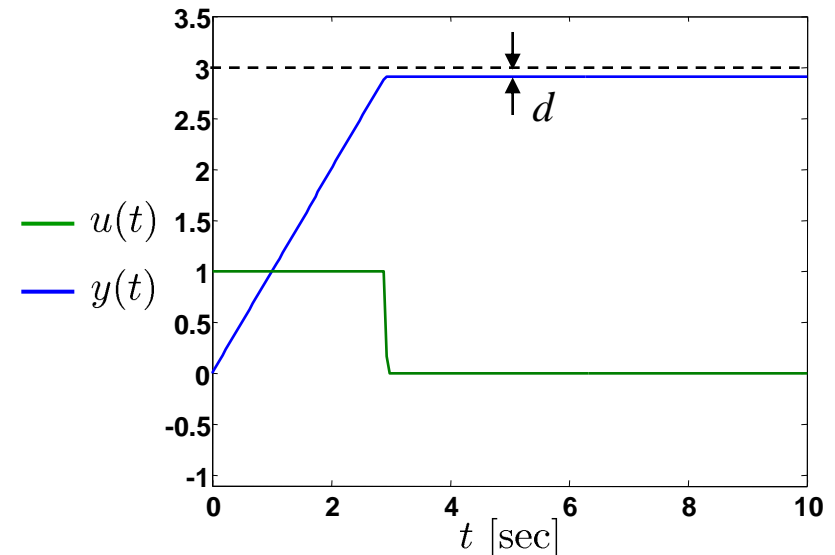
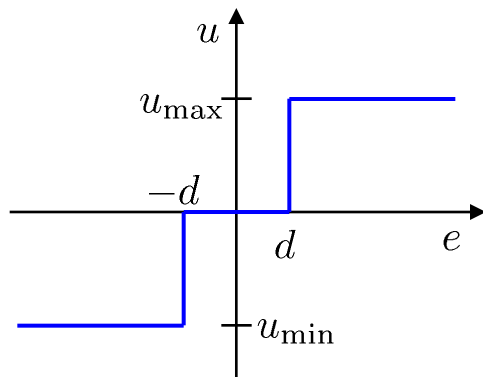


16.4 Kennlinienregler

Zweipunktregler mit toter Zone

Wie auf der Folie zuvor zu sehen ist, lässt sich zwar durch die Wahl der Hysteresebreite und maximalen Stellgröße die Amplitude und Frequenz der Dauerschwingung beeinflussen, eine Dauerschwingung tritt aber in jedem Fall auf. Das ist häufig aus zahlreichen Gründen unerwünscht (Abnutzung, Energieverbrauch, Akustik, etc.). Eine **tote Zone** im Zweipunktregler kann die Dauerschwingung auf Kosten einer verminderten stationären Genauigkeit verhindern. Solange sich die Regelabweichung innerhalb der toten Zone befindet ($-d < e < d$), gibt ein solcher Regler die Stellgröße Null aus. Deshalb funktioniert dies auch nur bei Strecken mit I-Verhalten, weil dort stationär $u = 0$ ist.

**Zweipunktschalter
mit toter Zone**

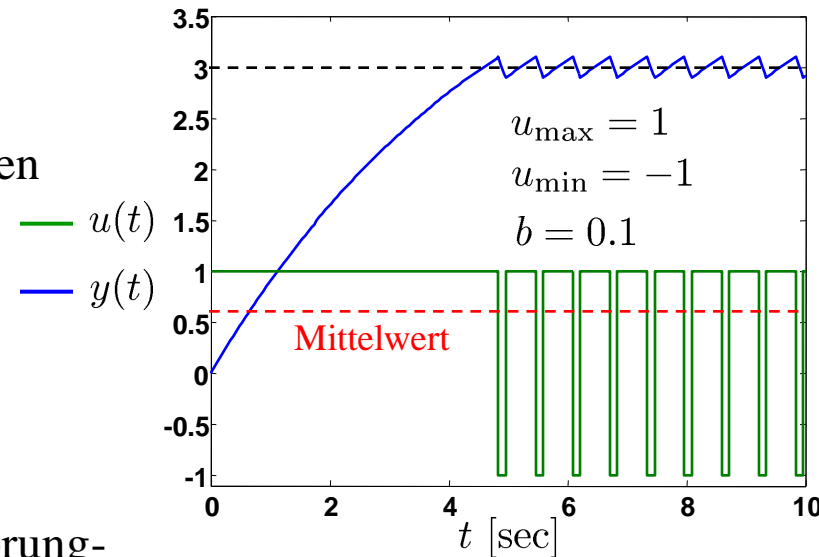
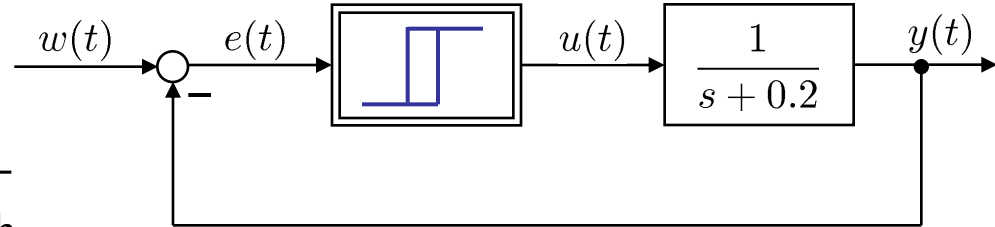


16.4 Kennlinienregler

Zweipunktregler für PT₁-Strecke

Wenn mit einem Zweipunktregler nicht eine I-Strecke sondern eine andere (phasenminimale) Strecke geregelt wird, ergibt sich ein prinzipiell ähnliches Verhalten.

Im Unterschied zu Strecken mit I-Verhalten, kommt es bei proportionalen Strecken nicht auf den absoluten Wert der Stellgröße an, sondern auf deren Abweichung vom Arbeitspunkt. Liegt der Arbeitspunkt bei $w = y = 3$, $u = 0.6$ (Streckenverstärkung ist 5), dann bedeutet ein Sprung auf $u_{\max} = 1$ ein $\Delta u = +0.4$ und ein Sprung auf $u_{\min} = -1$ ein $\Delta u = -1.6$. D.h. die Sprunghöhe nach unten ist 4 Mal so groß wie die Sprunghöhe nach oben. Daher dauert die Anstiegsphase ca. 4 Mal so lang und hat eine ca. 4 Mal kleinere Steigung als die Abstiegsphase. Weil die Amplitude der Dauerschwingung recht klein ist, erkennt man kaum einen Unterschied zwischen integralem Verhalten (linearer An-/Abstieg) und proportionalem Verhalten (1 – exponentieller An-/Abstieg).



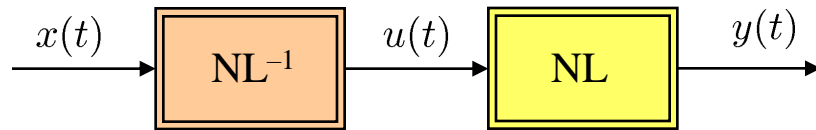
16.4 Kennlinienregler

Kompensation von Nichtlinearitäten

Die Kompensation unerwünschter Effekte ist eine häufig angewandte Idee in vielen Disziplinen. Wir haben Beispiele in Form der Vorsteuerung, Störgrößenaufschaltung und Entkopplung von Mehrgrößenregelkreisen gesehen, wo unerwünschtes dynamisches Verhalten durch Kompensation (näherungsweise) eliminiert wurde.

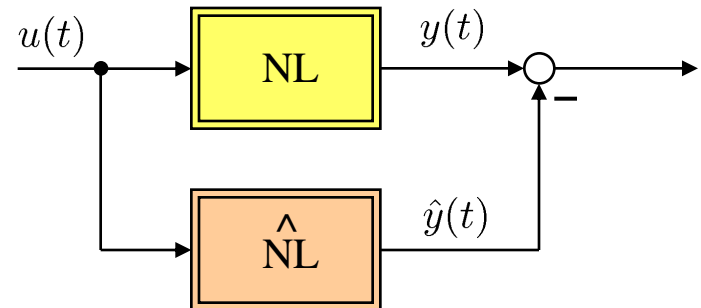
Diese Idee lässt sich auch auf Nichtlinearitäten anwenden. Prinzipiell kann man dabei zwischen einer Kompensation mittels Reihen- und mittels Parallelschaltung unterscheiden. Bei der Reihenschaltung muss die **inverse Nichtlinearität** berechnet werden, bei der Parallelschaltung reicht eine direkte Modellierung der Nichtlinearität. Die 1. Methode wird auf der nächsten Folie demonstriert; die 2. wird bei dem Verfahren *Feedback Linearization* genutzt.

Kompensation durch Reihenschaltung



$$y = f(u) \quad \rightarrow \quad y = f(f^{-1}(x)) = x$$
$$u = f^{-1}(x)$$

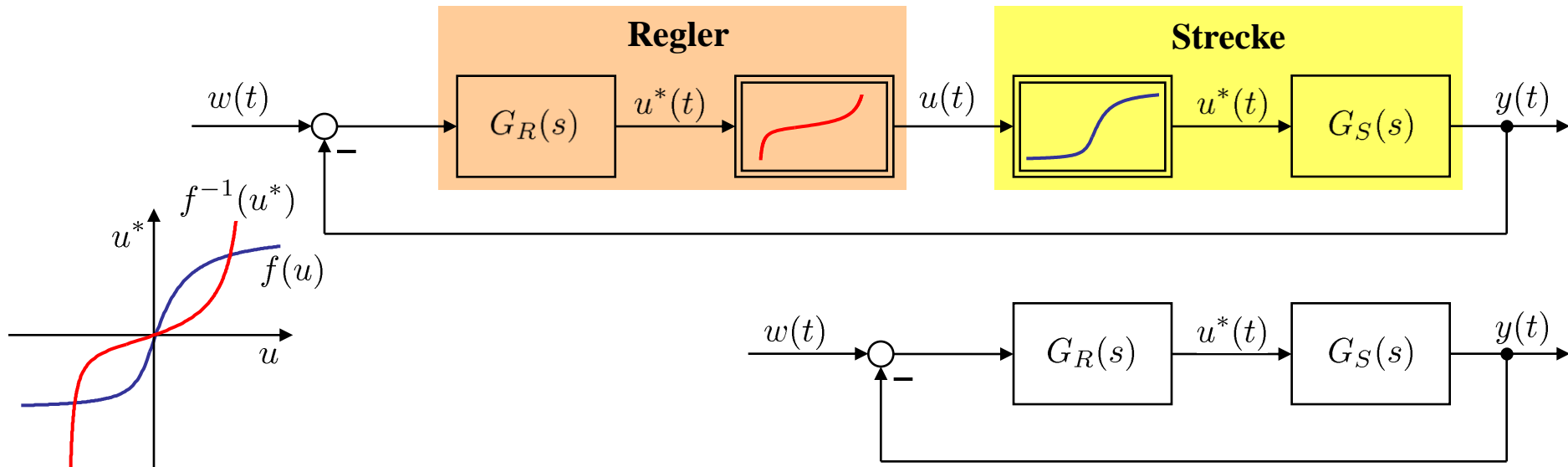
Kompensation durch Parallelschaltung



16.4 Kennlinienregler

Regelung mit inverser Kennlinie

Viele Regelstrecken haben eine Hammerstein-Struktur, weil die wesentliche Nichtlinearität oft im Stellglied (also am Streckeneingang) steckt. Dann kann diese **Nichtlinearität** leicht durch eine **inverse Kennlinie** kompensiert werden. Gelingt diese Kompensation perfekt ist jeglicher nichtlinearer Einfluss aus dem Regelkreis eliminiert. Es kann dann mit Hilfe der linearen Regelungstheorie ein linearer Regler $G_R(s)$ für den linearen Streckenanteil $G_S(s)$ entworfen werden. Die Analyse und Synthese des Regelkreises ist genauso einfach wie im rein linearen Fall.



16.5 Übersicht wichtiger nichtlinearer Regelungsansätze

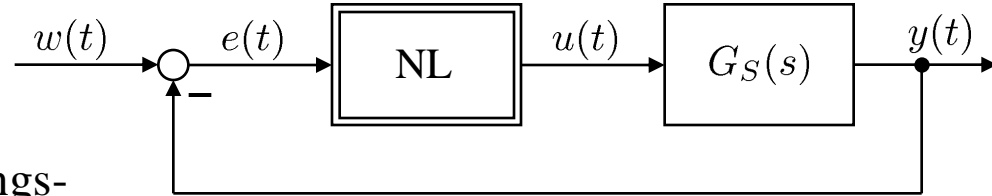
In der nichtlinearen Regelungstechnik gibt es keine allgemeine und ausgereifte Theorie wie für lineare Systeme. Vielmehr kommt ein Bündel an verschiedenen Methoden zum Einsatz, die jeweils für eine bestimmte Klasse von Systemen oder Anwendungen besonders geeignet sind. Da viele der Methoden auch mathematisch recht aufwändig sind und einige Übung und teilweise Kreativität im effizienten Umgang erfordern, beschränken wir uns hier auf eine Kurzübersicht über die wichtigsten Ansätze:

- *Beschreibungsfunktionen (Harmonische Balance)*: Besonders für die Analyse von Dauerschwingungen geeignet.
- *Zustandsebene*: Für die Analyse und Reglersynthese für Systeme 2. Ordnung geeignet. Die zeitoptimale Regelung lässt sich gut behandeln.
- *Direkte Methode von Ljapunow*: Sehr allgemeiner Ansatz zur Stabilitätsuntersuchung. Lässt sich auch für die Reglersynthese einsetzen. Sehr breites Anwendungsfeld.
- *Neuronale Netze und Fuzzy-Systeme*: Besonders für die experimentelle Modellierung nichtlinearer Systeme (Identifikation) aber auch für die Reglersynthese (direkt als Regler oder als Zwischenschritt zu einem "klassischen" Regler) geeignet. Sehr breit anwendbar und mit relativ wenig mathematischem Aufwand zu verstehen.

16.5 Übersicht wichtiger nichtlinearer Regelungsansätze

Beschreibungsfunktionen (Harmonische Balance)

Grundidee:

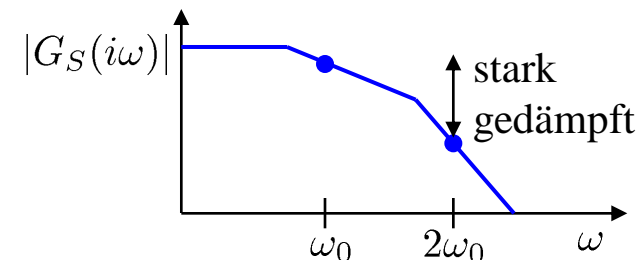


- Es wird angenommen, dass sich der geschlossener Regekreis im Schwingungsgleichgewicht befindet. Die sich einstellenden Dauerschwingungen sollen analysiert werden, d.h. man möchte Amplituden, Frequenzen und Stabilitätseigenschaften der möglichen Dauerschwingungen bestimmen.
- Jede Schwingung lässt sich als Fourier-Reihe schreiben, also in eine Summe aus einem Gleichwert A_0 und Sinus-Signalen der Grundfrequenz $A_1 \sin(\omega_0 t + \varphi_1)$ und Oberschwingungen (Vielfache der Grundfrequenz) $A_k \sin(k\omega_0 t + \varphi_k)$ zerlegen:

$$\text{Dauerschwingung} = A_0 + A_1 \sin(\omega_0 t + \varphi_1) + A_2 \sin(2\omega_0 t + \varphi_2) + A_3 \sin(3\omega_0 t + \varphi_3) + \dots$$

- Wenn das nichtlineare dynamische System genügend stark gedämpft ist (siehe Bild), werden die hohen Frequenzen im Vergleich zu den niedrigeren sehr stark unterdrückt. Dann kann man die Signale mit ihren Grundschwingungen approximieren:

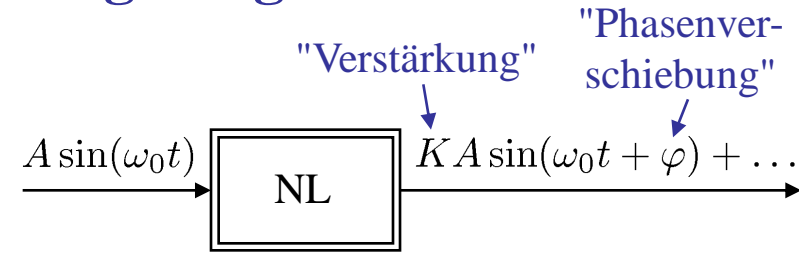
$$\text{Dauerschwingung} \approx A_0 + A_1 \sin(\omega_0 t + \varphi_1)$$



16.5 Übersicht wichtiger nichtlinearer Regelungsansätze

Beschreibungsfunktionen (Harmonische Balance)

- Die statische Nichtlinearität wird durch einen Verstärkungsfaktor K und eine Phasenverschiebung φ beschrieben, die in Bezug auf die Grundschiwingung definiert sind. Alle höheren Frequenzen werden vernachlässigt, weil ihre Amplituden sehr stark durch das dynamische System gedämpft werden (Annahme!). Die Gleichwerte werden zuvor über eine Transformation der Variablen eliminiert, so dass man mit Kleinsignalgrößen rechnet. Daher gilt im Folgenden auch $\omega = 0$.



- Die Nichtlinearität wird durch ihre sog. **Beschreibungsfunktion $N(A)$** dargestellt:

$$N(A) = K e^{i\varphi} = K \cos\varphi + i K \sin\varphi$$

Die Beschreibungsfunktion ist also etwas Ähnliches wie eine *Übertragungsfunktion*; sie hängt aber *nicht* von der *Frequenz* ω_0 des Signals, *sondern* von dessen *Amplitude* A ab!

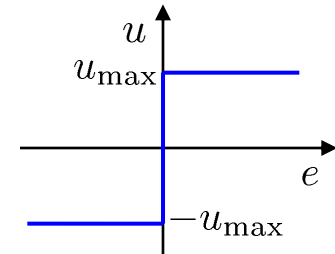
Die Beschreibungsfunktion kann aber nur genutzt werden, um die (näherungsweise) Antwort auf ein Sinus-Signal zu berechnen; sie hat nicht den universellen Charakter der Übertragungsfunktion, mit deren Hilfe man ja die Antwort auf jedes beliebige Eingangssignal berechnen kann.

16.5 Übersicht wichtiger nichtlinearer Regelungsansätze

Beschreibungsfunktionen (Harmonische Balance)

- Für eindeutige Kennlinien (für jedes e gibt es nur ein u) ist die Beschreibungsfunktion reell, d.h. $\varphi = 0$.
- Die Beschreibungsfunktion z.B. für ein Zweipunktglied lautet:

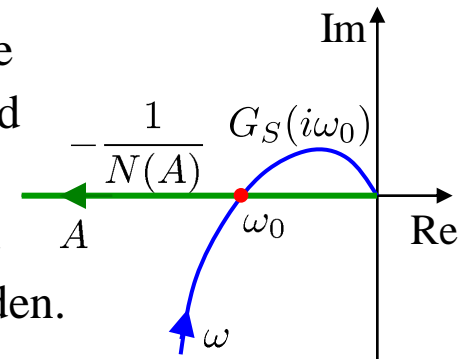
$$N(A) = \frac{4u_{\max}}{\pi A} \quad (\text{aus Tabelle entnommen})$$



- Zur Berechnung der Dauerschwingungseigenschaften erhalten wir für $\omega = 0$ folgende charakteristische Gleichung für den geschlossenen Regelkreis:

$$e = -G_S(i\omega_0)N(A)e \quad \boxed{G_S(i\omega_0) = -\frac{1}{N(A)}}$$

- Dies sind in Wirklichkeit 2 Gleichungen, eine für den Realteil, eine für den Imaginärteil. Daraus können die beiden Unbekannten A und ω_0 bestimmt werden. Geometrisch sind sie durch den **Schnittpunkt** der Ortskurven $G_S(i\omega)$ und $-1/N(A)$ gegeben. So können die Amplitude und die Frequenz der Dauerschwingung berechnet werden.



16.5 Übersicht wichtiger nichtlinearer Regelungsansätze

Zustandsebene und zeitoptimale Regelung

In der **Zustandsebene** (manchmal auch als **Phasenebene** bezeichnet) können nichtlineare dynamische Systeme 2. Ordnung behandelt werden, die folgender DGL genügen:

$$\ddot{x} = f(x, \dot{x}, u)$$

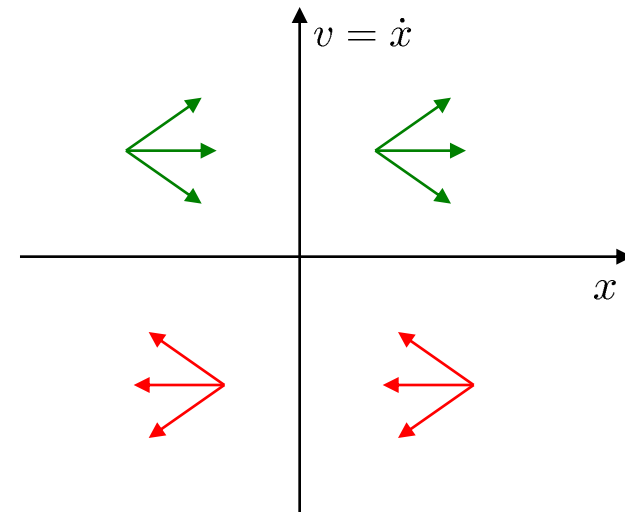
Führt man neben dem Zustand x den Zustand $v = \dot{x}$ ein, erhält man:

$$\dot{x} = v$$

$$\dot{v} = f(x, v, u)$$

In vielen Anwendungen bezeichnet x einen Weg und v eine Geschwindigkeit. Die Bahnen oder *Trajektorien* eines solchen Systems 2. Ordnung lassen sich in der Zustandsebene, d.h. dem 2-dimensionalen Zustandsraum, anschaulich darstellen.

Weil der 2. Zustand v gleich der Ableitung des 1. Zustands x ist, sind die Bewegungsrichtungen der Trajektorien in der Zustandsebene nicht völlig frei, sondern durch die mit den Pfeilen gekennzeichneten Richtungen in den 4 verschiedenen Quadranten vorgegeben.



16.5 Übersicht wichtiger nichtlinearer Regelungsansätze

Zustandsebene und zeitoptimale Regelung

Um die Trajektorien in der Zustandsebene zu berechnen, geht man wie folgt trickreich vor, um die explizite Zeitabhängigkeit zu eliminieren und damit eine DGL 2. Ordnung in der Zeit t umzuformen in eine DGL 1. Ordnung im Zustand x :

$$\frac{dv}{dx} = \frac{\frac{dv}{dt}}{\frac{dx}{dt}} = \frac{\dot{v}}{\dot{x}} = \frac{f(x, v, u)}{v} \quad \rightarrow \quad v dv = f(x, v, u) dx$$

Die sich daraus ergebende Trajektorienschar lässt sich meist leicht über Trennung der Variablen bestimmen. Dazu muss allerdings angenommen werden, dass die Stellgröße u stückweise konstant ist, damit die Zeit t nicht explizit in der Gleichung vorkommt. Damit kann man also die homogene Lösung ($u = 0$) berechnen und alle Fälle behandeln, wo u aus einer Schaltkennlinie (wie bei einem Zweipunktregler) erzeugt wird.

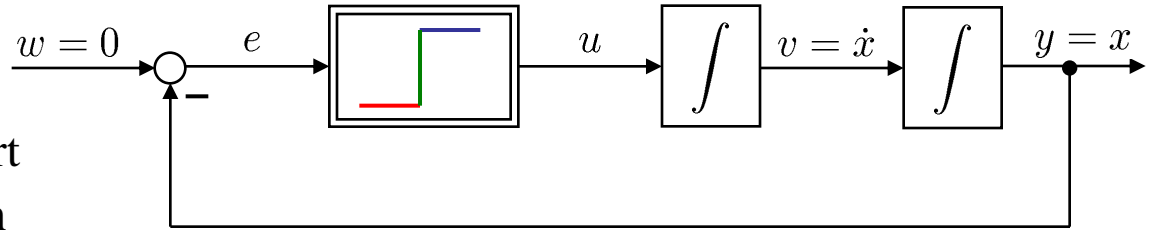
Beispiel: Zweipunktschalter mit Doppelintegrator

Ein doppelter Integrator beschreibt z.B. die einfachst mögliche Bewegungsgleichung, nämlich die ungedämpfte beschleunigte Bewegung: $F = m\ddot{x}$. Das kann z.B. für Steuerung der Antriebe (nur Ein/Aus) eines Satelliten im Weltall verwendet werden.

16.5 Übersicht wichtiger nichtlinearer Regelungsansätze

Zustandsebene und zeitoptimale Regelung

Die Antriebe des Satelliten können nur in Flugrichtung oder in Gegenflugrichtung gefeuert werden. Das lässt sich durch einen



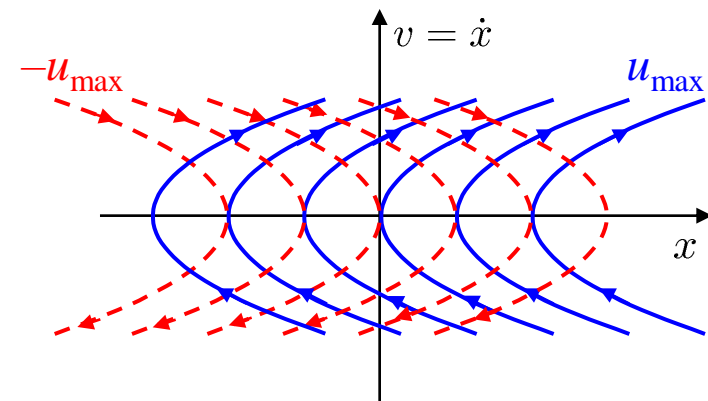
Zweipunktschalter modellieren $u = u_{\max}$ oder $u = u_{\min} = -u_{\max}$. Damit ergibt sich ($m = 1$):

$$\begin{aligned} \dot{x} &= v \\ \dot{v} &= \pm u_{\max} \end{aligned} \quad \rightarrow \quad v dv = \pm u_{\max} dx \quad \rightarrow \quad \int v dv = \int \pm u_{\max} dx \quad \rightarrow \quad \frac{v^2}{2} + c = \pm u_{\max} x$$

mit einer Integrationskonstanten c . Interpretiert als $x(v)$ sind das nach oben (für $u = u_{\max}$) bzw. nach unten (für $u = -u_{\max}$) geöffnete Parabeln.

Liefert der Zweipunktregler die Stellgröße $u = u_{\max}$, dann bewegt sich das System entlang der blauen, durchgezogenen Parabeln; bei der Stellgröße $u = -u_{\max}$, bewegt es sich entlang der roten, gestrichelten Parabeln.

Auf welcher Parabel der ganzen Parabelschar das System startet, hängt vom Anfangszustand (x_0, v_0) ab.



16.5 Übersicht wichtiger nichtlinearer Regelungsansätze

Zustandsebene und zeitoptimale Regelung

Wann bewegt sich das System nun auf welcher Parabel? Der Zweipunktregler schaltet nach folgender Strategie um:

$$e > 0 \rightarrow x < 0: u_{\max}$$

$$e < 0 \rightarrow x > 0: -u_{\max}$$

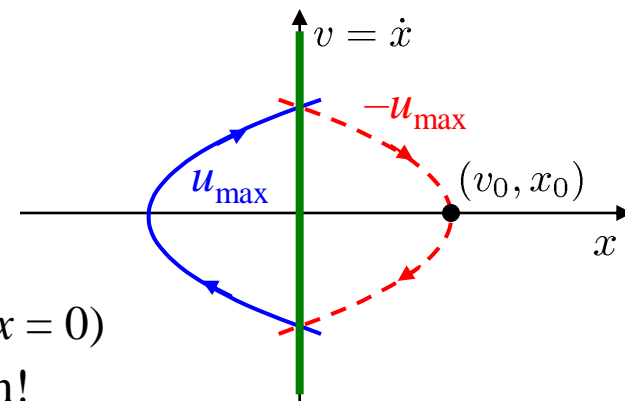
D.h. in der linken Halbebene ($x < 0$) bewegt sich das System entlang der blauen, durchgezogenen Parabeln und in der rechten Halbebene ($x > 0$) entlang der roten, gestrichelten.

D.h. der einfache Zweipunktregler ist nicht in der Lage den Regelkreis zur Ruhe zu bringen. Das System führt Schwingungen aus, deren Amplitude von dem Anfangszustand (x_0, v_0) abhängt.

Übrigens sind die Schwingungen *nicht* sinusförmig, sonst müssten die Kurven in der Zustandsebene Halbkreise oder -ellipsen statt Parabeln sein.

Wie kann der Regler verbessert werden?

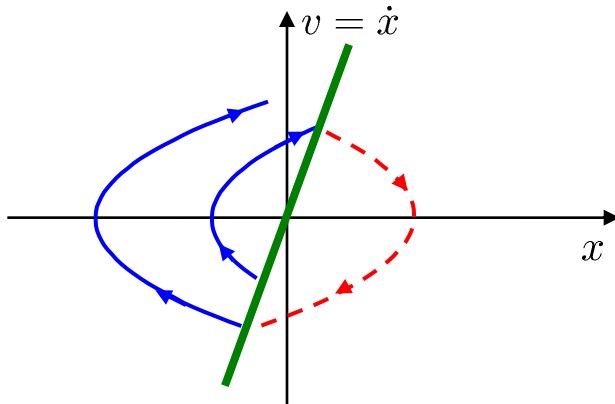
Eine nahe liegende Idee ist, nicht auf der senkrechten Achse ($x = 0$) umzuschalten, sondern die grüne dicke Schaltgerade zu drehen!



16.5 Übersicht wichtiger nichtlinearer Regelungsansätze

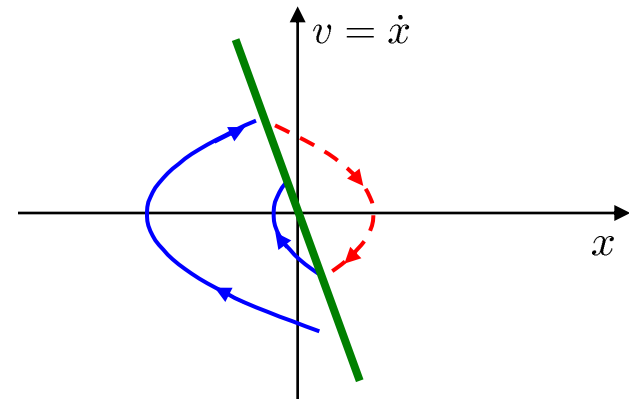
Zustandsebene und zeitoptimale Regelung

Drehen der Schaltgeraden nach rechts



Dauerschwingung wird instabil!

Drehen der Schaltgeraden nach links



Dauerschwingung klingt gegen Null ab!

Offensichtlich klingt die Dauerschwingung gegen die Ruhelage im Ursprung ab, wenn die Schaltgerade nach links geneigt wird, und dieses Abklingen geschieht um so schneller je stärker die Schaltgerade nach links geneigt wird. Diese stabilisierende Schaltgerade folgt der Gleichung:

$$v = \dot{x} = -k x \quad \text{mit } k > 0$$

16.5 Übersicht wichtiger nichtlinearer Regelungsansätze

Zustandsebene und zeitoptimale Regelung

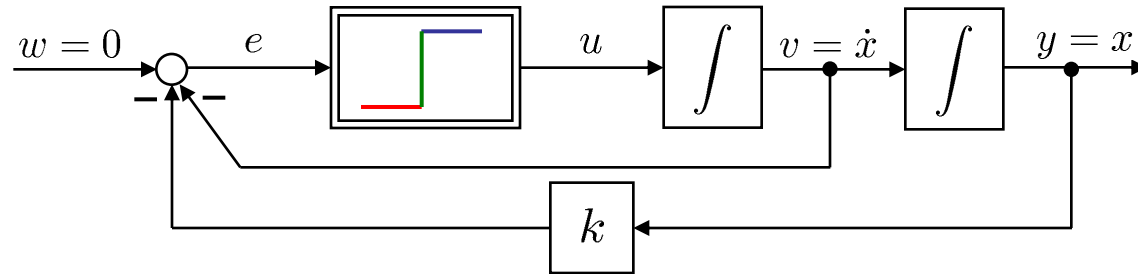
Um diese Schaltgerade zu realisieren, müssen wir die aktuelle Umschaltstrategie:

$$e > 0 \rightarrow x < 0: u_{\max}$$

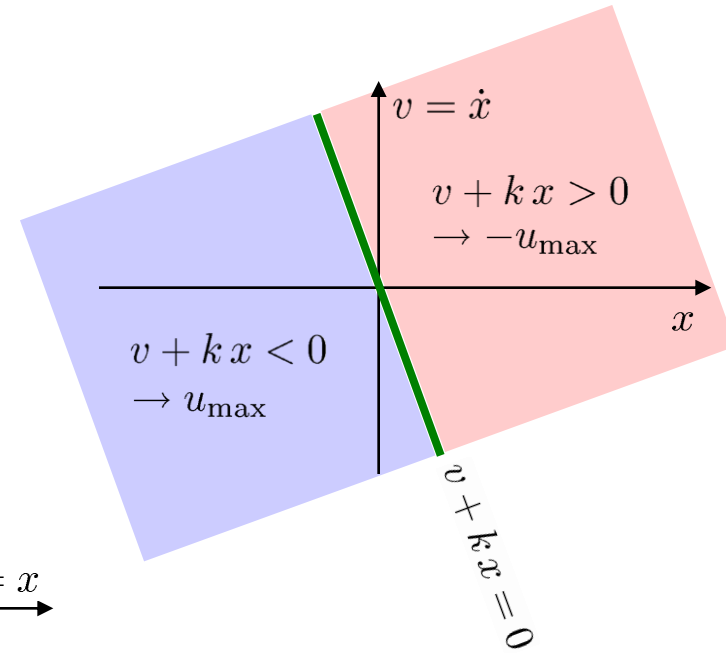
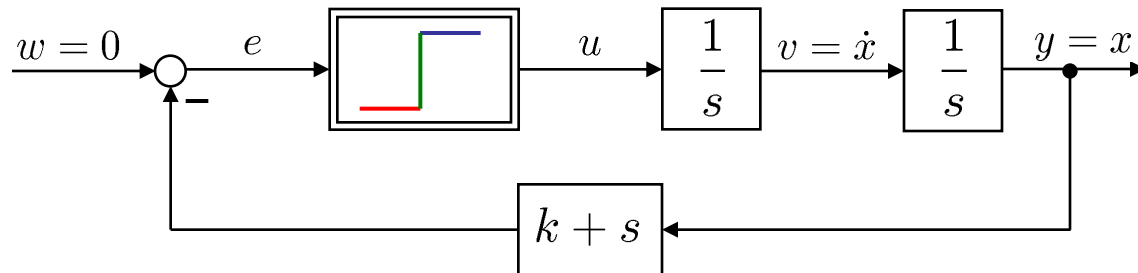
$$e < 0 \rightarrow x > 0: -u_{\max}$$

ändern, indem wir $e = -x$ durch $e = -(v + kx)$ ersetzen.

Damit ändert sich das Blockschaltbild des Regelkreises:



Lässt sich v nicht messen, so kann dieser Zustandsregler auch als PD-Regler realisiert werden:



16.5 Übersicht wichtiger nichtlinearer Regelungsansätze

Zustandsebene und zeitoptimale Regelung

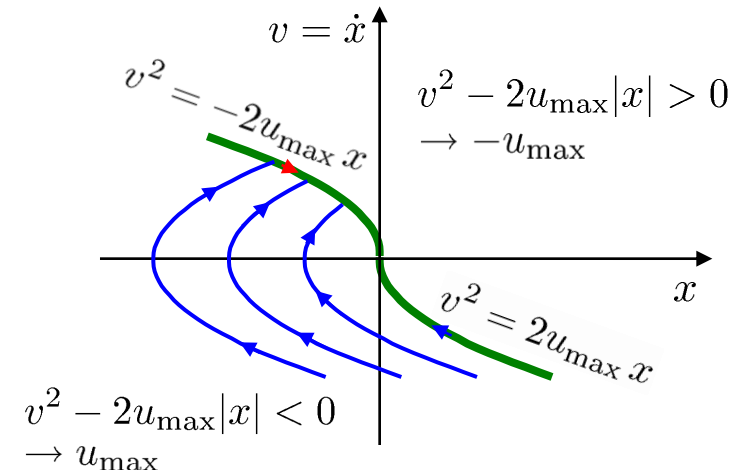
Diese Umschaltstrategie funktioniert sehr gut und ist recht robust. Es kommt gar nicht auf den exakten Verlauf der Trajektorien an, sondern nur auf deren prinzipielle Form.

Wie muss die Umschaltstrategie geändert werden, damit der Systemzustand *möglichst schnell* in den Ursprung gezwungen wird?

Es zeigt sich, dass bei optimaler Gestaltung der Schaltkurve (dann keine Gerade mehr!) ein einziges Mal Umschalten ausreicht, um den Systemzustand in den Ursprung zu regeln, und zwar für jede beliebige Anfangsbedingung. Dies ist möglich, wenn man die Trajektorien durch den Ursprung als Schaltkurven definiert:

$$v^2 = \pm 2u_{\max} x \quad \rightarrow \quad v^2 = 2u_{\max}|x|$$

Das ist die **zeitoptimale Regelung**. Generell erhält man eine zeitoptimale Regelung, wenn man zwischen den Stellgrößenextrema (u_{\max} und $-u_{\max}$) (**Pontrijaginsches Maximumsprinzip**) "geschickt" $n-1$ Mal umschaltet, bei einer Systemordnung von n (**Satz von Feldbaum**).



16.5 Übersicht wichtiger nichtlinearer Regelungsansätze

Direkte Methode von Ljapunow

Mit der direkten Methode von Ljapunow wird ein i.a. sehr schwieriges Problem angegangen, nämlich die Stabilitätsuntersuchung nichtlinearer dynamischer Systeme. Die Grundidee ist dabei, eine Funktion zu finden, die z.B. die *Gesamtenergie* des betrachteten Systems repräsentiert. Wenn man dann zeigen kann, dass diese Funktion (Gesamtenergie) ständig abnimmt, dann ist auch die Stabilität des Systems bewiesen. Denn bei ständig abnehmender Funktion (Gesamtenergie) muss das System einer Ruhelage zustreben.

Beispiel: Feder-Dämpfer-Masse-System

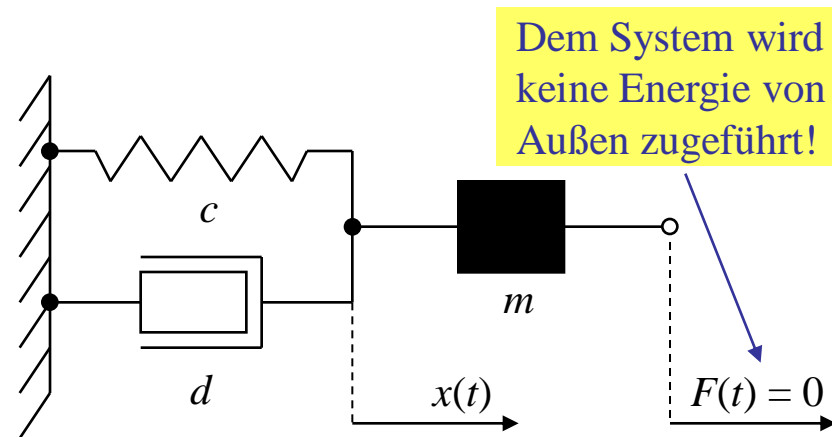
Kinetische Energie: $E_{\text{kin}} = \frac{1}{2} m \dot{x}^2$

Potentielle Energie: $E_{\text{pot}} = \frac{1}{2} c x^2$

Gesamtenergie: $V = E_{\text{kin}} + E_{\text{pot}}$

Modell des Systems: $m\ddot{x} + d\dot{x} + cx = 0$

Mit den Zuständen $x_1 = x$ und $x_2 = \dot{x}$:



Zustands-
gleichungen:

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -\frac{c}{m} x_1 - \frac{d}{m} x_2 \end{aligned}$$

16.5 Übersicht wichtiger nichtlinearer Regelungsansätze

Direkte Methode von Ljapunow

Mit den Zuständen ergibt sich folgende Gleichung für die Gesamtenergie:

$$V(x_1, x_2) = E_{\text{kin}} + E_{\text{pot}} = \frac{1}{2} m x_2^2 + \frac{1}{2} c x_1^2$$

Wie verändert sich die Gesamtenergie über die Zeit?

$$\dot{V}(x_1, x_2) = m x_2 \dot{x}_2 + c x_1 \dot{x}_1 = m x_2 \left(-\frac{c}{m} x_1 - \frac{d}{m} x_2 \right) + c x_1 x_2 = -d x_2^2$$

Das gleiche Ergebnis erhält man auch, wenn man mit einer beliebigen nichtlinearen Federkennlinie statt der Federkonstanten c rechnet.

Die Änderung der Gesamtenergie ist eine **negativ semidefinite** Funktion, d.h. sie ist für alle x_1 und x_2 negativ; erlaubt ist nur die Ausnahme $x_2 = 0$ (und auch $x_1 = 0$), wo $\dot{V} = 0$ sein darf. Das bedeutet, dass dem System so lange Energie entzogen (im Dämpfer in Wärme umgewandelt) wird bis $x_2 = 0$ ist. Dann ist aber auch $\dot{x}_2 = 0$ und damit $x_1 = 0$. D.h. das System strebt unaufhaltsam dem Ursprung des Zustandsraums ($x_1 = 0, x_2 = 0$) zu. Da dies für jeden beliebigen Anfangszustand ($x_1(0), x_2(0)$) gilt, ist das System **global asymptotisch stabil**! Diese Argumentation funktioniert nur für Systeme mit Dämpfer, also $d > 0$! Bei $d = 0$ verändert sich die Gesamtenergie nicht und das System oszilliert für immer.

16.5 Übersicht wichtiger nichtlinearer Regelungsansätze

Direkte Methode von Ljapunow

Diese Idee mit der Gesamtenergie kann auf beliebige **positiv definite** Funktionen (müssen überall positiv sein) erweitert werden. Eine *positiv definite Funktion* deren *zeitliche Ableitung negativ semidefinit* ist (und noch einige weitere Bedingungen erfüllt) nennt man **Ljapunow-Funktion**. Die Gesamtenergie ist nicht immer eine gute (oder auch nur mögliche) Wahl für eine Ljapunow-Funktion, weil:

- Die Gesamtenergie kann zeitlich schwingen (also zwischendrin auch zunehmen) aber dennoch im Mittel abnehmen. Dann ist die Gesamtenergie keine Ljapunow-Funktion.
- Es kann sehr aufwändig sein, die Gesamtenergie eines komplexen Systems zu berechnen.

Vereinfachend kann man 2 Situationen unterscheiden:

1. Eine Ljapunow-Funktion existiert im gesamten Zustandsraum, also für alle Werte von x_1, x_2, \dots . Dann ist das System *global asymptotisch stabil*.
2. Eine Ljapunow-Funktion existiert nur in einem Teilraum des gesamten Zustandsraums. Das kann passieren, weil z.B. die positive Definitheit der Funktion oder die negative Semidefinitheit von deren Ableitung außerhalb verloren geht. Dann gehört dieser Teilraum zum Einzugsgebiet der Ruhelage. (Das Stabilitätsgebiet kann aber größer sein!)

16.5 Übersicht wichtiger nichtlinearer Regelungsansätze

Direkte Methode von Ljapunow

Das bisher Gesagte war nur ungefähr richtig. Feinheiten wurden der Einfachheit halber weggelassen. Mathematisch exakt lautet das **Stabilitätskriterium nach Ljapunow** wie folgt:

Das dynamische System $\dot{x}_i = f_i(x_1, x_2, \dots, x_n)$, $i = 1, 2, \dots, n$, bzw. $\dot{\underline{x}} = \underline{f}(\underline{x})$ besitze die Ruhelage $\underline{x} = \underline{0}$. In einem Gebiet S des Zustandsraums sei $V(\underline{x})$ stetig differenzierbar, und es gelte dort:

- (i) V ist positiv definit.
- (ii) \dot{V} ist negativ semidefinit.
- (iii) Die Punktmenge, auf der $\dot{V} = 0$ ist, enthält außer $\underline{x} = \underline{0}$ keine Trajektorie.

B sei ein Gebiet aus S , in dem $V(\underline{x}) < c$ gilt und dessen Rand durch $V(\underline{x}) = c$ gebildet wird, wobei c eine positive Konstante darstellt. B sei beschränkt und enthalte die Ruhelage $\underline{x} = \underline{0}$.

Dann ist die Ruhelage $\underline{x} = \underline{0}$ *asymptotisch stabil*, und B gehört zu ihrem *Einzugsgebiet*.

Gelten die Bedingungen (i), (ii) und (iii) im gesamten Zustandsraum und strebt zusätzlich $V(\underline{x}) \rightarrow +\infty$ für $|\underline{x}| \rightarrow +\infty$, so ist die Ruhelage *global asymptotisch stabil*.

16.5 Übersicht wichtiger nichtlinearer Regelungsansätze

Direkte Methode von Ljapunow

Die Anwendung des Ljapunow-Stabilitätskriteriums auf *lineare Systeme* liefert einen neuen Stabilitätstest als Alternative zur Berechnung der Eigenwerte der Systemmatrix \underline{A} .

Untersuchen wir die Stabilität der linearen Zustandsgleichungen (externe Signale haben keinen Einfluss auf die Stabilität; wir können daher $u = 0$ setzen)

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

Üblicherweise setzt man hierfür eine quadratische Ljapunow-Funktion an:

$$V(\underline{x}) = \underline{x}^T(t) \underline{P} \underline{x}(t)$$

mit einer positiv definiten konstanten Matrix \underline{P} . Die zeitliche Ableitung von $V(\underline{x})$ ist:

$$\dot{V}(\underline{x}) = \underline{x}^T(t) \underline{P} \dot{\underline{x}}(t) + \dot{\underline{x}}^T(t) \underline{P} \underline{x}(t)$$

Setzt man nun die Zustandsgleichungen ein, erhält man:

$$\dot{V}(\underline{x}) = \underline{x}^T(t) \underline{P} \underline{A} \underline{x}(t) + \underline{x}^T(t) \underline{A}^T \underline{P} \underline{x}(t) = \underline{x}^T(t) (\underline{P} \underline{A} + \underline{A}^T \underline{P}) \underline{x}(t) < 0$$

Daraus folgt, dass für Stabilität die **Ljapunow-Gleichung** mit einer positiv definiten Matrix \underline{Q} erfüllt sein muss:

$$\underline{P} \underline{A} + \underline{A}^T \underline{P} = -\underline{Q}$$

16.5 Übersicht wichtiger nichtlinearer Regelungsansätze

Neuronale Netze und Fuzzy-Systeme

Oft ist es nicht möglich die nichtlinearen DGLs, die einen Prozess beschreiben, oder die einen Regler beschreiben sollen, geschlossen herzuleiten. Dann sind Ansätze gefragt, mit denen nichtlineare Zusammenhänge näherungsweise beschrieben werden können. Die wichtigsten sind:

- Kennlinien und Kennfelder.
- Polynome und Splines.
- Neuronale Netze.
- Fuzzy-Systeme.

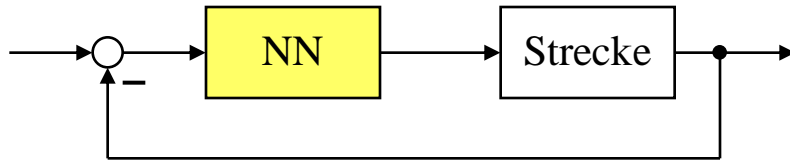
Kennlinien und Kennfelder sind die einfachste und weit verbreitetste Weise, nichtlineare Zusammenhänge zu beschreiben. Messwerte können direkt als Stützstellen gespeichert werden. Polynome und Splines bieten sich an, wenn der nichtlineare Zusammenhang möglichst kompakt, d.h. mit wenigen Parametern (Polynomkoeffizienten), beschrieben werden soll. Wenn hochdimensionale Probleme (d.h. viele Eingänge) vorliegen, sind besonders gut neuronale Netze geeignet. Und wenn man nichtlineare Zusammenhänge mit Hilfe von Regeln qualitativ beschreiben möchte, kommen Fuzzy-Systeme ins Spiel.

16.5 Übersicht wichtiger nichtlinearer Regelungsansätze

Neuronale Netze und Fuzzy-Systeme

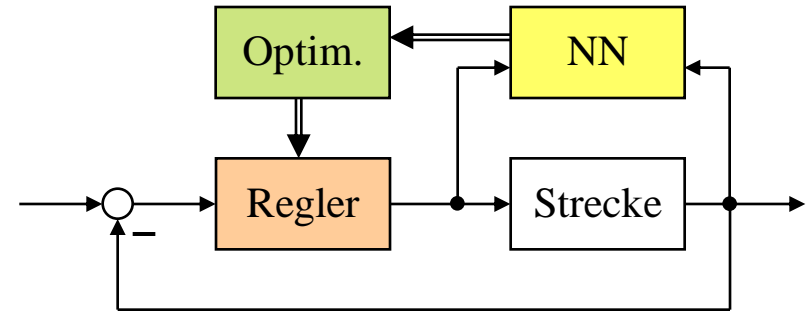
Grundsätzlich muss in der Regelungstechnik zwischen 2 Anwendungsklassen für neuronale Netze und Fuzzy-Systeme unterschieden werden:

Direkter Einsatz als Regler



- Neuronales Netz oder Fuzzy-System wird direkt als Regler eingesetzt.
- Besonders für Fuzzy-Systeme beliebt, da man das gewünschte Verhalten des Reglers in Wenn-Dann-Regeln vorgeben kann.
- Einfach und schnell für erste "brauchbare" Lösung, aber sehr schwer zu optimieren.

Indirekter Einsatz als Modell



- Neuronales Netz oder Fuzzy-System wird als Modell der Regelstrecke eingesetzt.
- Standardverfahren der modellbasierten Regelung können dann eingesetzt werden.
- Reglertyp kann frei gewählt werden: von einfachem PID bis zu prädiktiver Regelung.

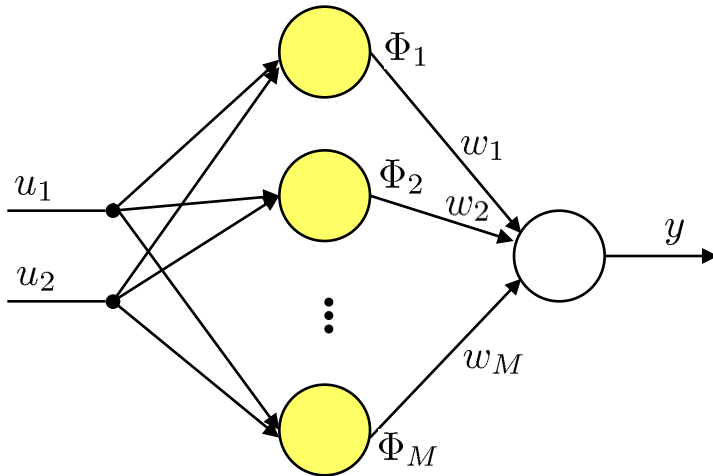
16.5 Übersicht wichtiger nichtlinearer Regelungsansätze

Neuronale Netze und Fuzzy-Systeme

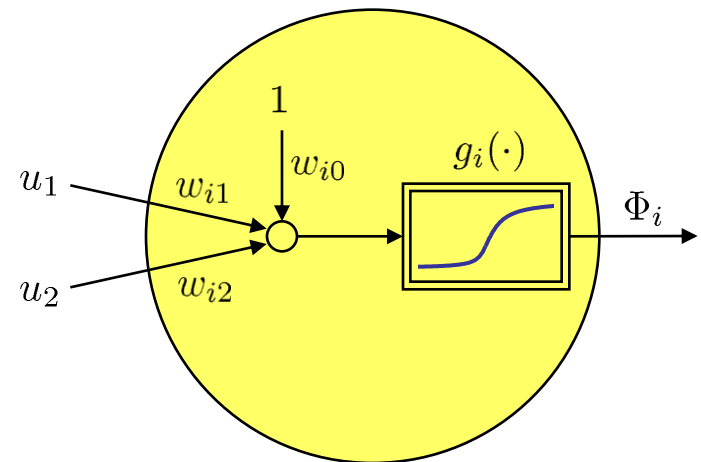
Viele Modellarchitekturen lassen sich als eine Summe von M Basisfunktionen Φ_i darstellen. Bei neuronalen Netzen sind diese Basisfunktionen alle vom gleichen Typ. Der bekannteste neuronale Netz Typ heißt Multilayer-Perzeptron (MLP) und verwendet eindimensionale sigmoide Basisfunktionen. Die vielen Eingänge eines Neurons werden durch ein Skalarprodukt mit einem Parametervektor zu einer skalaren Größe umgewandelt.

$$y = \sum_{i=0}^M w_i \Phi_i(\underline{u}) \quad \Phi_i(\underline{u}) = g_i \left(\sum_{j=0}^n w_{ij} u_j \right) \quad \text{mit } u_0 = 1 \text{ und } \underline{u} = (u_1 \ u_2 \ \dots \ u_n)^T$$

Neuronales Netz



i -tes Neuron



16.5 Übersicht wichtiger nichtlinearer Regelungsansätze

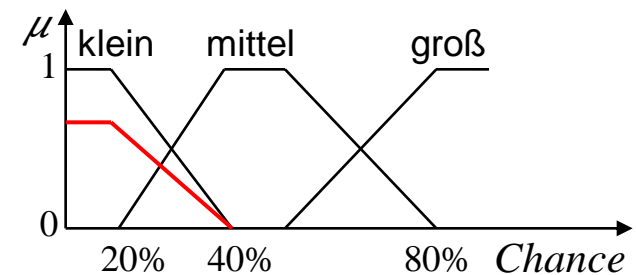
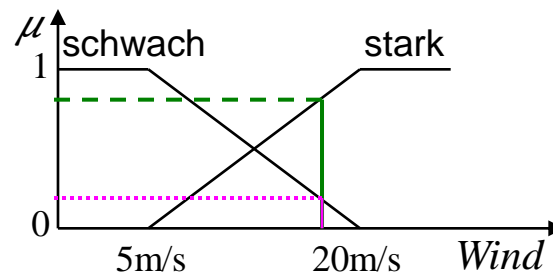
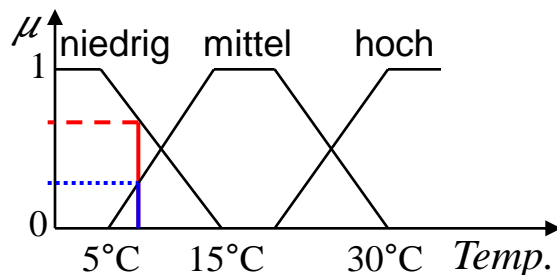
Neuronale Netze und Fuzzy-Systeme

Die *unscharfe Logik* (**Fuzzy-Logik**) erlaubt die Beschreibung nichtlinearer Zusammenhänge zwischen Ein- und Ausgangsgrößen mit Hilfe von Regeln. Dies kann nützlich sein, um Expertenwissen möglichst direkt umzusetzen bzw. um aus Daten gewonnene Modelle auf Stimmigkeit zu überprüfen.

Z.B. kann der sog. Wind-Chill-Effekt durch eine solche Regel beschrieben werden:

WENN *Temperatur* = niedrig UND *Wind* = stark DANN *Überlebenschance* = klein

Um solche Regeln auswerten zu können, müssen zunächst für die beiden Eingangsgrößen *Temperatur* und *Wind* und die Ausgangsgröße *Überlebenschance* **Zugehörigkeitsfunktionen** definiert werden, welche die genaue Bedeutung der Fuzzy-Mengen bestimmen. Dann können die Regeln einzeln ausgewertet werden und anschließend deren Ergebnisse zum Ausgangswert zusammengefasst werden.



17. *Digitale Regelung

Inhalt Kapitel 17

17. *Digitale Regelung

17.1 *Zeitdiskrete Systeme

17.2 *Differenzgleichungen

17.3 *Digitaler Regelkreis

17.4 *Digitaler PID-Regler

17.5 *Faltungssumme

17.6 *Abtasttheorem und Aliasing

17.1 *Zeitdiskrete Systeme

Digitale Regelung

Mit dem Begriff "**Digitale Regelung**" bezeichnet man die Regelung zeitabgetasteter Systeme. Synonym werden auch die Begriffe **zeitdiskrete Regelung** oder **Abtastregelung** verwendet.

Hauptvorteile der digitalen gegenüber der analogen Regelung

- Realisierung in Software statt Hardware ist kostengünstiger und flexibler.
- Größerer Funktionsumfang kann realisiert werden (komplexere Regelalgorithmen, Identifikation, Adaption, Überwachung, Diagnose, Lernfähigkeiten, ...).

Aus diesen Gründen sind heute fast nur noch digitale Regler im Einsatz. Die Behandlung digitaler Regelsysteme unterscheidet sich von den analogen wie folgt:

- Differenzgleichungen statt Differentialgleichungen. Summen statt Integralen. Differenzen statt Ableitungen.
- Z-Transformation statt Laplace-Transformation.
- Andere Stabilitätskriterien. Einfachere Behandlung von Totzeiten. Einfachere Simulation.
- Ganz neue Möglichkeiten (Deadbeat-Regler) und Einschränkungen (Abtasttheorem).

17.1 *Zeitdiskrete Systeme

Digitale oder analoge Regelung?

Beispiel: Stromregler für elektrische angesteuerte Ventile in einem elektronischen Motor- oder Getriebesteuergerät für Pkws (2005)

Analoge Realisierung

- Regler wird mit Operationsverstärkern, Widerständen und Kondensatoren als *Hardware* auf der Platine des Steuergeräts realisiert.
- Baueile und Platz auf der Platine kosten eine bestimmte Summe pro gebautem Steuergerät. → *Stückzahlabhängige* Kosten.
- Hohe Regelgüte, da kein Informationsverlust durch Abtastung.
- Für beliebig *schnelle Prozesse* geeignet (keine Rechenleistung notwendig).
- Nur für *einfache Reglerstrukturen* (wie PID) sinnvoll.
- Reglerstruktur und Reglerparameter können nach Abschluss des Hardware-Designs nicht mehr verändert werden. → *Frühzeitige Festlegung*, wenig zeitliche Flexibilität.
- *Kein Update* ohne Austausch (oder Änderung) der Hardware möglich.

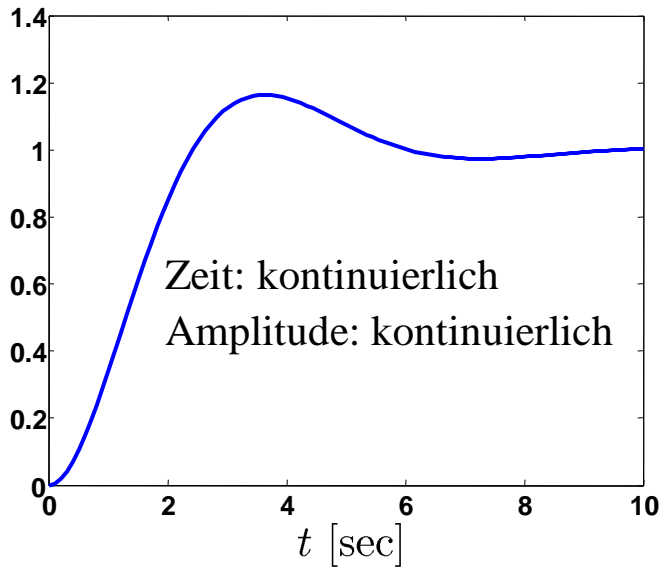
17.1 *Zeitdiskrete Systeme

Digitale Realisierung

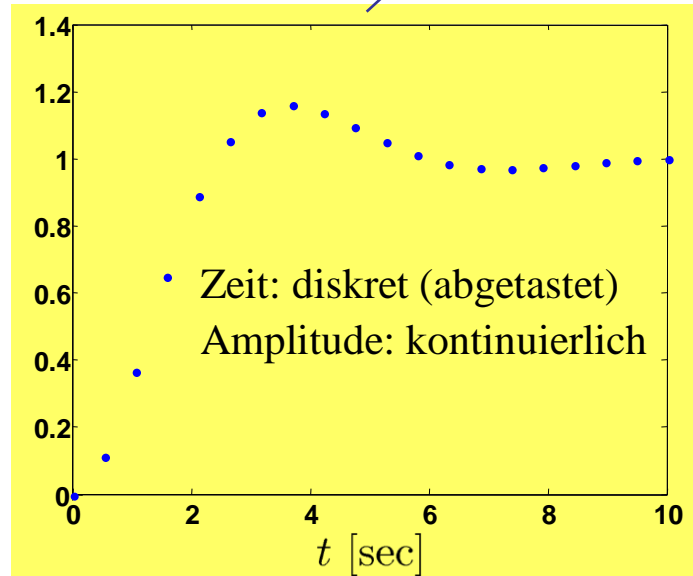
- Regler wird in *Software* realisiert; typischerweise in *C* programmiert.
- Einmalig fällt ein Entwicklungsaufwand an; Hardwareaufwand entsteht nicht (vorausgesetzt es ist noch genügend Rechenleistung des Mikrocontrollers frei). Entwicklungsaufwand kann evtl. auf mehrere Projekte umgelegt werden → *Fixkosten*.
- Benötigte *Rechenleistung steigt linear mit der Abtastfrequenz* an.
- Beliebige *komplexe Reglerstrukturen* können leicht realisiert werden. Auch beliebige Nichtlinearitäten lassen sich umsetzen.
- *Reglerstruktur* lässt sich durch *Umprogrammieren* ändern.
- *Reglerparameter* lassen sich durch *Kalibration*, d.h. auch nach Freigabe der Hard- und Software des Steuergerätes ganz kurzfristig (auch noch in der Fertigung) ändern.
→ Sehr große zeitliche Flexibilität (Änderung am Ende erhöhen allerdings das Risiko).
- *Updates* lassen sich *leicht durchführen*, sowohl während der Fertigung ab einer bestimmten Seriennummer als auch (im Notfall) als Rückruf. → Geringe Kosten.
- *Verschiedene Varianten* für verschiedene Baureihen lassen sich leicht erzeugen.

17.1 *Zeitdiskrete Systeme

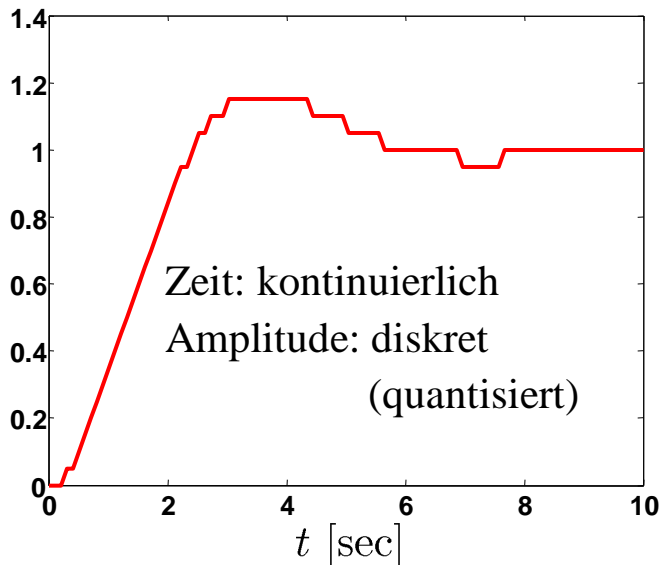
Digitale Regelung



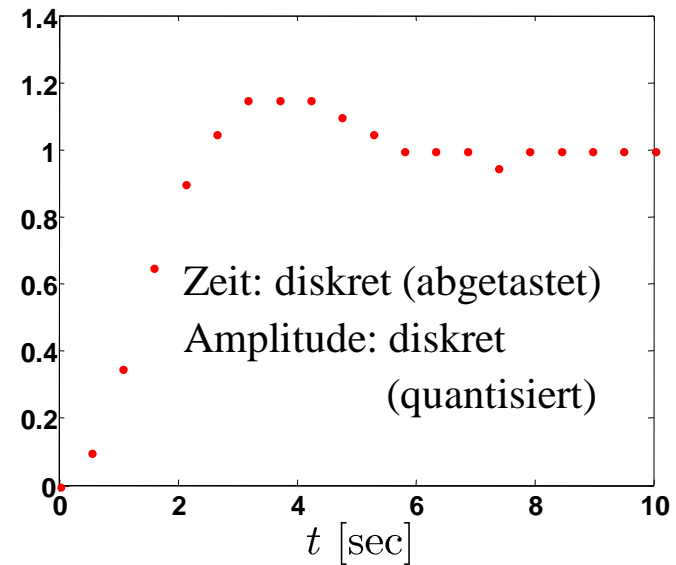
Abtastung



Quantisierung



Abtastung



Quantisierung

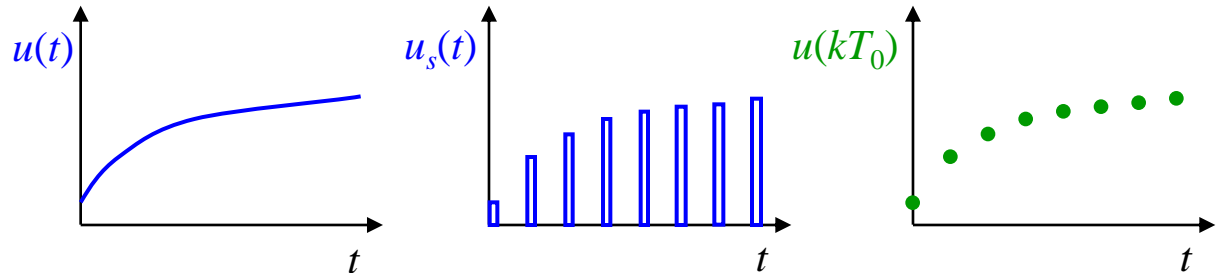
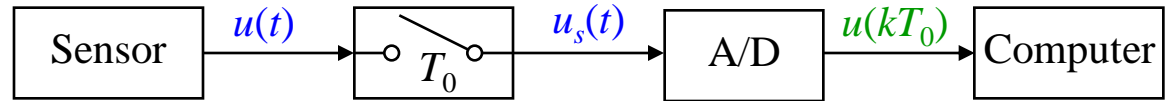
17.1 *Zeitdiskrete Systeme

Abkürzung: $u(k) = u(kT_0)$
 $y(k) = y(kT_0)$

Analog/Digital- und Digital/Analog-Wandlung

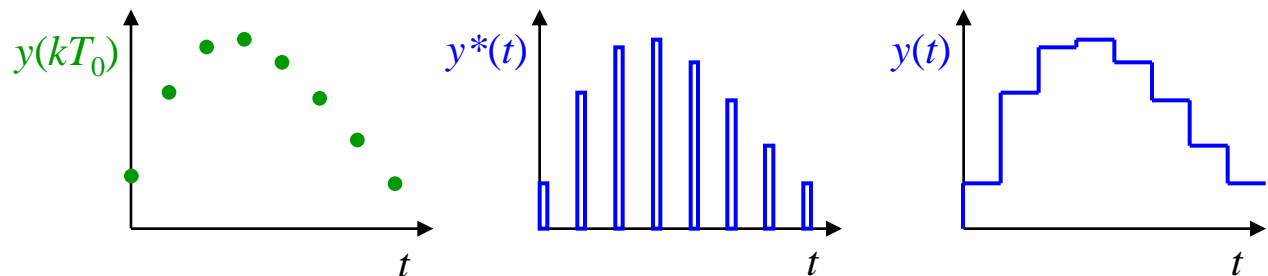
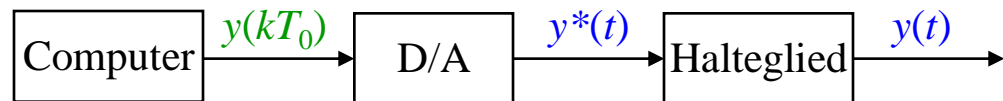
A/D-Wandlung

- Abtastzeit T_0 liegt zwischen μsec (Signalverarbeitung) und Stunden (thermisch, biologische Prozesse)
- Amplitudenauflösung von 8, 12 oder 16 Bit.



D/A-Wandlung

- Computer rechnet mit zeitdiskreten Folgen.
- Halteglied 0. Ordnung erzeugt stückweise konstante Signale.



17.2 *Differenzgleichungen

Beziehung zwischen Differential- und Differenzgleichungen

Der Computer verarbeitet **Signalfolgen**, die *zeitsynchron* mit der Abtastzeit T_0 abgetastet werden. Aus der Eingangsfolge $u(k)$ berechnet der Computer die Ausgangsfolge $y(k)$.

$$y(k) + a_1y(k - 1) + \dots + a_ny(k - n) = b_0u(k) + b_1u(k - 1) + \dots + b_mu(k - m)$$

Beispiele für die zeitkontinuierliche und zeitdiskrete Beschreibung dynamischer Systeme:

zeitkontinuierlich

zeitdiskret

P $y(t) = Ku(t)$

$y(k) = Ku(k)$

PT₁ $y(t) + a_1\dot{y}(t) = b_0u(t)$

$y(k) + \tilde{a}_1y(k - 1) = \tilde{b}_1u(k - 1)$

PT₂ $y(t) + a_1\dot{y}(t) + a_2\ddot{y}(t) = b_0u(t)$

$y(k) + \tilde{a}_1y(k - 1) + \tilde{a}_2y(k - 2) = \tilde{b}_1u(k - 1)$

D $y(t) = K_D\dot{u}(t)$

$y(k) = \tilde{b}_0(u(k) - u(k - 1))$

I $y(t) = K_I \int_0^t u(\tau)d\tau$

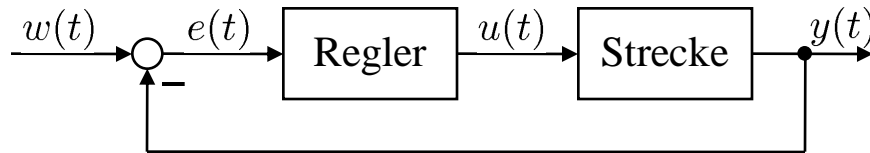
$y(k) = y(k - 1) + \tilde{b}_1u(k - 1)$

nicht
sprungfähig!

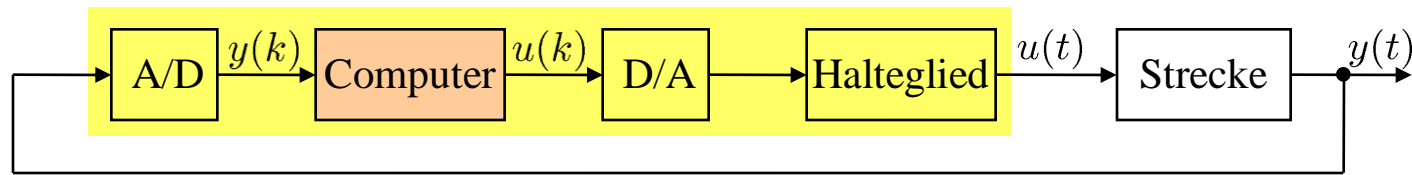
17.3 Digitaler Regelkreis

Aufbau eines digitalen Regelkreises

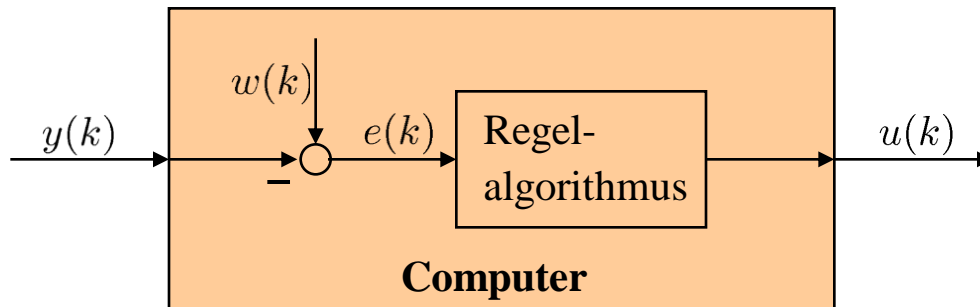
Analoger Regelkreis



Digitaler Regelkreis



Die Führungsgröße wird normalerweise im Computer direkt erzeugt und dort wird auch der Soll/Istwert-Vergleich durchgeführt:



17.3 *Digitaler Regelkreis

Beispiel: P-Regler mit PT₁-Glied

Zeitkontinuierlicher Fall:

P-Regler: $u(t) = K_R e(t)$

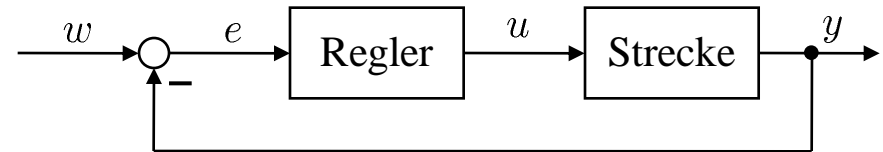
PT₁-Strecke: $T\dot{y}(t) + y(t) = u(t)$

Geschlossener Regelkreis: $T\dot{y}(t) + y(t) = K_R[w(t) - y(t)]$

$$\rightarrow T\dot{y}(t) + (K_R + 1)y(t) = K_R w(t) \quad \rightarrow \quad \boxed{\frac{T}{K_R + 1}\dot{y}(t) + y(t) = \frac{K_R}{K_R + 1}w(t)}$$

Der geschlossene Regelkreis hat also auch PT₁-Verhalten. Und zwar mit einer Zeitkonstanten von $T/(K_R+1)$, die umso kleiner wird, je größer die Reglerverstärkung ist und einer Verstärkung von $K_R/(K_R+1)$, die umso weniger von 1 abweicht (also eine umso kleinere bleibende Regelabweichung aufweist), je größer K_R ist.

Der geschlossene Regelkreis ist *strukturstabil*, d.h. K_R kann gegen unendlich gehen und die Stabilität ist (theoretisch) nicht gefährdet. (Mit Hilfe der Laplace-Transformation hätten wir das schneller herleiten können. Aber da wir die z-Transformation als deren zeitdiskretes Äquivalent noch nicht kennen, beschränken wir uns hier auf den Zeitbereich.)



17.3 *Digitaler Regelkreis

Beispiel: P-Regler mit PT₁-Glied

Zeitdiskreter Fall:

$$\text{P-Regler: } u(k) = K_R e(k)$$

$$\text{PT}_1\text{-Strecke: } y(k) + a_1 y(k-1) = b_1 u(k-1)$$

$$\text{Geschlossener Regelkreis: } y(k) + a_1 y(k-1) = b_1 K_R [w(k-1) - y(k-1)]$$

$$\rightarrow \boxed{y(k) = -(a_1 + b_1 K_R) y(k-1) + b_1 K_R w(k-1)}$$

$= \tilde{a}_1$. Muss betragsmäßig < 1 sein, damit Stabilität gesichert ist!

$$\text{Für } w = 0 \text{ gilt: } y(k) = \tilde{a}_1 y(k-1)$$

$$y(1) = \tilde{a}_1 y(0)$$

$$y(2) = \tilde{a}_1 y(1) = \tilde{a}_1^2 y(0)$$

\vdots

$$y(k) = \tilde{a}_1^k y(0)$$

Dieser Ausdruck strebt nur gegen 0, wenn $|\tilde{a}_1| < 1$ ist.

$$a_1 = -e^{-\frac{T_0}{T}}$$

$$b_1 = 1 - e^{-\frac{T_0}{T}}$$

Der geschlossene Regelkreis ist stabil für:

$$-\frac{1+a_1}{b_1} < K_R < \frac{1-a_1}{b_1}$$

Im Gegensatz zum analogen Regelkreis, ist der digitale Regelkreis *nicht strukturstabil!*

17.4 *Digitaler PID-Regler

$$\text{PID: } u(t) = K_P \left(e(t) + \frac{1}{T_I} \int_0^t e(\tau) d\tau + T_D \frac{d}{dt} e(t) \right)$$

Zeitdiskretisierung des PID-Reglers

Es gibt viele Möglichkeiten, ein zeitkontinuierliches System in ein zeitdiskretes System umzuwandeln. Am einfachsten ist die Approximation mittels Differenzenquotienten. Eine solche Approximation hat aber den schwerwiegenden Nachteil, dass sie nur für sehr kleine Abtastzeiten eine gute Näherung darstellt:

Approximation mittels Differenzenquotienten (nur für sehr kleine T_0)

$$\text{P-Anteil: } u(k) = K_P e(k)$$

$$\text{D-Anteil: } u(k) = T_D \frac{e(k) - e(k-1)}{T_0}$$

$$\text{I-Anteil: } u(k) = u(k-1) + \frac{T_0}{T_I} e(k-1)$$

$$u(k) = \frac{T_0}{T_I} \sum_{i=1}^k e(i-1)$$

$$u(k-1) = \frac{T_0}{T_I} \sum_{i=1}^{k-1} e(i-1)$$

$$\text{PID-Regler: } u(k) = u(k-1) + K_P \left(1 + \frac{T_D}{T_0} \right) e(k) + K_P \left(\frac{T_0}{T_I} - \frac{T_D}{T_0} \right) e(k-1)$$

17.4 *Digitaler PID-Regler

Übliche Variante des digitalen PID-Reglers

Die obigen Näherungen werden sehr schlecht, wenn die Abtastzeit nicht sehr klein gegenüber den relevanten Prozesszeitkonstanten gewählt wird. Darüber hinaus ist nicht gewährleistet, dass stabile Prozesse nach dieser Approximation auch im Zeitdiskreten stabil sind. Eine wesentlich bessere Transformation führt auf folgenden Regler:

Digitaler PID-Regler: $u(k) = u(k-1) + q_0 e(k) + q_1 e(k-1) + q_2 e(k-2)$

$$q_0 = K_P \left(1 + \frac{T_0}{2T_I} + \frac{T_D}{T_0} \right) \quad q_1 = K_P \left(-1 + \frac{T_0}{2T_I} - \frac{2T_D}{T_0} \right) \quad q_2 = K_P \frac{T_D}{T_0}$$

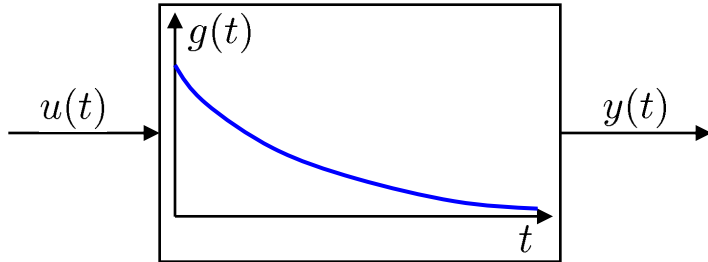
Die 3 Parameter des digitalen PID-Reglers q_0 , q_1 und q_2 können, wie im obigen Kasten angegeben, zu den Parametern des analogen PID-Reglers in Beziehung gesetzt werden. Das ist wichtig, wenn man einen analogen in einen digitalen Regler umwandeln möchte oder die Interpretation in Form eines P-, I- und D-Anteils benötigt.

Ansonsten können die 3 Parameter q_0 , q_1 und q_2 auch als unabhängige Werte gesehen werden und direkt durch beliebige Reglerentwurfsverfahren bestimmt werden.

17.5 *Faltungssumme

Faltungsintegral und Faltungssumme

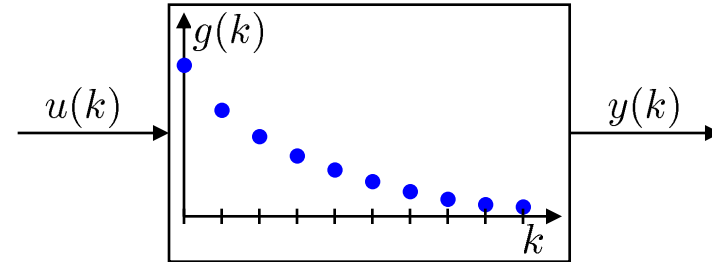
zeitkontinuierlich



$$y(t) = \int_0^t g(t - \tau)u(\tau)d\tau$$
$$= \int_0^t g(\tau)u(t - \tau)d\tau$$

- Die Gewichtsfunktion $g(t)$ enthält alle Informationen über das lineare System.
- Leichter rechnet es sich im Frequenzbereich mit der Laplace-Transformierten $G(s)$.

zeitdiskret



$$y(k) = \sum_{i=0}^k g(k - i)u(i)$$
$$= \sum_{i=0}^k g(i)u(k - i)$$

- Die Gewichtsfunktion $g(k)$ enthält alle Informationen über das lineare System.
- Leichter rechnet es sich im Frequenzbereich mit der z-Transformierten $G(z)$.

17.5 *Faltungssumme

$$y(k) = \sum_{i=0}^k g(k-i)u(i) = \sum_{i=0}^k g(i)u(k-i)$$

Eigenschaften der Faltungssumme

- Ausgeschrieben liest sich die Faltungssumme wie folgt:

$$y(k) = g(k)u(0) + g(k-1)u(1) + \dots + g(1)u(k-1) + g(0)u(k)$$

- Wie im Zeitkontinuierlichen ist die Gewichtsfunktion identisch mit der Systemantwort auf einen (zeitdiskreten) Dirac-Impuls am Eingang $u(k) = \delta(k)$:

$$u(0) = 1, u(1) = 0, u(2) = 0, \dots \rightarrow y(k) = g(k).$$

- Nur für sprungfähige Systeme existiert $g(0)$. Im Normalfall gilt also $g(0) = 0$.
- Hat die Gewichtsfunktion die Länge $l+1$, ist also endlich, d.h. alle $g(i) = 0$ für $i > l$, dann hat das System eine **endliche Impulsantwort (finite impulse response, FIR)**. Man spricht dann auch davon, dass $y(k)$ ein **gleitender Mittelwert (moving average, MA)** ist.
- Ist die Gewichtsfunktion unendlich lang, dann hat das System eine **unendliche Impulsantwort (infinite impulse response, IIR)**. Ähnlich wie im Zeitkontinuierlichen kann dann der Ausgang aber über eine endliche Gleichung aus vergangenen (im Zeitkontinuierlichen: abgeleiteten) Ausgangswerten berechnet werden:

$$y(k) = -a_1y(k-1) - \dots - a_ny(k-n) + b_0u(k) + b_1u(k-1) + \dots + b_mu(k-m)$$

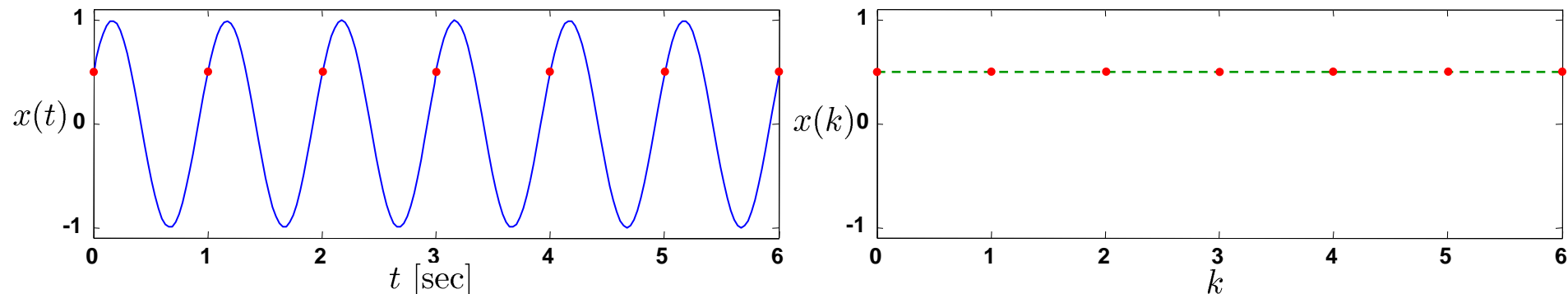
17.6 *Abtasttheorem und Aliasing

Abtastung zeitkontinuierlicher Signal

Jeder hat schon einmal drehende Speichenräder einer losfahrenden Kutsche im Film gesehen. Zuerst sieht man das Rad beschleunigen. Ab einer gewissen Drehzahl scheint das Rad plötzlich seine Drehrichtung zu ändern und schnell rückwärts zu drehen (obwohl die Kutsche weiter beschleunigt). Dann scheint das Rad immer langsamer zu werden bis es sogar stehen bleibt. Das ist ein offensichtlicher Widerspruch zur immer schneller fahrenden Kutsche.

Diese komischen Effekte werden durch das sog. **Aliasing** verursacht und treten bei allen abgetasteten Systemen auf. Offensichtlich gibt es Probleme, wenn wir Signale abtasten, deren Frequenz in den Bereich der Abtastfrequenz kommt. Der *Film* spielt dabei die Rolle des *Abtasters* mit einer Bildwiederholrate bzw. Abtastfrequenz von $f_0 = 25$ Hz.

Was passiert, wenn wir ein Signal der Frequenz $f = 1$ Hz mit $f_0 = 1$ Hz abtasten?

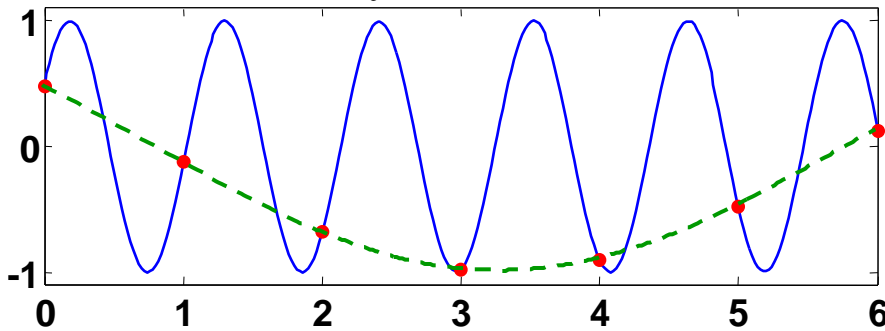


17.6 *Abtasttheorem und Aliasing

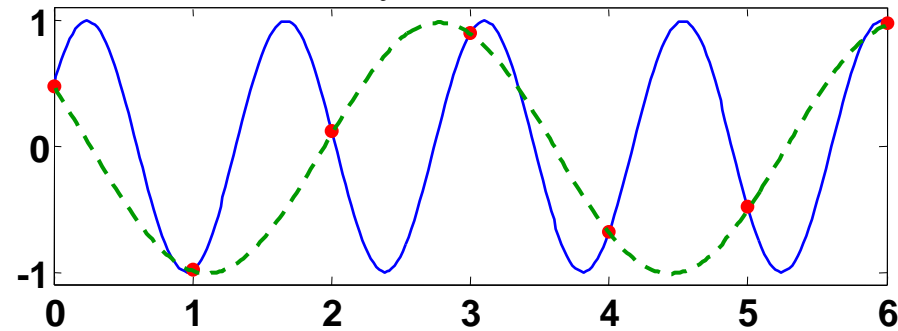
Aliasing

Offensichtlich geht die Schwingung komplett verloren. Wir erhalten ein Signal der Frequenz null (einen Gleichwert). Diese Tatsache ist unabhängig von der Phasenlage des Abtasters (nur die Höhe des Gleichwertes hängt davon ab). Zur Illustration noch ein paar weitere Beispiele mit $f = 0.9$ Hz, 0.7 Hz, 0.5 Hz und 0.3 Hz und jeweils $f_0 = 1$ Hz.

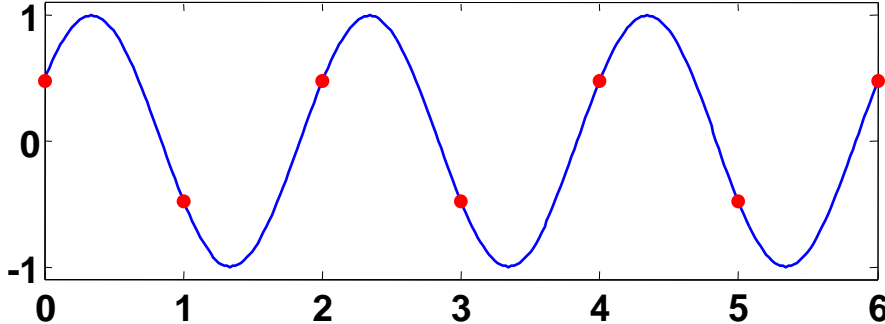
$f = 0.9$ Hz



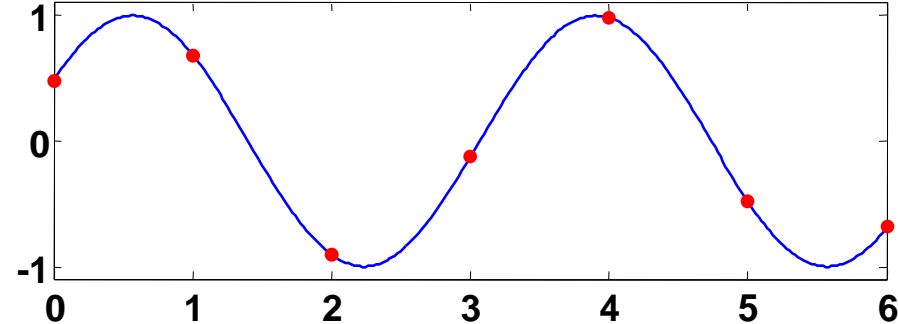
$f = 0.7$ Hz



$f = 0.5$ Hz



$f = 0.3$ Hz



17.6 *Abtasttheorem und Aliasing

Abtasttheorem

Aus den Beispielen der vorangegangenen Seite erkennen wir empirisch, dass wir mindestens mit der doppelten Signalfrequenz abtasten müssen ($f = 0.5 \text{ Hz}$, $f_0 = 1 \text{ Hz}$), um das Signal nach der Abtastung noch richtig wiedergeben zu können. Reale Signale bestehen aus einem Gemisch vieler Frequenzen. Dann bezieht sich die Forderung nach der doppelten Abtastfrequenz auf den Signalanteil mit der höchsten Frequenz f_{\max} .

Shannonsches Abtasttheorem

Das Signal $x(t)$ soll abgetastet werden. Die höchste Frequenzkomponente von $x(t)$ sei f_{\max} . Dann muss die Abtastfrequenz mindestens doppelt so groß gewählt werden, wie die höchste Frequenzkomponente von $x(t)$:

$$f_0 > 2f_{\max}$$

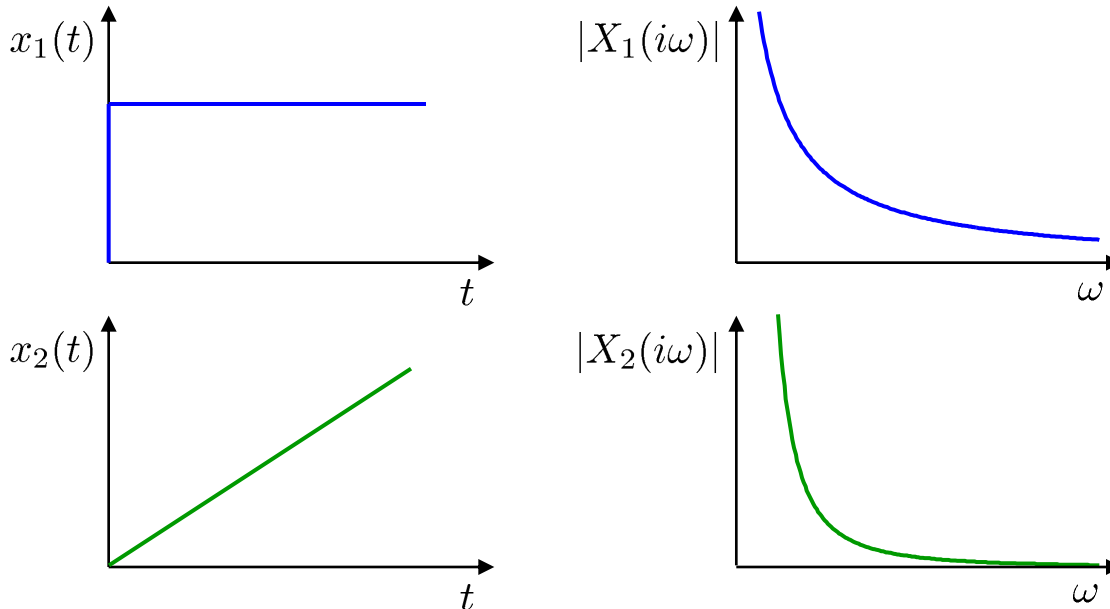
Ist dies nicht gewährleistet, kommt es zu Aliasing, d.h. die Frequenzkomponenten mit $f > \frac{1}{2}f_0$ werden in den Bereich niedrigerer Frequenzen gespiegelt. Auf diese Weise können hochfrequente Störsignale großen Schaden anrichten und niederfrequente Nutzsignale überlagern.

In der Praxis wählt man ca. $f_0 = 5 \dots 10 f_{\max}$

17.6 *Abtasttheorem und Aliasing

Veranschaulichung des Abtasttheorems und des Aliasing-Effekts

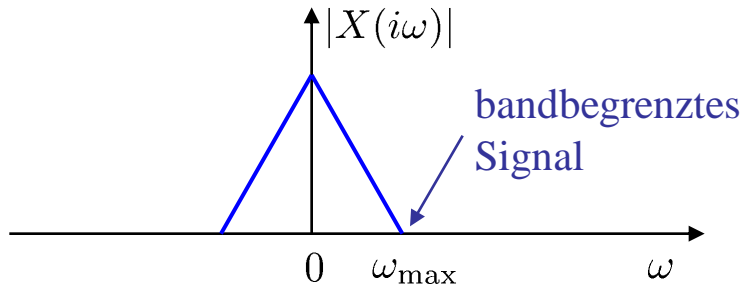
Wenn das Abtasttheorem eingehalten wird, kann das Originalsignal aus dem abgetasteten Signal rekonstruiert werden, d.h. es findet durch die Abtastung kein Informationsverlust statt. In Realität sind aber die meisten Signale nicht bandbegrenzt, haben also gar keine Komponente maximaler Frequenz f_{\max} bzw. $f_{\max} = \infty$. Vielmehr sind in vielen typischen Signalen (Sprünge, Rampen, Rechtecke, ...) alle Frequenzen von 0 bis unendlich enthalten. Solche Signale lassen sich daher aus dem abgetasteten Signal nicht perfekt rekonstruieren.



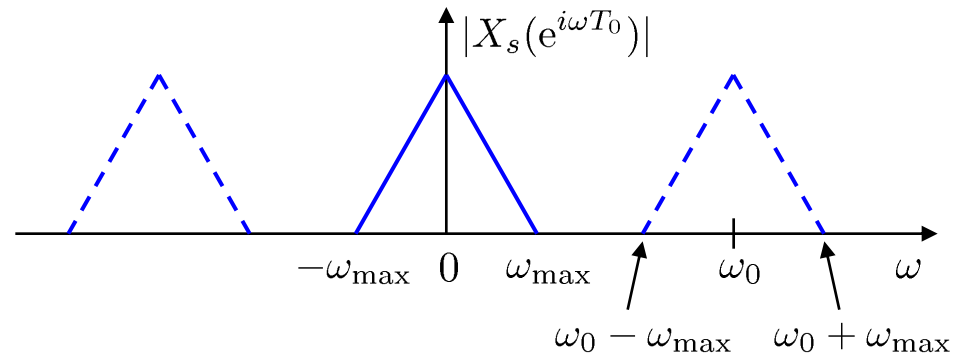
17.6 *Abtasttheorem und Aliasing

Veranschaulichung des Abtasttheorems und des Aliasing-Effekts

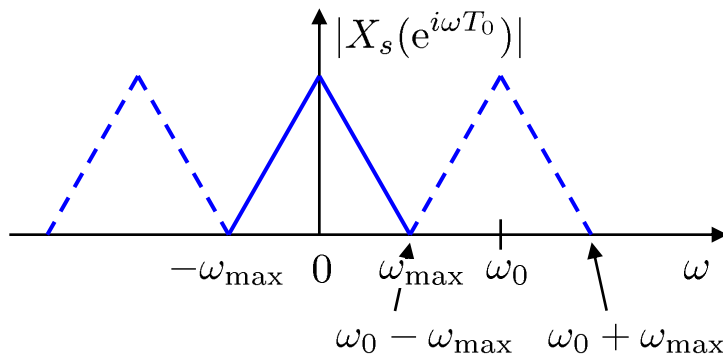
Spektrum des kontinuierlichen Signals



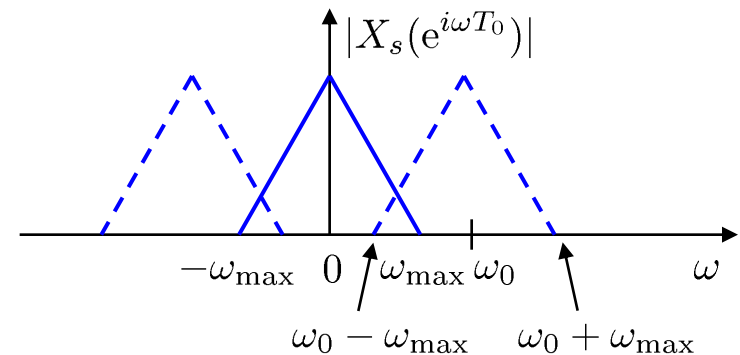
Spektrum des abgetasteten Signals $\omega_0 > 2\omega_{\max}$



Spektrum des abgetasteten Signals $\omega_0 = 2\omega_{\max}$

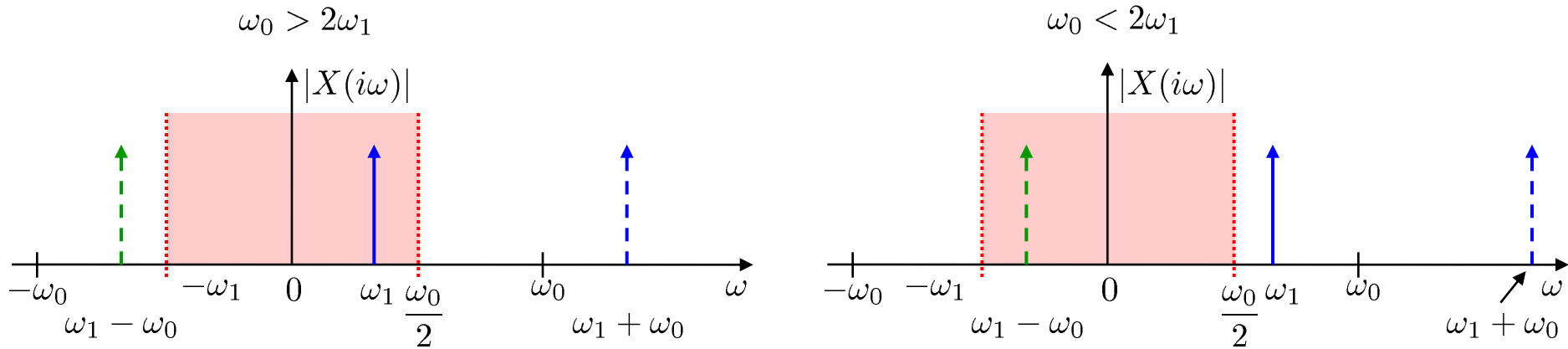


Spektrum des abgetasteten Signals $\omega_0 < 2\omega_{\max}$



17.6 *Abtasttheorem und Aliasing

Aliasing bei der Abtastung einer Sinus-Schwingung mit Kreisfrequenz ω_1



Jede Signalkomponente der Frequenz ω_1 wird durch das Abtasten an folgende Frequenzen gespiegelt:

$$\omega_l = \omega_1 + l\omega_0 \quad \text{mit } l = \dots, -2, -1, 0, 1, 2, \dots$$

Solange ω_1 im roten Bereich bleibt, also das Abtasttheorem nicht verletzt, liegen die gespiegelten Komponenten außerhalb des roten Bereichs (linkes Bild).

Sobald ω_1 den roten Bereich verlässt, wandert eine gespiegelte Komponente hinein und verursacht Aliasing (rechtes Bild)!

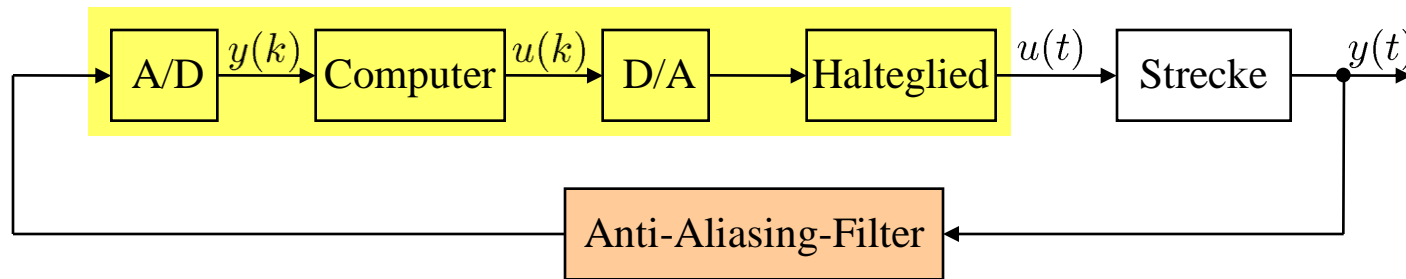
Wandert ω_1 bis ω_0 , so entsteht eine Alias-Komponente bei $\omega = 0$ (stehendes Speichenrad).

17.6 *Abtasttheorem und Aliasing

Anti-Aliasing-Filter

Um Aliasing zu verhindern, muss das analoge Signal *vor* der Abtastung mit einem **Anti-Aliasing-Filter** gefiltert werden. Dieses Filter ist notwendigerweise ein analoges Filter, denn nach dem Abtasten kann das Aliasing nicht mehr rückgängig gemacht werden.

Typischerweise werden sehr steile Filter mit einer Grenzfrequenz bei ω_0 eingesetzt, um die niederfrequenten Signalanteile möglichst wenig zu verzerren aber alle Frequenzen oberhalb von ω_0 sehr stark zu dämpfen.



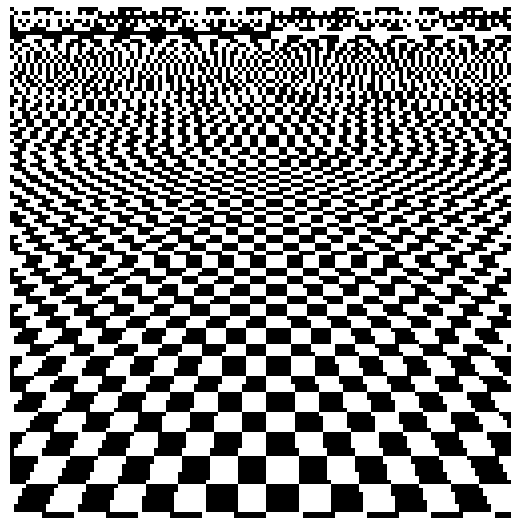
Mit einem Trick kann auch ein digitales Anti-Aliasing-Filter eingesetzt werden. Dazu muss zunächst mit viel höherer als der geplanten Abtastrate abgetastet werden, z.B. mit $8\omega_0$ (8-fach Oversampling). Die Frequenz muss so hoch sein, dass dort keine signifikanten Signalkomponenten mehr existieren. Die gespiegelten Frequenzen liegen dann weit außerhalb des Bereichs $|\omega| < \omega_0/2$. Danach wird digital gefiltert und downgesampelt.

17.6 *Abtasttheorem und Aliasing

Aliasing in der Bildverarbeitung

Die Signalverarbeitung taucht nicht nur im Zusammenhang mit *zeitlichen* Signalen auf. Auch für *räumliche* Signale wie z.B. Bilder (2-dim. räumliche Signale) gelten die selben Gesetze. So wie man zeitliche Signale filtern kann, ist Gleiches auch mit räumlichen Signalen möglich. In der Bildverarbeitung tritt deshalb ebenfalls der Aliasing-Effekt auf. Eine hohe (räumliche) Frequenz entspricht einer dichten Abfolge von abwechselnd hellen und dunklen Punkten. Wenn bei der Reduzierung der Auflösung (entspricht der Abtastung) kein Anti-Aliasing-Filter verwendet wird, spiegeln die hohen Frequenzen in die niederfrequenten Bereiche und stören das Bild. Ähnliches kennt man vom Fernsehen, wenn jemand Kleidung mit kleinem Karomuster trägt. In der Bildverarbeitung nennt man dies **Moiré-Effekt**.

Ohne Anti-Aliasing



Mit Anti-Aliasing

