

1 Systemdefinitionen

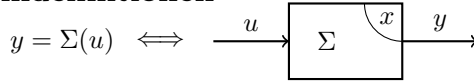


Abb. 1: System Σ mit Eingang u , Ausgang y und Zustandsvariablen x

Signal: Ein Signal ist eine Funktion der Zeit (z.B. $u(t), y(t)$)

System: Ein System ist ein Operator der ein Signal ändert. Σ in Abb. 1 transformiert einen Eingang u in einen Ausgang y . Bem.: Ein System kann interne Zustandsvariablen x haben, welche von aussen nicht zugänglich sind.

2 Systemklassifikation

SISO/MIMO: SISO (Single Input Single Output) sind Systeme mit genau einem Eingang und einem Ausgang. Systeme mit mehreren Eingängen oder Ausgängen heissen MIMO (Multiple Input Multiple Output).

Linear/Nichtlinear: Ein System Σ heisst linear, falls gilt:
 $\Sigma(\alpha \cdot u_1 + \beta \cdot u_2) = \alpha \cdot \Sigma(u_1) + \beta \cdot \Sigma(u_2)$

Linear

$$y(t) = \frac{d}{dt}u(t)$$

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

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

Nichtlinear

$$y(t) = \alpha \cdot u(t) + \beta$$

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

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

Eine Änderung des Ausgangs bei einem **Linearen** System ist proportional zur Änderung des Eingangs.

Kausal/Akausal: Ein kausales System hängt nicht von Eingängen in der Zukunft ab.

Kausal

$$y(t) = u(t - \tau) \quad \forall \tau \geq 0$$

$$y(t) = \int_{-\infty}^t u(\tau) d\tau$$

Akausal

$$y(t) = u(t + 5)$$

$$y(t) = \int_{-\infty}^{t+1} u(\tau) d\tau$$

Alle physikalischen Systeme sind kausal. Solche Systeme können **Dynamisch** oder **Statisch** sein

Dynamisch

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

$$y(t) = u(t - \tau) \quad \forall \tau \neq 0$$

Statisch

$$y(t) = 3 \cdot u(t)$$

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

Der Ausgang eines **statischen** Systems zur Zeit t^* hängt nur vom Eingang zur Zeit t^* ab.

Zeitinvariant/Zeitvariant: Zeitvariante Systeme geben bei gleichen Eingängen zu unterschiedlichen Zeitpunkten und gleicher Anfangsbedingung unterschiedliche Ausgänge.

Zeitinvariant

$$y(t) = 3 \cdot u(t)$$

$$y(t) = \frac{d}{dt}u(t)$$

Zeitvariant

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

$$y(t) = u(t) + t$$

3 Steuerung vs. Regelung vs. Vorsteuerung

In diesem Abschnitt wird der Unterschied zwischen Steuerung und Regelung mit Hilfe von Beispielen illustriert.

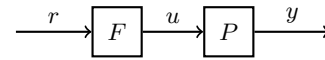


Abb. 2: Steuerungsstruktur

Betrachte Abb. 2. Das System stellt sich zusammen aus einer Steuerung F und einer Strecke P . Ein- und Ausgänge in die Systemblöcke sind Führungsgrösse r , Eingangsgrösse u und Ausgangsgrösse y . Bei einem Steuersystem wird die Steuergrösse nicht mit der Ausgangsgrösse verglichen. Diese Struktur hat ein Freiheitsgrad: Die Steuerung F .

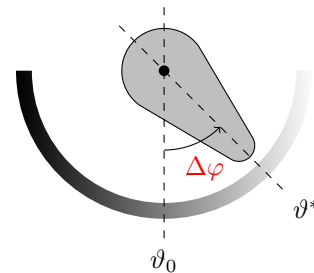


Abb. 3: Temperaturhebel deiner Dusche.

Beispiel: Betrachte Abb. 3. Du hast schon oft in deiner Dusche geduscht und ein affines Verhältnis zwischen dem Winkel des Temperaturhebels und der Ausgangstemperatur des Wassers identifiziert. Das heisst, eine Temperaturänderung ist Proportional zur Winkeländerung des Hebels: $\vartheta = \nu(\varphi - \varphi_0) + \vartheta_0 = P(\varphi)$, wobei ν die Proportionalitätskonstante und ϑ_0 die Referenztemperatur beim Referenzwinkel φ_0 ist. Dieses Modell ist die Steuerstrecke P . Du möchtest, dass die Ausgangsgrösse $y = \vartheta$ deiner Lieblingstemperatur $r = \vartheta^*$ entspricht. Da du die Proportionalitätskonstante kennst, kannst du die Steuerstrecke invertieren: $\varphi = \frac{1}{\nu}(\vartheta - \vartheta_0) + \varphi_0$ und erhältst somit die Steuerung $F = P^{-1}$. Setzt man die gewünschte Temperatur ϑ^* in die Steuerung, erhält man die Eingangsgrösse $u = \varphi^* = \frac{1}{\nu}(\vartheta^* - \vartheta_0) + \varphi_0$. Der Aktuator ist in diesem Fall deine Hand.

Als nächstes wird die folgende Regelstruktur betrachtet.

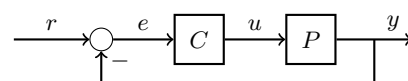


Abb. 4: Regelstruktur

Das System setzt sich immer noch aus einer Regelstrecke P und einem Regler C zusammen. Der Hauptunterschied zur Steuerung ist, dass nun die Ausgangsgrösse y zurückgeführt wird und mit der Sollgrösse r verglichen wird. Die Differenz dieser Grössen ist der Fehler e zwischen Soll- und Istwert. Diese Struktur hat einen Freiheitsgrad: Den Regler C .

Beispiel: Du hast dich für Tennis am Campus Irchel angemeldet und benutzt die Dusche der Garderobe. Du stellst fest, dass du diese Dusche noch nie verwendet hast, und weisst nicht welche Einstellung deine Lieblingstemperatur ϑ^* liefert. In anderen Worten kennst du die Regelstrecke P nicht.

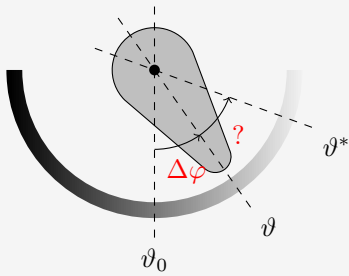


Abb. 5: Temperaturhebel einer Dusche am Campus Irchel.

Betrachte Abb. 5. Du hast den Hebel auf den arbiträren Winkel $\varphi = \Delta\varphi + \varphi_0$ gestellt. Nun fühlst du auf deiner Haut, dass die resultierende Temperatur ϑ nicht deiner Lieblingstemperatur ϑ^* entspricht. Das heisst, durch Rückführung der Temperatur, mittels deiner Haut stellst du eine Temperaturdifferenz $e = \vartheta^* - \vartheta$ fest. Intuitiv wirst du, wenn die Temperaturdifferenz sehr gross ist, den Hebel stark in die Richtung bewegen, die den Fehler reduziert. Falls die Temperaturdifferenz eher klein ist, musst du den Hebel nur leicht bewegen. Ein Beispiel eines solchen Reglers ist ein Proportionalregler: $\Delta\varphi = k_P \cdot e$, wobei k_P eine Proportionalitätskonstante ist, die du ungefähr schätzt. Ein grosser Fehler e , wird in einer grossen Winkeländerung $\Delta\varphi$ resultieren, wobei ein kleiner Fehler in einer kleinen Winkeländerung resultiert. Du wirst den Winkel basierend auf dem Fehler e kontinuierlich ändern, bis sich die gewünschte Temperatur eingestellt hat.

Als nächstes wird eine allgemeine Regelstruktur inklusive Störung und Rauschen erklärt.

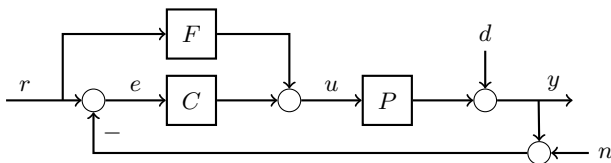


Abb. 6: Regelstruktur mit Vorsteuerung.

Die Rückführung der Ausgangsgrösse in Abb. 6 übernimmt immer noch die gleiche Rolle wie in der Regelstruktur in Abb. 4. Zusätzlich hat man nun eine Schätzung \hat{P} der Regelstrecke P . Die Vorsteuerung kann nun ähnlich wie bei der Steuerung als Inversion der ungefähren Regelstrecke gewählt werden: $F = \hat{P}^{-1}$. Das Ziel der Vorsteuerung ist, das System durch Vorwissen des Verhaltens der Regelstrecke in die Umgebung des gewünschten Operationspunkt zu steuern. Fehler der Vorsteuerung F , werden durch den Regler C korrigiert. Zusätzlich sind noch die Störung d und das Rauschen n eingezeichnet. Die Störung d wirkt sich auf die wahre Ausgangsgrösse aus, wobei das Rauschen n auf die Messung von y wirkt. Über die Rückführung wird n auch auf y wirken. Diese Struktur hat zwei Freiheitsgrade: C, F .

Beispiel: Du warst jetzt schon öfters im Training und weist welche Hebeleinstellung deine Lieblingstemperatur liefert (Vorsteuerung F). Da die Einstellung aber nicht genau stimmt, regelst du den Hebel noch auf die richtige Einstellung (Regler C). Leider duscht das Fussballteam auch noch. Wenn alle Duschen gleichzeitig laufen, entspricht die resultierende Temperatur nicht mehr deiner Lieblingseinstellung (Störung d). Da du dies mittels Rückführung merkst, regelst du den Hebel, bis sich wieder deine Lieblingstemperatur einstellt. Nun hat jemand die Tür zur Garderobe offen gelassen und du kriegst ab und zu Hühnerhaut (Rauschen n). Da du das Gefühl hast, dass sich die Temperatur geändert hat, obwohl sie noch gleich ist, regelst du die Dusche auf eine Temperatur, die sich gleich anfühlt wie deine Lieblingstemperatur.

4 Modellierung

Drei Methoden um die Differentialgleichung eines Systems herzuleiten werden beschrieben.

Impulserhaltung:
$$\frac{d}{dt}(m\dot{x}) = \sum_i F_i$$

Die momentane Beschleunigung \ddot{x} eines Körpers mit Masse m ist durch die Summe der momentan anliegenden Kräfte F_i bestimmt.

Drehimpulserhaltung:
$$\frac{d}{dt}(J_B \dot{\theta}) = \sum_i T_i$$

Die momentane Winkelbeschleunigung $\ddot{\theta}$ eines Körpers mit Trägheitsmoment J_B und Ruhepunkt B ist durch die Summe der momentan anliegenden Momente T_i bestimmt.

Speichermethode:

1. Identifiziere die Systemgrenze (Zuflüsse, Ausflüsse)
2. Identifiziere die relevanten Speicher im System (Masse, Energie, Ladung) und ihre zugehörigen Pegelvariablen
3. Formuliere die Differentialgleichung für alle relevanten Speicher

$$\frac{d}{dt}(\text{Speicherinhalt}) = \sum \text{Zuflüsse} - \sum \text{Ausflüsse}$$

Falls der Speicherinhalt die Energie des Systems ist, sind Zu- und Ausflüsse Leistungen P_+ und P_- . Die Leistung einer Kraft F in einem System hängt von der Geschwindigkeit \dot{x} ab. Die Leistung eines Momentes T hängt von der Winkelgeschwindigkeit $\dot{\varphi}$ ab.

$$P_{\text{Kraft}} = F\dot{x}$$

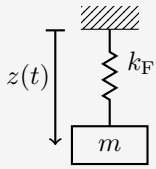
$$P_{\text{Moment}} = T\dot{\varphi}$$

4. Formuliere algebraische Relationen um Zuflüsse/Ausflüsse eines Speichers als Funktion der Pegelvariablen auszudrücken. Nach einsetzen sollten sowohl der Speicherinhalt wie auch die Zuflüsse/Ausflüsse als Funktion der Pegelvariablen ausgedrückt sein.

5. Identifiziere Systemparameter durch Experimente, Designspezifikationen oder Systemoptimierung

6. Validiere das Modell mit Experimenten

Beispiel: Betrachte Abb.7. Die Masse m ist an einer Feder mit Federkonstante k_F aufgehängt und wird von unten mit einem Gas mit Geschwindigkeit $v(t)$ angeströmt. Die Geschwindigkeit $v(t)$ entspricht der Eingangsgrösse $u(t)$. Die Kraft der Anströmung auf die Masse m sei proportional zum Quadrat der relativen Geschwindigkeit. Die Position des Systems ist der Ausgang $y(t)$ und wird mit einem Sensor gemessen.



Die Systemgleichungen können für dieses Beispiel am schnellsten mittels Impulserhaltung hergeleitet werden:

$$\begin{aligned} m\ddot{z} &= F_G - F_{\text{Feder}} - F_{\text{Wind}} \\ &= mg - k_F z - k(\dot{z} + v(t))^2 \end{aligned}$$

Abb. 7: Angeströmte, aufgehängte Masse

Das System lässt sich auch allgemeiner mit der Speichertheorie herleiten:

1. Die externe Kraft durch die Anströmung führt dem System Energie zu/ab.
2. Die Energie im System setzt sich aus potentieller Energie der Masse und der Feder, und der kinetischen Energie der Masse zusammen:

$$E(t) = -mgz(t) + \frac{1}{2}k_F z(t)^2 + \frac{1}{2}m\dot{z}(t)^2$$

Die Pegelvariablen des Systems sind: z, \dot{z}

3. Die Energiebilanz liefert:

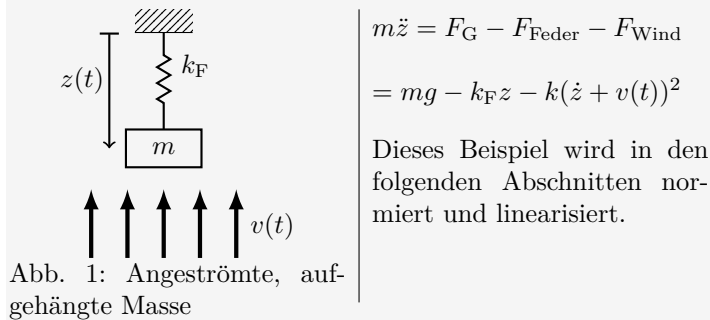
$$\frac{d}{dt}E(t) = -mg\dot{z} + k_F z\dot{z} + m\dot{z}\ddot{z} = P_+ - P_-$$

4. Die Kraft der Anströmung wirkt laut Abb. 7 der positiven z -Richtung entgegen und hat deshalb ein negatives Vorzeichen. Somit ist das Modell des Systems gegeben durch:

$$\begin{aligned} -mg\dot{z} + k_F z\dot{z} + m\dot{z}\ddot{z} &= -F_{\text{Wind}} \cdot \dot{z} \\ &= -\underbrace{k(\dot{z} + v(t))^2}_{\text{Kraft}} \cdot \dot{z} \end{aligned}$$

$$\Rightarrow m\ddot{z} = mg - k_F z - k(\dot{z} + v(t))^2$$

Beispiel: Betrachte Abb.1. Die Masse m ist an einer Feder mit Federkonstante k_F aufgehängt. Man kann ein Gasfluss mit Geschwindigkeit $v(t)$ von unten einstellen. Die Kraft der Anströmung auf die Masse m sei proportional zum Quadrat der relativen Geschwindigkeit. Die Gleichgewichtsposition des Objekt bei gegebenem v_e sei bei p_e . Die Position des Systems ist die Ausgangsgrösse und wird mit einem Sensor gemessen.



$$m\ddot{z} = F_G - F_{\text{Feder}} - F_{\text{Wind}}$$

$$= mg - k_F z - k(\dot{z} + v(t))^2$$

Dieses Beispiel wird in den folgenden Abschnitten normiert und linearisiert.

1 Zustandsgleichung erster Ordnung

In diesem Abschnitt wird erklärt, wie man eine nichtlineare Differentialgleichung (Gl. (1)) in eine Zustandsgleichung erster Ordnung (Gl. (2)) umwandelt. Aus dem Modellierungsschritt erhält man generell nichtlineare Differentialgleichungen der Ordnung n :

$$q(z^{(n-1)}, \dots, z^{(1)}, z^{(0)}, v) = z^{(n)} \quad (1)$$

$$\dot{z} = f(z, v), \quad w = g(z, v), \quad (2)$$

wobei $z(t)$ der Zustandsvektor, $v(t)$ die Eingangsgrösse, und $w(t)$ die Ausgangsgrösse ist. $f(\cdot)$ und $g(\cdot)$ sind generell nichtlineare Funktionen.

Beispiel $\frac{1}{m}(mg - k_F z - k(\dot{z} + v(t))^2) = \ddot{z}.$

Durch die Wahl

$$z = \begin{bmatrix} z^{(0)} & z^{(1)} & \dots & z^{(n-2)} & z^{(n-1)} \end{bmatrix}^T$$

$$= \begin{bmatrix} z_1 & z_2 & \dots & z_{n-1} & z_n \end{bmatrix}^T \quad (3)$$

lässt sich Gl. (1) in eine Zustandsgleichung erster Ordnung umwandeln. Aus Gl. (3) und Gl. (1) folgt

$$\dot{z} = \begin{bmatrix} z^{(1)} & z^{(2)} & \dots & z^{(n-1)} & z^{(n)} \end{bmatrix}^T$$

$$= \begin{bmatrix} f_1 & f_2 & \dots & f_{n-1} & q(z_n, \dots, z_2, z_1, v) \end{bmatrix}^T$$

$$= f(z, v).$$

Beispiel

$$z = \begin{bmatrix} z \\ \dot{z} \end{bmatrix} = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$$

$$\dot{z} = \begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \end{bmatrix} = \begin{bmatrix} z_2 \\ \frac{1}{m}(mg - k_F z_1 - k(z_2 + v(t))^2) \end{bmatrix}$$

2 Normierung

Die Grössen im Zustandsvektor z weisen verschiedene Einheiten auf. Die Normierung dient einerseits einer vereinfachten Interpretation und andererseits zur Vorbeugung von numerischen Problemen. Falls der Betrachter die physikalischen Grössen bevorzugt und die Numerik kein Problem darstellt, kann man den Normierungsschritt überspringen.

$$z_i(t) = z_{i,0} \cdot x_i(t), \quad z_{i,0} \in \mathbb{R} \setminus \{0\}$$

In Vektornotation

$$z = T \cdot x, \quad T = \text{diag}(z_{1,0}, \dots, z_{n,0})$$

Die Ein- und Ausgangsgrösse werden analog normiert:

$$v(t) = v_0 \cdot u(t) \quad v_0 \in \mathbb{R} \setminus \{0\}$$

$$w(t) = w_0 \cdot y(t) \quad w_0 \in \mathbb{R} \setminus \{0\}$$

Generell gilt

$$T \cdot \dot{x} = \dot{z} = f(z, v)$$

$$w_0 \cdot y = w = g(z, v)$$

Nun normiert man das System:

$$\dot{x} = T^{-1} \cdot f(T \cdot x, v_0 \cdot u) = f_0(x, u) \quad (4)$$

$$y = w_0^{-1} \cdot g(T \cdot x, v_0 \cdot u) = g_0(x, u) \quad (5)$$

Die Einheit der normierten Gleichung (4) ist $\left[\frac{1}{s}\right]$.

Beispiel:

$$\dot{z} = \begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \end{bmatrix} = \begin{bmatrix} z_2 \\ \frac{1}{m}(mg - k_F z_1 - k(z_2 + v(t))^2) \end{bmatrix}$$

Es liege im Interesse des Betrachters, dass die normierte Position x_1 der Masse m im Gleichgewichtszustand p_e (bei Wind v_e), $x_{1,e} = 1$ entspreche. Ausserdem weiss der Betrachter, dass gilt: $0 < v(t) < v_{\max}$. Dies ist hilfreich, um die Eingangsgrösse in die Region $0 < u(t) < 1$ zu normieren. Die maximale Geschwindigkeit \dot{z}_{\max} sei h :

$$T = \begin{bmatrix} p_e & 0 \\ 0 & h \end{bmatrix}, \quad v(t) = v_{\max} \cdot u(t)$$

$$\dot{x} = \begin{bmatrix} \frac{h}{p_e} \cdot x_2 \\ \frac{1}{h \cdot m}(mg - k_F \cdot p_e \cdot x_1 - k(h \cdot x_2 + v_{\max} \cdot u(t))^2) \end{bmatrix}$$

3 Linearisierung von nichtlinearen Systemen

Das normierte nichtlineare System wird nun um den Gleichgewichtspunkt linearisiert. Die Linearisierung vereinfacht die Analyse um den Gleichgewichtspunkt.

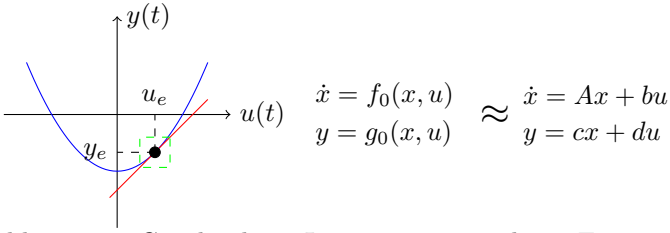


Abb. 2: Graphische Interpretation der Eingangs-Ausgangslinearisierung. Nichtlineares System in blau. Linearisiertes System in rot.

Berechnung der Gleichgewichtslage

Um das System zu linearisieren wird zuerst die Gleichgewichtslage berechnet. Im Gleichgewichtszustand gilt per Definition:

$$\dot{x} = [\dot{x}_1 \quad \dots \quad \dot{x}_n]^T = [0 \quad \dots \quad 0]^T = f(x_e, u_e).$$

Dies ist ein lineares Gleichungssystem in x_1, x_2, \dots, x_n und u . Für einen gewünschten Zustand x_e lässt sich ein u_e berechnen. Dabei ist zu beachten, dass nicht alle gewünschten x_e möglich sind. Umgekehrt lässt sich für ein konstantes u_e der resultierende Gleichgewichtszustand x_e berechnen.

Beispiel:

$$\begin{aligned} \dot{x}_e &= \begin{bmatrix} \frac{h}{p_e} \cdot x_{2,e} \\ \frac{1}{h \cdot m} (mg - k_F \cdot p_e \cdot x_{1,e} - k(h \cdot x_{2,e} + v_{\max} \cdot u_e)^2) \end{bmatrix} \\ &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow x_{2,e} = 0, \quad u_e = \frac{1}{v_{\max}} \sqrt{\frac{1}{k} (mg - k_F p_e \cdot x_{1,e})}, \end{aligned}$$

mit $x_{1,e} = 1$.

Linearisierung

Um das System um den Gleichgewichtspunkt zu linearisieren, werden die Matrizen A, b, c, d berechnet:

$$\begin{aligned} A &\stackrel{\text{def}}{=} \left[\begin{array}{ccc|c} \frac{\partial f_{0,1}}{\partial x_1} & \frac{\partial f_{0,1}}{\partial x_2} & \dots & \frac{\partial f_{0,1}}{\partial x_n} \\ \frac{\partial f_{0,2}}{\partial x_1} & \frac{\partial f_{0,2}}{\partial x_2} & \dots & \frac{\partial f_{0,2}}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_{0,n}}{\partial x_1} & \frac{\partial f_{0,n}}{\partial x_2} & \dots & \frac{\partial f_{0,n}}{\partial x_n} \end{array} \right] \bigg|_{x_e, u_e} & b &\stackrel{\text{def}}{=} \left[\begin{array}{c} \frac{\partial f_{0,1}}{\partial u} \\ \vdots \\ \frac{\partial f_{0,n}}{\partial u} \end{array} \right] \bigg|_{x_e, u_e} \\ c &\stackrel{\text{def}}{=} \left[\begin{array}{ccc|c} \frac{\partial g_0}{\partial x_1} & \frac{\partial g_0}{\partial x_2} & \dots & \frac{\partial g_0}{\partial x_n} \end{array} \right] \bigg|_{x_e, u_e} & d &\stackrel{\text{def}}{=} \left[\frac{\partial g_0}{\partial u} \right] \bigg|_{x_e, u_e} \end{aligned}$$

Die Matrizen sind eine direkte Konsequenz der Taylorentwicklung¹ erster Ordnung von $\dot{x} = f_0(\delta x + x_e, \delta u + u_e)$ um x_e, u_e

$$\delta \dot{x} \approx \underbrace{f_0(x_e, u_e)}_{=0} + \underbrace{\frac{\partial f_0(x_e, u_e)}{\partial x}}_A \delta x + \underbrace{\frac{\partial f_0(x_e, u_e)}{\partial u}}_B \delta u$$

¹Analog für $y = g_0(\delta x + x_e, \delta u + u_e)$

Wichtig: Bei linearisierten Systemen betrachtet man nur die Dynamik der Differenzen mit Ursprung x_e, u_e und lässt das δ weg ($\delta x \stackrel{\text{def}}{=} x, \delta u \stackrel{\text{def}}{=} u, \delta y \stackrel{\text{def}}{=} y$)

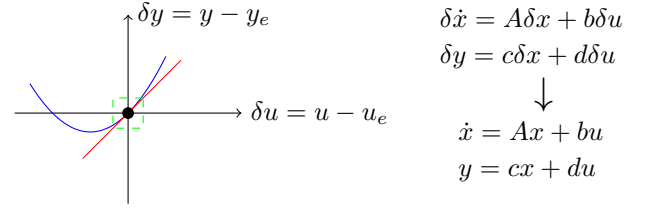


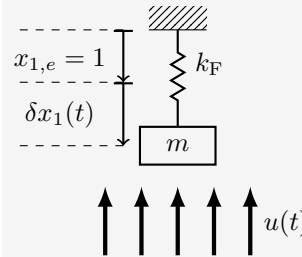
Abb. 3: Graphische Interpretation der Delta-Formulierung.

Beispiel: Das System wird nun um u_e, x_e linearisiert

$$\begin{aligned} A &= \begin{bmatrix} 0 & h/p_e \\ -\frac{1}{h \cdot m} \cdot k_F \cdot p_e & -\frac{1}{m} \cdot 2k \cdot (h \cdot x_{2,e} + v_{\max} \cdot u_e) \end{bmatrix}, \\ B &= \begin{bmatrix} 0 \\ -\frac{1}{h \cdot m} \cdot 2k \cdot v_{\max} \cdot (h \cdot x_{2,e} + v_{\max} \cdot u_e) \end{bmatrix} \end{aligned}$$

Die Ausgangsgröße ist die Position $z_1 = z$ der Masse. Wählt man die normierte Ausgangsgröße $p_e \cdot y = z$, folgt

$$y = \frac{z}{p_e} = x_1 = \underbrace{\begin{bmatrix} 1 & 0 \end{bmatrix}}_c \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad d = 0.$$



Das normierte System hat die Gleichgewichtslage bei $x_{1,e} = 1$. Das linearisierte System beschreibt das originale nichtlineare Systemverhalten sehr genau für kleine Abweichungen $\delta x_1(t)$ um den linearisierten Punkt.

1 Allgemeine Lösung eines LZI Systems

Ein lineares zeitinvariantes SISO System hat die allgemeine Form:

$$\dot{x}(t) = A \cdot x(t) + b \cdot u(t) \quad A \in \mathbb{R}^{n \times n}, b \in \mathbb{R}^{n \times 1}$$

$$y(t) = c \cdot x(t) + d \cdot u(t) \quad c \in \mathbb{R}^{1 \times n}, d \in \mathbb{R}$$

$$x(0) = x_0$$

Die allgemeine Lösung der Zustandsgrösse $x(t)$ ist:

$$x(t) = e^{A \cdot t} \cdot x_0 + \int_0^t e^{A \cdot (t-\rho)} \cdot b \cdot u(\rho) d\rho \quad (1)$$

Setzt man die allgemeine Lösung in die Gleichung der Ausgangsgrösse $y(t)$ ein, erhält man die Superposition dreier Grössen:

$$y(t) = \underbrace{c \cdot e^{A \cdot t} \cdot x_0}_I + \underbrace{c \cdot \int_0^t e^{A \cdot (t-\rho)} \cdot b \cdot u(\rho) d\rho}_{II} + \underbrace{d \cdot u(t)}_{III}$$

Die natürliche Antwort des Systems (*I*) ist **unabhängig** von der Eingangsgrösse u . Der Eingang u trägt einerseits zum Beitrag der Systemdynamik (*II*) bei, und andererseits zum Feedthrough Term (*III*).

2 Stabilitätseigenschaften

Für die Stabilitätseigenschaften betrachtet man die natürliche Antwort des Systems. Aus der allgemeinen Lösung Gl. (1) folgt:

$$x(t) = e^{A \cdot t} \cdot x_0, \quad y(t) = c \cdot e^{A \cdot t} \cdot x_0$$

$$e^{A \cdot t} \in \mathbb{R}^{n \times n}, \quad x(0) = [x_{0,1} \quad x_{0,2} \quad \dots \quad x_{0,n}]^T$$

Im Allgemeinen sind die einzelnen Zustände gekoppelt

$$x_i(t) = f(x_{0,1}, \dots, x_{0,n}, t),$$

da $e^{A \cdot t}$ im Allgemeinen nicht diagonal ist. Man wählt die Transformation

$$\tilde{x} = V^{-1} \cdot e^{A \cdot t} \cdot V \cdot \tilde{x}(0) = e^{\tilde{A} \cdot t} \cdot \tilde{x}(0),$$

sodass \tilde{A} diagonal ist. Daraus folgt, dass $e^{\tilde{A} \cdot t}$ diagonal ist, und die transformierten Zustände entkoppelt sind:

$$\tilde{x}(t) = \begin{bmatrix} e^{\lambda_1 \cdot t} & 0 & \dots & 0 \\ 0 & e^{\lambda_2 \cdot t} & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & e^{\lambda_n \cdot t} \end{bmatrix} \tilde{x}_0 = \begin{bmatrix} e^{\lambda_1 \cdot t} \cdot \tilde{x}_{0,1} \\ e^{\lambda_2 \cdot t} \cdot \tilde{x}_{0,2} \\ \vdots \\ e^{\lambda_n \cdot t} \cdot \tilde{x}_{0,n} \end{bmatrix},$$

wobei λ_i die Eigenwerte von A und \tilde{A} sind¹. Die Eigenwerte sind generell komplexwertig $\lambda_i = \sigma_i + j\omega_i$, wobei $\text{Re}(\lambda_i) = \sigma_i$, $\text{Im}(\lambda_i) = \omega_i$, und $j^2 = -1$.

¹Es gilt: $\text{eig}(A) = \text{eig}(\tilde{A})$

Setzt man die allgemeinen Eigenwerte in die obere Gleichung ein, folgt:

$$\tilde{x}(t) = \begin{bmatrix} e^{\lambda_1 \cdot t} \cdot \tilde{x}_{0,1} \\ e^{\lambda_2 \cdot t} \cdot \tilde{x}_{0,2} \\ \vdots \\ e^{\lambda_n \cdot t} \cdot \tilde{x}_{0,n} \end{bmatrix} = \begin{bmatrix} e^{\sigma_1 t} \cdot (\cos(\omega_1 t) + j \cdot \sin(\omega_1 t)) \cdot \tilde{x}_{0,1} \\ e^{\sigma_2 t} \cdot (\cos(\omega_2 t) + j \cdot \sin(\omega_2 t)) \cdot \tilde{x}_{0,2} \\ \vdots \\ e^{\sigma_n t} \cdot (\cos(\omega_n t) + j \cdot \sin(\omega_n t)) \cdot \tilde{x}_{0,n} \end{bmatrix}$$

In dieser Form sieht man, dass die Geschwindigkeit mit der ein Zustand divergiert/konvergiert von den Realteilen σ_i abhängt ($e^{\sigma_i \cdot t}$ -Term). Ausserdem sieht man, dass der Imaginärteil ω_i bestimmt, ob und wie schnell ein Zustand oszilliert.

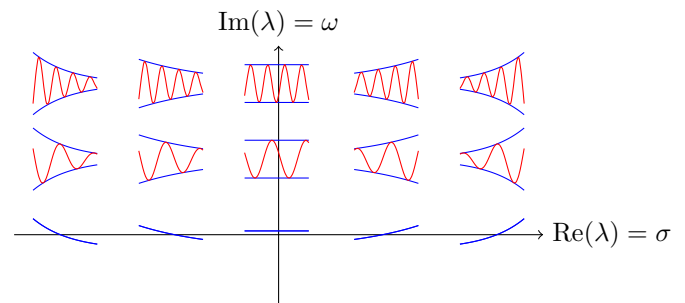


Abb. 1: Systemantwort als Funktion der Eigenwerte. Rot: Einfluss Imaginärteil, Blau: Einfluss Realteil

In physikalischen Systemen haben komplexe Eigenwerte λ_i immer ein komplex-konjugiertes Gegenstück λ_i^* . Für lineare Systeme kann das Verhalten der natürlichen Antwort anhand der Eigenwerte der Systemmatrix A charakterisiert werden. Dies erlaubt die Analyse der Stabilität von Gleichgewichtspunkten nach Lyapunov:

Lyapunov Stabilität

Stabilität nach Lyapunov erlaubt die Stabilitätsanalyse von Gleichgewichtspunkten (GGWP) von linearen und linearisierten Systemen. Falls die Stabilität des GGWP eines linearisierten Systems den Fällen (i) oder (iii) entspricht, gilt dies auch für die Stabilität desselben GGWP im nichtlinearen System.

i. Asymptotisch stabil: ($\lim_{t \rightarrow \infty} \|x(t)\| = 0$), falls alle Eigenwerte λ_i ($i = 1, \dots, n$) einen negativen Realteil haben $\sigma_i < 0$.

ii. Stabil: ($\|x(t)\| < \infty \quad \forall t \in [0, \infty]$), falls mindestens ein Eigenwert λ_k $k \in \{1, \dots, n\}$ keinen Realteil hat $\sigma_k = 0$ und kein anderer Realteil positiv ist $\sigma_i \not> 0$

iii. Instabil: ($\lim_{t \rightarrow \infty} \|x(t)\| = \infty$), falls mindestens ein Eigenwert λ_k einen positiven Realteil hat $\sigma_k > 0$.

Wichtig! Falls ein GGWP eines linearisierten Systems Eigenwerte mit $\sigma_i = 0$ vorweist, lässt sich mit dieser Analyse keine Aussage über die Stabilität desselben GGWP des nichtlinearen Systems machen.

Beispiel: Betrachte Abb. 2. Die Masse m ist mit einer Feder mit Federkonstante k_F und einem Dämpfer mit Dämpferkonstante c_D verbunden. Das System kann mit einer Kraft $u(t)$ aktuiert werden.

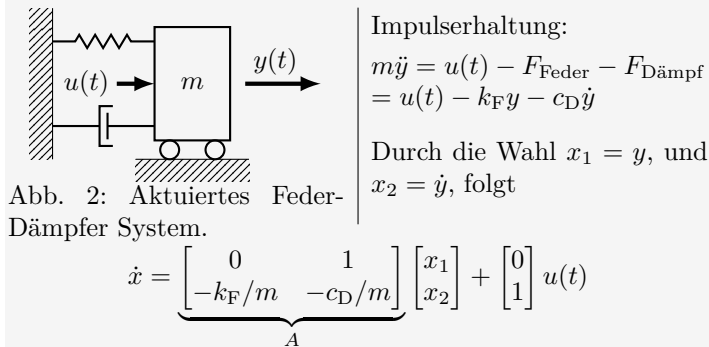


Abb. 2: Aktuiertes Feder-Dämpfer System.

Die Entwicklung des Systems kann nun mittels der Eigenwerte λ_i von A analysiert werden:

$$\text{eig}(A) = \lambda_i = -\frac{c_D}{2m} \pm \sqrt{\frac{c_D^2}{4m^2} - \frac{k_F}{m}}$$

Fall (1): $\frac{c_D^2}{4m^2} - \frac{k_F}{m} \geq 0 \Rightarrow c_D^2 \geq 4k_F m \Rightarrow \sigma_i < 0, \omega_i = 0$

Falls der Dämpfer relativ zur Feder und zur Masse genug stark ist, sind alle Eigenwerte reellwertig negativ. Das heisst, das System konvergiert ohne Oszillationen zum GGWP.

Asymptotisch stabil nach Lyapunov.

Fall (2): $c_D^2 < 4k_F m \Rightarrow \sigma_i < 0, \omega_i \neq 0$

Falls der Dämpfer schwach ist, werden die Eigenwerte komplex, mit negativem Realteil. Das heisst, das System oszilliert um den GGWP mit abnehmender Amplitude.

Asymptotisch stabil nach Lyapunov.

Fall (3): $c_D = 0 \Rightarrow \sigma_i = 0, \omega_i \neq 0$

Falls kein Dämpfer im System ist, werden alle Eigenwerte komplex mit Realteil $\sigma_i = 0$. Das System oszilliert für immer um den GGWP.

Stabil nach Lyapunov.

Fall (4): $k_F = 0 \Rightarrow \lambda_1 = 0, \sigma_2 < 0, \omega_2 = 0$

Falls keine Feder im System ist, wird ein Eigenwert 0 und der andere wird reell negativ. Falls das System mehrmals angestossen wird, kommt es jedesmal ohne Oszillation zum Stillstand, jedoch nicht zwingend im GGWP.

Stabil nach Lyapunov.

3 Testsignale auf Systeme erster Ordnung

In diesem Abschnitt werden spezielle Eingänge für Systeme erster Ordnung mit Zeitkonstante τ und Eingangsstärke k betrachtet

$$\dot{x}(t) = -\frac{1}{\tau} \cdot x(t) + \frac{k}{\tau} u(t), \quad y(t) = x(t)$$

3.1 Impulsantwort

$$u(t) = \delta(t) = \begin{cases} +\infty, & t = 0 \\ 0, & t \neq 0 \end{cases}$$

Aus Gl. (1) folgt die allgemeine Lösung :

$$y_\delta(t) = e^{-t/\tau} \cdot (x_0 + k/\tau). \quad (2)$$

Aus Gl. (2) folgt, dass ein Impuls die Anfangsbedingung x_0 um k/τ ändert. Das heisst, ein Impuls induziert eine Anfangsbedingung. Das System entwickelt sich von der neuen Anfangsbedingung, als ob $u(t) = 0$ wäre.

3.2 Sprungantwort

$$u(t) = h(t) = \begin{cases} 1, & t \geq 0 \\ 0, & t < 0 \end{cases}$$

Aus Gl. (1) folgt die allgemeine Lösung:

$$y_h(t) = e^{-t/\tau} \cdot x_0 + k \cdot (1 - e^{-t/\tau}) \quad (3)$$

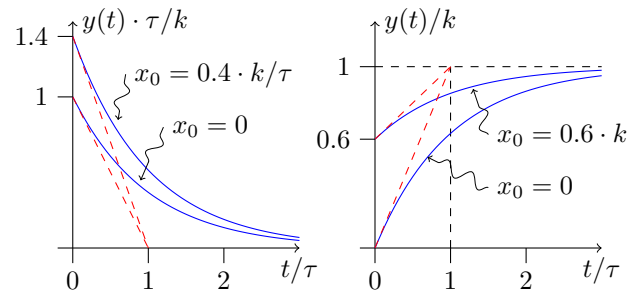


Abb. 3: Impulsantwort links und Sprungantwort rechts.

Die Tangente an die Impulsantwort zum Zeitpunkt $t = 0$ schneidet die Zeitachse zum Zeitpunkt $t = \tau$. Die Tangente an die Sprungantwort zum Zeitpunkt $t = 0$ schneidet den Sprung $k \cdot h(t)$ auch zum Zeitpunkt $t = \tau$. Je kleiner τ ist, desto schneller konvergiert das System.

Beispiel: Betrachte Abb. 2. Die Systemgleichung ohne Feder, mit $x = \dot{y}$ ist

$$\dot{x}(t) = \frac{-c_D}{m} x(t) + \frac{1}{m} u(t)$$

Es folgt: $\tau = m/c_D$ und $k = \tau/m = 1/c_D$. Ein Impuls auf das ruhende System induziert eine Anfangsgeschwindigkeit der Magnitude $k/\tau = 1/m$. Ein Sprung $h(t)$ auf das ruhende System stellt die Geschwindigkeit der Masse auf $\frac{1}{c_D}$.

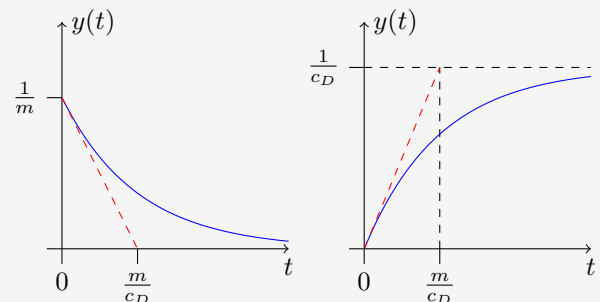


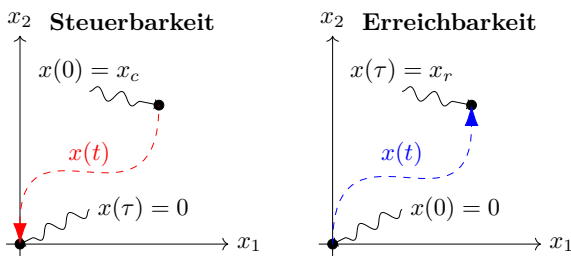
Abb. 4: Impulsantwort links und Sprungantwort rechts

Gegeben sei ein Lineares Zeit Invariantes (LZI) System:

$$\begin{aligned}\dot{x}(t) &= A \cdot x(t) + b \cdot u(t) & A \in \mathbb{R}^{n \times n}, b \in \mathbb{R}^{n \times 1} \\ y(t) &= c \cdot x(t) + d \cdot u(t) & c \in \mathbb{R}^{1 \times n}, d \in \mathbb{R} \\ x(0) &= x_0\end{aligned}\quad (1)$$

1 Steuerbarkeit / Erreichbarkeit

Ein Punkt $x_c \in \mathbb{R}^n$ ist steuerbar, falls ein Eingangssignal $u(t)$ existiert, das den Zustandsvektor des Systems von $x(0) = x_c$ zum Zustand $x(\tau) = 0$ (zum Ursprung) in endlicher Zeit τ bringt. Falls alle Punkte in \mathbb{R}^n steuerbar sind, heisst das System vollständig steuerbar. Ein System ist potentiell stabilisierbar, falls alle nicht-steuerbaren Zustände asymptotisch stabil sind.



Ein Punkt $x_r \in \mathbb{R}^n$ ist erreichbar, falls ein Eingangssignal $u(t)$ existiert, das den Zustandsvektor des Systems von Zustand $x(0) = 0$ zum Zustand $x(\tau) = x_r$ in endlicher Zeit τ bringt. Falls alle Punkte in \mathbb{R}^n erreichbar sind, heisst das System vollständig erreichbar.

Wichtig: Für LZI Systeme sind die Teilräume der erreichbaren und steuerbaren Zustände identisch.

Das System ist vollständig steuerbar/erreichbar, wenn die **Steuerbarkeitsmatrix** \mathcal{R} (reachability/controlability matrix) vollen Rang hat.

$$\mathcal{R} = [b, A \cdot b, A^2 \cdot b, \dots, A^{n-1} \cdot b]$$

Beispiel:

$$A = \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \Rightarrow \mathcal{R} = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

Da $\text{rank}(\mathcal{R}) = 1 < n$, ist das System nicht vollständig erreichbar. Aus $\{A, b\}$ folgt, dass der zweite Zustand nicht direkt vom Eingang beeinflusst wird (b) und dass er nur abhängig von sich selber ist (A): $\dot{x}_2 = A_{22} \cdot x_2$, mit Eigenwert $\lambda_2 = A_{22} = -1$. Daraus folgt, dass der zweite Zustand unabhängig stabil ist und somit ist das System $\{A, b\}$ stabilisierbar.

2 Beobachtbarkeit

Ein System ist vollständig beobachtbar wenn man mit der Messung des Ausgangssignals $y(t)$, $t \in [0, \tau]$, $\tau > 0$ eindeutig auf den Anfangszustand $x(0)$ des Systems schliessen kann.

Ein LZI System ist dann vollständig beobachtbar, wenn die **Beobachtbarkeitsmatrix** \mathcal{O} (observability matrix) vollen Rang hat.

$$\mathcal{O} = \begin{bmatrix} c \\ c \cdot A \\ c \cdot A^2 \\ \vdots \\ c \cdot A^{n-1} \end{bmatrix}$$

Beispiel:

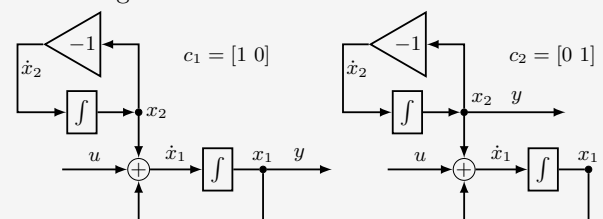
$$A = \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix}, \quad c_1 = [1 \ 0], \quad c_2 = [0 \ 1],$$

Es sei möglich entweder x_1 , oder x_2 zu messen.

$$\mathcal{O}_1 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \quad \mathcal{O}_2 = \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix}$$

Falls man x_1 misst, ist das System vollständig beobachtbar: $\text{rank}(\mathcal{O}_1) = 2$. D.h. man kann durch messen von x_1 auf die Anfangsbedingungen $x_1(0)$ und $x_2(0)$ schliessen. Falls man nur x_2 misst, erhält man $\text{rank}(\mathcal{O}_2) = 1$. Das System ist somit nicht vollständig beobachtbar.

Die grafische Darstellung der zwei beschriebenen Systeme macht deren Eigenschaften bzgl. Steuerbarkeit und Beobachtbarkeit direkt ablesbar durch den Informationsfluss der Variablen im Diagramm.



3 Koordinatentransformationen

Ein Zustandsraum mit Koordinaten x kann durch eine (von unendlich vielen) Koordinatentransformation in anderen Koordinaten \tilde{x} beschrieben werden:

$$x(t) = T \cdot \tilde{x}(t) \quad T \in \mathbb{R}^{n \times n}, \det(T) \neq 0$$

$$\frac{d}{dt} \tilde{x}(t) = T^{-1} \cdot A \cdot T \cdot \tilde{x}(t) + T^{-1} \cdot b \cdot u(t)$$

$$y(t) = c \cdot T \cdot \tilde{x}(t) + d \cdot u(t)$$

Fundamentale Systemeigenschaften (Stabilität, Steuerbarkeit, Beobachtbarkeit, I/O-Verhalten) sind transformationsinvariant (unabhängig von der Wahl der Koordinaten x), d.h. sie bleiben nach einer Koordinatentransformation erhalten.

4 Input/Output (I/O) Darstellung

Eine Zustandsraumdarstellung $\{A, b, c, d\}$ beschreibt das gesamte System (die Zustände $x(t)$ und den Ausgang $y(t)$) für gegebene $x(0)$ und $u(t)$. Oftmals ist man nur am I/O Zusammenhang $u(t) \rightarrow y(t)$ interessiert. Dann eignet sich eine I/O-Beschreibung wie folgt:

$$y^{(n)}(t) + a_{n-1} \cdot y^{(n-1)}(t) + \dots + a_1 \cdot y^{(1)}(t) + a_0 \cdot y(t) = b_m \cdot u^{(m)}(t) + \dots + b_1 \cdot u^{(1)}(t) + b_0 \cdot u(t) \quad (2)$$

wobei $y^{(r)}(t)$ die r -te zeitliche Ableitung von $y(t)$ ist.

Eine I/O Darstellung hat keine physikalischen Koordinaten mehr, weshalb alle Anfangsbedingungen auf null gesetzt werden: $y^{(n)} = 0 \forall n$.

5 Zustandsraum Normalformen

Falls man eine I/O Systembeschreibung (Gl. (2)) in eine Zustandsraumdarstellung (Gl. (1)) umwandeln möchte, müssen Koordinaten x gewählt werden. Dafür gibt es unendlich viele Möglichkeiten. Zustandsraumdarstellung in Koordinaten x , welche sich für Analyse-Methoden besonders eignen, werden als Normalformen oder kanonische Formen bezeichnet.

Eine wichtige Normalform ist die **Reglernormalform**. Für ein I/O-System nach Gl. (2) ist diese definiert als:

$$\left[\begin{array}{c|c} A & b \\ \hline c & d \end{array} \right] = \left[\begin{array}{ccccc|c} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \ddots & 0 & \vdots \\ \dots & \dots & \dots & \ddots & 0 & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ \hline -a_0 & -a_1 & -a_2 & \dots & -a_{n-1} & 1 \\ b_0 & \dots & b_m & 0 & \dots & d \end{array} \right],$$

und erlaubt u.A. das direkte Ablesen der Differentialgleichungen (Gl. (2)) aus den Systemmatrizen.

6 Zustandsraumzerlegung

Die Sets von erreichbaren (\mathcal{R}) und/oder beobachtbaren (\mathcal{O}) Punkten sind invariante Unterräume im Zustandsraum. Durch eine geeignete Koordinatentransformation $x = T \cdot \tilde{x}$ kann der gesamte Zustandsraum in die invarianten Unterräume $\{\tilde{x}_1, \tilde{x}_2, \tilde{x}_3, \tilde{x}_4\}$ in Abb. 1 zerlegt werden.

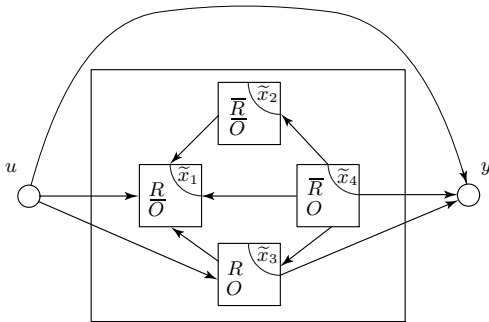


Abb. 1: Invariante Unterräume eines LZI Systems: R erreichbar, \bar{R} nicht erreichbar, O beobachtbar, \bar{O} nicht beobachtbar.

Wie ersichtlich in Abb. 1 sind zur Beschreibung des I/O-Verhaltens eines Systems nur die Zustände \tilde{x}_3 relevant (gleichzeitig erreichbar und beobachtbar)¹, denn alle anderen Zustände werden entweder nicht von Eingang beeinflusst oder beeinflussen den Ausgang nicht. Die Anzahl n Zustände im Unterraum \tilde{x}_3 entspricht der minimalen Anzahl Zustände, die zur Beschreibung des I/O-Verhaltens nötig sind, d.h. $\tilde{x}_3 \in \mathbb{R}^{n_{\min} \times 1}$. Deshalb wird die Darstellung des Systems in den Koordinaten \tilde{x}_3 **minimale Zustandsraumdarstellung** genannt.

Für gegebene Koordinaten mit $n > n_{\min}$, enthält die Zustandsraumdarstellung nicht steuerbare oder nicht beobachtbare Zustände. Falls diese im realen physikalischen System existieren, können sie sehr wichtig sein, da ihre Initialwerte den Systemausgang direkt beeinflussen können und sie, falls sie instabil sind, nicht stabilisiert werden können.

Beispiel: Bestimme, ob das System $\{A, b\}$ steuerbar ist, ohne die Erreichbarkeitsmatrix \mathcal{R} zu berechnen.

$$A = \begin{bmatrix} -1 & -2 & 0 \\ -2 & -1 & 0 \\ 0 & 0 & -2 \end{bmatrix}, b = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Eine einfache Darstellung von *diagonalisierbaren*² Systemen ergibt sich durch eine Hauptachsentransformation, $A = V \cdot D \cdot V^{-1}, x = V^{-1} \cdot \tilde{x} \rightarrow \tilde{x} = D \cdot \tilde{x} + V^{-1} \cdot b \cdot u$:

$$\begin{bmatrix} \tilde{x}_1 \\ \tilde{x}_2 \\ \tilde{x}_3 \end{bmatrix} = \begin{bmatrix} -3 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \tilde{x}_1 \\ \tilde{x}_2 \\ \tilde{x}_3 \end{bmatrix} + \begin{bmatrix} \sqrt{2} \\ 1 \\ 0 \end{bmatrix} \cdot u$$

Dabei ist V die Matrix der Eigenvektoren von A und D eine Matrix mit den Eigenwerten von A auf der Diagonalen: $D = \text{diag}(\lambda_1, \dots, \lambda_n)$.

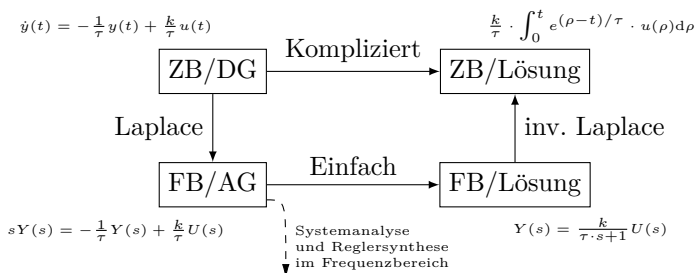
In diesen Koordinaten wird klar, dass das System einen Zustand \tilde{x}_3 hat, der instabil ($\lambda_3 = 1$) und nicht steuerbar ($b_3 = 0$) ist.

¹Zur Erinnerung: in der I/O-Beschreibung gilt $\tilde{x}_i(0) = 0 \forall i$

²Eine Matrix $A \in \mathbb{R}^{n \times n}$ ist diagonalisierbar falls sie Rang n hat \Leftrightarrow falls sie n linear unabhängige Eigenvektoren hat.

1 Laplace-Transformation

Die Laplace-Transformation ermöglicht es, eine Differentialgleichung (DG) im Zeitbereich (ZB) in eine algebraische Gleichung (AG) im Frequenzbereich (FB) umzuwandeln. Die Algebraische Gleichung im Frequenzbereich lässt sich oft einfacher lösen als die Differentialgleichung im Zeitbereich. Die Lösung im Frequenzbereich lässt sich mit der inversen Laplace-Transformation in den Zeitbereich zurückwandeln. Die inverse Laplace-Transformation ist jedoch oft nicht nötig, da Aussagen über das System auch im Frequenzbereich möglich sind. Zudem können Regler direkt im FB ausgelegt werden.



Die Laplace-Transformation eines Zeitsignals $x(t)$ ist:

$$\mathcal{L}\{x(t)\} = X(s) = \int_0^\infty x(t) \cdot e^{-s \cdot t} dt,$$

mit $s = \sigma + j \cdot \omega$. Die Laplace-Transformation ist eine Generalisierung der Fourier-Transformation. Setzt man $s = j \cdot \omega$, erhält man die Fourier-Transformation.

Wichtige Eigenschaften

- Linearität : $\mathcal{L}\{a \cdot x_1(t) + b \cdot x_2(t)\} = a \cdot X_1(s) + b \cdot X_2(s)$
- Ähnlichkeit : $\mathcal{L}\{\frac{1}{a} \cdot x(\frac{t}{a})\} = X(s \cdot a)$
- Verschiebung : $\mathcal{L}\{x(t - T)\} = e^{-T \cdot s} \cdot X(s)$
- Dämpfung : $\mathcal{L}\{x(t) \cdot e^{a \cdot t}\} = X(s - a)$
- Ableitung t : $\mathcal{L}\{\frac{d}{dt}x(t)\} = s \cdot X(s) - x(0)$ (1)
- n -te Abl. t : $\mathcal{L}\{\frac{d^n x(t)}{dt^n}\} = s^n \cdot X(s) \left(\frac{d^k x(t=0)}{dt^k} = 0 \forall k \right)$ (2)
- Ableitung s : $\mathcal{L}\{t \cdot x(t)\} = -\frac{d}{ds}X(s)$
- Integration t : $\mathcal{L}\{\int_0^t x(\tau) d\tau\} = \frac{1}{s} \cdot X(s)$
- Integration s : $\mathcal{L}\{\frac{1}{t} \cdot x(t)\} = \int_s^\infty X(\sigma) d\sigma$
- Faltung t : $\mathcal{L}\{x_1(t) * x_2(t)\} = X_1(s) \cdot X_2(s)$ ¹
- Faltung s : $\mathcal{L}\{x_1(t) \cdot x_2(t)\} = X_1(s) * X_2(s)$
- Anfangswert : $\lim_{t \rightarrow 0+} x(t) = \lim_{s \rightarrow \infty} s \cdot X(s)$ (3)
- Endwert : $\lim_{t \rightarrow \infty} x(t) = \lim_{s \rightarrow 0+} s \cdot X(s)$ (4)

¹Der Operator $*$ repräsentiert die Faltung (convolution) zweier Signale, definiert als: $x_1 * x_2(t) = \int_0^t x_1(t - \rho) \cdot x_2(\rho) d\rho$.

Wichtige Signaltransformationen

$x(t)$	$X(s)$
$\delta(t)$	1
$h(t)$	$\frac{1}{s}$
$h(t) \cdot t^n \cdot e^{\alpha \cdot t}$	$\frac{n!}{(s - \alpha)^{n+1}}$
$h(t) \cdot \sin(\omega \cdot t)$	$\frac{\omega}{s^2 + \omega^2}$
$h(t) \cdot \cos(\omega \cdot t)$	$\frac{s}{s^2 + \omega^2}$
$h(t) \cdot \sinh(\omega \cdot t)$	$\frac{\omega}{s^2 - \omega^2}$
$h(t) \cdot \cosh(\omega \cdot t)$	$\frac{s}{s^2 - \omega^2}$

Als nächstes werden Eigenschaften der Laplace-Transformation mit Beispielen erklärt.

Eigenschaft Gl. (1): Ableitungen im Zeitbereich werden zu algebraischen Grössen im Frequenzbereich. Insbesondere lässt sich mit Gl. (1) die Lösung einer Zustandsgleichung erster Ordnung finden:

$$\dot{x}(t) = A \cdot x(t) + b \cdot u(t), \quad y(t) = c \cdot x(t), \quad x(0) = 0$$

Unter Anwendung von Gl. (1) gilt:

$$s \cdot X(s) = A \cdot X(s) + b \cdot U(s), \quad Y(s) = c \cdot X(s)$$

was umgeformt werden kann zu:

$$Y(s) = c \cdot (sI - A)^{-1} \cdot b \cdot U(s) = \Sigma(s) \cdot U(s), \quad (5)$$

wobei $\Sigma(s)$ Übertragungsfunktion heisst. $\Sigma(s)$ ist im Allgemeinen ein Bruch rationaler Funktionen, wobei der Nenner bei physikalischen Systemen mindestens die Ordnung des Zählers hat ($n \geq m$):

$$\begin{aligned} \Sigma(s) &= \frac{Y(s)}{U(s)} = \frac{c \cdot \text{Adj}(sI - A) \cdot b}{\det(sI - A)} \\ &= \frac{b_m \cdot s^m + \dots + b_1 \cdot s + b_0}{s^n + a_{n-1} \cdot s^{n-1} + \dots + a_1 \cdot s + a_0} \end{aligned}$$

Der Nenner $\det(sI - A)$ entspricht der charakteristischen Gleichung der Matrix A . D.h. Stabilitätseigenschaften der GGWP lassen sich am Nenner ablesen.

Beispiel: Betrachte Abb. 1. Die Masse m ist mit einer Feder mit Federkonstante k_F und einem Dämpfer mit Dämpferkonstante c_D verbunden. Das System kann mit einer Kraft $u(t)$ aktuiert werden.

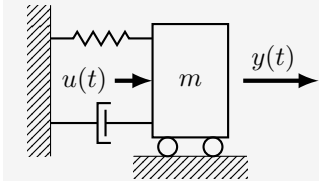


Abbildung 1: Aktuiertes Feder-Dämpfer System.

Impulserhaltung:

$$m\ddot{y} = u(t) - F_{\text{Feder}} - F_{\text{Dämpf}} = u(t) - k_F y - c_D \dot{y} \quad (6)$$

Durch die Wahl $x_1 = y$, und $x_2 = \dot{y}$, folgt:

$$\dot{x} = \begin{bmatrix} 0 & 1 \\ -k_F/m & -c_D/m \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t)$$

Durch die Wahl $k_F = 1$, $c_D = 1$, $m = 1$, und Ausgangsgrösse x_1 , erhält man folgende Systemmatrizen:

$$A = \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix}, \quad b = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad c = [1 \quad 0], \quad d = [0]$$

Mit Gl. (5) lässt sich die Übertragungsfunktion finden:

$$\begin{aligned} \text{es gilt: } K &= \begin{bmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{bmatrix} \rightarrow K^{-1} = \frac{1}{\det(K)} \cdot \begin{bmatrix} k_{22} & -k_{12} \\ -k_{21} & k_{11} \end{bmatrix} \\ \Rightarrow c \cdot (s \cdot I - A)^{-1} \cdot b &= [1 \quad 0] \cdot \begin{bmatrix} s & -1 \\ 1 & s+1 \end{bmatrix}^{-1} \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ &= \frac{1}{s(s+1)+1} \cdot [1 \quad 0] \cdot \begin{bmatrix} s+1 & 1 \\ -1 & s \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ \Rightarrow \Sigma(s) &= \frac{1}{s^2 + s + 1} \end{aligned} \quad (7)$$

Die Übertragungsfunktion lässt sich auch direkt durch Transformation von Gl. (6) (mit Gl. (2)), berechnen:

$$\begin{aligned} m \cdot s^2 \cdot Y(s) &= U(s) - k_F \cdot Y(s) - c_D \cdot s \cdot Y(s) \\ \frac{Y(s)}{U(s)} &= \Sigma(s) = \frac{1}{ms^2 + c_D s + k_F} = \frac{1}{s^2 + s + 1} \end{aligned}$$

Übertragungsfunktionen haben die allgemeine Form:

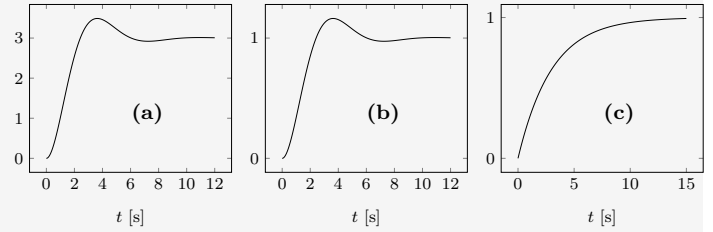
$$\Sigma(s) = b_m \cdot \frac{\prod_{j=1}^m (s - \xi_j)}{\prod_{i=1}^n (s - \pi_i)}, \quad \xi_j, \pi_i \in \mathbb{C}$$

wobei π_i Pole und ξ_j Nullstellen genannt werden. Jeder Pol π_i entspricht einem Eigenwert λ_i von A .

Vorsicht: Nicht alle Eigenwerte von A sind Pole π_i von $\Sigma(s)$, da sich Nullstellen und Pole kürzen können! Wenn die Übertragungsfunktion aus einer minimalen Systemrealisierung $\{A, b, c, d\}$ berechnet wird, lassen sich keine Pole und Nullstellen aufheben. Das Kürzen von Termen ist somit ein Hinweis dafür, dass das System $\{A, b, c, d\}$ nicht beobachtbare oder nicht steuerbare Zustände enthält. Diese Zustände beeinflussen das I/O Verhalten nicht.

Eigenschaften Gl. (3) und Gl. (4): Der Anfangs- und Endwert der Lösung im Zeitbereich lässt sich im Frequenzbereich berechnen. **Vorsicht:** Der Limes in Gl. (4) kann existieren obwohl das System nie zum GGWP konvergiert.

Beispiel: Gegeben sei das System in Gl. (7). Welches Diagramm zeigt die korrekte Sprungantwort?



Zunächst wird die Lösung im FB berechnet:

$$\begin{aligned} Y(s) &= \Sigma(s) \cdot U(s), \quad U(s) = \mathcal{L}\{h(t)\} = \frac{1}{s} \\ \Rightarrow Y(s) &= \frac{1}{s(s^2 + s + 1)} \end{aligned}$$

Unter Anwendung von Gl. (3) und Gl. (4):

$$\begin{aligned} y(0_+) &\stackrel{\text{def}}{=} \lim_{s \rightarrow \infty} s \cdot Y(s) = \lim_{s \rightarrow \infty} \frac{1}{s^2 + s + 1} = 0 \\ y(\infty) &\stackrel{\text{def}}{=} \lim_{s \rightarrow 0_+} s \cdot Y(s) = \lim_{s \rightarrow 0_+} \frac{1}{s^2 + s + 1} = 1 \end{aligned}$$

Diagramme (b) und (c) haben beide einen Anfangswert 0 und einen Endwert 1. Diagramm (a) kann ausgeschlossen werden, da es gegen 3 konvergiert. Um zwischen (b) und (c) zu unterscheiden kann man die Pole von Gl. (7) betrachten. Für komplex-konjugierte Pole erwartet man eine oszillierende Ausgangsgrösse (b). Für rein reelle Pole erwartet man keine Schwingungen in der Ausgangsgrösse (c). Die Pole sind:

$$s_{1,2} = -\frac{1}{2} \pm \frac{\sqrt{3}}{2} \cdot j \Rightarrow \text{richtige Sprungantwort ist (b)}$$

Beispiel: Endwert der Impulsantwort von $\sin(t)$:

$$\frac{Y(s)}{U(s)} = \Sigma_{\sin}(s) = \frac{1}{s^2 + 1} \quad \text{mit} \quad u(t) = \delta(t)$$

Aus $U(s) = \mathcal{L}\{\delta(t)\} = 1$, folgt $Y(s) = \Sigma_{\sin}(s)$ und somit gilt $y(t) = \mathcal{L}^{-1}\{\Sigma_{\sin}(s)\} = \sin(t)$ ¹. Dieses Signal oszilliert für immer um den GGWP, was sich auch aus den Polen von $\Sigma_{\sin}(s)$ ablesen lässt: $s = \pm j$.

Mit Gl. (4) berechnet sich jedoch fälschlicherweise ein diskreter Endwert von $y(t)$:

$$y(\infty) \stackrel{\text{def}}{=} \lim_{s \rightarrow 0} \frac{s}{s^2 + 1} = 0$$

¹Dieses Resultat gilt allgemein und ist sehr wichtig: Die Impulsantwort ist die Laplaceinverse des Systems \Leftrightarrow die Laplacetransformation der Impulsantwort ist die Übertragungsfunktion des Systems. Die Impulsantwort kann oft experimentell gemessen werden.

1 Inverse Laplace-Transformation

In Kapitel 6 wurde die Laplace-Transformation eingeführt. Es wurde erwähnt, dass sich im Frequenzbereich leicht eine Lösung finden lässt:

$$Y(s) = \Sigma(s) \cdot U(s) \quad (1)$$

Um die Antwort im Zeitbereich zu erhalten, könnte man die inverse Laplace-Transformation berechnen:

$$y(t) = \mathcal{L}^{-1}\{Y(s)\} = \frac{1}{2 \cdot \pi \cdot j} \cdot \oint Y(s) \cdot e^{s \cdot t} ds \quad t \geq 0$$

Dieses Linienintegral ist jedoch sehr schwer zum Ausrechnen. Man kann ausnützen, dass Lösungen im Frequenzbereich Brüche rationaler Funktionen sind:

$$Y(s) = b_m \cdot \frac{\prod_{j=1}^m (s - \xi_j)}{\prod_{i=1}^n (s - \pi_i)}, \quad \xi_j, \pi_i \in \mathbb{C}$$

Insbesondere kann man mit der Partialbruchzerlegung den Bruch in eine Linearkombination von Brüchen tieferer Ordnung zerteilen:

$$Y(s) = \sum_{i=1}^p \sum_{k=1}^{\phi_i} \frac{\rho_{i,k}}{(s - \pi_i)^k} \quad \rho_{i,k} \in \mathbb{C},$$

wobei $\rho_{i,k}$ die Residuen sind, und ϕ_i die Vielfachheit von π_i ist. Die inverse Laplace-Transformation der einzelnen Brüche kann allgemein hergeleitet werden:

$$\mathcal{L}^{-1}\left\{\frac{1}{(s - \pi_i)^k}\right\} = \frac{1}{(k-1)!} \cdot t^{k-1} \cdot e^{\pi_i \cdot t} \cdot h(t)$$

Beispiel: Die inverse Laplace-Transformation von:

$$Y(s) = \frac{s^2 + s + 2}{s^3 + s^2 + s + 1} = \frac{s^2 + s + 2}{(s - j) \cdot (s + j) \cdot (s + 1)} \quad (2)$$

Mit der Partialbruchzerlegung:

$$Y(s) = \frac{1}{s + 1} + \frac{\frac{1}{2 \cdot j}}{s - j} + \frac{\frac{-1}{2 \cdot j}}{s + j} \quad (3)$$

Da die Laplace-Transformation eine lineare Operation ist, kann man nun die inverse Laplace-Transformation auf die einzelnen Brüche anwenden:

$$\mathcal{L}^{-1}(Y(s)) = y(t) = e^{-t} + \underbrace{\frac{1}{2 \cdot j} (e^{j \cdot t} - e^{-j \cdot t})}_{\sin(t)} = e^{-t} + \sin(t)$$

Bemerkung: Komplexe Residuen sind oft sehr mühsam auszurechnen (Gl. (3)). Man kann bei Gl. (2) abkürzen, indem man erkennt, dass zwei rein konjugiert-komplexe Pole vorhanden sind. D.h. bevor dass man überhaupt etwas berechnet, kann man vorhersagen, dass entweder ein Cosinus oder

ein Sinus vorhanden sein muss. Um diesen Fakt auszunutzen fasst man die beiden Pole bei der Partialbruchzerlegung zusammen:

$$Y(s) = \frac{s^2 + s + 2}{(s^2 + 1) \cdot (s + 1)} = \underbrace{\frac{1}{s + 1}}_{\mathcal{L}\{e^{-t}\}} + \underbrace{\frac{1}{s^2 + 1}}_{\mathcal{L}\{\sin(t)\}} \quad (4)$$

In diesem Beispiel sind alle Pole entweder auf der rein reellen Achse ($s = -1$), oder auf der rein komplexen Achse ($s = \pm j$). Um eine Intuition für allgemeine Pole der Form $s = \sigma + j \cdot \omega$ (und deren inversen Laplace-Transformation) zu erhalten, wird das System zweiter Ordnung eingeführt.

2 Systeme 2. Ordnung

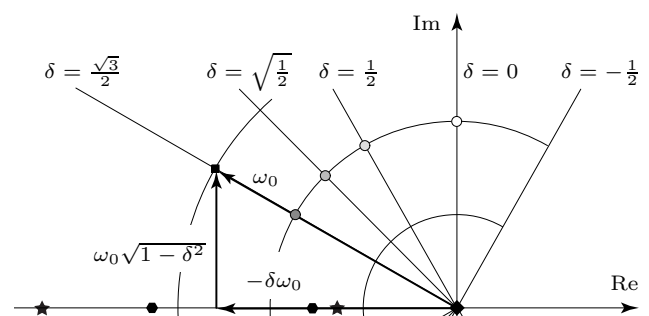
Die Übertragungsfunktion eines Systems zweiter Ordnung mit statischer Verstärkung von 1 hat folgende Form:

$$\Sigma(s) = \frac{\omega_0^2}{s^2 + 2 \cdot \delta \cdot \omega_0 \cdot s + \omega_0^2} \quad \Sigma(0) = 1 \quad (5)$$

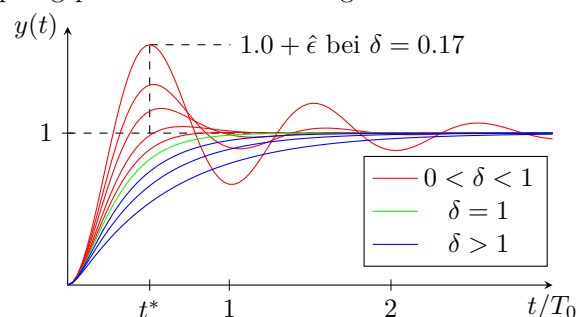
Dieses System hat zwei Pole:

$$s_{1,2} = \pi_{1,2} = \left(-\delta \pm \sqrt{\delta^2 - 1}\right) \cdot \omega_0 \quad (6)$$

Anordnung der Pole auf der real-imaginären Ebene:



Der Parameter δ wird als Dämpfungsparameter bezeichnet. Für $|\delta| < 1$ wird das Argument der Wurzel in Gl. (6) negativ. Somit werden die Pole komplex. Für $|\delta| > 1$ werden die Pole rein reell. Der Parameter $\omega_0 = 2\pi/T_0$ wird als natürliche Frequenz des Systems bezeichnet, wobei T_0 die natürliche Periode ist. Die zeitnormierte Sprungantwort, für verschiedene Dämpfungsparameter sieht wie folgt aus:



Vorsicht! Die Zeit $t = t^*$ ist nur für $\delta = 0.17$ eingezeichnet. Generell gilt $t^* = t^*(\delta, \omega_0)$. Ausserdem überschiesst das System nur für $\delta < 1$. $\delta \geq 1$ wird später behandelt.

Abhängig von der Dämpfung gibt es drei grundsätzlich unterschiedliche Fälle des Systemverhaltens:

$0 < \delta < 1$ Da für $\delta < 1$ die Pole komplex werden, beinhaltet die Sprungantwort Schwingungen. Die erste Schwingung überschiesst das Ziel $u(t) = h(t)$ um den Überschuss $\hat{\epsilon} = e^{-\delta \cdot \pi / \sqrt{1-\delta^2}}$, bei der Zeit $t^* = \frac{\pi}{\omega_0 \cdot \sqrt{1-\delta^2}}$.

$\delta > 1$ Für übergedämpfte Systeme konvergiert das System ähnlich wie bei einem System erster Ordnung zum Endwert. Die Antwort überschiesst in diesem Fall nicht, da alle Pole rein reell sind:

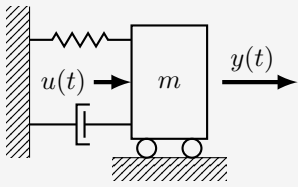
$$\Sigma(s) = \frac{\omega_0^2}{(s + \pi_1) \cdot (s + \pi_2)} \quad \pi_1, \pi_2 \in \mathbb{R}$$

Falls ein Pol viel schneller konvergiert ($\pi_2 \gg \pi_1$) gilt:

$$\Sigma(s) \approx \frac{\pi_1}{s + \pi_1}$$

$\delta = 1$ Dieser Fall heisst kritisch gedämpft. Mit Gl. (6) ist ersichtlich, dass für $\delta = 1$ beide Pole an der gleichen Stelle sind: $\pi_{1,2} = \omega_0$. Dieser Fall entspricht der schnellstmöglichen Konvergenz ohne Überschwinger.

Beispiel für $0 < \delta < 1$: Betrachte Abb. 1. Die Masse m ist mit einer Feder mit Federkonstante k_F und einem Dämpfer mit Dämpferkonstante c_D verbunden. Das System kann mit einer Kraft $u(t)$ aktuiert werden.



Impulserhaltung:

$$m\ddot{y} = u(t) - F_{\text{Feder}} - F_{\text{Dämpf}} = u(t) - k_F y - c_D \dot{y}$$

$$\Rightarrow \Sigma(s) = \frac{1/m}{s^2 + \frac{c_D}{m}s + \frac{k_F}{m}} \quad (7)$$

Abbildung 1: Aktuiertes Feder-Dämpfer System.

Fall 1: $m = 1$, $k_F = 1$, und $c_D = 1$:

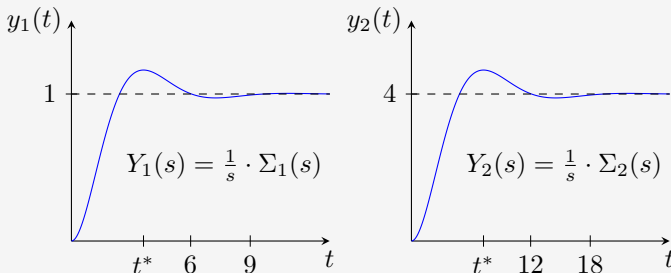
$$\Sigma_1(s) = \frac{1}{s^2 + s + 1}.$$

Vergleicht man nun mit der Standardform Gl. (5), erhält man: $\omega_0 = 1$, $\delta = 0.5$, $t^* = 3.628$, $\Sigma_1(0) = 1$, $\hat{\epsilon} = 0.163$.

Fall 2: $m = 1$, $k_F = 0.25$, und $c_D = 0.5$:

$$\Sigma_2(s) = \frac{1}{s^2 + 0.5s + 0.25} = 4 \cdot \frac{0.25}{s^2 + 0.5s + 0.25}$$

Es folgt: $\omega_0 = 0.5$, $\delta = 0.5$, $t^* = 7.255$, $\Sigma_2(0) = 4$, $\hat{\epsilon} = 0.163 \cdot \Sigma_2(0) = 0.652$.



Obwohl beide Systeme die gleiche Dämpfung haben, ist der Überschuss des zweiten Systems 4 mal grösser als beim ersten System. Ausserdem konvergiert die Sprungantwort des zweiten Systems zum Endwert $y(\infty) = 4$. Aus Gl. (5) ist ersichtlich, dass gilt: $\Sigma(0) = \omega_0^2 / \omega_0^2 = 1$. Dasselbe gilt für das erste System: $\Sigma_1(0) = 1/1 = 1$. Das zweite System, jedoch, hat eine grössere statische Verstärkung: $\Sigma_2(0) = 1/0.25 = 4$. Nun stellt sich die Frage, wann das allgemeine Beispielsystem (Gl. (7)) eine statische Verstärkung von 1 hat ($\Sigma(0) = 1$). Aus Gl. (5) ist leicht ersichtlich, dass gelten muss: $\frac{1}{m} = \frac{k_F}{m}$. D.h. das System konvergiert zum angewandten Sprung nur, falls gilt: $k_F = 1 \frac{N}{m}$! Dies macht Sinn, da eine Feder mit $k_F = 1$ sich bei einem Newton um einen Meter auslenkt. Eine Feder mit $k_F = 0.25$ lenkt sich bei einem Newton um 4 Meter aus. Wie schnell sich die Auslenkung einpendelt hängt vom Dämpfer ab.

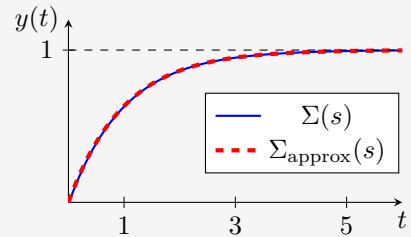
Beispiel für $\delta > 1$: Betrachte Abb. 1. Wählt man: $m = 0.01$, $k_F = 1$, und $c_D = 101/100$, folgt:

$$\Sigma(s) = \frac{100}{s^2 + 101s + 100} = \frac{100}{(s + 100) \cdot (s + 1)} \quad (8)$$

In Gl. (8) ist im Nenner ersichtlich, dass die Impulsantwort im ZB eine Superposition der Form, $y(t) = \alpha \cdot e^{-100 \cdot t} + \beta \cdot e^{-t}$, sein wird. Da $e^{-100 \cdot t}$ viel schneller abklingt, kann man das System approximieren:

$$\Sigma_{\text{approx}}(s) = \frac{1}{s + 1} \quad (9)$$

Man muss den Zähler anpassen, um die statische Verstärkung des approximierten Systems gleich zu halten: $\Sigma(0) = \Sigma_{\text{approx}}(0)$. Die Simulation beider Systeme liefert:



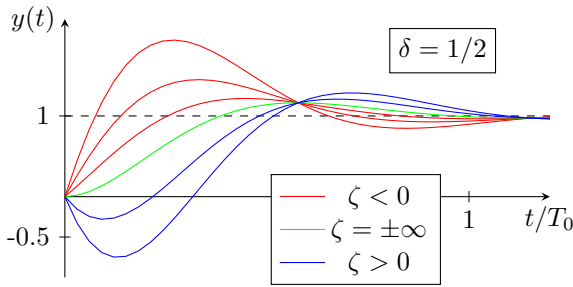
Die beiden Sprungantworten sind fast identisch.

3 Nullstelleneinfluss

Bis jetzt wurden nur die Pole betrachtet. Nullstellen sind jedoch auch sehr wichtig für das Systemverhalten. Um Nullstellen besser zu verstehen, wird das System zweiter Ordnung (Gl. (5)) mit einer rein reellen Nullstelle an der Stelle $s = \zeta$ erweitert:

$$\Sigma(s) = \frac{(-s/\zeta + 1) \cdot \omega_0^2}{s^2 + 2 \cdot \delta \cdot \omega_0 \cdot s + \omega_0^2} \quad (10)$$

Eine Nullstelle weit weg vom Ursprung ($\zeta = \pm \infty$) hat keinen Einfluss auf das Systemverhalten. Je näher die Nullstelle am Ursprung ist, desto stärker ist der Einfluss dieser Nullstelle. Dies widerspiegelt sich an einem stärkeren Überschuss in der Systemantwort. Wie bereits in Kapitel 5 behandelt, kann eine Nullstelle nahe an einem Pol, den Einfluss des Pols schwächen. Für $\zeta > 0$ hat die Systemantwort einen "Undershoot". In anderen Worten reagiert das System zuerst in die "falsche" Richtung. Ein System mit einer Nullstelle der Form $\zeta > 0$ wird als nicht-minimalphasig bezeichnet.



Nicht-minimalphasige Nullstellen folgen aus der Sensor- und Aktuatorwahl. D.h. sie sind eine Eigenschaft aus der Kombination der Systemvektoren b und c mit der Systemmatrix A (im Gegensatz zu den Polen, die nur von der Systemmatrix A abhängen.) Durch Änderung der Messgrösse (z.B. durch Änderung der Sensorenkonfiguration), kann eine nicht-minimalphasige Nullstelle minimalphasig werden.

Bemerkung: Die initiale Sprungantwort in die "falsche" Richtung tritt bei einer ungeraden Anzahl positiver Nullstellen ($\text{Re}(\zeta > 0)$) auf.

4 BIBO Stabilität

Als nächstes wird das Konzept von Bounded Input Bounded Output (BIBO) Stabilität eingeführt, welche sich auf Übertragungsfunktionen $\Sigma(s)$ (das I/O Verhalten) bezieht.

Es wurde bereits im Zeitbereich die Stabilität vom GG-WP nach Lyapunov eingeführt. Diese wird für den allgemeinen Fall mit Anfangsbedingung $x(0) = x_0 \neq 0$ betrachtet. Bei Übertragungsfunktionen $\Sigma(s)$ hingegen wird $x(0) = 0$ angenommen. Zudem wurde gezeigt, dass sich nicht beobachtbare und nicht steuerbare Zustände in $\Sigma(s)$ kürzen. Falls alle Zustände beobachtbar und steuerbar sind, entsprechen die Nullstellen des Nenners der Übertragungsfunktion im FB den Eigenwerten der Matrix A im ZB. Die Stabilität im I/O Setting braucht jedoch eine neue Interpretation, da es keinen Zustandsvektor mehr gibt. → **BIBO Stabilität** bezieht sich auf das **I/O Verhalten** von $\Sigma(s)$, und **Lyapunov Stabilität** bezieht sich auf das Gleichgewicht der **Zustände**.

Ein System ist BIBO stabil, falls für die Impulsantwort $\sigma(t)$ folgendes gilt:

$$\int_0^\infty |\sigma(t)| dt < \infty \quad (11)$$

Für ein System mit Übertragungsfunktion $\Sigma(s)$, gilt:

- Das System ist BIBO stabil falls alle Pole π_i negativen Realteil haben.
- Das System ist nicht BIBO stabil in allen anderen Fällen.

Dabei ist wichtig, dass nicht beobachtbare Zustände und nicht steuerbare Zustände die BIBO Stabilität nicht beeinflussen, da sich diese in $\Sigma(s)$ wegekürzen. BIBO Stabilität und Lyapunov Stabilität scheinen zunächst sehr ähnlich zu sein, man muss aber die Konzepte auseinanderhalten. Ein BIBO stabiles System kann Lyapunov instabil sein, und ein Lyapunov stabiles System kann BIBO instabil sein.

Beispiel: Lyapunov stabil und BIBO instabil



Das System in (a) wird betrachtet. Ein Körper mit Masse

$m = 1$ sei ohne Reibung frei beweglich. Die einzige Kraft, die wirkt sei der Eingang $u(t)$. Der Ausgang sei die Geschwindigkeit des Körpers $\dot{z}(t)$. Durch Impulserhaltung folgt:

$$\ddot{z}(t) = u(t), \quad y(t) = \dot{z}(t)$$

Durch die Wahl des Zustands $x(t) = \dot{z}(t)$, folgt:

$$\dot{x}(t) = u(t), \quad y(t) = x(t) \quad (12)$$

Vergleicht man Gl. (12) mit der Standardform, folgt:

$$\dot{x} = Ax + bu \quad \leftrightarrow \quad \dot{x} = 0 \cdot x + u \quad (13)$$

D.h. das System hat einen Eigenwert $\lambda = 0$. Der GG-WP ist stabil nach Lyapunov (nicht asymptotisch). Die I/O Übertragungsfunktion des Systems lautet:

$$\Sigma(s) = \frac{Y(s)}{U(s)} = \frac{1}{s} \quad (\text{Integrator})$$

Es gibt verschiedene Methoden, wie man zeigen kann, dass (a) nicht BIBO stabil ist. Die Übertragungsfunktion hat einen Pol bei $s = 0$. Per Definition ist das System somit nicht BIBO stabil. Rechnet man die Impulsantwort aus, folgt: $\sigma(t) = \mathcal{L}^{-1}\{1/s\} = 1$. Somit ist die Bedingung in Gl. (11) auch nicht erfüllt. Ausserdem lässt sich leicht ein Bounded Input finden ($u(t) = c$), der einen Unbounded Output ($y(t) = c \cdot t$) liefert.

Bemerkung: Das System in (a) wäre BIBO stabil, falls auf die Kugel eine geschwindigkeitsabhängige Widerstandskraft $F_{\text{Dämpf}} = \dot{z}$ wirkte, wie dargestellt in (b).

$$\ddot{z}(t) = -\dot{z}(t) + u(t), \quad y(t) = \dot{z}(t)$$

Durch dieselbe Zustandswahl $x = \dot{z}$ folgt:

$$\dot{x} = -x(t) + u(t), \quad y(t) = x(t)$$

Dies ist ein System erster Ordnung mit Eigenwert $\lambda = -1$ und ist somit asymptotisch stabil nach Lyapunov. Die Impulsantwort ergibt sich zu:

$$\Sigma(s) = \frac{Y(s)}{U(s)} = \frac{1}{s+1} \Rightarrow \sigma(t) = e^{-t},$$

somit ist das System (b) auch BIBO stabil.

Beispiel: Lyapunov instabil und BIBO stabil:

Das folgende System wird auf Lyapunov und BIBO Stabilität geprüft:

$$\dot{x} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} x + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u, \quad y = \begin{bmatrix} 1 & 0 \end{bmatrix} \cdot x = x_1 \quad (14)$$

In Gl. (14) stehen die Eigenwerte direkt auf der Diagonalen: $\lambda_1 = -1$, $\lambda_2 = 1$. Das System ist somit instabil nach Lyapunov. Die Instabilität nach Lyapunov ist nachvollziehbar wenn $x_2(0) \neq 0$, dann divergiert der Zustand x_2 für $t \rightarrow \infty$. Da der instabile Zustand weder steuerbar noch beobachtbar ist, ist das System trotzdem BIBO stabil. Durch ausrechnen der Übertragungsfunktion folgt:

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

Der Pol hat einen negativen Realteil. Das System ist somit gleichzeitig BIBO stabil und Lyapunov instabil.

1 Frequenzantworten

In Kapitel 4 wurden zwei verschiedene Eingangsgrößen eingeführt: Der Sprung $h(t)$ und der Impuls $\delta(t)$. Als nächstes wird die harmonische Eingangsgrösse eingeführt:

$$u(t) = \alpha \cdot \cos(\omega \cdot t + \phi),$$

wobei α die Amplitude, ω die Frequenz in $\frac{\text{rad}}{\text{s}}$, und ϕ die Phasenverschiebung ist. Der Ausgang eines Systems $\Sigma(s)$ mit harmonischen Eingang hat die Form:

$$y(t) = y_{\text{transient}}(t) + y_{\infty}(t)$$

Unter der Annahme, dass das System $\Sigma(s)$ linear, zeitinvariant und asymptotisch stabil ist, gilt:

$$\lim_{t \rightarrow \infty} y_{\text{transient}}(t) \rightarrow 0 \Rightarrow y(t) \rightarrow y_{\infty}$$

Die asymptotische Systemantwort ist ein verstärkter und phasenverschobener cosinus bei derselben Frequenz wie der Eingang:

$$y_{\infty}(t) = m(\omega) \cdot \alpha \cdot \cos(\omega \cdot t + \phi + \varphi(\omega))$$

Die Verstärkung $m(\omega)$ und die Phasenverschiebung $\varphi(\omega)$ sind systemabhängig. Es gilt:

$$y_{\infty}(t) = |\Sigma(j\omega)| \cdot \alpha \cdot \cos(\omega \cdot t + \phi + \angle \Sigma(j\omega)) \quad (1)$$

D.h. die Verstärkung entspricht der Systemverstärkung bei der Eingangsfrequenz und die Phasenverschiebung entspricht der Systemphase bei der Eingangsfrequenz.

Wichtig! Lineare Systeme generieren keine neue Frequenzen. Dies ist in Gl. (1) ersichtlich. Es kommen immer dieselben Frequenzen aus dem System, wie in das System gehen. D.h. die Phasenverschiebung und Verstärkung ist nur frequenzabhängig und somit eine Eigenschaft des Systems.

Beispiel 1 Gegeben sei:

$$\Sigma(s) = \frac{s+8}{s^2+6s+8} = \frac{(s+8)}{(s+2) \cdot (s+4)} \quad (2)$$

Berechne die asymptotische Frequenzantwort y_{∞} auf den Eingang:

$$u(t) = \cos(8 \cdot t)$$

Ohne etwas zu rechnen, kennt man die Form von y_{∞}

$$y_{\infty} = |\Sigma(8j)| \cdot \cos(8 \cdot t + \angle \Sigma(8j))$$

Einsetzen von $s = 8j$, liefert:

$$\Sigma(8j) = \frac{(8j+8)}{(8j+2) \cdot (8j+4)} \quad (3)$$

Die Magnitude ist:

$$|\Sigma(8j)| = \frac{|(8j+8)|}{|(8j+2)| \cdot |(8j+4)|} = \frac{\sqrt{8^2+8^2}}{\sqrt{8^2+2^2} \cdot \sqrt{8^2+4^2}}$$

= 0.1534, und die Phase ist:

$$\begin{aligned} \angle \Sigma(8j) &= \angle(8j+8) - \angle(8j+2) - \angle(8j+4) \\ &= \arctan\left(\frac{8}{8}\right) - \arctan\left(\frac{8}{2}\right) - \arctan\left(\frac{8}{4}\right) \end{aligned}$$

= -1.648 rad. Die asymptotische Frequenzantwort ist folglich:

$$y_{\infty}(t) = 0.1534 \cdot \cos(8 \cdot t - 1.648)$$

Beispiel 2 Als nächstes wird mit einem Beispiel illustriert, wie man viele Frequenzen eines Systems auf einmal testen könnte. Betrachte Abb. 1. Die Masse m ist mit einer Feder mit Federkonstante k_F und einem Dämpfer mit Dämpferkonstante c_D verbunden. Das System kann mit einer Kraft $u(t)$ aktuiert werden.

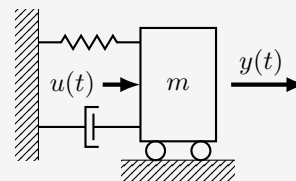


Abb. 1: Aktuiertes Feder-Dämpfer System.

Impulserhaltung:

$$\begin{aligned} m\ddot{y} &= u(t) - F_{\text{Feder}} - F_{\text{Dämpf}} \\ &= u(t) - k_F y - c_D \dot{y} \end{aligned}$$

$$\Rightarrow \Sigma(s) = \frac{1/m}{s^2 + \frac{c_D}{m}s + \frac{k_F}{m}} \quad (4)$$

Mit $k_F = 1$, $c_D = 1$, $m = 0.25$, folgt:

$$\Sigma(s) = \frac{4}{s^2 + s + 4} = \frac{4}{(s + \frac{1}{2} + \frac{\sqrt{15}}{2}j)(s + \frac{1}{2} - \frac{\sqrt{15}}{2}j)}$$

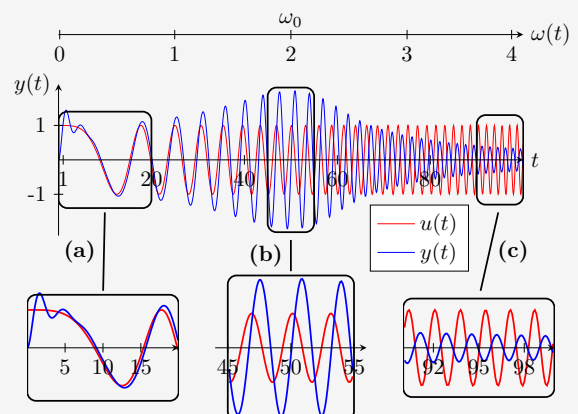
Das System ist *asymptotisch stabil* nach Lyapunov, und erfüllt somit die Bedingung um mit Methoden der Frequenzanalyse betrachtet zu werden.

Als Eingangssignal wird $u(t) = \cos(\omega(t) \cdot t)$ gewählt, wobei:

$$\omega(t) = \frac{4}{100} \frac{\text{rad}}{\text{s}} \cdot t.$$

D.h. über 100 Sekunden hinweg werden alle Frequenzen von 0 bis $4 \frac{\text{rad}}{\text{s}}$ getestet. Die Systemantwort sieht wie folgt aus:

Abb. 2: Rot: input $u(t)$, blau: output $y(t)$



Man bemerke folgende Eigenschaften des Systems:

(a) Bei $\omega = 0$ springt der Wert des Eingangssignals von 0 auf 1. Das Feder-Dämpfer System erfährt etwas ähnliches wie eine Sprungantwort. Aufgrund der Eigenwerte mit negativem Realteil, konvergiert das System schnell zur harmonischen Systemantwort. Ausserdem ist bei tiefen Frequenzen ersichtlich, dass die Phase des Ausgangs quasi der Phase des Eingangs entspricht.

(b) Die natürliche Frequenz des Systems ist bei $\omega_0 = \sqrt{4} = 2$. Da $\omega(t)$ linear gewählt wurde ($\omega \in [0, 4]$), folgt: $\omega_0 = \omega(t = 50)$. Aus der Antwort ist ersichtlich, dass $y(t)$ in der Nähe der natürlichen Frequenz die grösste Magnitude hat. Ausserdem ist der Ausgang gegenüber dem Eingang phasenverschoben.

(c) Bei sehr hohen Frequenzen kann das Feder-Dämpfer System nicht mehr schnell genug auf die Anregung reagieren, und die Amplitude seiner Schwingung nimmt ab. Zusätzlich ist bei hohen Frequenzen ersichtlich, dass die Phase des Ausgangs um ca. $-\pi$ verschoben ist.

Es gibt verschiedene Methoden, wie man die Magnitude $m(\omega)$ und Phasenantwort $\varphi(\omega)$ von Systemen $\Sigma(s)$ repräsentieren kann. Im folgenden werden zwei Methoden beschrieben: Bode Diagramme und Nyquist Diagramme.

2 Bode Diagramme

Bei Bodediagrammen werden die Magnitude $m(\omega) = |\Sigma(j \cdot \omega)|$ und die Phase $\varphi(\omega) = \angle \Sigma(j \cdot \omega)$ gegenüber einer logarithmischen Frequenzskala eingezeichnet. Dabei ist die Magnitude üblicherweise in Dezibel, und die Phase in Grad dargestellt.

Umrechnung zwischen dezimal und dezibel

$$|\Sigma(s)|_{\text{dB}} = 20 \cdot \log_{10} |\Sigma(s)|$$

$$|\Sigma(s)| = 10^{\frac{|\Sigma(s)|_{\text{dB}}}{20}}$$

Dezimalskala	Dezibelskala
100	40
10	20
5	13.97...
2	6.02...
1	0
$1/\sqrt{2}$	-3.0103
0.1	-20
0.01	-40
0	-Inf

Tabelle 1: Einige wichtige Skalenumrechnungen

Beispiel: System erster Ordnung

Viele reale Regelkreise können als Systeme erster Ordnung approximiert werden. Ein solches System reagiert auf tiefe Frequenzen, die kleiner als die Eckfrequenz (Cutoff-Frequency) $\omega_c = 1/\tau$ sind. Bei Anregungsfrequenzen höher als ω_c verhindert die Trägheit des Systems eine starke Änderung des Ausgangs. Ausserdem reagiert das System für $\omega > \omega_c$ zunehmend verzögert, wie an der Phase des Bode-Diagramms ersichtlich ist. Ein wichtiges Merkmal: die Magnitude bei $\omega_c = \frac{1}{\tau}$ ist immer bei $\frac{1}{\sqrt{2}} \approx -3$ dB.

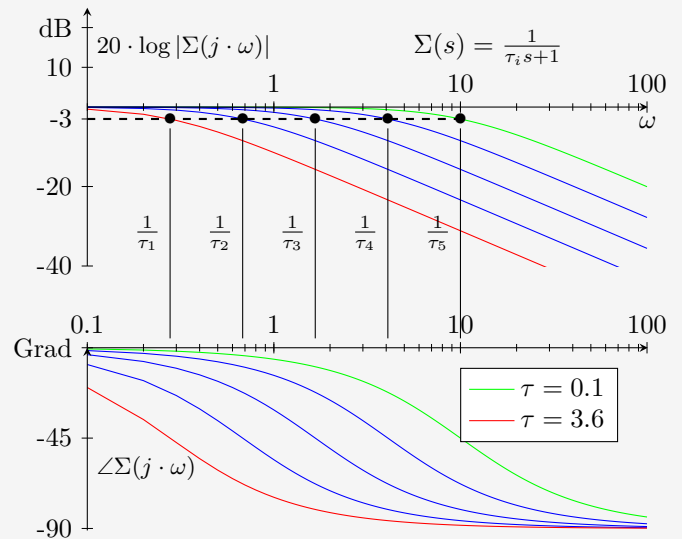


Abb. 3: Bode-Diagramme von Systemen erster Ordnung.

Beispiel: System zweiter Ordnung

Viele mechanischen Systeme zeigen resonantes Verhalten (grössere Verstärkung bei mittleren Frequenzen als bei tiefen - eine sogenannte Resonanzüberhöhung). Solche Systeme werden oft als Systeme zweiter Ordnung approximiert.

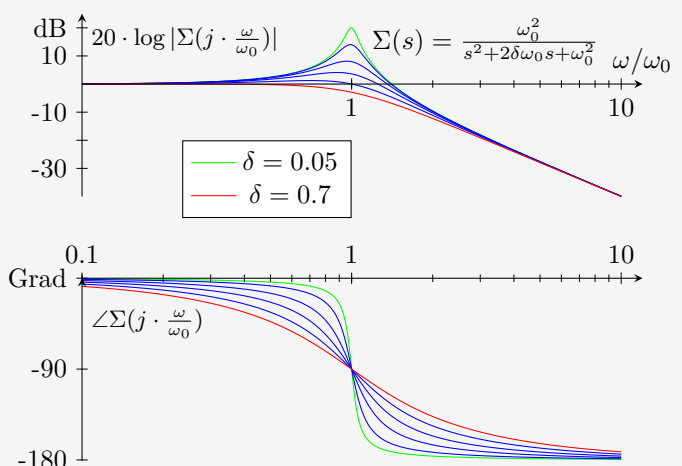


Abb. 4: Frequenznormiertes Bode-Diagramm für ein System 2. Ordnung bei unterschiedlichen Dämpfungsparametern.

Vorsicht! Die resonante Frequenz (maximale Verstärkung) ist nicht bei der natürlichen Frequenz $\omega/\omega_0 = 1$, sondern bei:

$$\omega_{\max} = \omega_0 \cdot \sqrt{1 - 2 \cdot \delta^2}, \quad 0 < \delta < \frac{1}{\sqrt{2}}$$

Jedoch gilt für kleine Dämpfungsparameter $\omega_{\max} \approx \omega_0$. Ausserdem zeigen Systeme 2. Ordnung für $\delta > \frac{1}{\sqrt{2}}$ kein resonantes Verhalten.

Vergleich Abb. 2 und Abb. 4 | Beide Abbildungen beinhalten dieselbe Information, jedoch dient das Bode-Diagramm zur vereinfachten Nutzung. Bei tiefen Frequenzen wurde in Abb. 2 identifiziert, dass quasi keine Phasenverschiebung und eine Verstärkung von 1 vorzufinden ist. Dies lässt sich direkt am Bode-Diagramm in Abb. 4 ablesen. Um die natürliche Frequenz ist in Abb. 2 eine starke Verstärkung ersichtlich. Zusätzlich könnte man eine Phase von $-\pi/2$ errahnen. Am Bode-Diagramm ist die Phase von -90 Grad direkt ersichtlich.

Einfluss von Polen und Nullstellen auf das Bode Diagramm

Standardelemente	Änderung im Bode Diagramm	
	Verstärkung	Phase
Stabiler Pol	-20 dB/dec bei ω_c	-90° bei ω_c
Instabiler Pol	-20 dB/dec bei ω_c	$+90^\circ$ bei ω_c
Minimalphasige NS	$+20 \text{ dB/dec}$ bei ω_c	$+90^\circ$ bei ω_c
Nicht-Minimalphasige NS	$+20 \text{ dB/dec}$ bei ω_c	-90° bei ω_c
Delay um τ [s] ($\forall \omega [\text{rad/s}]$)	0 dB/dec	$-180/\pi \cdot \omega \cdot \tau^\circ$

Für Pole oder Nullstellen mit Vielfachheit grösser als eins, muss sowohl die Änderung der Verstärkung wie auch die gesamte Phasenänderung mit der Vielfachheit multipliziert werden. Beispielsweise führt eine doppelte minimalphasige Nullstelle zu einer Steigung von $+40 \text{ dB/Dekade}$ und eine Phasenänderung von $+180^\circ$.

Allgemeine Übertragungsfunktionen

Gegeben sei eine Übertragungsfunktion als Serienschaltung:

$$\Sigma(s) = \Sigma_1(s) \cdot \Sigma_2(s)$$

Das Bode-Diagramm der Serienschaltung $\Sigma(s)$ lässt sich einfach aus den Bode-Diagrammen der einzelnen Subsysteme $\Sigma_1(s)$ und $\Sigma_2(s)$ konstruieren.

Die **Magnitude** der Serienschaltung $\Sigma(s)$ in Bode Diagram ergibt sich aus der Summe der Magnituden der Subsysteme $\Sigma_1(s)$ und $\Sigma_2(s)$. Dies, weil dB eine logarithmische Einheit ist und somit gilt: $\log(a \cdot b) = \log(a) + \log(b)$.

$$\begin{aligned} |\Sigma(j\omega)| &= |\Sigma_1(j\omega)| \cdot |\Sigma_2(j\omega)| \\ (|\Sigma(j\omega)|)_{\text{dB}} &= (|\Sigma_1(j\omega)| \cdot |\Sigma_2(j\omega)|)_{\text{dB}} \\ |\Sigma(j\omega)|_{\text{dB}} &= |\Sigma_1|_{\text{dB}} + |\Sigma_2|_{\text{dB}} \end{aligned}$$

Die **Phase** der Übertragungsfunktion $\angle \Sigma(j\omega)$ wird in Grad oder Bogenmass angegeben. Da $e^{j\varphi_1} \cdot e^{j\varphi_2} = e^{j(\varphi_1 + \varphi_2)}$ gilt, ergibt sich auch die Phase der Serienschaltung als Summe der Phasen der einzelnen Subsysteme:

$$\angle \Sigma = \angle \Sigma_1 + \angle \Sigma_2$$

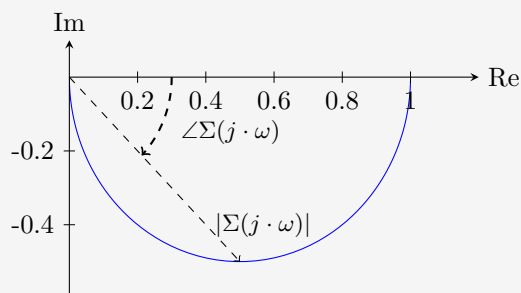
3 Nyquist Diagramme

Das Nyquist Diagramm ist die Darstellung des Frequenzganges (Magnitude und Phase über die Anregungsfrequenz ω) direkt in der komplexen Ebene.

Beispiel: System erster Ordnung

Ein allgemeines System 1. Ordnung bei der Frequenz $s = j\omega$ hat folgende Magnituden und Phasen:

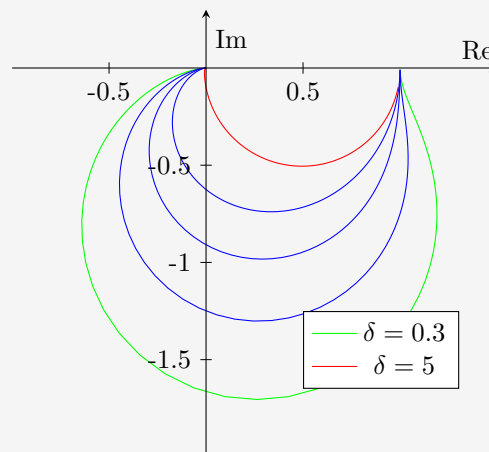
$$\begin{aligned} |\Sigma(j\omega)| &= \frac{1}{\sqrt{\tau^2 \omega^2 + 1}} \\ \angle \Sigma(j\omega) &= -\arctan(\tau \cdot \omega) \end{aligned}$$



Beispiel: System zweiter Ordnung

Ein allgemeines System 2. Ordnung bei der Frequenz $s = j\omega$ hat folgende Magnituden und Phasen:

$$\begin{aligned} |\Sigma(j \cdot \omega)| &= \frac{\omega_0^2}{\sqrt{(\omega_0^2 - \omega^2)^2 + 4 \cdot \delta^2 \cdot \omega_0^2 \cdot \omega^2}} \\ \angle \Sigma(j \cdot \omega) &= -\arctan\left(\frac{2 \cdot \delta \cdot \omega_0 \cdot \omega}{\omega_0^2 - \omega^2}\right) \end{aligned}$$



Das Nyquist Diagramm wird im Kapitel 9 im Zusammenhang mit der Stabilitätsanalyse eines Regelkreises nützlich. Wir werden sehen, dass Aussagen zum geschlossenen Regelkreis anhand des Nyquist Diagramms des offenen Regelkreises getroffen werden können.

1 Asymptotische Eigenschaften von Frequenzantworten

Betrachte die folgende Struktur einer allgemeinen Übertragungsfunktion:

$$\Sigma(s) = \frac{b_m \cdot s^m + \dots + b_1 \cdot s + b_0}{s^k \cdot (s^{n-k} + a_{n-1-k} \cdot s^{n-1-k} + \dots + a_1 \cdot s + a_0)}$$

Zwei wichtige Eigenschaften zur Frequenzantwort werden im folgenden beschrieben. Einerseits kann man aus der Übertragungsfunktion direkt die Phase bei $\omega = 0$ aus dem Systemtyp k bestimmen. Andererseits kann man das Asymptotische Verhalten der Magnitude $|\Sigma(j\omega)|$ für $\omega \rightarrow \infty$ direkt aus dem relativen Grad $r = n - m$ bestimmen.

Systemtyp k

Der Systemtyp k entspricht der Vielfachheit offener Integratoren ($\frac{1}{s^k}$) des Systems. Die Phase bei $\omega = 0$ ist als folgende Funktion definiert:

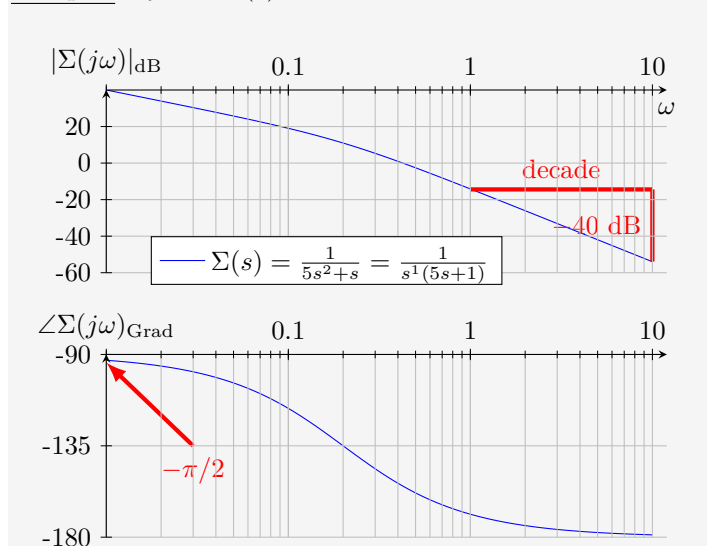
$$\angle \Sigma(0) = \begin{cases} -k \cdot \frac{\pi}{2} & , \operatorname{sgn}\left(\frac{b_0}{a_0}\right) > 0 \\ -\pi - k \cdot \frac{\pi}{2} & , \operatorname{sgn}\left(\frac{b_0}{a_0}\right) < 0 \end{cases}$$

Relativer Grad $r = n - m$

Die Steigung des Magnitudenverlauf im Bode-Diagramm konvergiert asymptotisch zu:

$$\frac{\partial |\Sigma(j\omega)|_{\text{dB}}}{\partial \log(\omega)} = -r \cdot 20 \text{ dB/decade}$$

Beispiel: System $\Sigma(s)$ mit $k = 1$, $r = 2 - 0 = 2$



2 Systemidentifikation

Generell wird zwischen drei Arten von Systemmodellen unterschieden:

White box model: Es existiert eine explizite Darstellung der Physik des Systems mit bekannten Parameterwerten.

Grey box model: Es existiert eine explizite Darstellung der Physik des Systems mit unbekannten Parameterwerten.

Black box model: Es existiert keine explizite Darstellung der Physik des Systems. Das Systemverhalten muss durch empirische Datenanalyse ermittelt werden.

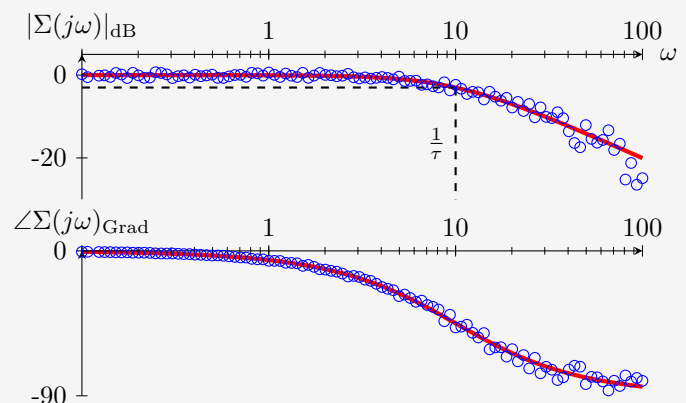
Black box Modelle werden verwendet wenn die Physik des Systems nicht genau bekannt oder zu komplex ist um einfach modelliert zu werden. Durch experimentelle Modellbildung (Systemidentifikation) kann ein Modell des Systems abgeleitet werden, das für die Reglerentwicklung erforderlich ist.

Systemidentifikation mittels Frequenzgang

Ein unbekanntes System kann wie folgt identifiziert werden:

1. Das System wird mit einem bekannten harmonischen Eingangssignal angeregt: $u(t) = \cos(\omega t)$, $\omega \in [0, \infty)$.
2. Die Verstärkung $|\Sigma(j\omega)|$ und die Phase $\angle \Sigma(j\omega)$ der Systemantwort $y_\infty = |\Sigma(j\omega)| \cdot \cos(\omega t + \angle \Sigma(j\omega))$ werden gemessen. Somit kann das Bodediagramm des Systems experimentell bestimmt werden.
3. Eine Übertragungsfunktion wird an die Daten angepasst. Hier ist es wichtig, die Auswirkung verschiedener Standardelemente auf Magnitude und Phase zu verstehen.

Beispiel zu Black-box System Identifikation



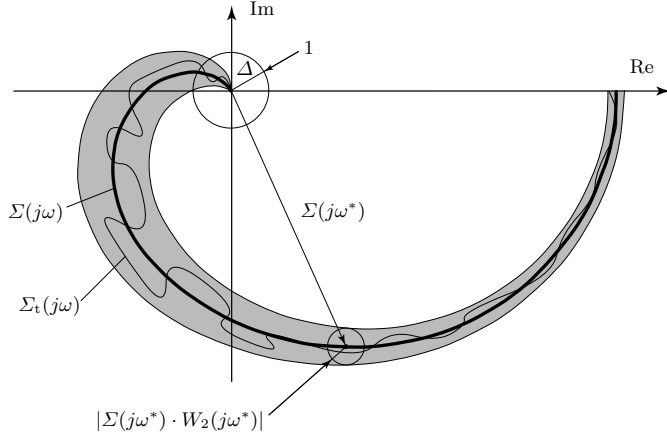
Extensives testen bei verschiedenen Frequenzen ergibt die blauen Datenpunkte. Du erkennst die Form als System erster Ordnung. Durch Auslesen der Zeitkonstante bei $\omega_c = \frac{1}{\tau} = 10$, schätzt du das System als (rot eingezeichnet):

$$\Sigma(s) = \frac{1}{\tau s + 1} = \frac{1}{0.1s + 1}$$

3 Modellunsicherheit

Ein Modell eines physikalischen Systems kann das wahre System nicht perfekt reproduzieren. Durch die Berücksichtigung der maximal zu erwartenden Modellierungsunsicherheit beim Entwurf eines Regelsystems kann robustes Verhalten garantiert werden.

3.1 Nichtparametrische Unsicherheit



Annahme: es existiert eine lineare, zeitinvariante wahre Übertragungsfunktion $\Sigma_t(s)$, die das System exakt beschreibt, die jedoch wegen Modellunsicherheiten nicht bekannt ist. Die wahre Übertragungsfunktion $\Sigma_t(s)$ liegt in der Menge \mathcal{S} :

$$\mathcal{S} = \left\{ \Sigma(s) \cdot (1 + \Delta \cdot W_2(s)) \mid |\Delta| \leq 1, \angle \Delta \in [-\pi, \pi] \right\}$$

$\Sigma(s)$: Nominelle Übertragungsfunktion, durch (imperfekte) Systemmodellierung gefunden.

Δ : Unsicherheitsgenerator: Kreis in der komplexen Ebene.

$W_2(s)$: Übertragungsfunktion der Unsicherheit: quantifiziert die frequenzabhängige Unsicherheit des Modells

Bei jeder Frequenz ω^* liegt die wahre Übertragungsfunktion $\Sigma_t(j\omega^*)$ innerhalb von einem Kreis mit Radius $|\Sigma(j\omega^*) \cdot W_2(j\omega^*)|$ um die nominelle Übertragungsfunktion $\Sigma(j\omega^*)$.

3.2 Unsicherheitsübertragungsfunktion $W_2(s)$

Es gibt mehrere Methoden, um ein Modell für die Unsicherheitsübertragungsfunktion zu bestimmen.

3.2.1 Unsicherheitsschätzung mittels Messdaten

Diese Methode verwendet Messungen am realen System, um die Unsicherheitsgrenzen des Modells davon zu bestimmen.

1. Es werden $k = 1, \dots, K$ Messungen des Frequenzgangs durchgeführt. Für jede Messung bei Frequenz ω_i , $i = 1, \dots, I$ werden die Werte $|\Sigma(j\omega_{i,k})|$ und $\angle \Sigma(j\omega_{i,k})$ identifiziert.

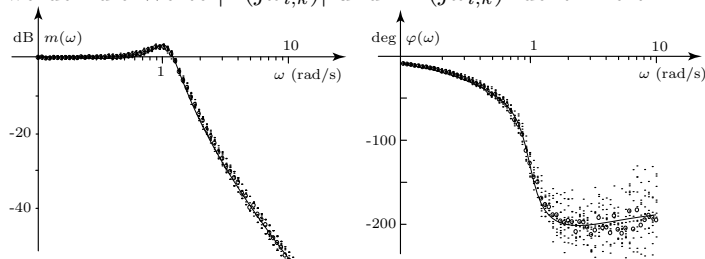


Abb. 1: Repräsentation der Messung des Ausgangs für verschiedene Frequenzen.

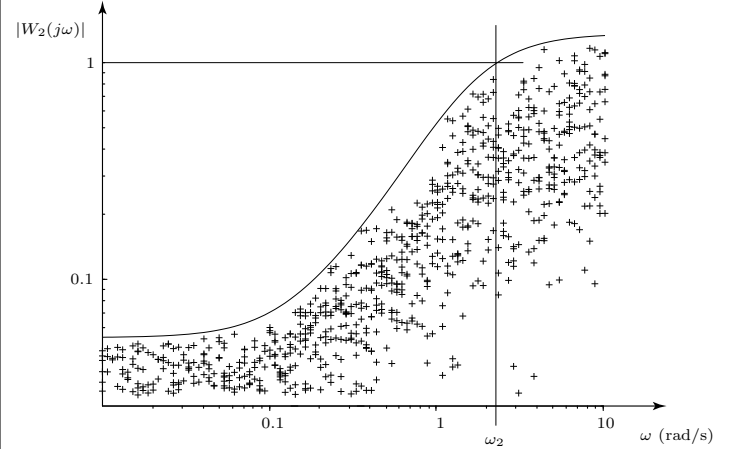
2. Eine nominelle Übertragungsfunktion $\Sigma(s)$ wird an die experimentellen Daten angepasst (analog zur Methode der Systemidentifikation).

$$\Sigma(j\omega_i) = m_i \cdot e^{j \cdot \varphi_i}$$

3. Bei jeder Frequenz ω_i sind die Werte der K Messungen von $|\Sigma(j\omega_{i,k})|$ verteilt um den Wert der nominellen Übertragungsfunktion $|\Sigma(j\omega_i)|$. Die Unsicherheitsübertragungsfunktion bildet einen Kreis mit Radius $|W_2(j\omega_i)|$ um $|\Sigma(j\omega_i)|$ so dass alle Messpunkte von $|\Sigma(j\omega_{i,k})|$ darin enthalten sind:

$$\left| \frac{\Sigma(j\omega_{i,k})}{\Sigma(j\omega_i)} - 1 \right| < |W_2(j\omega_i)| \quad k \in [1, K] \quad i \in [1, I] \quad (1)$$

Die Ungleichung Gl. (1) definiert eine Bedingung bei jeder Frequenz ω_i . Wird die Linke Seite der Ungleichung als Funktion der Frequenz dargestellt, ergibt sich:

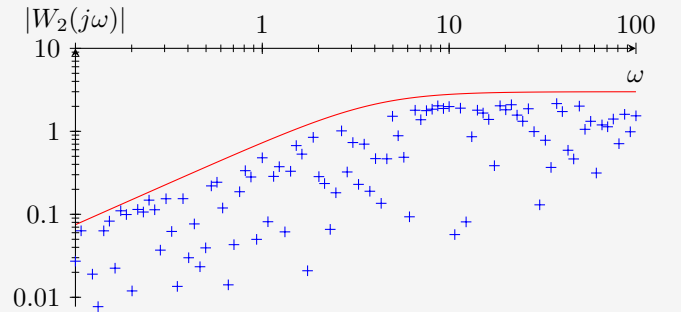


Wie erwartet, steigt die Unsicherheit bei höheren Frequenzen. Den Daten kann eine Unsicherheitsübertragungsfunktion $W_2(s)$ zugeordnet werden.

Bemerkung: Die Unsicherheitsübertragungsfunktion $W_2(s)$ enthält keine Phaseninformation.

Beispiel zu Unsicherheit

Um eine Unsicherheitsschranke W_2 für das identifizierte System aus dem vorherigen Beispiel zu erhalten, wendest du Gl. (1) an:



Du findest eine Übertragungsfunktion $W_2(s)$, die die Datenpunkte umhüllt:

$$W_2(s) = \frac{3/4 \cdot s}{1/4 \cdot s + 1}$$

1 Definitionen

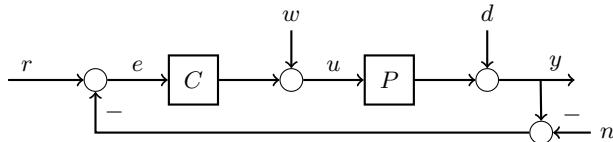
Signale im Regelkreis:

$r(t)$: Referenz (Sollzustand des Systems)

$w(t)$: Störung der Eingangsgrösse u

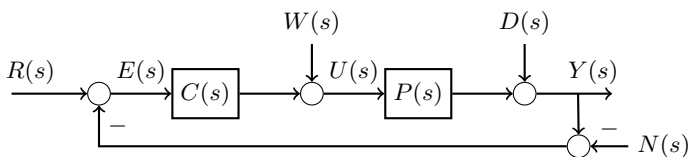
$d(t)$: Störung der Ausgangsgrösse y

$n(t)$: Sensorrauschen

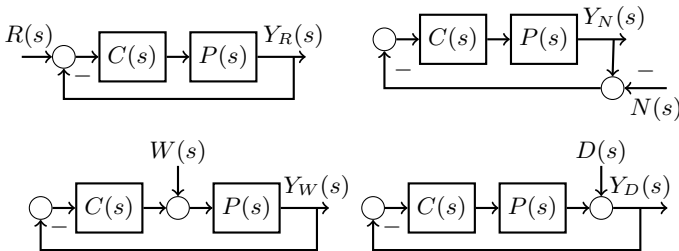


Standardregelstruktur mit Störgrössen w und d .

Das Konzept einer Übertragungsfunktion kann erweitert werden, um die Beziehung zwischen zwei beliebigen Signalen zu beschreiben. Dafür wird der Regelkreis zunächst in den Frequenzbereich transformiert:



Ziel ist, die Ausgangsgrösse $Y(s)$ als Funktion des Reglers $C(s)$, der Regelstrecke $P(s)$ und der Eingänge $R(s)$, $N(s)$, $D(s)$, und $W(s)$ zu schreiben. Durch die Linearität der Regelstruktur und unter Annahme von unkorrelierten Eingängen können die Eingangsbeiträge einzeln betrachtet werden:



Daraus folgt:

$$Y_R(s) = \frac{P(s)C(s)}{1 + P(s)C(s)} \cdot R(s), \quad Y_N(s) = \frac{P(s)C(s)}{1 + P(s)C(s)} \cdot N(s)$$

$$Y_W(s) = \frac{P(s)}{1 + P(s)C(s)} \cdot W(s), \quad Y_D(s) = \frac{1}{1 + P(s)C(s)} \cdot D(s)$$

Die gesamte Ausgangsgrösse $Y(s)$ lautet somit:

$$Y(s) = Y_R(s) + Y_N(s) + Y_W(s) + Y_D(s) \quad (1)$$

Es werden Komponenten von Gl. (1) neu definiert:

Kreisverstärkung : $L(s) = P(s) \cdot C(s)$ ($e \rightarrow y$)

Sensitivität : $S(s) = \frac{1}{1+L(s)}$ ($d \rightarrow y, r \rightarrow e$)

Komplementäre : $T(s) = \frac{L(s)}{1+L(s)}$ ($r \rightarrow y, n \rightarrow y$)

Sensitivität

Mit diesen Definitionen können wir die Beziehung zwischen allen Eingangssignalen und dem Ausgang kompakt schreiben:

$$Y(s) = S(s) \cdot [D(s) + P(s) \cdot W(s)] + T(s) \cdot [R(s) + N(s)]$$

2 Stabilität des geschlossenen Regelkreises

Für geschlossene Regelkreise muss das Konzept der Stabilität erweitert werden. Ein System ist *intern stabil*, wenn alle Übertragungsfunktionen, welche die Eingänge w, d, r in den Regelkreis auf die Ausgänge u, y, e abbilden, asymptotisch stabil sind ($\operatorname{Re}(\lambda_i) < 0, i = 1, \dots, n$). Die Beziehungen zwischen diesen Signalen sind gegeben durch:

$$\begin{bmatrix} U(s) \\ Y(s) \\ E(s) \end{bmatrix} = \begin{bmatrix} S(s) & -S(s) \cdot C(s) & S(s) \cdot C(s) \\ S(s) \cdot P(s) & S(s) & T(s) \\ -S(s) \cdot P(s) & -S(s) & S(s) \end{bmatrix} \cdot \begin{bmatrix} W(s) \\ D(s) \\ R(s) \end{bmatrix}$$

D.h. die Übertragungsfunktionen $S(s)$, $T(s)$, $S(s) \cdot C(s)$, und $S(s) \cdot P(s)$ dürfen nur asymptotisch stabile Pole haben.

Falls $P(s)$ und $C(s)$ nur asymptotisch stabile Pole haben, genügt es also, die asymptotische Stabilität von $S(s)$ und $T(s)$ zu überprüfen um die interne Stabilität zu garantieren.

3 Nyquist Theorem

Durch das Nyquist-Theorem kann die asymptotische Stabilität eines geschlossenen Regelkreissystems $T(s) = \frac{L(s)}{1+L(s)}$ durch Analyse seiner Kreisverstärkung $L(s)$ (offener Regelkreis) bestimmt werden. Dabei wird angenommen, dass keine Modellunsicherheit $W_2(s)$ vorhanden ist.

Nominelles Stabilitätskriterium von Nyquist:

Der geschlossene Regelkreis, mit Übertragungsfunktion $T(s)$, ist asymptotisch stabil, falls für $L(s)$ gilt:

$$n_c = \frac{n_0}{2} + n_+ \quad (2)$$

n_c : Anzahl Umrundungen von $L(j\omega)$ um den Punkt $(-1 + j0)$, wenn ω zwischen $(-\infty, \infty)$ variiert wird.

n_0 : Anzahl Pole von $L(s)$ mit Realteil = 0

n_+ : Anzahl Pole von $L(s)$ mit Realteil > 0

Wichtig: Stabilität nach Nyquistkriterium gilt nur, falls keine Kürzungen von instabilen Polen mit nicht-minimalphasigen Nullstellen auftreten in $L(s) = C(s) \cdot P(s)$. Andernfalls kann nicht von $L(s)$ auf die interne Stabilität des geschlossenen Regelkreises geschlossen werden.

Vorgehen zur Auswertung des Stabilitätskriteriums

1. Betrachte das Nyquistdiagramm von $L(j\omega)$ in der komplexen Ebene mit $\omega \in [0, \infty)$.

2. Spiegle das Diagramm um die reelle Achse. Die gespiegelte Kurve entspricht dem Bereich $\omega \in (-\infty, 0]$. Die kombinierte Kurven entspricht also $L(j\omega)$, $\omega \in (-\infty, \infty)$

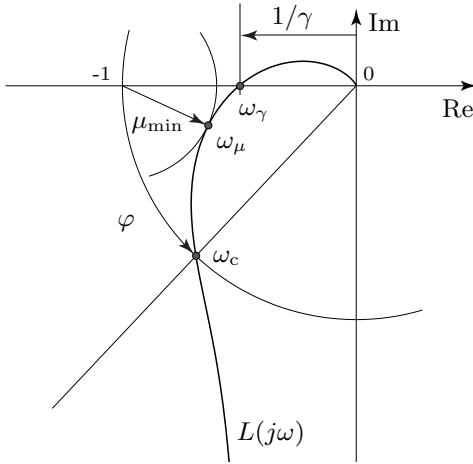
3. Zähle n_c , die Anzahl Umrundungen von $L(j\omega)$ um den Punkt $(-1 + j0)$, wenn ω von $-\infty$ bis ∞ variiert wird. Ein Umlauf in \odot wird positiv gezählt, und in \ominus negativ.

Detaillierte Anleitung zum Zählen der Umrundungen

Betrachten einen Zeiger dessen Basis auf dem Punkt $(-1 + j0)$ liegt und dessen Kopf auf $L(j\omega)$, $\omega \rightarrow -\infty$ zeigt. Folge mit dem Kopf des Zeigers der Kurve $L(j\omega)$ von $\omega \rightarrow -\infty$ bis $\omega \rightarrow +\infty$ und zähle die Anzahl Umdrehungen um den Punkt $(-1 + j0)$, die der Zeiger dabei vollzieht.

4 Verstärkungsreserve und Phasenreserve

Falls $L(s)$ eines Systems Gl. (2) nicht erfüllt, hat das resultierende $T(s) = \frac{L(s)}{1+L(s)}$ instabile Pole. Falls ein $L(s)$ Gl. (2) erfüllt, ist ein Mass für die Robustheit der Stabilität also: "Wie nahe sind wir an einer weiteren Umdrehung?"



Es werden drei Robustheitsmasse eingeführt: die Verstärkungsreserve γ , die Phasenreserve φ , und der kritische Abstand μ

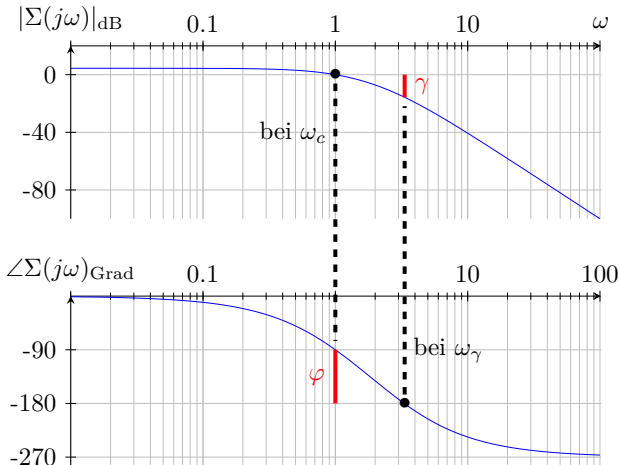
- γ : Verstärkungsreserve zu $(-1 + j0)$ bei $\angle L(j\omega) = -180^\circ$
- φ : Phasenabstand zu -180° bei der Durchtrittsfrequenz ω_c
- μ : kleinste Distanz zwischen $(-1 + j0)$ und $L(j\omega)$
- $\mu = \min_{\omega} |1 + L(j\omega)| = \frac{1}{\max_{\omega} |S(j\omega)|}$

Beispiel Das nominelle System $L(s)$ hat Phasenfehler α oder Verstärkungsfehler k gegenüber dem wahren System:

$$L_{t,\alpha}(s) = e^{-\alpha \cdot s/\omega_c} \cdot L(s), \quad L_{t,k}(s) = k \cdot L(s)$$

$L_{t,\alpha}$ ist stabil für $\alpha < \varphi$ und $L_{t,k}$ für $k < \gamma$. Falls beide Fehler gleichzeitig vorhanden sind, sind γ und φ keine guten Robustheitsmasse, sie messen beide nur eindimensional.

Auslesen der Reserven bei Bode-Diagrammen



5 Robustes Nyquist Theorem

Annahme: ein wahres Modell der linearen und zeitinvarianten Regelstrecke $P_t(s)$ existiert, ist aber wegen Modellierungsunsicherheit nicht bekannt. Stattdessen wurde ein Nominalmodell der Regelstrecke $P(s)$ und eine zugehörige multiplikative Unsicherheitsübertragungsfunktion $W_2(s)$ gefunden. Der Regler $C(s)$ ist exakt bekannt. Die wahre Kreisverstärkung des Systems $L_t(s) = P_t(s) \cdot C(s)$ ist teil der Menge $\mathcal{S}_{\mathcal{L}}$:

$$\mathcal{S}_{\mathcal{L}} = \{P(s) \cdot C(s) \cdot (1 + \Delta \cdot W_2(s)) \mid |\Delta| \leq 1, \angle \Delta \in [-\pi, \pi]\} \quad (3)$$

Es wird angenommen, dass $L(s)$ und $L_t(s)$ dieselbe Anzahl instabile (n_+) und stabile (n_0) Pole haben.

Robustes Stabilitätskriterium von Nyquist:

Das robuste Stabilitätskriterium von Nyquist wird aus dem nominalen Nyquist-Stabilitätskriterium abgeleitet: Falls nämlich das nominale Nyquist-Stabilitätskriterium Gl. (2) für jedes $L_t(j\omega) \in \mathcal{S}_{\mathcal{L}}$ in Gl. (3) erfüllt ist, dann ist der Regelkreis garantiert asymptotisch stabil.

Umsetzung des robusten Stabilitätskriteriums: Durch betrachten der grössten Modellunsicherheit, d.h. mit $|\Delta| = 1$ und für alle möglichen Richtungen (Phasen) von Δ beschränken wir unsere Überlegung auf den schlimmstmöglichen Fall, d.h.

$$L_t(s) = L(s) + L(s) \cdot W_2(s) \quad (4)$$

Betrachte Abb. 2 und vergleiche mit Gl. (4). Um zusätzliche Umrundungen des Punktes $(-1 + j0)$ garantiert zu verhindern, darf der Unsicherheitsradius $|L(j\omega) \cdot W_2(j\omega)|$ (in rot) nie grösser als $|1 + L(j\omega)|$ (in blau) werden. Daraus folgt das robuste Stabilitätskriterium nach Nyquist:

$$|L(j\omega) \cdot W_2(j\omega)| < |1 + L(j\omega)|, \quad \forall \omega \in [0, \infty)$$

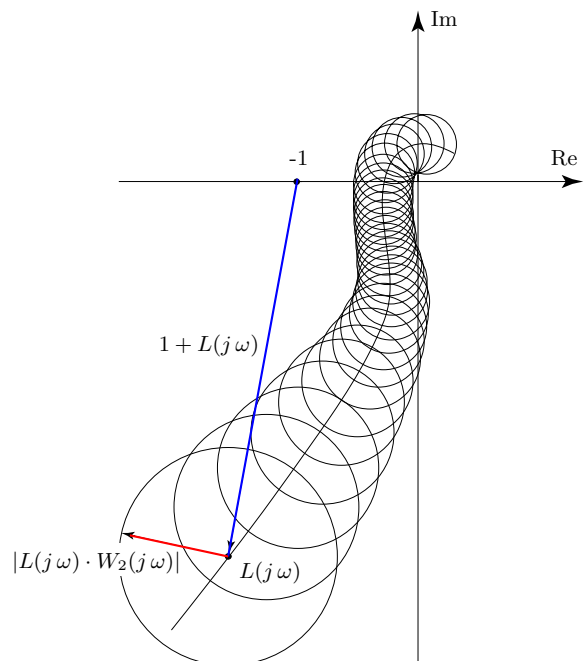


Abb. 2: Robustes Stabilitätskriterium nach Nyquist.

1 Frequenzbedingungen des geschlossenen Regelkreises

Frequenzeigenschaften von Störungen und Rauschen

Die Übertragungsfunktionen des geschlossenen Regelkreises $S(s)$ und $T(s)$ sind intrinsisch gekoppelt:

$$T(s) + S(s) = \frac{L(s)}{1 + L(s)} + \frac{1}{1 + L(s)} = 1, \forall s \in \mathbb{C} \quad (1)$$

Gl. (1) setzt voraus, dass bei gegebener Frequenz ω nur entweder $|T(j\omega)|$ oder $|S(j\omega)|$ viel kleiner als 1 sein kann.

Die Einflüsse des Rauschens (N für Noise) und der Störung (D für Disturbance) auf den Ausgang sind wie folgt:

$$Y(s) = Y_D(s) + Y_N(s) = S(s) \cdot D(s) + T(s) \cdot N(s) \quad (2)$$

Störungen werden also mit $S(s)$ auf den Ausgang übertragen und Rauschen mit $T(s)$.

Die generelle Aufgabe eines Reglers ist die gleichzeitige Unterdrückung von Rauschen und Störungen. Dies ist laut Gl. (1) nur möglich, wenn die Signale in unterschiedlichen Frequenzbändern auftreten. Glücklicherweise tritt Rauschen normalerweise bei hohen Frequenzen ($\omega > \omega_n$) auf und Störungen normalerweise bei tiefen Frequenzen ($\omega < \omega_d$). Daraus ergibt sich, dass $S(j\omega)$ für tiefe Frequenzen klein, und $T(j\omega)$ für hohe Frequenzen klein sein muss.

Niedrige Frequenzen $\omega < \omega_d$

$$|S(j\omega)| = \left| \frac{1}{1 + L(j\omega)} \right| \ll 1 \Rightarrow |L(j\omega)| \gg 1 \quad (3)$$

Hohe Frequenzen $\omega > \omega_n$

$$|T(j\omega)| = \left| \frac{L(j\omega)}{1 + L(j\omega)} \right| \ll 1 \Rightarrow |L(j\omega)| \ll 1 \quad (4)$$

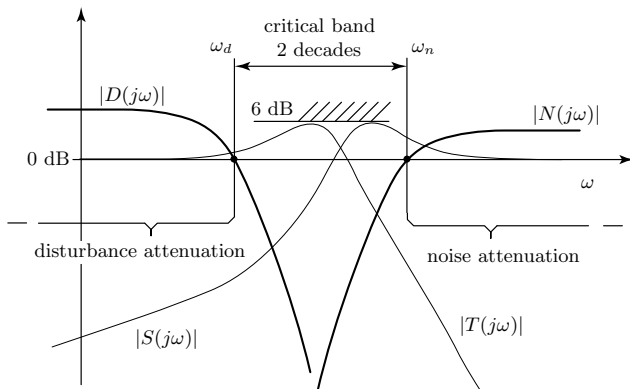


Abb. 1: Störung und Rauschen in unterschiedlichen Frequenzbändern. $S(s)$ und $T(s)$, die Gl. (3) und Gl. (4) erfüllen.

Beschränkungen der Sensitivität

Der Frequenzgang der Sensitivität $S(j\omega)$ kann durch Einstellen des Reglers $C(j\omega)$ lokal beeinflusst werden.

Global betrachtet, über alle ω , muss die Sensitivität für alle stabilen geschlossenen Regelkreise (d.h. Stabilität durch das Nyquist Theorem bestimmt) folgende Gleichung erfüllen:

$$\int_0^\infty \ln |S(j\omega)| d\omega = \pi \cdot \sum_{i=1}^{n_+} \pi_i^+, \quad (5)$$

wobei n_+ die Anzahl der instabilen Pole π^+ der Kreisverstärkung $L(s)$ ist. Falls die Kreisverstärkung $L(s)$ keine instabile Pole hat ($n_+ = 0$) vereinfacht sich Gl. (5) zu:

$$\int_0^\infty \ln |S(j\omega)| d\omega = 0 \quad (6)$$

Gleichungen Gl. (5) und Gl. (6) implizieren, dass eine Verringerung von $|S(j\omega)|$ in einem Frequenzband durch eine Erhöhung in einem anderen Frequenzband kompensiert wird.

2 Beschränkungen der Durchtrittsfrequenz

Definitionen Durchtrittsfrequenz und Bandbreite

Bei der **Durchtrittsfrequenz** ω_c , schneidet das Bode Diagramm von $L(j\omega)$ die 0 dB-Linie: $|L(j\omega_c)| = 0 \text{ dB} = 1$.

Zusätzlich definiert man die **Bandbreite des geschlossenen Regelkreises** ω_b : $|T(j\omega_b)| = -3 \text{ dB} \approx 0.7$. Die Bandbreite ist ein Mass für die höchste Frequenz des Eingangssignals, die der geschlossene Regelkreis verfolgen kann.

Die Bandbreite entspricht ungefähr der Durchtrittsfrequenz.

$$\omega_b \approx \omega_c$$

Beschränkungen der Durchtrittsfrequenz

Im Folgenden werden verschiedene Grössen eingeführt, welche die Durchtrittsfrequenz und somit die Bandbreite (oder umgekehrt) des geschlossenen Regelkreises limitieren. Die quantitativen Werte der Beschränkungen basieren auf Erfahrungswerten und sind sinnvolle Faustregeln.

Beschränkungen durch Modellunsicherheiten W_2

Aus dem robusten Stabilitätskriterium folgt:

$$\begin{aligned} |L(j\omega) \cdot W_2(j\omega)| &< |1 + L(j\omega)|, \forall \omega \in [0, \infty) \\ \Rightarrow \left| \frac{L(j\omega)}{1 + L(j\omega)} \right| &< \left| \frac{1}{W_2(j\omega)} \right| \\ \Rightarrow |T(j\omega)| &< |W_2^{-1}(j\omega)| \end{aligned} \quad (7)$$

Da die Unsicherheit $|W_2(j\omega)|$ tendenziell mit der Frequenz zunimmt (und $W_2^{-1}(j\omega)$ somit mit der Frequenz abnimmt), ist Gl. (7) eine obere Beschränkung der Bandbreite, und somit eine Beschränkung der Durchtrittsfrequenz von $|L(j\omega)|$.

Man will die Unsicherheit auf jeden Fall vermeiden. Deswegen setzt man als obere Schranke für die Durchtrittsfrequenz eine Dekade kleiner als die Unsicherheitsdurchtrittsfrequenz.

$$\omega_c \stackrel{!}{<} \frac{1}{10} \cdot \omega_2 \quad |W_2(j\omega_2)| = 1$$

Beschränkung durch eine Totzeit τ

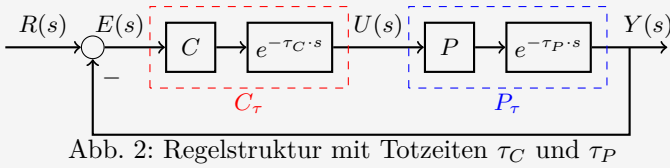


Abb. 2: Regelstruktur mit Totzeiten τ_C und τ_P

Totzeit im Regler:

Durch die Verarbeitung des Signals im Regler vergeht Zeit. Die vergangene Zeit wird als Totzeit τ_C modelliert:

$$C_\tau(s) = C(s) \cdot e^{-\tau_C \cdot s}$$

Totzeit in der Regelstrecke

Die Totzeit τ_P der Regelstrecke ist identisch definiert wie die Totzeit des Reglers:

$$P_\tau(s) = P(s) \cdot e^{-\tau_P \cdot s}$$

Ein Beispiel einer Totzeit in der Regelstrecke ist die Kommunikation zu einem Satellit. Es vergeht Zeit τ_P zwischen dem Senden der Eingangsgrösse $u(t^*)$ zur Zeit t^* und der Reaktion des Satelliten $y(t^* - \tau_P)$.

Die Übertragungsfunktion der Kreisverstärkung mit Verzögerung im Regler und der Regelstrecke ist gegeben durch:

$$L_\tau(s) = C(s) \cdot P(s) \cdot e^{-(\tau_C + \tau_P) \cdot s} = C(s) \cdot P(s) \cdot e^{-\tau \cdot s}$$

Die Totzeit induziert eine obere Grenze für die Durchtrittsfrequenz. Um die Totzeitfrequenz gut zu vermeiden wird als Grenze die halbe Totzeitfrequenz gewählt.

$$\rightarrow \omega_c \stackrel{!}{<} \frac{1}{2} \cdot \omega_\tau = \frac{1}{2} \cdot \frac{1}{\tau}$$

Der Faktor von $\frac{1}{2}$ ist ein Erfahrungswert. Falls möglich sollte man einen konservativen Vorfaktor von $\frac{1}{5}$ wählen.

Beschränkungen durch nicht-minimalphasige (NMP) Nullstellen $\omega_{\zeta+}$

Gegeben sei eine Regelstrecke $P(s) = \frac{n(s)}{d(s)}$ mit mindestens einer nicht-minimalphasigen Nullstelle. Um die Wirkung der Nullstellen zu veranschaulichen, wählt man einen konstanten Regler $C(s) = k_p$, $k_p \in \mathbb{R}$.

$$S(s) = \frac{d(s)}{d(s) + k_p \cdot n(s)}, \quad T(s) = \frac{k_p \cdot n(s)}{d(s) + k_p \cdot n(s)}$$

Wenn $k_p \rightarrow \infty$ strebt, nähern sich die Pole von $S(s)$ und $T(s)$, gegeben durch $d(s) + k_p \cdot n(s) = 0$, an die Lösung von $n(s) = 0$. Da $n(s)$ mindestens eine nicht-minimalphasige Nullstelle hat, wird das System bei $k_p = k_{p,\text{crit}}$ instabil. Dies impliziert, dass die Bandbreite durch einen Regler k_p nicht beliebig hoch gewählt werden kann. Daraus folgt:

$$\omega_c \stackrel{!}{<} \frac{1}{2} \cdot \omega_{\zeta+} \quad (\text{konservativer mit } \frac{1}{5} \text{ als Faktor}) \quad (8)$$

wobei $\omega_{\zeta+}$ der "langsamsten" nicht-minimalphasigen Nullstelle entspricht (kleinster positiver Realteil).

Beschränkungen durch instabile Pole π^+

I: Instabile Pole π^+ ohne Modellierungsunsicherheit

Die Wirkung von instabilen Polen wird anhand eines einfachen Beispiels gezeigt. Gegeben sei eine Regelstrecke $P(s) = \frac{1}{s - \pi^+}$, $\pi^+ \in \mathbb{R}_+$ welche mit einem konstanten Regler $C(s) = k_p$, $k_p \in \mathbb{R}$ stabilisiert werden soll.

Aus $n_c = \frac{n_0}{2} + n_+$ mit $n_0 = 0$ und $n_+ = 1$ folgt das eine Umkreisung von $(-1 + j0)$ in \odot notwendig ist um das System zu stabilisieren. Dies kann durch einen Regler $k_p > |\pi^+|$ erreicht werden.

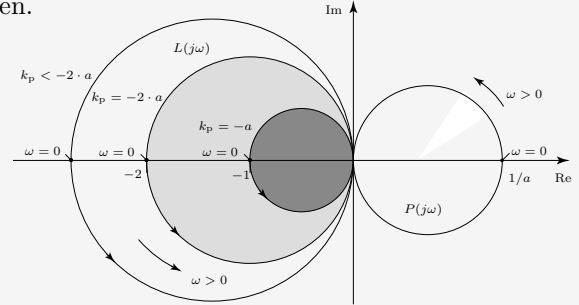


Abb. 3: Nyquistdiagramm der Regelstrecke $P(j\omega)$ und Kreisverstärkung $L(j\omega)$

Daraus folgt eine untere Schranke für die Durchtrittsfrequenz:

$$\omega_c > 2 \cdot \omega_{\pi^+} \quad (\text{konservativer mit 5 als Faktor}) \quad (9)$$

wobei ω_{π^+} der "schnellste" instabile Pol von $L(s)$ ist (Pol mit dem grössten positiven Realteil).

II: Instabile Pole π^+ mit Modellierungsunsicherheit

Zusätzlich zu Gl. (9) müssen wir folgende Bedingung für alle instabilen Pole π_i^+ prüfen:

$$|W_2(\pi_i^+)| < 1 \quad \forall i \quad (10)$$

3 Zusammenfassung der Beschränkungen

Alle oben beschriebenen Einschränkungen können als eine Reihe von Bedingungen an die Durchtrittsfrequenz zusammengefasst werden. Da es mehrere untere Schranken und obere Schranken für die Durchtrittsfrequenz gibt, muss man das engstmögliche Frequenzband nehmen, sodass alle Bedingungen gleichzeitig erfüllt sind:

$$\max \left\{ 10 \cdot \omega_d, 2 \cdot \omega_{\pi^+} \right\} < \omega_c < \min \left\{ \frac{1}{10} \cdot \omega_n, \frac{1}{10} \cdot \omega_2, \frac{1}{2} \cdot \omega_\tau, \frac{1}{2} \cdot \omega_{\zeta+} \right\}$$

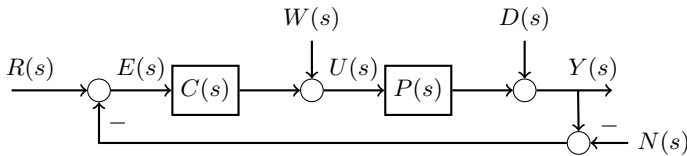
\uparrow konservativer mit 5 \uparrow konservativer mit $\frac{1}{5}$

Wichtig:

1. Es handelt sich bei den besprochenen Beschränkungen nicht um Design-Spezifikationen an den Regler, sondern um eine Überprüfung, ob das Regelproblem überhaupt sinnvoll gelöst werden kann. Wenn es kein gültiges Frequenzband für ω_c gibt, existiert kein Regler, der das Problem sinnvoll löst.
2. Falls $D(j\omega)$ oder $N(j\omega)$ mit 20 dB/dek abfallen oder ansteigen, garantiert der Sicherheitsfaktor 10 eine Verstärkung der Signale nicht grösser als 10% (\leftrightarrow Unterdrückung um 90%).
3. Bei instabilen Polen oder NMP Nullstellen geht es nicht um die Unterdrückung von Signalen, sondern um die Garantie von Stabilität. Diese wird nominell auch mit einem kleineren Sicherheitsabstand von Faktor von 2 oder 5 erreicht.

1 Statischer Nachlauffehler

Bis jetzt wurde jeweils $Y(s)$ als Funktion der Eingänge $R(s)$, $W(s)$, $D(s)$, und $N(s)$ betrachtet. Im Folgenden wird besprochen, wie sich der Fehler $e(t)$ im eingeschwungenen Zustand verhält. Dazu wird der Regelkreis im Frequenzbereich betrachtet:



Ähnlich wie bei der Ausgangsgrösse $Y(s)$ kann man $E(s)$ als Funktion der Eingänge beschreiben¹:

$$\begin{aligned} E(s) &= E_R(s) + E_N(s) + E_D(s) + E_W(s) \\ &= S(s)R(s) + S(s)N(s) - S(s)D(s) - S(s)P(s)W(s) \\ &= S(s) \cdot [R(s) + N(s) - D(s) - P(s) \cdot W(s)] \end{aligned}$$

Man betrachtet Referenzen und Störungen die als Sprünge $h(t)$ auf den Fehler abgebildet werden²:

$$h(t) = 1, t > 0 \rightarrow H(s) = \frac{1}{s}$$

Mit dem Endwerttheorem (final value theorem) kann man den Fehler auf eine Sprungantwort nach langer Zeit berechnen:

$$e_{\infty}^h = \lim_{t \rightarrow \infty} e^h(t) = \lim_{s \rightarrow 0^+} s \cdot S(s) \cdot \frac{1}{s} = \lim_{s \rightarrow 0^+} S(s) = S(0)$$

Man schreibt $S(0)$ als Funktion des offenen Regelkreises $L(0)$:

$$e_{\infty}^h = S(0) = \frac{1}{1 + L(0)} \quad (1)$$

$L(0)$ hängt vom Systemtyp k der Kreisverstärkung $L(s)$ ab:

$$L(s) = \frac{b_m \cdot s^m + \dots + b_1 \cdot s + b_0}{s^k \cdot (s^{n-k} + a_{n-1-k} \cdot s^{n-1-k} + \dots + a_1 \cdot s + a_0)} \quad (2)$$

Aus Gl. (1) und Gl. (2) ist ersichtlich, dass für $L(0)$ und somit e_{∞}^h zwei Fälle vorliegen:

$$\begin{aligned} k = 0 &\Rightarrow L(0) \rightarrow \frac{b_0}{a_0} \Rightarrow e_{\infty}^h = \frac{a_0}{a_0 + b_0} \\ k > 0 &\Rightarrow L(0) \rightarrow \infty \Rightarrow e_{\infty}^h \rightarrow 0 \end{aligned}$$

In anderen Worten kann ein System mit Systemtyp $k > 0$ einem Sprung nach langer Zeit fehlerfrei folgen, und somit Sprungartige Störungen unterdrücken. Für ein Systemtyp $k = 0$ wird die Antwort $y(t)$ des Systems vom angewandten Sprung abweichen.

¹Die Eingänge werden dabei als unkorreliert behandelt und ihr Einfluss wird individuell betrachtet.

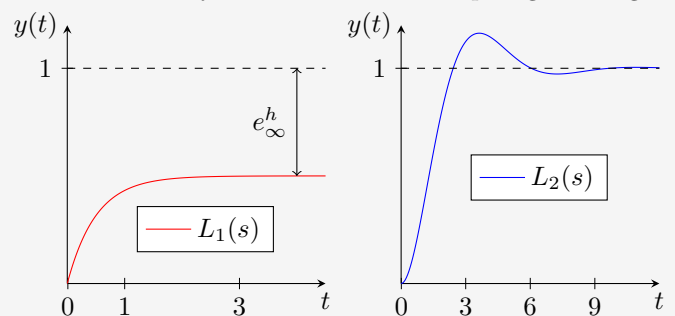
² $R(s)$ und $D(s)$ werden durch $S(s)$ abgebildet. Das Rauschen $n(t)$ hat in der Regel Mittelwert 0 und induziert dadurch im Mittel keinen Fehler.

Beispiel: Sprungantworten und Systemtyp:

$$L_1(s) = \frac{1}{s+1} \quad (k=0), \quad L_2(s) = \frac{1}{s(s+1)} \quad (k=1)$$

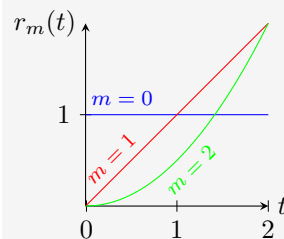
$$T_1(s) = \frac{s+1}{s+2}, \quad T_2(s) = \frac{s(s+1)}{s^2+s+1}$$

Das erste System hat Systemtyp $k=0$ und weist somit einen Fehler in der Sprungantwort auf: $S(0) = \frac{1}{1+L_1(0)} = \frac{1}{2}$. Der zweite offene Regelkreis $L_2(s)$ ist Systemtyp $k=1$. ($L_2(s)$ strebt für $s \rightarrow 0$ linear gegen ∞ .) Daraus folgt, dass das zweite System fehlerfrei zum Sprung konvergiert.



Beispiel: Das Konzept des statischen Nachlauffehlers kann auch auf Referenzen höherer Ordnung³ erweitert werden:

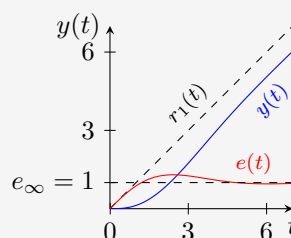
$$r_m(t) = \frac{1}{m!} \cdot t^m, t \geq 0, m \in \{0, 1, 2, \dots\}.$$



Ähnlich wie beim speziellen Fall der Sprungantwort ergibt sich mit Hilfe der Ordnung der Referenz m , und der Anzahl der offenen Integratoren von $L(s)$ (Systemtyp k) eine Fallunterscheidung:

$$\begin{aligned} m < k &\rightarrow e_{\infty} = \lim_{s \rightarrow 0^+} \frac{1}{s^m \cdot (1+L(s))} = 0 \\ m > k &\rightarrow e_{\infty} = \lim_{s \rightarrow 0^+} \frac{1}{s^m \cdot (1+L(s))} = \infty \\ m = k &\rightarrow e_{\infty} = \lim_{s \rightarrow 0^+} \frac{S(s)}{s^k} \notin \{0, \infty\} \end{aligned}$$

Beispielhaft wird eine Rampenantwort ($m=1$) des zweiten Systems $T_2(s)$ ($k=1$) aus dem vorherigen Beispiel betrachtet. Da $k=m$ gilt, hat die Rampenantwort einen statischen Nachlauffehler.



$$S_2(s) = \frac{s(s+1)}{s^2+s+1}$$

$$\Rightarrow e_{\infty} = \lim_{s \rightarrow 0^+} \frac{S(s)}{s^k} = 1$$

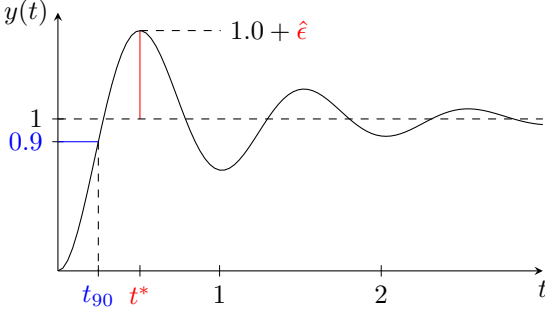
³Die Referenz $r_0(t)$ entspricht der Sprungantwort $h(t)$, $r_1(t)$ entspricht der Rampe $p(t)$, $r_2(t)$ einer quadratisch ansteigenden Funktion, usw.

2 Spezifikationen basierend auf Systemen 2. Ordnung

Es wird angenommen, dass der geschlossene Regelkreis $T(s)$ einem System zweiter Ordnung entspricht:

$$T(s) = \frac{\omega_0^2}{s^2 + 2 \cdot \delta \cdot \omega_0 \cdot s + \omega_0^2}.$$

Dies ist für sinnvolle geschlossene Regelkreise eine gute Annahme, es verlangt asymptotische Stabilität und erlaubt ein Überschwingen. Der geschlossene Regelkreis $T(s)$ zweiter Ordnung soll Spezifikationen in der Anstiegszeit t_{90} und im relativen Überschwingen $\hat{\epsilon}$ erfüllen:



Die Spezifikationen von $\hat{\epsilon}$ und t_{90} können erfüllt werden, in dem man Anforderungen an die typischen Parameter eines Systems 2ter Ordnung aufstellt:

$$\delta = \frac{-\ln(\hat{\epsilon})}{\sqrt{\pi^2 + \ln^2(\hat{\epsilon})}}, \quad \omega_0 = (0.14 + 0.4 \cdot \delta) \cdot \frac{2 \cdot \pi}{t_{90}}$$

Im nächsten Schritt werden die Spezifikationen an $T(s)$ in Anforderungen an die Kreisverstärkung $L(s)$ umgewandelt, unter Anwendung von $T(s) = \frac{L(s)}{1+L(s)}$.

Die Anforderungen des geschlossenen Regelkreises können in Anforderungen an die Durchtrittsfrequenz ω_c und die Phasenreserve φ der Kreisverstärkung $L(s)$ umformuliert werden:

$$\omega_c = \omega_0 \cdot \sqrt{\sqrt{4 \cdot \delta(\hat{\epsilon})^4 + 1} - 2 \cdot \delta(\hat{\epsilon})^2}$$

$$\varphi = \frac{\pi}{2} - \arctan \left(\frac{\sqrt{4 \cdot \delta(\hat{\epsilon})^4 + 1} - 2 \cdot \delta(\hat{\epsilon})^2}{2 \cdot \delta(\hat{\epsilon})} \right)$$

Die obigen Gleichungen können für $0.45 < \delta < 1$ sehr gut mit den folgenden vereinfachten Zusammenhängen angenähert werden:

$$\omega_c \approx \frac{1.7}{t_{90}}$$

$$\varphi \approx 71^\circ - 117^\circ \cdot \hat{\epsilon}$$

Somit sind Spezifikationen des geschlossenen Regelkreises $T(s)$ in Anforderungen an den offenen Regelkreis $L(s)$ umformuliert worden.

Geschlossene Regelkreise mit Charakteristiken, welche einer Dämpfung ausserhalb des Bereiches $0.45 < \delta < 1$ entsprechen sind in der Praxis nicht relevant, weil sie entweder sehr stark überschwingen oder extrem langsam sind.

3 Frequenzbereich - Spezifikationen

Die Störung $D(s)$ und das Rauschen $N(s)$ werden durch die Sensitivität $S(s)$ und durch die komplementäre Sensitivität $T(s)$ auf den Ausgang abgebildet:

$$Y(j\omega) = S(j\omega) \cdot D(j\omega) + T(j\omega) \cdot N(j\omega)$$

Um die Auswirkung von Störungen und Rauschen um die Durchtrittsfrequenz ω_c zu minimieren, beschränkt man den Maximalwert von $S(s)$ und $T(s)$.

$$\|S\|_\infty < S_{\max}, \quad \|T\|_\infty < T_{\max}, \quad S_{\max}, T_{\max} > 1, \quad (3)$$

wobei per Definition $\|\Sigma\|_\infty = \max_\omega |\Sigma(j\omega)|$.

Die Bedingungen in Gl. (3) werden in Anforderungen an die Kreisverstärkung $L(s)$ umgewandelt:

$$\|S\|_\infty < S_{\max} \Leftrightarrow L(j\omega) \notin \left\{ |1+z| \leq \frac{1}{S_{\max}} \mid z \in \mathbb{C} \right\} \quad (4)$$

$$\|T\|_\infty < T_{\max} \Leftrightarrow L(j\omega) \notin \left\{ \left| \frac{T_{\max}^2}{T_{\max}^2 - 1} + z \right| \leq \frac{T_{\max}}{T_{\max}^2 - 1} \mid z \in \mathbb{C} \right\} \quad (5)$$

Die geometrische Interpretation von Gl. (4) ist, dass $L(j\omega)$ nicht in einen in -1 zentrierten Kreis mit Radius $\frac{1}{S_{\max}}$ eintreten darf.

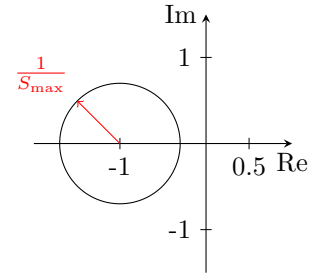


Abb. 1: Darstellung der nicht zulässigen Region aus Gl. (4)

Die geometrische Interpretation von Gl. (5) ist, dass $L(j\omega)$ nicht in einen in $\frac{-T_{\max}^2}{T_{\max}^2 - 1}$ zentrierten Kreis mit Radius $\frac{T_{\max}}{T_{\max}^2 - 1}$ eintreten darf.

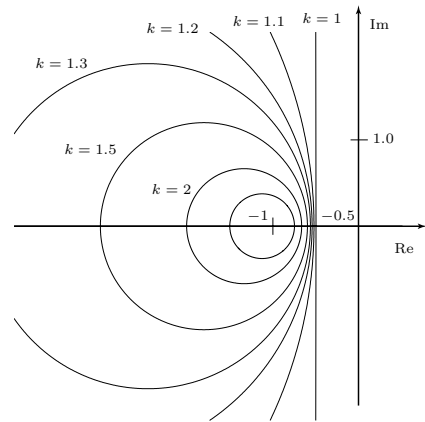


Abb. 2: Darstellung der nicht zulässigen Region aus Gl. (5) mit $k = T_{\max}$

Wenn S_{\max} und T_{\max} verringert werden, wird eine zunehmende Teilmenge der komplexen Ebene für $L(j\omega)$ nicht zulässig. Für $T_{\max} \rightarrow 1$ ist die gesamte komplexe Ebene links von $-\frac{1}{2}$ ausgeschlossen.

1 PID-Regler

Bisher wurde in Beispielen jeweils ein Proportionalregler (P-Regler) $C(s) = k_p$ verwendet. In diesem Kapitel wird der P-Regler durch einen Integral-Term (I-Teil), und durch einen Derivative-Term (D-Teil) erweitert. Dieser PID-Regler wird in der Praxis sehr häufig verwendet.

Die Reglerstruktur wird zunächst im Zeitbereich betrachtet:

$$u_{\text{PID}}(t) = k_p \cdot \left(\underbrace{e(t)}_{\text{P-Term}} + \underbrace{\frac{1}{T_i} \cdot \int_0^t e(\tau) d\tau}_{\text{I-Term}} + \underbrace{T_d \cdot \frac{d}{dt} e(t)}_{\text{D-Term}} \right)$$

Der I-Term wird betragsmässig grösser, je länger ein einseitiger Fehler (z.B. $e(t) > 0$) vorhanden ist. Der D-Term wirkt auf schnelle Änderungen im Fehlersignal. Ein Nachteil des D-Terms ist, dass er Rauschen auf dem Fehlersignal $e(t)$ verstärkt.

Eine Transformation in den Frequenzbereich ergibt:

$$C_{\text{PID}}(s) = k_p \cdot \left(1 + \frac{1}{T_i \cdot s} + T_d \cdot s \right) = \frac{U(s)}{E(s)}$$

Wie bereits erwähnt wird $u(t)$ durch den D-Term sehr empfindlich auf Rauschen. Man kann im Frequenzbereich hohe Frequenzen ganz einfach unterdrücken, indem man eine hochfrequente doppelte Nullstelle an den Regler hängt. Dieser Term wird roll-off Term genannt:

$$C_{\text{PID}}(s) = k_p \cdot \left(1 + \frac{1}{T_i \cdot s} + T_d \cdot s \right) \cdot \underbrace{\frac{1}{(\tau \cdot s + 1)^2}}_{\text{roll-off}}$$

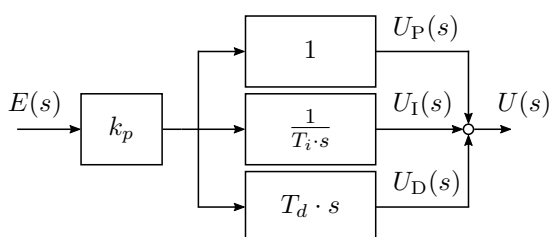
Als Übertragungsfunktion ergibt sich:

$$C_{\text{PID}}(s) = k_p \cdot \left(\overbrace{\frac{T_d \cdot T_i \cdot s^2 + T_i \cdot s + 1}{T_i \cdot s}}^{\text{kausal}} \right) \cdot \frac{1}{(\tau \cdot s + 1)^2}$$

nicht kausal

Ohne den roll-off Term wäre die Übertragungsfunktion $\frac{U(s)}{E(s)}$ nicht kausal und entsprechend nicht praktisch realisierbar. Um $u(t)$ ohne roll-off zu berechnen, bräuchte man Zukunftswerte des Fehlersignals $e(t)$.

PID-Regler in Standardform im Frequenzbereich



Proportionales Verhalten (P-Term)

$$u_P(t) = k_p \cdot e(t), \quad U_P(s) = k_p \cdot E(s)$$

Der P-Term reagiert auf den momentanen Wert des Fehlers $e(t)$. Die Stärke der Reaktion ist proportional zur Grösse des momentanen Fehlers.

Integratives Verhalten (I-Term)

$$u_I(t) = \frac{k_p}{T_i} \cdot \int_0^t e(\tau) d\tau, \quad U_I(s) = \frac{k_p}{T_i} \cdot \frac{1}{s} \cdot E(s)$$

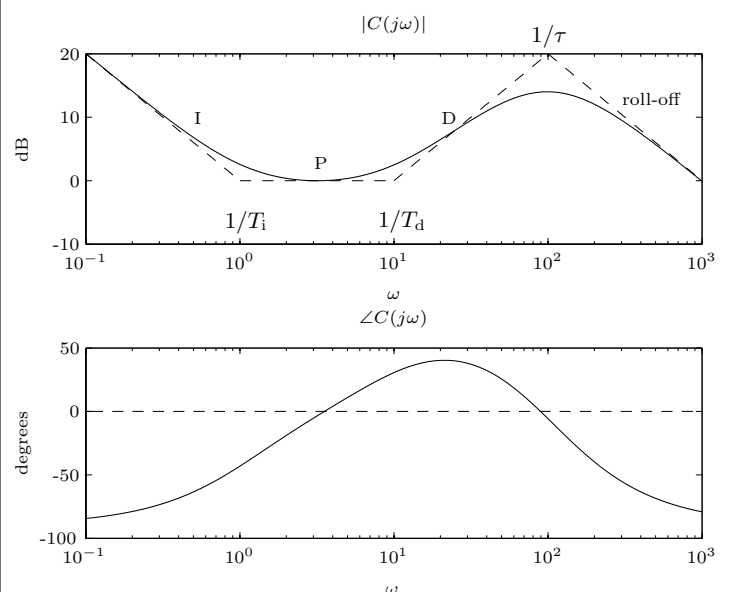
Der I-Term reagiert zum Zeitpunkt t proportional auf den kumulierten Fehler, für $t \in [0, t]$. Falls ein statischer Nachlauffehler vorhanden ist, wird dieser aufintegriert, und der Reglerausgang wird immer grösser, bis kein Fehler mehr vorhanden ist. Ein Nachteil des Integrators ist, dass der Reglerausgang theoretisch beliebig gross werden kann.

Derivatives Verhalten (D-Term)

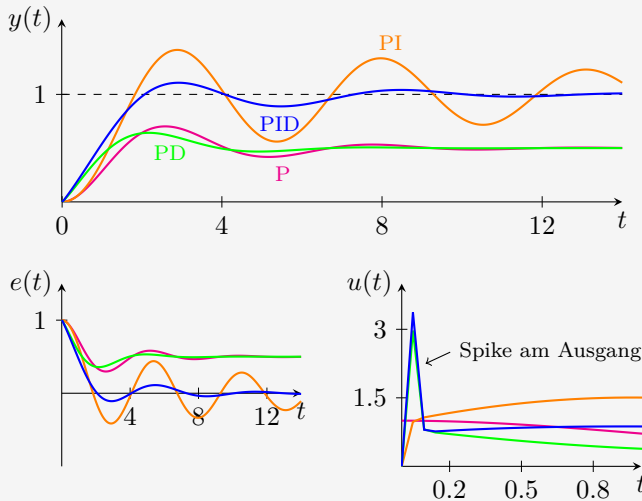
$$u_D(t) = k_p \cdot T_d \cdot \frac{d}{dt} e(t), \quad U_D(s) = k_p \cdot T_d \cdot s \cdot E(s)$$

Der D-Term wirkt antizipierend, er reagiert zum Zeitpunkt t auf die momentane Änderungsrate des Fehlers. Der D-Term wirkt wie ein Dämpfer gegen ein schnelles Erhöhen oder Verringern des Fehlers. Eine starke Änderung im Fehler resultiert in einem erhöhten Reglerausgang. Falls die Änderung zu stark ist, kann der gewünschte Reglerausgang grösser als der grösstmögliche Eingang eines Systems werden.

Bodediagramm eines PID-Reglers mit roll-off



Beispiel: Regelung eines Systems zweiter Ordnung



Es ist ersichtlich, dass ein statischer Fehler durch den Integrator eliminiert werden kann. Der D-Term ermöglicht es, schneller auf die Fehleränderung zu reagieren, jedoch wird dadurch der Reglerausgang auch grösser.

PID-Regler Parameter Tuning nach Ziegler Nichols

Die Parameter k_p , T_i , und T_d können durch extensives Testen des Systems bestimmt werden. Ein anderer Ansatz ist, der von Ziegler-Nichols. Hier geht man davon aus, dass das System $P(s)$ ein System erster Ordnung mit zusätzlicher relativ kleiner Totzeit ist:

$$P(s) \approx \frac{k}{\tau \cdot s + 1} \cdot e^{-T \cdot s}, \quad \text{wobei: } \frac{T}{T + \tau} \stackrel{!}{<} 0.3$$

Zur Bestimmung der Ziegler-Nichols Parameter startet man mit einem reinen P-Regler und erhöht die Verstärkung k_p soweit, bis der geschlossene Regelkreis grenzstabil wird bei der Verstärkung k_p^* (Pole von $T(s)$ auf der imaginären Achse). Falls die Modellannahme ungefähr stimmt, oszilliert das grenzstabile System bei k_p^* mit einer Periode von T^* . Man kann diese Parameter des grenzstabilen System in folgende Tabellen einsetzen um Reglerparameter für verschiedene PID Kombinationen zu erhalten.

Regler	k_p	T_i	T_d
P	$0.5 \cdot k_p^*$	$\infty \cdot T^*$	$0 \cdot T^*$
PI	$0.45 \cdot k_p^*$	$0.85 \cdot T^*$	$0 \cdot T^*$
PD	$0.55 \cdot k_p^*$	$\infty \cdot T^*$	$0.15 \cdot T^*$
PID	$0.60 \cdot k_p^*$	$0.50 \cdot T^*$	$0.125 \cdot T^*$

2 Iterative Loop Shaping

Ein System, das mit einem PID-Regler ausgelegt wird, erfüllt unter Umständen nicht alle Designspezifikationen. Um gewisse Frequenzbänder nach Wunsch abzuändern, kann man einen beliebigen Regler zum Beispiel mit Lead/Lag-Elemente erweitern oder einen Regler von Grund auf neu erstellen.

Lead-Lag Elemente erster Ordnung

Der Term 'Lead-Lag' bezeichnet zwei Arten von Systemen mit gleicher Struktur und den zwei Parametern α und T :

$$C(s) = \frac{T \cdot s + 1}{\alpha \cdot T \cdot s + 1}, \quad \alpha, T \in \mathbb{R}_+ \quad (1)$$

Der Wert von α definiert ob es sich um ein Lead- oder ein Lag-Element handelt:

$$\begin{aligned} 0 < \alpha < 1 &\Leftrightarrow \text{Lead-Element} \\ 1 < \alpha &\Leftrightarrow \text{Lag-Element} \end{aligned}$$

Die Parameter α und T werden gezielt gewählt, sodass bei der Frequenz $\hat{\omega}$ eine maximale Phasenänderung von $\hat{\varphi}$ vorliegt:

$$\alpha = \left(\sqrt{\tan^2(\hat{\varphi}) + 1} - \tan(\hat{\varphi}) \right)^2, \quad T = \frac{1}{\hat{\omega} \cdot \sqrt{\alpha}}$$

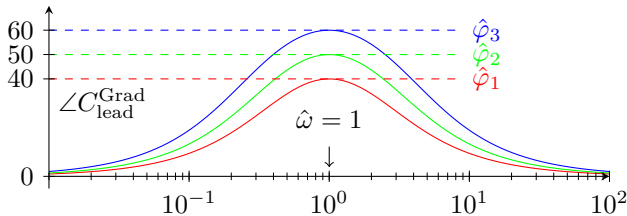
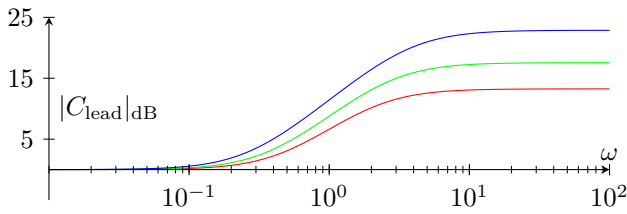


Abb. 1: Lead-Element für $\hat{\omega} = 1$ und $\hat{\varphi} = \{40, 50, 60\}$

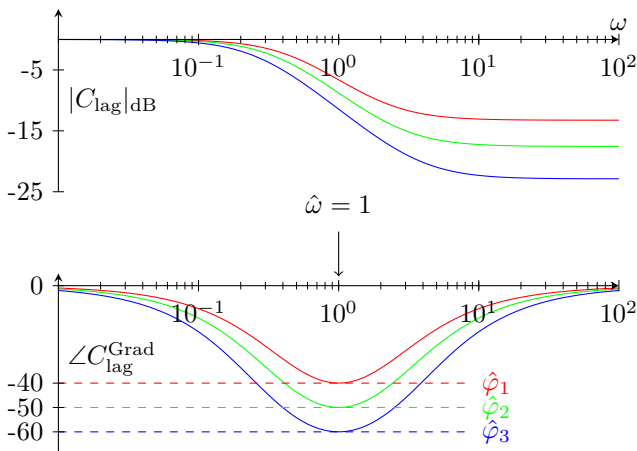


Abb. 2: Lag-Element für $\hat{\omega} = 1$ und $\hat{\varphi} = \{-40, -50, -60\}$

Ein Lag-Element mit $-\hat{\varphi}$ entspricht einer Spiegelung des Magnituden- und des Phasendiagramms des Lead-Elements mit $\hat{\varphi}$.

Lead-Lag Elemente zweiter Ordnung

Die Verwendung eines Elements erster Ordnung beeinflusst Frequenzen in der grösseren Umgebung von $\hat{\omega}$. Die Idee eines Elements zweiter Ordnung ist, dass der gewünschte Effekt an einer bestimmten Frequenz besser isoliert ist. Die Struktur erfordert die Parameter κ, ϵ , und ω_0 :

$$C(s) = k \cdot \frac{s^2 + 2 \cdot \kappa \cdot \epsilon \cdot (1 - \epsilon) \cdot \omega_0 \cdot s + (1 - \epsilon)^2 \cdot \omega_0^2}{s^2 + 2 \cdot \kappa \cdot \epsilon \cdot (1 + \epsilon) \cdot \omega_0 \cdot s + (1 + \epsilon)^2 \cdot \omega_0^2} \quad (2)$$

Zusätzlich zur Wahl der mittleren Frequenz $\hat{\omega}$ und der maximalen Phasenverschiebung $\hat{\varphi}$ kann man nun zusätzlich die Breite des Frequenzbands durch den Parameter ϵ wählen:

$$k = \frac{(1 + \epsilon)^2}{(1 - \epsilon)^2}, \quad \kappa = \frac{\cot(\hat{\varphi}/2)}{\sqrt{1 - \epsilon^2}}, \quad \omega_0 = \frac{\hat{\omega}}{\sqrt{1 - \epsilon^2}}$$

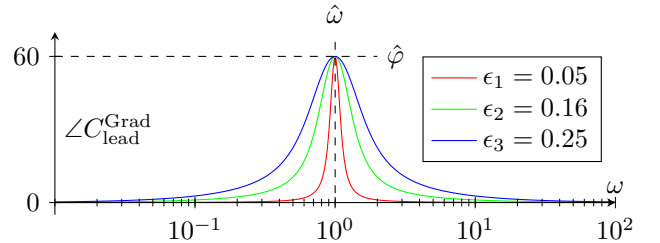


Abb. 3: Lead-Element zweiter Ordnung für $\hat{\omega} = 1$, $\hat{\varphi} = 60$, und $\epsilon = \{0.05, 0.16, 0.25\}$.

3 Inversion der Regelstrecke

Wenn die Regelstrecke $P(s) = \frac{n_p(s)}{d_p(s)}$ mit relativem Grad r asymptotisch stabil ist und nur minimalphasige Nullstellen enthält, kann ein Regler $C(s)$ gewählt werden, der die Dynamik der Regelstrecke exakt kompensiert und gleichzeitig in einer gewünschten Übertragungsfunktion $L(s)$ des offenen Regelkreises resultiert:

$$\begin{aligned} L(s) &= P(s) \cdot P(s)^{-1} \cdot \underbrace{\frac{1}{T_i \cdot s} \cdot \frac{1}{(\tau \cdot s + 1)^{r-1}}}_{C(s)} \\ \Rightarrow C(s) &= P(s)^{-1} \cdot \frac{1}{T_i \cdot s} \cdot \frac{1}{(\tau \cdot s + 1)^{r-1}} \\ C(s) &= \frac{d_p(s)}{n_p(s)} \cdot \frac{1}{T_i \cdot s} \cdot \frac{1}{(\tau \cdot s + 1)^{r-1}} \end{aligned}$$

Der Regler invertiert die Dynamik der Regelstrecke, und somit haben die Pole und Nullstellen von $P(s)$ keinen Einfluss auf die Übertragungsfunktion des offenen Regelkreises $L(s)$. Die übrigen Elemente von $C(s)$ stellen sicher, dass die gewünschte Übertragungsfunktion $L(s)$ des offenen Regelkreises resultiert:

$$L(s) = \frac{1}{T_i \cdot s} \cdot \frac{1}{(\tau \cdot s + 1)^{r-1}}$$

Mit der Verstärkung $T_i = \omega_c^{-1}$ kann die gewünschte Durchtrittsfrequenz ω_c eingestellt werden. Zusätzlich wählen wir $\tau < T_i$ und $\omega_c < \omega_2$.