

Digitale Regelung

Professor Dr.-Ing. Ferdinand Svaricek

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

Vorwort

Diese Arbeitsblätter beschreiben den wesentlichen Stoffinhalt des Moduls "DIGITALE REGELUNG", (V 2, SÜ 1), im Master-Studiengang "Luft- und Raumfahrttechnik" an der *Universität der Bundeswehr München* ab FT 2010. Das Modul beginnt jedes Studienjahr jeweils im Frühjahrstrimester. Als Startzeitpunkt ist das Frühjahrstrimester im 1. MA-Studienjahr vorgesehen.

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

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

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

Neubiberg, im April 2012

F. Svaricek

Inhaltsverzeichnis

Literatur	4
1 Einführung	5
2 Diskrete Signale und Systeme	8
2.1 Elementare diskrete Signale	14
2.2 Eigenschaften diskreter Systeme	15
2.3 Spektrum diskreter Signale	18
2.4 Das Abtasttheorem	20
2.5 Frequenzfaltung oder Aliasing	22
3 Die z-Transformation	24
3.1 Definition der z -Transformation	24
3.2 Existenz der z -Transformation	25
3.3 Eigenschaften der z -Transformation	26
3.4 Rücktransformation in den Zeitbereich	28
3.5 Übertragungsfunktion diskreter Systeme	32
3.6 Pole und Nullstellen der z -Übertragungsfunktion	33
3.7 z -Übertragungsfunktionen von Abtastsystemen	35
3.8 z -Übertragungsfunktion und Differenzgleichung	38
4 Pole und Nullstellen von Abtastsystemen	40
4.1 Zusammenhang zwischen den Polen des kontinuierlichen und des Abtastsystems	43
4.2 Zusammenhang zwischen den Nullstellen des kontinuierlichen und des Abtastsystems	46
4.3 Stabilitätsanalyse und w -Transformation	51
5 Beschreibung und Analyse im Zustandsraum	53
5.1 Die zeitdiskrete Zustandsraumdarstellung	53
5.2 Ableitung des Zustandsraummodells aus der Differenzgleichung	54
5.3 Lösung der Zustandsgleichung	55
5.4 Zusammenhang zwischen der kontinuierlichen und der zeitdiskreten Zustandsraumdarstellung	57
5.5 Steuer- und Erreichbarkeit zeitdiskreter Systeme	60
5.6 Beobachtbarkeit und Rekonstruierbarkeit zeitdiskreter Systeme	63
5.7 Erreichbarkeit und Beobachtbarkeit des kontinuierlichen und des zeitdiskreten Systems	64
6 Entwurf digitaler Regler	68
6.1 Quasikontinuierliche digitale Regelung	68

A Die Fourier–Transformation	73
A.1 Fourier–Reihe	73
A.2 Kontinuierliche Fourier–Transformation	77
A.3 Diskrete Fourier–Transformation	79
A.3.1 Approximation der Fourier–Reihe mit Hilfe der DFT	82
A.4 Schnelle Fourier–Transformation, FFT	82

Literatur

- Braun, A.** 1997. *Digitale Regelungstechnik*. München: R. Oldenbourg Verlag.
- Franklin, G.F., J.D. Powell** und **M. Workman**. 1998. *Digital Control of Dynamic Systems*. Menlo Park: Addison Wesley Longman. (3. Auflage)
- Grünigen von, D.C.** 2002. *Digitale Signalverarbeitung*. München: Carl Hanser Verlag.
- Kammeyer, K.D.** und **K. Kroschel**. 1998. *Digitale Signalverarbeitung: Filterung und Spektralanalyse mit MATLAB-Übungen*. Stuttgart: B.G. Teubner.
- Lunze, J.** 2002. *Regelungstechnik 2. Mehrgrößensysteme, Digitale Regelung*. Berlin: Springer-Verlag. (2. Auflage).
- Meyer, M.** 2003. *Signalverarbeitung: Analoge und digitale Signale, Systeme und Filter*. Wiesbaden: Vieweg-Verlag. (3. Auflage).
- Unbehauen, H.** 2000. *Regelungstechnik. Band II*. Braunschweig: Vieweg und Sohn (8. Auflage).

1 Einführung

Die digitale Signalverarbeitung gibt es historisch gesehen, schon seit der Zeit, als Astronomen aus Zahlenkolonnen, die sie bei der Beobachtung der Bewegung der Himmelskörper gewannen, analytische Aussagen über deren Bahnkurven machten. Den Beginn der modernen digitalen Signalverarbeitung kann man in den fünfziger Jahre ausmachen als man sich Gedanken darüber machte, ob man im Bereich der Systemanalyse und –simulation die damals gebräuchlichen Analogrechner im Prinzip durch Digitalrechner ersetzen kann.

Die Grundlagen der Theorie der zeitdiskreten Systeme wurden im Laufe des 2. Weltkrieges unabhängig voneinander in der USA (Hurewicz, MacColl), in England (Barker), Rußland (Tsypkin) und Deutschland (Oldenbourg und Sartorius) entwickelt. Bei den Arbeiten bis Ende der fünfziger Jahre standen die z -Transformation und die Beschreibung und Analyse diskreter Systeme im Frequenzbereich im Mittelpunkt des Interesses.

Die ersten Bücher über digitale Regelungstechnik wurden 1958 von E.I. Jury bzw. von Ragazzini und Franklin veröffentlicht. Der Begriff *Digitale Signalverarbeitung* tauchte demgegenüber erstmals 1969 in dem Buch von Gold und Rader auf.

Etwa um 1960 fand der Digitalrechner als Prozeßrechner auch Einsatz bei der direkten digitalen Regelung (Direct Digital Control DDC) in komplexen regelungstechnischen Prozessen, wie sie bei technischen Großanlagen (Mehrgrößenregelsystemen) zur Prozeßführung erforderlich sind. Diese sogenannten Prozeßrechner ermöglichten die Verarbeitung der zahlreichen anfallenden Meßwerte und übernahmen dann die optimale Führung des gesamten Prozesses. Dieser Rechnereinsatz beim Realzeitbetrieb geregelter technischer Prozesse erfuhr anfänglich manchen Rückschlag, doch Anfang der siebziger Jahre gehörte bei vielen komplexen Regelanlagen der Prozeßrechner als Instrument zur Überwachung, Protokollierung, Regelung und Steuerung technischer Prozesse bereits zur Standardausrüstung.

Die Entwicklung relativ preiswerter Mikroprozessoren ab etwa 1975 führte schließlich in den achtziger Jahren zur digitalen Gerätetechnik und damit zu leistungsfähigen Prozeßleitsystemen mit dezentraler verteilter Rechnerkapazität. Diese dezentralen Prozeßleitsysteme ersetzen ab 1985 weitgehend den zentralen Prozeßrechner bei der Automatisierung technischer Prozesse.

Seit 1983, als das erste digital realisierte Anti-Blockier-System (ABS) von Teves in Serie ging, hat sich in Kraftfahrzeugen die Anzahl der digitalen Steuergeräte stetig erhöht. Moderne Luxusfahrzeuge wie der Maybach oder der Phaeton haben inzwischen bis zu 70 elektronische Steuergeräte an Bord. Moderne Kraftfahrzeuge und erst recht die kommende X-by-Wire Technologie sind ohne die digitale Signalverarbeitung und die digitale Regelung nicht zu realisieren. Da Fly-by-Wire in der Luftfahrttechnik bereits Stand der Technik ist, gilt diese Aussage auch für moderne Flugzeuge.

Die Gründe für den zunehmenden Einsatz digitaler Regler sind vielfältig:

- Einfache Realisierung komplexer Regelungs-, Steuerungs- und Überwachungsalgorithmen.
- Flexibilität (einfache Anpassung und Änderung der Algorithmen).
- Kosten (Einsatz standardisierter Mikroprozessoren).
- Kürzere Entwicklungszeiten (Rapid Prototyping).

In dem BA-Modul "Steuer- und Regelungstechnik wurden nur lineare und kontinuierliche Systeme entsprechend Bild 1.1 behandelt.

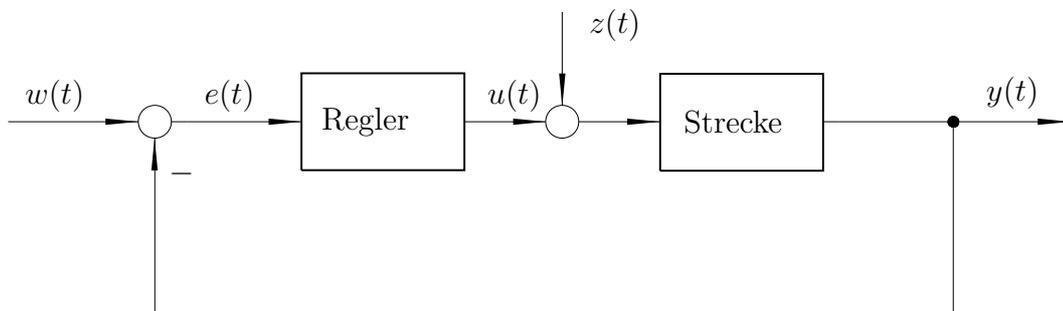


Bild 1.1: Blockschaltbild des linearen einschleifigen Regelkreises

Kontinuierliche Systeme haben Variable (Eingangs-, Zustands- und Ausgangsvariable), die alle kontinuierliche Zeitfunktionen sind. Die Werte $w(t)$, $e(t)$, $u(t)$, $z(t)$ und $y(t)$ sind für jeden Zeitpunkt t bestimmt.

Sind in einem Regelkreis Elemente enthalten, die Signale nur zu diskreten Zeitpunkten übertragen, sprechen wir von **zeitdiskreten Regelungssystemen** oder auch **Abtastregelung**.

Diskontinuierliche Signalübertragung tritt auf, wenn

- a) die Regelgröße nicht kontinuierlich meßbar ist (Rundsicht-Radargerät zur Flugüberwachung, Analyse von Stichproben in der chemischen und der metallurgischen Industrie);
- b) die Stellgröße nur zu bestimmten Zeitpunkten geändert werden kann (Multiplexbetrieb – Bild 1.2);

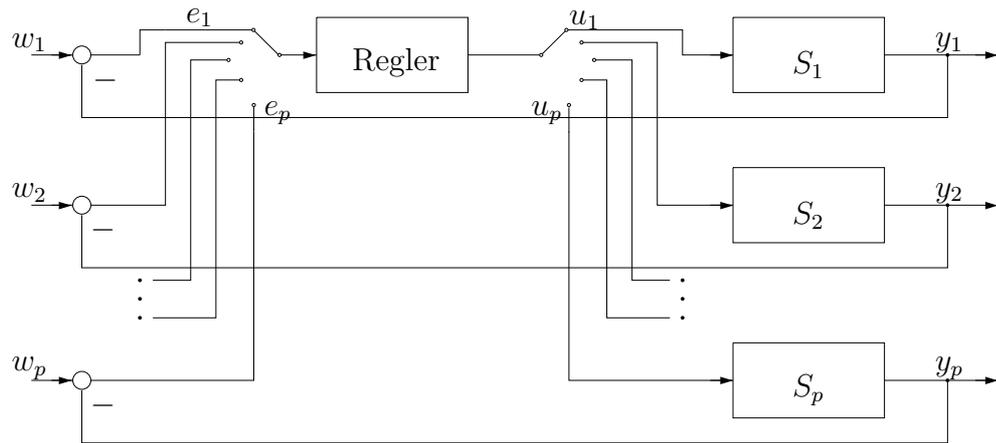


Bild 1.2: Schematische Darstellung einer Multiplexanordnung

- c) die Regeleinrichtung nur zeitdiskrete Signale verarbeiten kann, wie z. B. bei Einsatz eines digitalen Prozeß- bzw. Mikrorechners. (Bild 1.3)

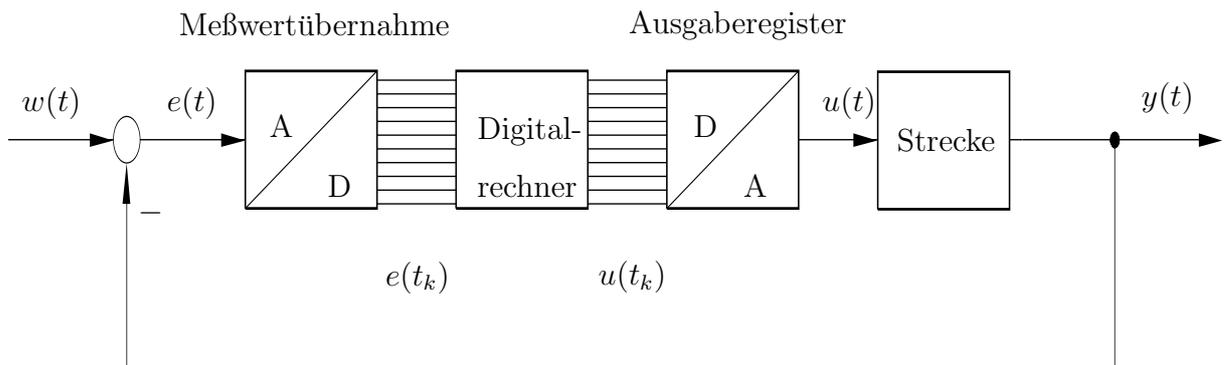


Bild 1.3: Regelkreis mit Prozeßrechner

2 Diskrete Signale und Systeme

Ein diskretes Zeitsignal erhält man meist durch *Abtastung* eines kontinuierlichen Signals.

Definition 2.1 Kontinuierliches Signal

Ein kontinuierliches Signal ist eine Zeitfunktion, die **zu jedem** Zeitpunkt t beliebige definierte Werte annehmen kann (Bild 2.1).

$$y = y(t) \in \mathbb{R}, \quad \forall t \quad . \quad (2.1)$$

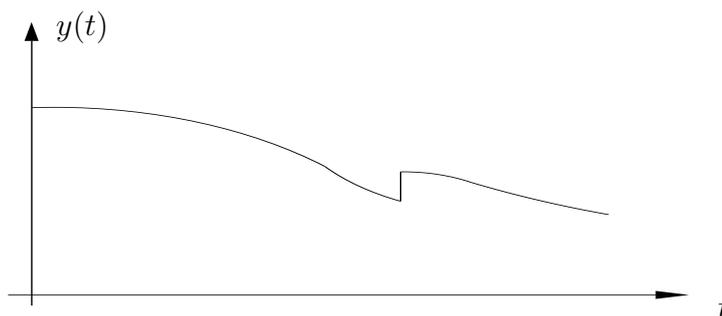


Bild 2.1: Kontinuierliches Zeitsignal

Die Abtastung kann äquidistant, zufällig oder nach einem anderen Gesetz erfolgen.

Definition 2.2 Zeitdiskretes Signal

Ein *diskretes* oder auch *zeitdiskretes* Signal wird durch eine Funktion eines diskreten Arguments beschrieben (Bild 2.2).

$$y = y(t_k), \quad k = 1, 2, \dots, n \quad . \quad (2.2)$$

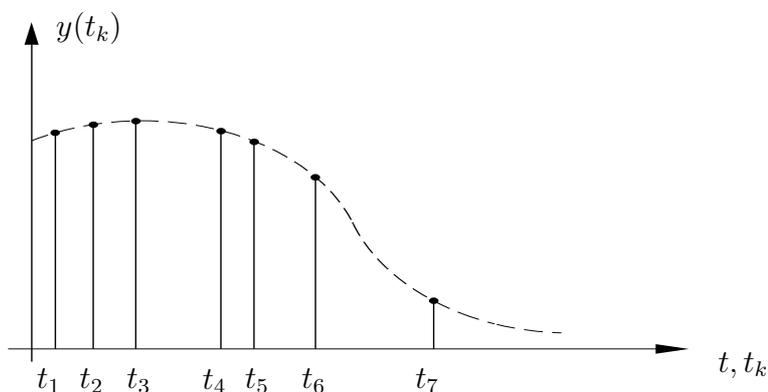


Bild 2.2: Darstellung eines zeitdiskreten Signalverlaufes

Im weiteren wird stets eine äquidistante Abtastung vorausgesetzt, die zudem so erfolgen soll, daß aus den Abtastwerten das ursprüngliche Signal fehlerfrei rekonstruiert werden kann. Das bedeutet, es gilt $t_k = kT_A$, wobei $T_A = 1/f_A := T$ das Abtastintervall und f_A die Abtastfrequenz ist.

Definition 2.3 Quantisiertes Signal

Ein quantisiertes Signal kann nur eine endliche Anzahl unterschiedlicher Werte annehmen (Bild 2.3).

$$y = y_k(t), \quad \forall t \quad k = 1, 2, \dots, n \quad . \quad (2.3)$$

□

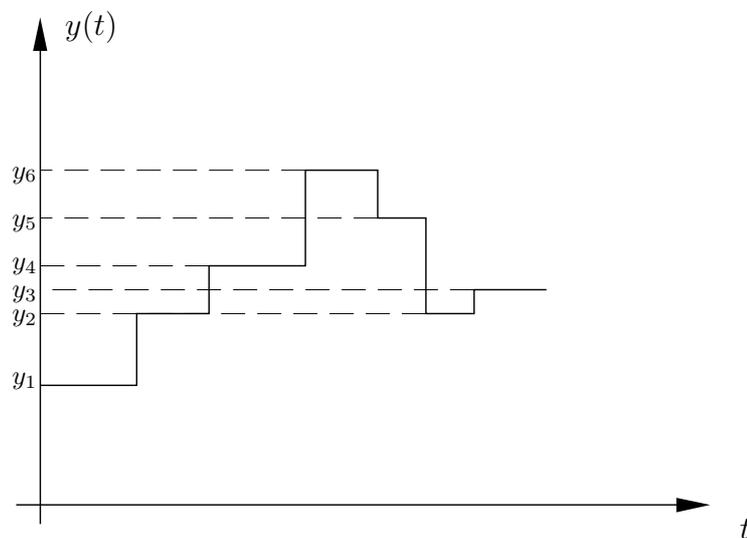


Bild 2.3: Darstellung eines quantisierten Signals

Werden Prozeß- und Mikrorechner als Regler eingesetzt, so müssen die überwiegend kontinuierlichen Regelgrößen durch A/D-Umsetzer (ADU) *zeitdiskretisiert* und *amplitudenquantisiert* werden. Bei regelungstechnischen Aufgaben ist der Einfluß der Amplitudenquantisierung auf das dynamische Systemverhalten von untergeordneter Bedeutung, wenn ADU mit einer Auflösung $< 1\%$, also 7 bit verwendet werden. Dagegen kann die Zeitdiskretisierung das Systemverhalten ganz entscheidend beeinflussen. Somit beziehen sich die weiteren Darstellung unter Vernachlässigung der Quantisierungseffekte des ADU ausschließlich auf diskrete Signale und Systeme. Bei den Systemen sollen im weiteren folgende Begriffe unterschieden werden:

Definition 2.4 Kontinuierliches System

In einem kontinuierlichen System sind die Eingangs- und Ausgangssignale sowie die Zustandsvariablen kontinuierliche Zeitfunktionen. Das dynamische Verhalten wird durch Dgln. beschrieben (Bild 2.4). □

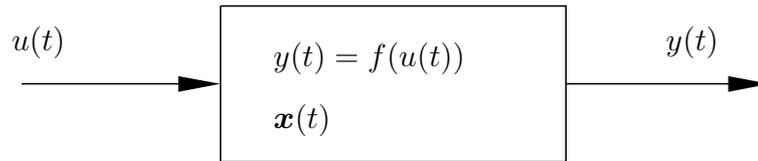


Bild 2.4: Blockbild eines kontinuierlichen Systems

Definition 2.5 Abtastsystem

Bei einem kontinuierlichen Abtastsystem sind die Zustandsvariablen kontinuierliche Funktionen, die Eingangs- und/oder Ausgangssignale diskrete Signale (Bild 2.5). □

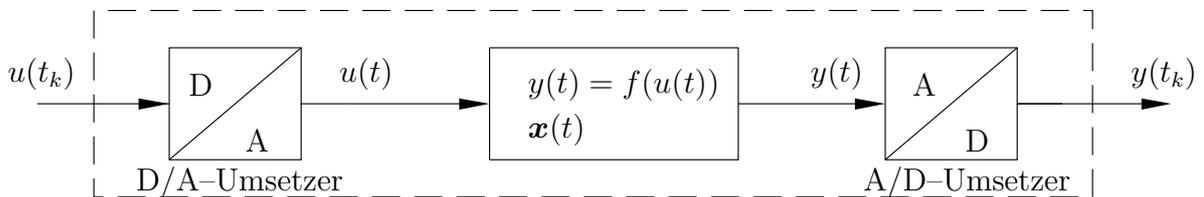


Bild 2.5: Blockschaltbild eines Abtastsystems

Definition 2.6 Diskretes System

Bei einem diskreten System sind die Eingangs-, Ausgangs- **und** Zustandsvariablen diskrete Zeitfunktionen (Bild 2.6). □

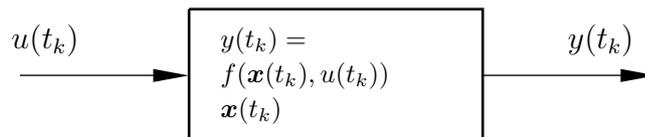


Bild 2.6: Zeitdiskretes System

Vernachlässigt man die Quantisierungseffekte, so läßt sich der ADU durch einen idealen Abtaster darstellen (Bild 2.7).

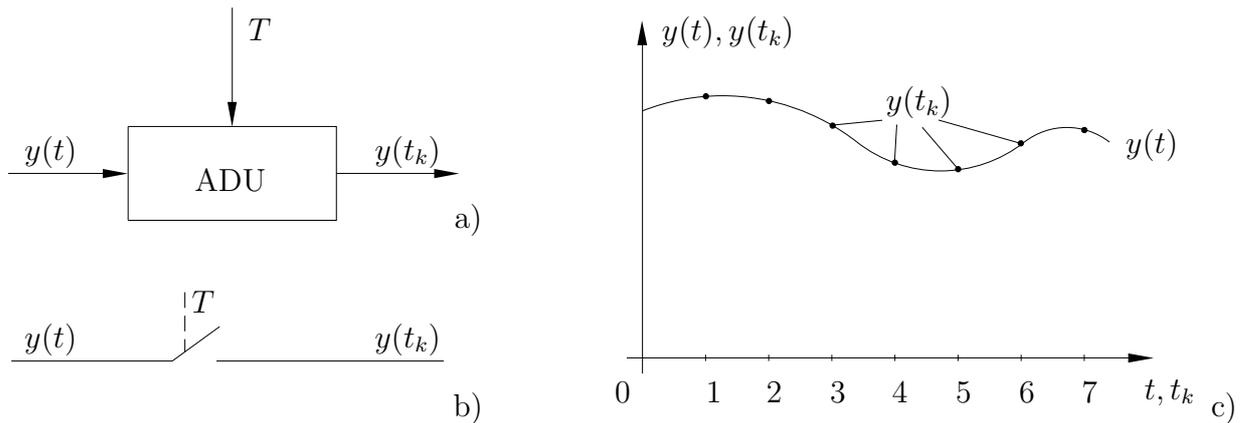


Bild 2.7: a) Blockbild eines Analog-/Digital-Umsetzers (ADU), der durch einen Takt T getaktet wird
 b) vereinfachte Darstellung des ADU als Abtaster
 c) kontinuierliches Signal $y(t)$ als Eingangssignal und zeitdiskrete Signalfolge $y(t_k)$ als Ausgangssignal eines ADU

Beispiele zu Konstruktionsprinzipien von A/D-Wandlern¹

a) **Struktur und Zeitverhalten des zählenden ADU**

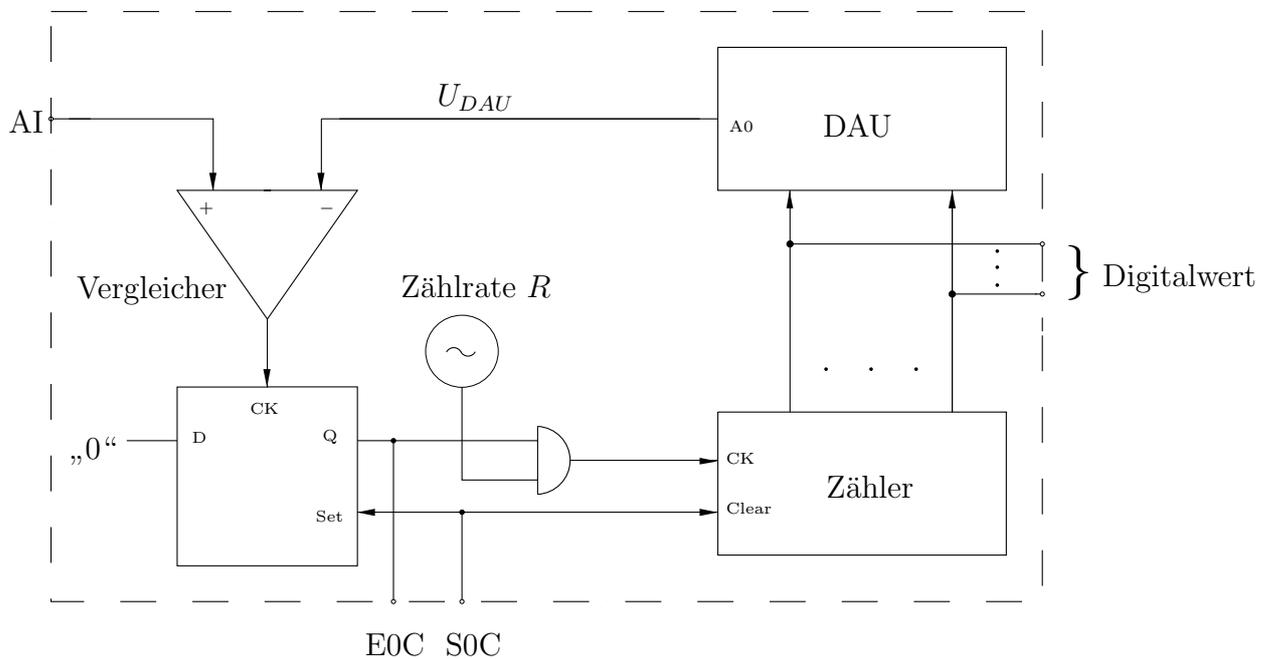


Bild 2.8: Geräteplan eines zählenden ADU

¹ Vergl. auch Färber, G. : Prozeßrechentchnik. Berlin: Springer 1979.

Ein Zähler, an dessen Ausgang ein Digital/Analogconverter angeschlossen ist, wird solange inkrementiert, bis ein nachgeschalteter Vergleicher feststellt, daß das zu messende Signal mit dem DAU-Signal übereinstimmt: Der Zählerstand entspricht dann genau dem digitalisierten Analogwert. Es handelt sich um ein einfaches und

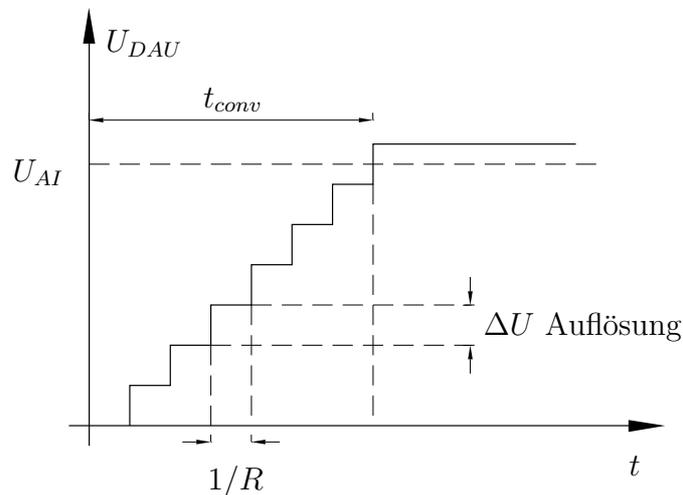


Bild 2.9: Ausgangssignal eines Digitalen–Analog–Umsetzers (DAU)

verhältnismäßig langsames Verfahren. Bei 12 bit Genauigkeit muß bis maximal $2^{12} = 4096$ gezählt werden, bei einer Zählrate von 1 MHz entsteht also eine maximale Wandlungszeit von 4 ms. Die Wandlungsrate ist im übrigen nicht konstant, sondern hängt von dem gemessenen Wert ab.

b) Struktur und Zeitverhalten des Tracking–ADU

Eine spezielle Form des zählenden ADU stellt der Tracking ADU dar, dessen Geräteplan Bild 2.10a) zeigt. In Bild 2.10b) ist dargestellt, daß dieser ADU langsam sich ändernden kontinuierlichen Signalen folgen kann.

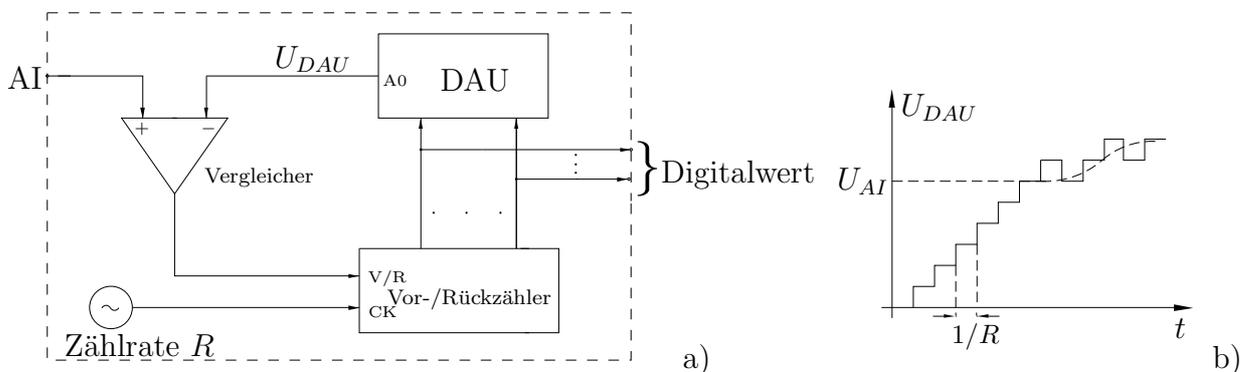


Bild 2.10: Geräteplan und Ausgangssignal des Tracking ADU

c) **Prinzip des Wäge-Codierers (Sukzessive Approximation)**

Ein Register wird zunächst auf den Wert $100\dots 0$ gesetzt, so daß der nachgeschaltete Digital/Analogumsetzer genau in die Mitte des Spannungsbereichs gesetzt wird. Die Spannung wird mit dem Meßwert verglichen; ist sie größer, dann wird der Digitalwert $01\dots 0$ erzeugt, sonst $110\dots 0$.

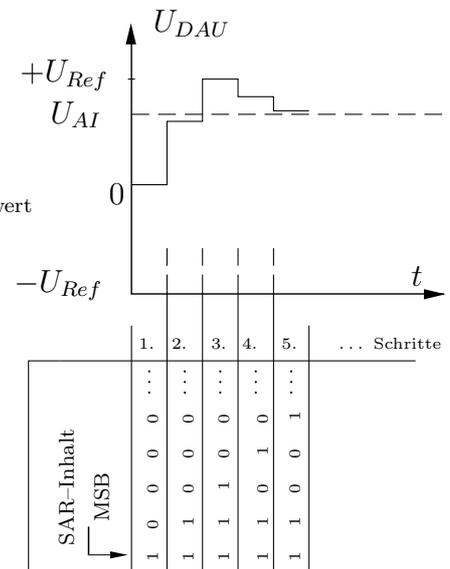
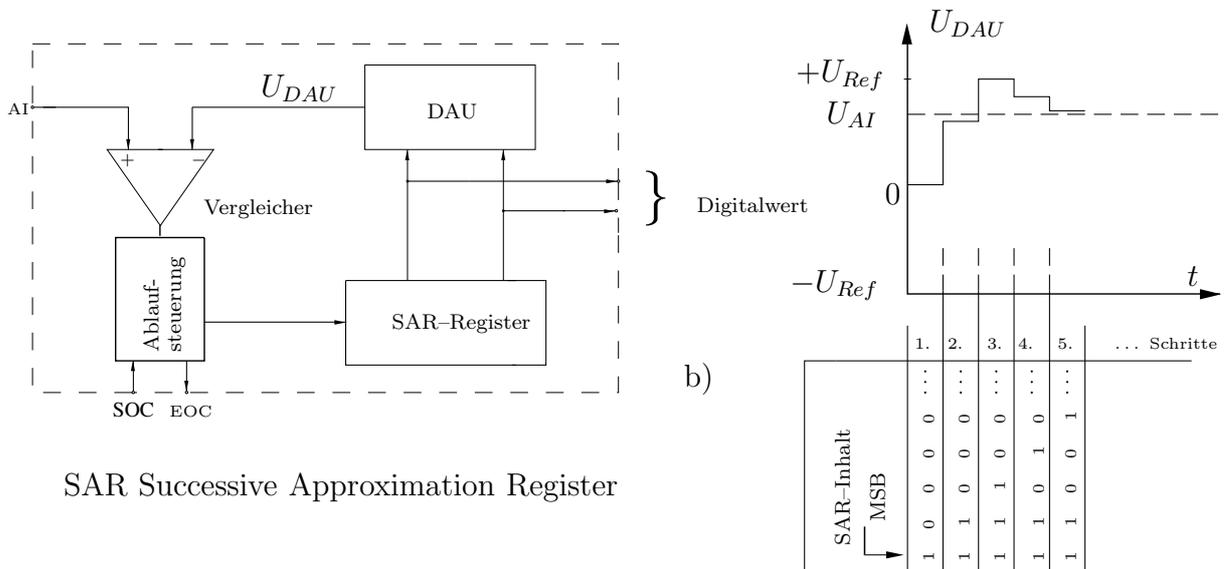
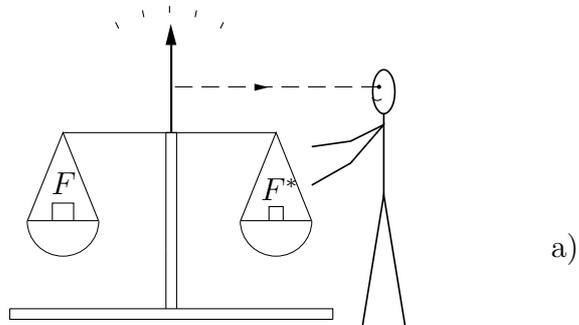


Bild 2.11: Zum Wäge-Codierverfahren a) Prinzip der manuellen Wägung b) Geräteplan des ADU c) Darstellung der Codierung

In Abhängigkeit vom Vergleichsergebnis wird also der der nächstniederen Binärstufe entsprechende Spannungswert entweder addiert oder subtrahiert. Mit jedem Schritt wird also die am Digital/Analogumsetzer anliegende Spannung sukzessive der Meßspannung angenähert – der entsprechende Digitalwert repräsentiert den Analogwert mit immer höherer Genauigkeit. Im Gegensatz zu dem zuvor erwähnten Verfahren steigt hier die Wandlungszeit nicht proportional zu der Anzahl der darstellbaren Amplitudenstufen, sondern nur zum Logarithmus dieser Anzahl. Es handelt sich also um ein schnelles Wandlungsverfahren.

In einem A/D–Umsetzer wird ein kontinuierliches Signal **zeitdiskretisiert** und **amplitudenquantisiert**. Bei hinreichend großer Auflösung (Fehler $< 1\%$) ist der Amplitudenquantisierungseffekt für die Dynamik des Regelkreises vernachlässigbar. Wie bereits zuvor erwähnt, wird in dieser einführenden Darstellung äquidistante Abtastung vorausgesetzt. Die Tasterperiode T ist dabei ein wesentlicher Analyse– und Syntheseparameter, der vom Prozeßrechner bzw. dessen Uhr gesteuert wird und der wesentlich die Dynamik des Systems bestimmt.

2.1 Elementare diskrete Signale

Im folgenden werden die Grundzüge der Theorie linearer zeitdiskreter Signale und Systeme dargestellt. Diese Theorie läuft weitgehend parallel zu der in den Modulen "Steuer– und Regelungstechnik" bzw. "Regelungstechnik" behandelten zeitkontinuierlichen Systeme. Diskrete Signale bzw. die sie repräsentierenden Zahlenfolgen können reell oder komplex sein. Die wichtigsten elementaren Signale, die auch hier z.B. zum Test der Eigenschaften dynamischer Systeme eingesetzt werden können sind:

a) Einheitsimpuls oder Impulsfolge

Der **Einheitsimpuls** (Bild 2.12a)

$$\delta(k - k_0) = \begin{cases} 1 & \text{für } k = k_0 \\ 0 & \text{für } k \neq k_0 \end{cases} \quad k = 0, 1, \dots \quad (2.4)$$

b) Einheitssprung oder Sprungfolge

Der **Einheitssprung** (Bild 2.12b)

$$1(k - k_0) = \begin{cases} 1 & \text{für } k \geq k_0 \\ 0 & \text{für } k < k_0 \end{cases} \quad k = 0, 1, 2, \dots \quad (2.5)$$

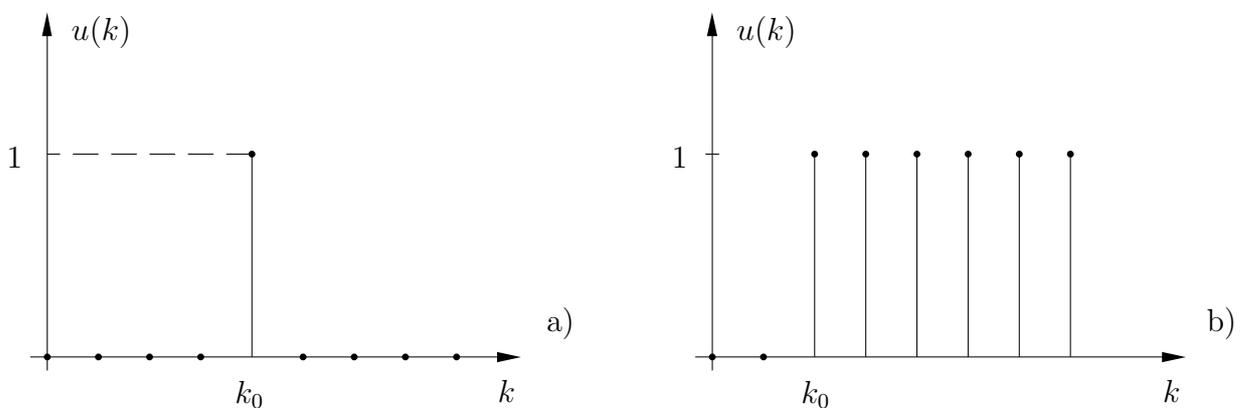


Bild 2.12: Zeitdiskrete Testsignale a) Einheitsimpuls b) Einheitssprung

2.2 Eigenschaften diskreter Systeme

Das Übertragungsverhalten des zeitdiskreten Systems in Bild 2.13, gekennzeichnet durch den Operator $L\{\cdot\}$, wird zunächst durch die zuvor definierten diskreten Testsignale beschrieben. Dabei werden die betrachteten Signale bzw. ihre Argumente weitgehend wie folgt abgekürzt, z. B. :

$$u_k = u(k) = u(t_k) = u(t)|_{t=t_k} \quad . \quad (2.6)$$

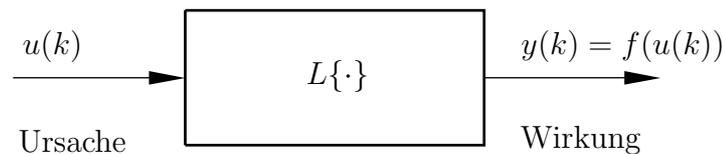


Bild 2.13: Blockbild des zeitdiskreten Übertragungssystems

Hierbei ist $u(k)$ für einen bestimmten Wert k eine feste Zahl. Für ein laufendes k steht $u(k)$ aber für eine Folge von Zahlen, so daß für das diskrete Signal eigentlich die Bezeichnung $\{u(k)\}$ angemessener wäre, da es sich um eine Zahlenfolge handelt. Zur Vereinfachung der Schreibweise ist es allerdings üblich in beiden Fällen die Bezeichnung $u(k)$, bzw. u_k zu verwenden.

a) Lineares System

Ein zeitdiskretes System ist **linear** wenn das Superpositionsprinzip gilt. Wenn man das System mit beliebigen Signalfolgen $u_i(k)$ erregt und dann die Systemantwort

$$y_i(k) = L\{u_i(k)\}, \quad i = 1, 2, \dots, n \quad (2.7)$$

erhält, so muß die Linearkombination der Eingangssignale $u_i(k)$ stets die Systemantwort

$$y(k) = L\left\{\sum_{i=1}^n \alpha_i u_i(k)\right\} = \sum_{i=1}^n \alpha_i L\{u_i(k)\} = \sum_{i=1}^n \alpha_i y_i(k) \quad (2.8)$$

liefern.

b) Zeitinvarianz

Bei *Zeitinvarianz* hängt die Systemantwort nicht vom Zeitpunkt der Erregung ab. Gilt also bei Erregung zum Zeitpunkt $k = 0$

$$y(k) = L\{u(k)\}, \quad (2.9)$$

so erhält man bei Erregung zum Zeitpunkt $k = k_0$

$$y(k - k_0) = L\{u(k - k_0)\}. \quad (2.10)$$

c) Kausalität

Ein System ist *Kausal*, wenn das Ausgangssignal $y(k)$ zu einem Zeitpunkt $k = k_0$ unabhängig von künftigen Werten des Eingangssignals $u(k)$ ist. Das bedeutet, die Antwort eines Systems erscheint bei Kausalität nicht vor der Erregung.

d) Beschreibung durch Gewichtsfolge

Genauso wie bei den kontinuierlichen Systemen läßt sich das Zeitverhalten diskreter Systeme durch die Impulsantwort oder Gewichtsfolge $g(k)$, der Antwort auf den Einheitsimpuls $\delta(k)$ beschreiben (Bild 2.14b).

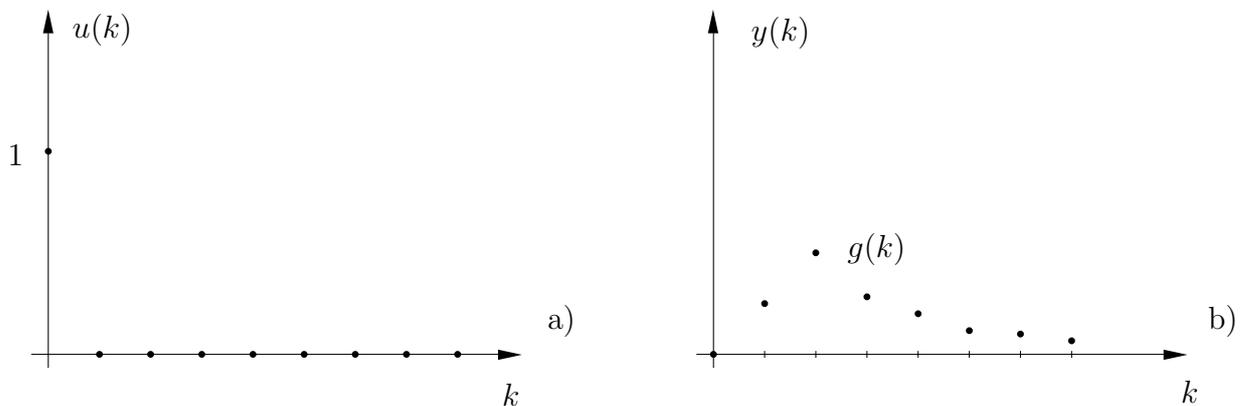


Bild 2.14: a) Einheitsimpuls als Systemerregung
b) Impulsantwort $g(k)$ des Systems in Bild 2.13

Die Antwort des linearen Systems mit der Gewichtsfolge $g(k)$ auf eine beliebige Eingangssfolge $u(k)$ kann mit der *Faltungssumation* bestimmt werden:

$$y(k) = \sum_{i=0}^k g(k-i)u(i) \quad ; \quad k = 0, 1, 2, \dots \quad (2.11)$$

oder mit der Substitution $k - i = r$:

$$y(k) = \sum_{r=0}^k g(r)u(k-r) \quad ; \quad k = 0, 1, 2, \dots \quad (2.12)$$

$$=: g(k) * u(k). \quad (2.13)$$

e) **Beschreibung durch Differenzgleichungen**

Das Signalübertragungsverhalten zeitdiskreter Systeme kann analog zu der Vorgehensweise bei kontinuierlichen Systemen durch Gleichungen erfolgen, die Eingangs- und Ausgangsgröße verknüpfen. Waren dies bei kontinuierlichen Systemen *Differentialgleichungen*, so sind es bei diskreten Systemen **Differenzgleichungen**.

Für das System in Bild 2.13 gilt dann:

$$\begin{aligned} y(k) + a_1y(k-1) + a_2y(k-2) + \dots + a_ny(k-n) &= \\ &= b_0u(k) + b_1u(k-1) + \dots + b_nu(k-n) \quad ; \quad k = 1, 2, 3, \dots \end{aligned} \quad (2.14)$$

mit den n Anfangsbedingungen:

$$y_0 = y(0) \quad , \quad y(-1) = \dots = y(-n+1) = 0 \quad (2.15)$$

oder in kompakterer Notierung mittels Summenzeichen:

$$y(k) = \sum_{j=0}^n b_j u(k-j) - \sum_{i=1}^n a_i y(k-i) \quad , \quad k = 1, 2, 3, \dots \quad (2.16)$$

Bei diesen Systemen ist der momentane Ausgangsabtastwert (Zeitpunkt k) gleich einer Linearkombination von vergangenen Ausgangs- und Eingangsabtastwerten plus dem mit b_0 gewichteten aktuellen Eingangsabtastwert. Diese Differenzgleichungen stellen also einen Rekursionsalgorithmus dar, der mit einem Digitalrechner schrittweise gelöst werden kann.

Beispiel:

$$y(k) + a_1y(k-1) + a_2y(k-2) = u(k)$$

$$k = 1 \quad :$$

$$y(1) = u(1) - a_1y(0) - a_2y(-1)$$

$$y(1) = u(1) - a_1y(0)$$

$$k = 2 \quad :$$

$$y(2) = u(2) - a_1y(1) - a_2y(0)$$

$$\vdots \quad \vdots$$

2.3 Spektrum diskreter Signale

Ein weitere mathematische Beschreibung für ein zeitdiskretes Signal erhält man aus der Multiplikation des zugehörigen kontinuierlichen Signals $f(t)$ mit einer Folge von Dirac-Impulsen:

$$f^*(t) = f(t) \sum_{k=-\infty}^{\infty} \delta(t - kT) \quad (2.17)$$

$$= \sum_{k=-\infty}^{\infty} f(kT) \delta(t - kT) \quad (2.18)$$

Dabei kann $f^*(t)$ als kontinuierliche Darstellung eines mit $f_A = 1/T$ abgetasteten Signals angesehen werden. Die Abtastwerte $f(kT)$ treten als Gewichte der Dirac-Impulse auf.

Das Spektrum der diskreten Signalfolge $f^*(t)$ berechnet sich aus der Faltung des Spektrums des kontinuierlichen Signals, $F(j\omega)$, mit dem Spektrum der Dirac-Impulsfolge, die eine in T periodische Folge von δ -Impulsen ist und daher in eine komplexe *Fourierreihe* entwickelt werden kann:

$$\sum_{k=-\infty}^{\infty} \delta(t - kT) = \sum_{\nu=-\infty}^{\infty} c_\nu e^{j\nu\omega_A t}, \quad (2.19)$$

mit der *Abtastkreisfrequenz* $\omega_A = 2\pi/T$.

Die Fourier-Koeffizienten C_ν ergeben sich als Integral über eine Periode zu

$$c_\nu = \frac{1}{T} \int_{-T/2}^{T/2} \sum_{k=0}^{\infty} \delta(t - kT) e^{-j\nu\omega_A t} dt. \quad (2.20)$$

Der einzige Term in der Summe von Impulsen, der innerhalb der Integrationsgrenzen liegt, ist der Impuls im Ursprung. Damit reduziert sich (2.20) zu

$$c_\nu = \frac{1}{T} \int_{-T/2}^{T/2} \delta(t) e^{-j\nu\omega_A t} dt. \quad (2.21)$$

Aufgrund der Ausblendeigenschaft des Dirac-Impuls wird das Integral in (2.21) zu Eins, so daß für alle ν -Werte

$$c_\nu = \frac{1}{T} \quad (2.22)$$

gilt.

Damit ergibt sich die Fourierreihe der Summe der Dirac-Impulse zu

$$\sum_{k=-\infty}^{\infty} \delta(t - kT) = \frac{1}{T} \sum_{\nu=-\infty}^{\infty} e^{j\nu\omega_A t} \quad (2.23)$$

und für Gl. (2.17)

$$f^*(t) = f(t) \frac{1}{T} \sum_{\nu=-\infty}^{\infty} e^{j\nu\omega_A t} . \quad (2.24)$$

Die Anwendung der Fourier–Transformation auf Gl. (2.24) liefert

$$\mathcal{F}\{f^*(t)\} = \frac{1}{T} \sum_{\nu=-\infty}^{\infty} \mathcal{F}\{f(t)e^{j\nu\omega_A t}\} . \quad (2.25)$$

Mit Hilfe des Verschiebungssatzes im Frequenzbereich der Fourier–Transformation (entspricht dem Dämpfungssatz der Laplace–Transformation) erhält man dann für die rechte Seite der Gl. (2.25)

$$\mathcal{F}\{f(t)e^{j\nu\omega_A t}\} = F(j(\omega - \nu\omega_A)) \quad (2.26)$$

und damit die Fourier–Transformierte von $f^*(t)$ zu

$$F^*(j\omega) = \frac{1}{T} \sum_{\nu=-\infty}^{\infty} F(j(\omega - \nu\omega_A)) . \quad (2.27)$$

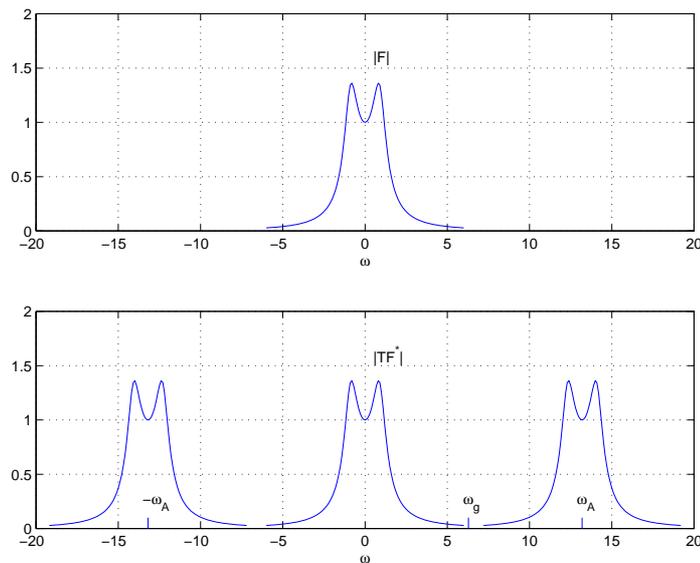


Bild 2.15: Spektren eines bandbegrenzten kontinuierlichen und eines mit $\omega_A > 2\omega_g$ abgetasteten Signals (ω_g ist die Grenzfrequenz des kontinuierlichen Signals)

Das Spektrum eines zeitdiskreten Signals ist periodisch mit der Periode ω_A , der Abtastkreisfrequenz, und entsteht aus dem mit $1/T$ gewichteten Spektrum $F(j\omega)$ des kontinuierlichen Signals $f(t)$ durch eine Überlagerung entsprechend Gl. (2.27). Das bedeutet, die Multiplikation des kontinuierlichen Signals $f(t)$ mit der Dirac–Impulsfolge erzeugt eine unendliche Serie von sogenannten *Seitenbänder* in positiver und negativer Richtung (vgl. Bild 2.15).

2.4 Das Abtasttheorem

Durch den Abtastvorgang mittels eines idealen Tasters (A/D–Umsetzers) wird aus einem kontinuierlichen Signal $y(t)$ eine Zahlenfolge $y(k)$ erzeugt (Bild 2.7). Liegen die Abtastwerte hinreichend dicht beieinander, ist also die Abtastfrequenz $f_A = 1/T$ hinreichend groß, geht durch den Abtastvorgang nur wenig an Informationen verloren. Dies ist aber nicht mehr gegeben, wenn von einem kontinuierlichen Signal zu wenig Abtastwerte vorliegen.

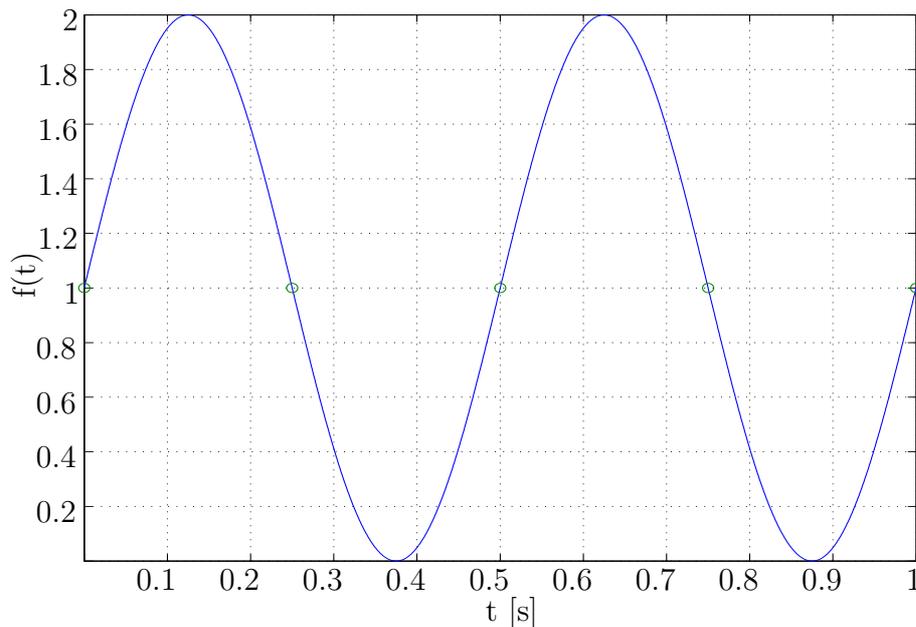


Bild 2.16: Abgetastetes Sinussignal

In Bild 2.16 ist zu erkennen, daß die abgetasteten Werte eines Sinussignals nicht von einem Gleichspannungssignal zu unterscheiden sind, wenn die Abtastfrequenz doppelt so hoch wie die Frequenz des Sinussignals ist. Wenn ein harmonisches Signal mit der Frequenz f_N aus den Abtastwerten rekonstruiert werden soll, so muß die Abtastfrequenz f_A mehr als doppelt so groß sein:

$$f_A > 2f_N . \quad (2.28)$$

Dieser Zusammenhang wurde bereits von *Nyquist* erkannt, und f_N wird daher als *Nyquistfrequenz* bezeichnet. Die exakten Bedingungen, unter denen ein Signal aus den äquidistanten Abtastwerten rekonstruiert werden kann, wurden 1949 von *Shannon* formuliert.

Satz 2.1 Abtasttheorem von Shannon

- a) Ein kontinuierliches, bandbegrenztes Signal
- $f(t)$
- dessen Fouriertransformierte

$$F(j\omega) = \int_{-\infty}^{\infty} f(t)e^{-j\omega t} dt \quad (2.29)$$

außerhalb des Intervalls $(-\omega_g, \omega_g)$ verschwindet, ist eindeutig durch die äquidistanten Abtastwerte bestimmt, wenn für die Abtast(kreis)frequenz $\omega_A = 2\pi f_A$ gilt:

$$\omega_A > 2\omega_g . \quad (2.30)$$

- b) Das kontinuierliche Signal
- $f(t)$
- kann dann aus der Abtastfolge
- $f(k)$
- mittels der
- Shannon'schen Interpolationsformel*
- berechnet werden:

$$f(t) = \sum_{k=-\infty}^{\infty} f(kT) \frac{\sin \omega_A(t - kT)/2}{\omega_A(t - kT)/2} = \sum_{k=-\infty}^{\infty} f(kT) \frac{\sin \omega_g(t - kT)}{\omega_g(t - kT)} . \quad (2.31)$$

Es ist bemerkenswert, daß das Shannon'sche Rekonstruktionsfilter zunächst nur von theoretischen Wert ist, da es *nicht kausal* ist: zur Berechnung eines konkreten Signalwertes $f(t_1)$ werden nicht nur vergangene, sondern auch zukünftige Abtastwerte benötigt. Bei der Signalauswertung in der Nachrichtentechnik ist dies allerdings häufig kein Problem, da die Signalfolge zunächst gespeichert und dann später mit Hilfe des Shannonfilter ausgewertet werden kann. In einem geschlossenen Regelkreis ist diese Vorgehensweise nicht möglich, so daß andere Rekonstruktionsverfahren notwendig sind.

2.5 Frequenzfaltung oder Aliasing

Wenn ein kontinuierliches Signal $f(t)$ mit dem zugehörigen Spektrum $F(j\omega) = \mathcal{F}\{f(t)\}$ mit der Frequenz $f_A = 1/T$ abgetastet wird, werden diskrete Signale mit periodisch sich wiederholenden Spektren erzeugt (vgl. Bild 2.15).

In Abschnitt 2.3 wurde gezeigt, daß die Spektren $F(j\omega)$ kontinuierlicher und $F^*(j\omega)$ diskreter Signale durch die Beziehung

$$F^*(j\omega) = \frac{1}{T} \sum_{\nu=-\infty}^{\infty} F(j(\omega - \omega_A \nu)) \quad (2.32)$$

miteinander verknüpft sind. Daraus folgt, daß das Spektrum diskreter Signale die Überlagerung gegeneinander versetzter und normierter Spektren der zugehörigen kontinuierlichen Signale ist, Bild 2.17.

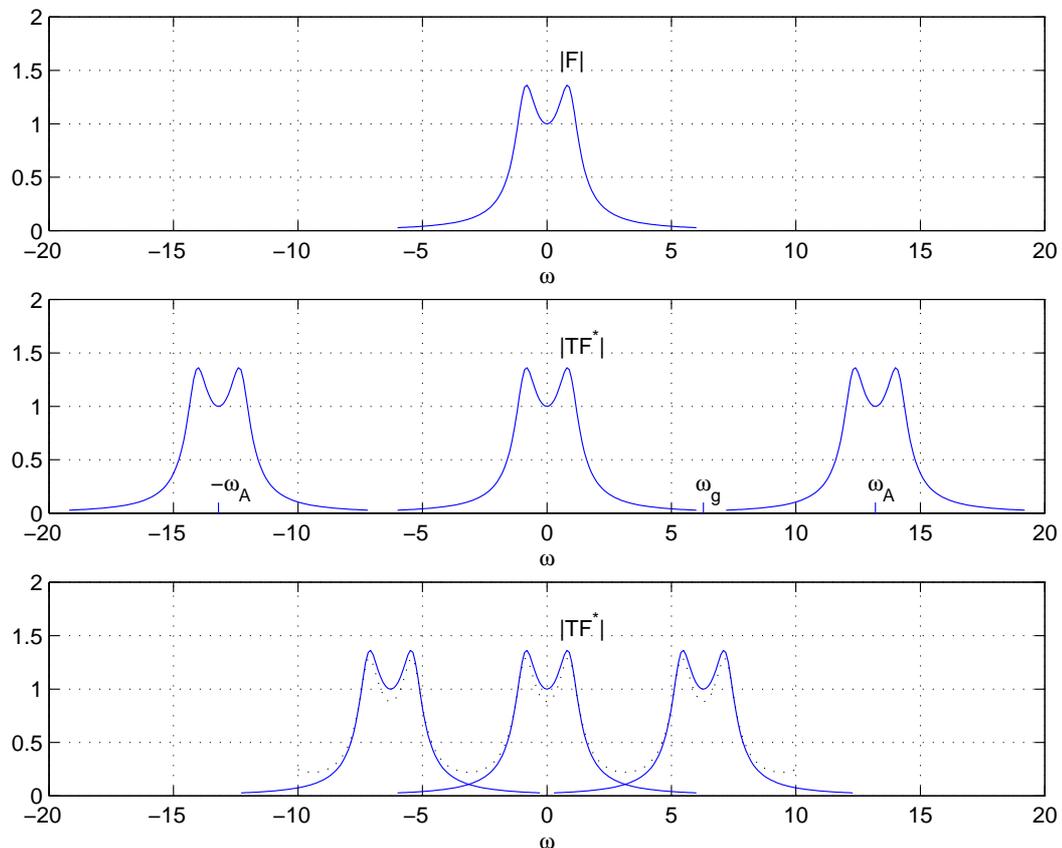


Bild 2.17: Spektren eines kontinuierlichen Signals und zweier durch verschiedene Abtastfrequenzen gewonnener diskreter Signale

Die Spektren des kontinuierlichen und des diskreten Signals stimmen offensichtlich im Intervall $(-\pi/T \leq \omega \leq \pi/T)$ überein, wenn die folgenden Forderungen eingehalten werden:

- $F(j\omega)$ muß bandbegrenzt sein, also im Frequenzbereich $|\omega| \geq \omega_g$ identisch verschwinden.
- Die Abtastkreisfrequenz $\omega_A = 2\pi f_A = 2\pi/T$ muß mindestens doppelt so groß sein, wie die maximale Frequenz ω_g von $F(j\omega)$.

Durch die Frequenzüberlappung, die auch als Faltung der Frequenzspektren interpretierbar ist, können erhebliche Probleme bei der Signalrekonstruktion auftreten. Das Spektrum des abgetasteten Signals wird durch diese Überlagerung verfälscht, so daß selbst bei idealer Rekonstruktion ein *anderes* Signal $f(t)$ gefunden wird, daß sich erheblich von dem nicht abgetasteten Signal unterscheiden kann. Dieser Effekt wird im Englischen als *aliasing* bezeichnet (abgeleitet von dem lat. alias: anders). In Bild 2.18 ist verdeutlicht, daß zwei kontinuierliche Signale mit sehr unterschiedlichen Zeitverlauf die gleichen Abtastwerte haben können.

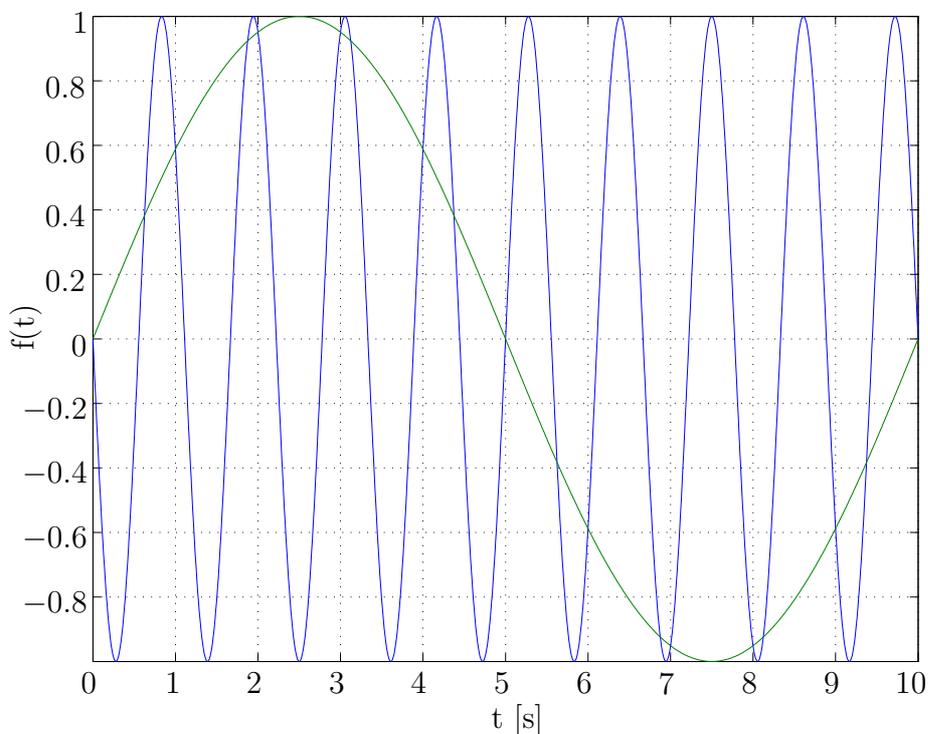


Bild 2.18: Zwei periodische Signale mit den Frequenzen 0,1 Hz bzw. 0,9 Hz, die bei einer Abtastperiode $T = 1$ die gleichen abgetasteten Werte haben.

In der praktischen Anwendung müssen deshalb immer durch ein dem A/D–Umsetzer vorgeschaltetes Tiefpaßfilter (*Anti–Aliasing–Filter*) mit einer Grenzfrequenz f_g die höheren Frequenzanteile hinreichend abgeschwächt werden. Hochfrequente Störsignale können bei einer unglücklichen Wahl der Abtastfrequenz auch im niederfrequenten Bereich erhebliche Signalverfälschungen bewirken. Damit der bandbegrenzende Tiefpaß keine zu steile Flanke zwischen Durchlaß– und Sperrbereich aufweisen muß, wählt man in der Praxis für f_A einen Wert, der deutlich über dem theoretischen Wert $2f_g$ liegt.

3 Die z -Transformation

Im Bereich der Systemtheorie kontinuierlicher Signale wird zur Transformation eines Signals $f(t)$ in den Frequenzbereich die *Laplace-Transformation* immer dann eingesetzt, wenn bei der Fourier-Transformation Konvergenzprobleme auftreten. Solche Probleme treten bereits bei elementaren Funktionen wie z.B. der kontinuierlichen Sprungfunktion auf. Entsprechendes gilt natürlich auch für die zeitdiskrete Sprungfolge.

Die Laplace-Transformation könnte man prinzipiell auch auf zeitdiskrete Signale und Systeme anwenden. Allerdings kann man zeigen, daß dabei Mehrdeutigkeiten auftreten, die die Anwendung erschweren. Bei diskreten Signalen und Systemen wird daher eine angepaßte Transformation, die z -Transformation, verwendet, die solche Mehrdeutigkeiten nicht aufweist.

3.1 Definition der z -Transformation

Die Analyse und Synthese diskreter Signale und Systeme läßt sich durch die sogenannte z -Transformation wesentlich erleichtern. Zur Ableitung dieser Transformation stellt man ein abgetastetes kontinuierliches Signal $f(t)$ als Folge gewichteter Dirac-Impulse dar:

$$f^*(t) = \sum_{k=0}^{\infty} f(kT)\delta(t - kT) . \quad (3.1)$$

Durch Laplace-Transformation erhält man aus (3.1) die komplexe Funktion

$$F^*(s) = \sum_{k=0}^{\infty} f(kT)e^{-kTs} . \quad (3.2)$$

Ersetzt man nun in (3.2) den Ausdruck e^{Ts} durch die komplexe Variable z , so erhält man:

Definition 3.1 z -Transformation

Die z -Transformation eines diskreten Signals $f(k)$ mit $k = 0, 1, 2, \dots$ ist als unendliche Potenzreihe in der komplexen Variablen z^{-1} erklärt. Die Koeffizienten dieser Reihe sind die Werte der Signalfolge:

$$F(z) = \mathcal{Z}\{f(kT)\} = \mathcal{Z}\{f(k)\} \quad (3.3)$$

$$= \sum_{k=0}^{\infty} f(k)z^{-k} . \quad (3.4)$$

Durch

$$z = e^{sT} \quad (3.5)$$

wird die s -Ebene eindeutig und *nichtlinear* auf die z -Ebene abgebildet (vgl. Bild 3.1). Würde man eine andere Transformationen verwenden, z.B. eine, bei der im Exponent von (3.5) ein negatives Vorzeichen steht, so würde dies bedeuten, daß dann die linke komplexe s -Halbebene, in der sich die Pole stabiler kontinuierlicher Systeme befinden, ins *Äußere* des Einheitskreises abgebildet würde.

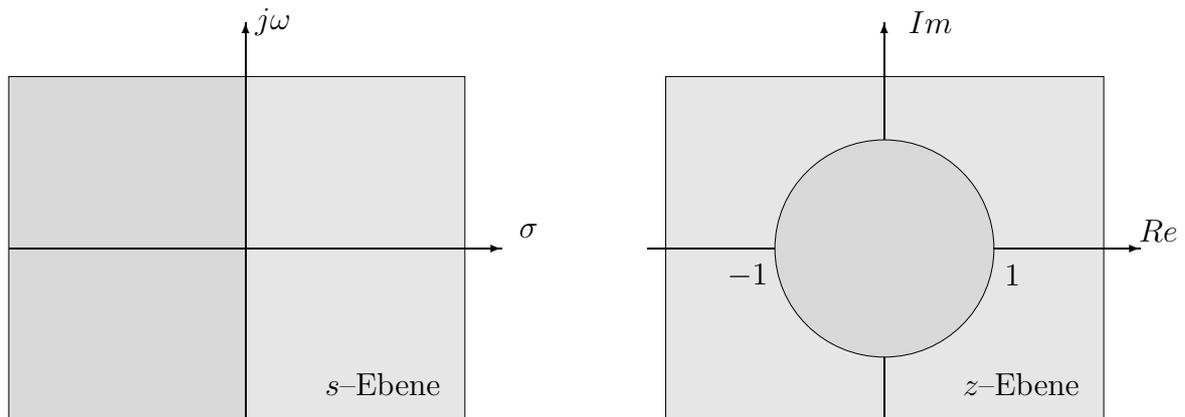


Bild 3.1: Abbildung der s -Ebene auf die z -Ebene

Aus der in (3.5) gegebenen Definition ergeben sich die in Tabelle 3.1 dargestellten Zusammenhänge.

s -Ebene	z -Ebene
linke komplexe Ebene	Inneres des Einheitskreises
imaginäre Achse	Peripherie des Einheitskreises
rechte komplexe Ebene	Äußeres des Einheitskreises
Ursprung ($s = 0$)	$z = 1$

Tabelle 3.1: Zusammenhänge zwischen Laplace- und z -Transformation

3.2 Existenz der z -Transformation

Die z -Transformierte existiert, wenn die Potenzreihe in (3.4) konvergiert, d.h. wenn

$$\sum_{k=0}^{\infty} |f(k)z^{-k}| < \infty. \tag{3.6}$$

gilt.

Diese Reihe konvergiert für sämtliche Werte von z mit

$$|z| > R \tag{3.7}$$

absolut, falls die Folge $f(k)$ für alle $k = 0, 1, 2, \dots$ die Ungleichung

$$|f(k)| < MR^k \tag{3.8}$$

erfüllt. Hierbei sind K und R positive Konstanten. Die z -Transformation existiert also außerhalb der Kreisscheibe mit Mittelpunkt 0 und Radius R .

3.3 Eigenschaften der z -Transformation

Die Eigenschaften der z -Transformation sind in der Tabelle 3.2 zusammengestellt.

	$f(k)$	$F(z)$
z -Transformation	$f(k)$	$\mathcal{Z}\{f(k)\} = \sum_{k=0}^{\infty} f(k)z^{-k}$
inverse z -Transformation	$f(k) = \frac{1}{2\pi j} \oint F(z)z^{k-1}dz$	$F(z)$
Linearität	$af_1(k) + bf_2(k)$	$aF_1(z) + bF_2(z)$
Rechtsverschiebung	$f(k-i)$	$z^{-i}F(z)$
Linksverschiebung	$f(k+i)$	$z^iF(z) - z^i \sum_{j=0}^{i-1} z^{-j}f(j)$
Differenzensatz	$\Delta f(k) = f(k) - f(k-1)$	$\frac{z-1}{z}F(z)$
Summensatz	$f_{\Sigma}(k) = \sum_{\nu=0}^k f(\nu)$	$\frac{z}{z-1}F(z)$
Faltungssatz	$\sum_{i=0}^{\infty} f_1(k-i)f_2(i)$	$F_1(z)F_2(z)$
Dämpfungssatz	$e^{akT}f(k)$	$F(e^{-aT}z)$
Ähnlichkeitssatz	$a^k f(k)$	$F\left(\frac{z}{a}\right)$
Anfangswertsatz	$f(0) = \lim_{z \rightarrow \infty} F(z)$	
Endwertsatz	$f(\infty) = \lim_{z \rightarrow 1} (z-1)F(z)$	

Tabelle 3.2: Definitionen und Eigenschaften der z -Transformation

Die Korrespondenzen für einige elementare Zeitfunktionen enthält Tabelle 3.3.

Nr.	Zeitfunktion $f(t)$	$F(s) = \mathcal{L}\{f(t)\}$	$F(z) = \mathcal{Z}\{f(kT)\}$
1	δ -Impuls $\delta(t)$	1	1
2	Einheitssprung $1(t)$	$\frac{1}{s}$	$\frac{z}{z-1}$
3	t	$\frac{1}{s^2}$	$\frac{Tz}{(z-1)^2}$
4	t^2	$\frac{2}{s^3}$	$\frac{T^2z(z+1)}{(z-1)^3}$
5	e^{-at}	$\frac{1}{s+a}$	$\frac{z}{z-e^{-aT}}$
6	te^{-at}	$\frac{1}{(s+a)^2}$	$\frac{T e^{-aT} z}{(z-e^{-aT})^2}$
7	$t^2 e^{-at}$	$\frac{2}{(s+a)^3}$	$\frac{T^2 e^{-aT} z(z+e^{-aT})}{(z-e^{-aT})^3}$
8	$1 - e^{-at}$	$\frac{a}{s(s+a)}$	$\frac{(1-e^{-aT})z}{(z-1)(z-e^{-aT})}$
9	$\sin \omega_0 t$	$\frac{\omega_0}{s^2 + \omega_0^2}$	$\frac{z \sin \omega_0 T}{z^2 - 2z \cos \omega_0 T + 1}$
10	$\cos \omega_0 t$	$\frac{s}{s^2 + \omega_0^2}$	$\frac{z^2 - z \cos \omega_0 T}{z^2 - 2z \cos \omega_0 T + 1}$
11	$1 - (1+at)e^{-at}$	$\frac{a^2}{s(s+a)^2}$	$\frac{z}{z-1} - \frac{z}{z-c} - \frac{acTz}{(z-c)^2}; \quad c = e^{-aT}$
12	$1 + \frac{be^{-at} - ae^{-bt}}{a-b}$	$\frac{ab}{s(s+a)(s+b)}$	$\frac{z}{z-1} + \frac{bz}{(a-b)(z-c)} - \frac{az}{(a-b)(z-d)}$ $c = e^{-aT}, \quad d = e^{-bT}$
13	$e^{-at} \sin \omega_0 t$	$\frac{\omega_0}{(s+a)^2 + \omega_0^2}$	$\frac{cz \sin \omega_0 T}{z^2 - 2cz \cos \omega_0 T + c^2}; \quad c = e^{-aT}$
14	$e^{-at} \cos \omega_0 t$	$\frac{s+a}{(s+a)^2 + \omega_0^2}$	$\frac{z^2 - cz \cos \omega_0 T}{z^2 - 2cz \cos \omega_0 T + c^2}; \quad c = e^{-aT}$
15	$a^{\frac{t}{T}}$	$\frac{1}{s - (\frac{1}{T}) \ln a}$	$\frac{z}{z-a}$

Tabelle 3.3: Korrespondenzentabelle

3.4 Rücktransformation in den Zeitbereich

Im BA-Modul "Steuer- und Regelungstechnik" sind die Vorteile der Laplace-Transformation für kontinuierliche Systeme aufgezeigt worden. Eine ähnliche Bedeutung hat die z -Transformation für die diskreten Systeme, da auch hier die Lösung der Regelkreisprobleme algebraisch im Bildbereich erfolgen kann. Man erhält als Lösung einen analytischen Ausdruck für den Regler im Bildbereich, etwa in Form einer oder mehrerer z -Übertragungsfunktionen. Für die Programmierung auf den Digitalrechner benötigt man jedoch immer die dazugehörigen Differenzgleichungen, die man durch Rücktransformation der analytischen Bildbereichsfunktionen in den Zeitbereich gewinnt.

Für das allgemeine Problem der Rücktransformation

$$F(z) \longrightarrow f(k)$$

gibt es drei Möglichkeiten:

1. Die *Polynomdivision*;
2. die *Residuenmethode* und
3. die Benutzung von Korrespondenztabelle nach geeigneter *Partialbruchzerlegung*.

1. Polynomdivision

Die zur gebrochenen rationalen Funktion

$$F(z) = \frac{b_m z^m + b_{m-1} z^{m-1} + \dots + b_1 z + b_0}{z^n + a_{n-1} z^{n-1} + \dots + a_1 z + a_0} \quad (3.9)$$

gehörende Zeitfolge kann durch die einfache Polynomdivision

$$\begin{aligned} (b_m z^m + \dots + b_1 z + b_0) : (z^n + a_{n-1} z^{n-1} + \dots + a_1 z + a_0) \\ = \sum_{k=0}^{\infty} f(k) z^{-k} = F(z) \end{aligned} \quad (3.10)$$

berechnet werden. Diese Methode ist besonders dann geeignet, wenn eine bestimmte Anzahl der ersten Folgenglieder von Interesse ist. Die Polynomdivision lässt sich auch leicht programmtechnisch realisieren. Zur Veranschaulichung sei ein Beispiel aufgeführt.

Die gegebene z -Funktion sei

$$F(z) = \frac{1.08z - 0.4z^2}{(z + 0.5)(z - 0.3)^2} = \frac{-0.4z^2 + 1.08z}{z^3 - 0.1z^2 - 0.21z + 0.045} \quad (3.11)$$

Die elementare Polynomdivision ergibt für die ersten fünf Koeffizienten

$$\begin{array}{r}
 (-0.4z^2 + 1.08z) : (z^3 - 0.1z^2 - 0.21z + 0.045) = 0 \cdot z^0 - 0.4z^{-1} \\
 \underline{-(-0.4z^2 + 0.04z + 0.084 - 0.018z^{-1})} \\
 1.04z - 0.084 + 0.018z^{-1} \qquad \qquad \qquad + 1.04z^{-2} \\
 \underline{-(1.04z - 0.104 - 0.2184z^{-1} + 0.0468z^{-2})} \\
 0.02 + 0.2364z^{-1} - 0.0468z^{-2} \qquad \qquad \qquad + 0.02z^{-3} \\
 \underline{-(0.02 - 0.002z^{-1} - \dots)} \\
 0.2384z^{-1} \dots \qquad \qquad \qquad + 0.2384z^{-4} \\
 \dots \qquad \qquad \qquad \dots \qquad \dots
 \end{array} \tag{3.12}$$

Man kann dann schreiben

$$F(z) = 0 \cdot z^0 - 0.4z^{-1} + 1.04z^{-2} + 0.02z^{-3} + 0.2384z^{-4} + \dots \tag{3.13}$$

und

$$\begin{aligned}
 f(k) &= \{f(0), f(1), f(2), f(3), f(4), \dots\} \\
 &= \{0, -0.4, 1.04, 0.02, 0.2384, \dots\}
 \end{aligned} \tag{3.14}$$

Für die weiteren Folgeglieder muss die elementare Polynomdivision (3.12) entsprechend fortgesetzt werden.

2. Residuenmethode:

Die Auswertung des Umkehrintegrals

$$f(k) = \frac{1}{2\pi j} \oint F(z)z^{k-1}dz \tag{3.15}$$

führt zu

$$f(k) = \sum_{(i)} Res [F(z)z^{k-1}]_{z=z_i} . \tag{3.16}$$

In (3.16) bedeutet „Res“ eine abgekürzte Schreibweise für „Residuum“.

Das Residuum für einen einfachen Pol bei $z = z_i$ erhält man durch

$$Res [F(z)z^{k-1}]_{z=z_i} = \lim_{z \rightarrow z_i} (z - z_i)F(z)z^{k-1} . \tag{3.17}$$

Das Residuum für den ν -fachen Pol lautet

$$Res [F(z)z^{k-1}]_{z=z_i} = \frac{1}{(\nu - 1)!} \lim_{z \rightarrow z_i} \frac{d^{\nu-1}}{dz^{\nu-1}} (z - z_i)^\nu F(z)z^{k-1} . \tag{3.18}$$

Für das Beispiel

$$F(z) = \frac{1.08z - 0.4z^2}{(z + 0.5)(z - 0.3)^2} \tag{3.19}$$

erhält man

$$\begin{aligned}
 f(k) &= \sum_{i=1}^2 \operatorname{Res} [F(z)z^{k-1}]_{z=z_i} \\
 &= \left. \frac{1.08z - 0.4z^2}{(z - 0.3)^2} z^{k-1} \right|_{z=-0.5} + \left. \frac{d}{dz} \frac{1.08z - 0.4z^2}{z + 0.5} z^{k-1} \right|_{z=0.3} \\
 &= 2 \cdot (-0.5)^k + (4k - 2) \cdot (0.3)^k .
 \end{aligned} \tag{3.20}$$

Im Gegensatz zur *Polynomdivision* führt die *Residuenmethode* zu einem analytischen Ausdruck für ein allgemeines Folgenglied $f(k)$, aus dem sich beliebige Glieder durch Einsetzen von k berechnen lassen.

3. Partialbruchzerlegung

Die Partialbruchzerlegung ist im Modul "Steuer- und Regelungstechnik" ausführlich behandelt worden. Deshalb sei hier nur das Beispiel ausgeführt.

Mit dem Partialbruchansatz

$$\begin{aligned}
 F(z) &= \frac{1.08z - 0.4z^2}{(z + 0.5)(z - 0.3)^2} \\
 &= \frac{Az}{(z + 0.5)} + \frac{Bz}{(z - 0.3)} + \frac{Cz}{(z - 0.3)^2}
 \end{aligned} \tag{3.21}$$

erhält man durch Zählerpolynomvergleich bei Nennergleichheit

$$Az(z - 0.3)^2 + Bz(z + 0.5)(z - 0.3) + Cz(z + 0.5) = (1.08 - 0.4z)z \tag{3.22}$$

und anschließendem Koeffizientenvergleich

$$\begin{array}{rcl}
 A + & B & = 0 \\
 -0.6A + 0.2 & B + & C = -0.4 \\
 0.09A - 0.15 & B + 0.5 & C = 1.08
 \end{array} \tag{3.23}$$

das Resultat

$$A = 2, \quad B = -2, \quad C = 1.2 . \tag{3.24}$$

Mit Hilfe der Korrespondenzen

(Nr. 5 und Nr. 6 in Bild 6)

$$\frac{z}{z - e^{-aT}} \bullet \circ e^{-aTk} , \tag{3.25}$$

$$\frac{Tz e^{-bT}}{(z - e^{-bT})^2} \bullet \circ kT e^{-bTk} \tag{3.26}$$

ergibt sich nach Einsetzen der numerischen Werte für die Polstellen

$$\frac{z}{z + 0.5} \quad \bullet \text{---} \circ \quad (-0.5)^k$$

$$\frac{z}{z - 0.3} \quad \bullet \text{---} \circ \quad (-0.3)^k$$

$$\frac{z}{(z - 0.3)^2} \quad \bullet \text{---} \circ \quad k \frac{1}{0.3} (0.3)^k$$

und die Werte für A, B und C aus (3.24) das bekannte Resultat (3.20)

$$f(k) = 2 \cdot (-0.5)^k - 2 \cdot (0.3)^k + 4k(0.3)^k .$$

3.5 Übertragungsfunktion diskreter Systeme

Für die Beschreibung und die Analyse diskreter Systeme haben zwei Eigenschaften der z -Transformation besondere Bedeutung: zum einen können Gleichungen, die im Zeitbereich das dynamische Systemverhalten beschreiben, mit Hilfe der Verschiebungssätze in einen Bildbereich überführt werden, d.h. Differenzgleichungen werden in komplexe Übertragungsfunktionen überführt. Zum anderen ermöglicht der Faltungssatz die Anwendung der Blockschaltbildalgebra auf diskrete Systeme.

Die z -Transformation wird im Bereich der zeitdiskreten Systeme völlig analog so angewendet, wie die Laplace-Transformation im Bereich der kontinuierlichen Systeme. Ein lineares diskretes System n -ter Ordnung wird entsprechend Gl. (2.16) durch die Differenzgleichung

$$y(k) + \sum_{i=1}^n a_i y(k-i) = \sum_{j=0}^m b_j u(k-j), \quad k = 1, 2, 3, \dots \quad (3.27)$$

beschrieben. Wendet man hierauf den Rechtsverschiebungssatz der z -Transformation an, so erhält man

$$Y(z)(1 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_n z^{-n}) = U(z)(b_0 + b_1 z^{-1} + b_2 z^{-2} + \dots + b_m z^{-m}),$$

woraus direkt als Verhältnis der z -Transformierten von Eingangs- und Ausgangsfolge die z -Übertragungsfunktion des diskreten Systems

$$G(z) = \frac{Y(z)}{U(z)} = \frac{(b_0 + b_1 z^{-1} + b_2 z^{-2} + \dots + b_m z^{-m})}{(1 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_n z^{-n})} \quad (3.28)$$

definiert werden kann. Dabei sind die Anfangsbedingungen der Differenzgleichung als Null vorausgesetzt. In Analogie zu den kontinuierlichen Systemen kann die z -Übertragungsfunktion $G(z)$ auch als z -Transformierte der Gewichtsfolge $g(k)$ definiert werden:

$$G(z) = \mathcal{Z}\{g(k)\}. \quad (3.29)$$

Mit der Definition der z -Übertragungsfunktion kann man nun diskrete Systeme formal ebenso behandeln wie kontinuierliche Systeme. Beispielsweise erhält man die Gesamtübertragungsfunktion zweier hintereinandergeschalteter z -Übertragungsfunktionen $G_1(z)$ und $G_2(z)$ zu

$$G(z) = G_1(z) \cdot G_2(z). \quad (3.30)$$

Ebenso ergibt sich für eine Parallelschaltung

$$G(z) = G_1(z) + G_2(z). \quad (3.31)$$

Bei Systemen mit P-Verhalten (Systeme mit Ausgleich) kann mit Hilfe der z -Übertragungsfunktion und dem Endwertsatz leicht die *Systemverstärkung* K eines diskreten Übertragungssystems bestimmt werden. Bei einer sprungförmigen Eingangsfolge

$$u(k) = 1 \quad \text{für } k \geq 0 \quad (3.32)$$

mit der z -Transformierten

$$U(z) = \frac{z}{z-1} \quad (3.33)$$

ergibt sich die Systemverstärkung als stationärer Endwert der Ausgangsfolge:

$$K = \lim_{k \rightarrow \infty} y(k) . \quad (3.34)$$

Nach dem Endwertsatz in Tabelle 3.2 gilt

$$\begin{aligned} \lim_{k \rightarrow \infty} y(k) &= \lim_{z \rightarrow 1} [(z-1)Y(z)] \\ &= \lim_{z \rightarrow 1} [(z-1)G(z)U(z)] \\ &= \lim_{z \rightarrow 1} [(z-1)G(z)\frac{z}{z-1}] \\ &= \lim_{z \rightarrow 1} [G(z)z] \\ &= G(1) , \end{aligned}$$

also unter Berücksichtigung von Gl. (3.28)

$$K = G(1) = \frac{b_0 + b_1 + \dots + b_m}{1 + a_1 + a_2 + \dots + a_n} . \quad (3.35)$$

3.6 Pole und Nullstellen der z -Übertragungsfunktion

Zur Definition der Pole und Nullstellen eines diskreten Übertragungssystems wird zunächst die z -Übertragungsfunktion (3.28) mit z^n erweitert, so daß man eine gebrochene rationale Funktion

$$G(z) = \frac{b_0 z^m + b_1 z^{m-1} + b_2 z^{m-2} + \dots + b_m}{z^n + a_1 z^{n-1} + a_2 z^{n-2} + \dots + a_n} \quad (3.36)$$

mit Polynomen in z erhält, es gilt also

$$G(z) = \frac{Z(z)}{N(z)} . \quad (3.37)$$

Wie bei den kontinuierlichen Systemen wird die Darstellung (3.36) *Polynomform* der z -Übertragungsfunktion genannt. Auch $G(z)$ kann in eine andere Form überführt werden, wenn man die Polynome im Zähler und Nenner als Produkte von Linearfaktoren schreibt (Fundamentalsatz der Algebra). Es gilt

$$b_0 z^m + b_1 z^{m-1} + \dots + b_m = b_0 \prod_{i=1}^m (z - n_i) \quad (3.38)$$

und

$$z^n + a_1 z^{n-1} + \cdots + a_n = \prod_{i=1}^n (z - p_i), \quad (3.39)$$

wobei n_i und p_i die Nullstellen des Zähler- bzw. des Nennerpolynoms von $G(z)$ sind. Bei technisch-physikalischen Systemen müssen alle a_i und b_i reell sein, daher sind die *Polstellen* p_i bzw. die *Nullstellen* n_i von $G(s)$ entweder reell oder konjugiert komplex.

Das auf der linken Seite von (3.39) stehende Polynom nennt man auch hier das *charakteristische Polynom* des Systems.

Pol-Nullstellen-Form der z -Übertragungsfunktion.

Unter Verwendung der Pole und Nullstellen kann die Übertragungsfunktion in der Form

$$G(z) = k_0 \cdot \frac{\prod_{j=1}^m (z - n_j)}{\prod_{i=1}^n (z - p_i)} \quad (3.40)$$

dargestellt werden, wobei

$$k_0 = b_0 \quad (3.41)$$

gilt. Die Darstellung der Übertragungsfunktion in Gl. (3.40) bezeichnet man als *Pol-Nullstellen-Form*.

Wie bei den kontinuierlichen Systemen sind die Pole und Nullstellen der Übertragungsfunktion wichtige Kenngrößen eines diskreten Systems. Dabei sind auch die Nullstellen diskreter Systeme invariant gegenüber Ausgangs- und Zustandsrückführungen. Die Polstellen der z -Übertragungsfunktion charakterisieren das Stabilitätsverhalten des diskreten Systems und können durch geeignete Rückführungen verändert werden.

3.7 z -Übertragungsfunktionen von Abtastsystemen

Für die Ableitung des Zusammenhanges zwischen der Übertragungsfunktion eines kontinuierlichen Systems und der Übertragungsfunktion des zugehörigen zeitdiskreten Systems wird der in Bild 3.2 dargestellte Abtastregelkreis betrachtet.

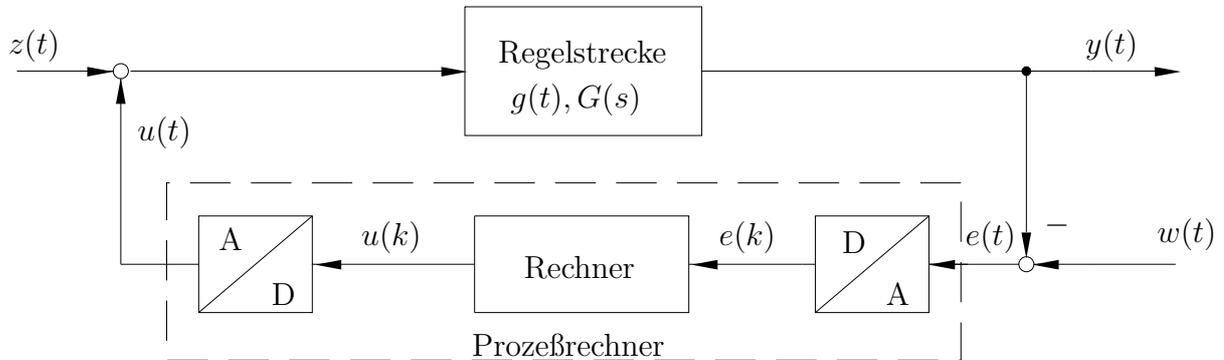


Bild 3.2: Regelkreis mit Prozeßrechnerregelung

Ein Regelkreis ist ein Abtastregelkreis, wenn neben kontinuierlich arbeitenden Systemteilen – meistens die Regelstrecke – mindestens ein Systemteil enthalten ist, der nur zeitdiskrete Signale verarbeitet. Von besonderer Bedeutung sind die Abtastregelkreise in der Form der vom Prozeß- oder Mikrorechner geregelten Systeme (Bild 3.2).

Da die Amplitudenquantisierungseffekte praktisch keine Rolle spielen, wird der A/D-Umsetzer in Bild 3.2 als idealer Abtaster (s. a. Abschnitt 2) und der D/A-Umsetzer als Halteglied 0-ter Ordnung angesehen und behandelt.

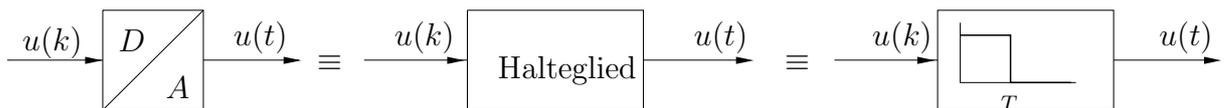


Bild 3.3: Behandlung des DAU als Halteglied 0-ter Ordnung

Für das kontinuierliche Ausgangssignal $u(t)$ des Haltegliedes soll dabei gelten:

$$u(t) = u(t_k) \quad \forall t \in [t_k, t_{k+1}] \quad k = 0, 1, 2, \dots \quad (3.42)$$

Dieses Signal $u(t)$ ist in Bild 3.4 dargestellt.

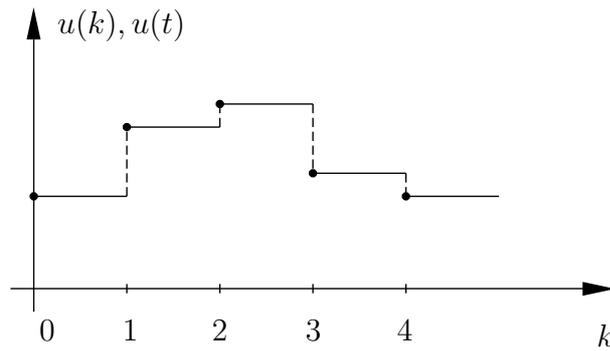


Bild 3.4: Verlauf des kontinuierlichen Signals $u(t)$ am Ausgang eines durch $u(k)$ erregten Haltegliedes 0-ter Ordnung (D/A-Umsetzer)

Abtastregelkreise werden zweckmäßigerweise als zeitdiskrete Systeme analysiert und synthetisiert, indem man der kontinuierlichen Regelstrecke ein *äquivalentes* zeitdiskretes System $G(z)$ zuordnet. Zu diesem Zweck werden der A/D- und der D/A-Umsetzer mit der Regelstrecke zusammengefaßt (Bild 3.5).

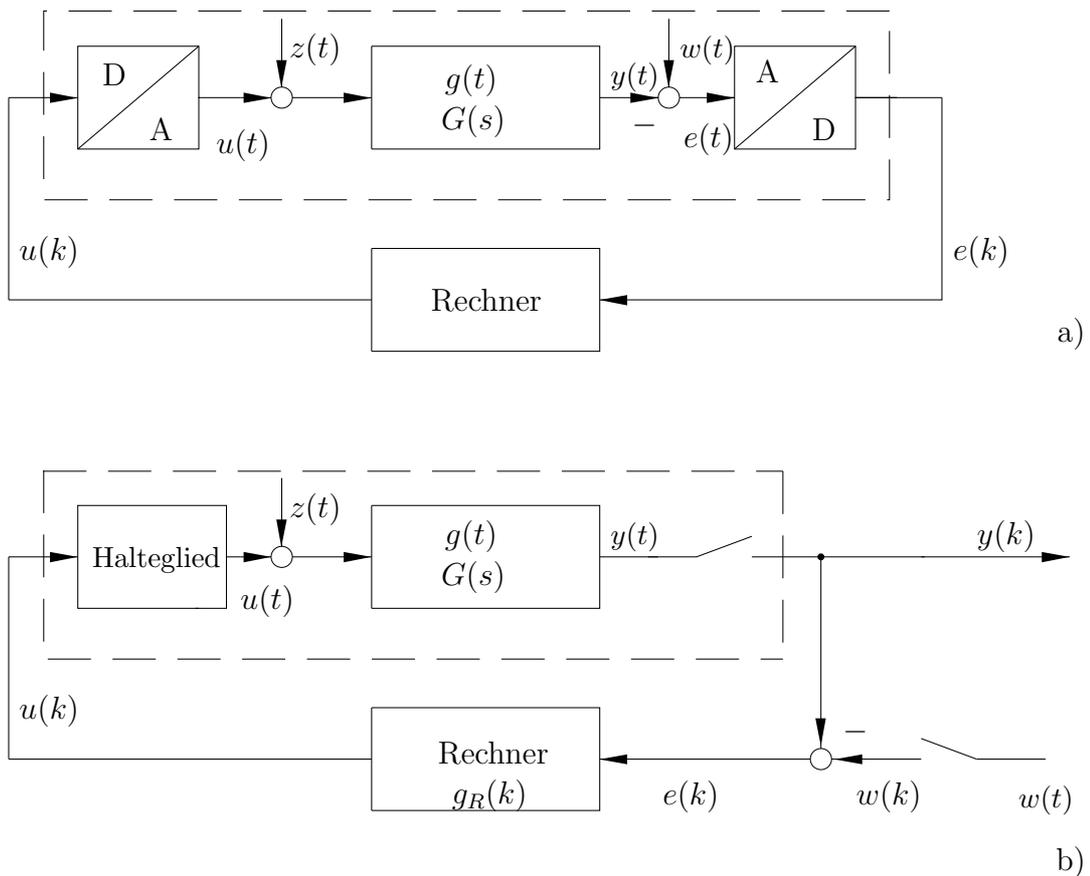


Bild 3.5: Behandlung des Abtastregelkreises als zeitdiskretes System

a) Umformung des Systems aus Bild 3.2

b) Ersatz des DAU bzw. ADU durch ein Halteglied bzw. idealer Tasting

Gesucht ist nun das Übertragungsverhalten des äquivalente zeitdiskreten Systems, das aus einer Reihenschaltung von Halteglied und kontinuierlicher Regelstrecke besteht. Zunächst wird die Antwort des Haltegliedes 0-ter Ordnung auf einen Dirac-Impuls, also die Gewichtsfunktion des Haltegliedes, betrachtet:

$$g_H(t) = 1(t) - 1(t - T), \quad (3.43)$$

zu der die Laplacetransformierte

$$G_H(s) = \frac{1}{s} - \frac{e^{-sT}}{s} = \frac{1 - e^{-sT}}{s} \quad (3.44)$$

gehört.

Die Impulsantwort im s -Bereich der Reihenschaltung von Halteglied und Regelstrecke lautet dann

$$Y_\delta(s) = G_H(s)G(s) = \frac{1 - e^{-sT}}{s}G(s). \quad (3.45)$$

Die zugehörige Antwort im Zeitbereich, die Gewichtsfunktion der Reihenschaltung ergibt sich damit zu

$$g_{HG}(t) = y_\delta(t) = \mathcal{L}^{-1}\left\{\frac{1 - e^{-sT}}{s}G(s)\right\}. \quad (3.46)$$

Durch ideale Abtastung der Gewichtsfunktion erhält man die Gewichtsfolge

$$g_{HG}(kT) = y_\delta(kT) = \mathcal{L}^{-1}\left\{\frac{1 - e^{-sT}}{s}G(s)\right\}_{t=kT}. \quad (3.47)$$

Damit ergibt sich als z -Transformierte dieser Gewichtsfolge die Beziehung

$$G(z) = \mathcal{Z}\{g_{HG}(kT)\} = \mathcal{Z}\left\{\mathcal{L}^{-1}\left\{\frac{1 - e^{-sT}}{s}G(s)\right\}_{t=kT}\right\} \quad (3.48)$$

$$= \mathcal{Z}\left\{\mathcal{L}^{-1}\left\{\frac{G(s)}{s} - \frac{G(s)}{s}e^{-sT}\right\}_{t=kT}\right\}. \quad (3.49)$$

In dieser Beziehung ist

$$h(t) = \mathcal{L}^{-1}\left\{\frac{G(s)}{s}\right\} \quad (3.50)$$

die *Übergangsfunktion* der kontinuierlichen Regelstrecke. Durch Anwendung des Verschiebungssatzes der Laplace-Transformation erhält man für den zweiten Term in Gl. (3.49) die Beziehung

$$h(t - T) = \mathcal{L}^{-1}\left\{\frac{G(s)}{s}e^{-sT}\right\}. \quad (3.51)$$

Durch Abtastung von $h(t)$ bzw. $h(t - T)$ entstehen die Folgen

$$h(kT) = \mathcal{L}^{-1}\left\{\frac{G(s)}{s}\right\}\Big|_{t=kT} \quad (3.52)$$

und

$$h(kT - T) = \mathcal{L}^{-1}\left\{\frac{G(s)}{s}e^{-Ts}\right\}\Big|_{t=kT} . \quad (3.53)$$

Die gesuchte z -Übertragungsfunktion von Strecke samt Halteglied ist nun durch

$$G(z) = \mathcal{Z}\{h(kT) - h(kT - T)\}. \quad (3.54)$$

gegeben. Durch Anwendung des Rechtsverschiebungssatzes der z -Transformation und unter Berücksichtigung von Gl. (3.52) erhält man dann diesen Zusammenhang:

$$G(z) = (1 - z^{-1})\mathcal{Z}\{h(kT)\} \quad (3.55)$$

$$= \left(\frac{z-1}{z}\right)\mathcal{Z}\left\{\mathcal{L}^{-1}\left\{\frac{G(s)}{s}\right\}\Big|_{t=kT}\right\} . \quad (3.56)$$

Dieses Ergebnis läßt sich wie folgt zusammenfassen: Die z -Übertragungsfunktion der Reihenschaltung eines Haltegliedes 0-ter Ordnung und eines kontinuierlichen Systems erhält man wie folgt: Zunächst wird der Ausdruck $G(s)/s$ (Übergangsfunktion des kontinuierlichen Systems) mit Hilfe z.B. der Korrespondenztabelle 3.3 in den z -Bereich transformiert. Anschließend wird der so erhaltene z -Ausdruck noch mit $(z - 1)/z$ multipliziert.

3.8 z -Übertragungsfunktion und Differenzengleichung

Ist eine z -Übertragungsfunktion

$$G(z) = \frac{b_m z^m + b_{m-1} z^{m-1} + b_{m-2} z^{m-2} + \dots + b_1 z + b_0}{z^n + a_{n-1} z^{n-1} + a_{n-2} z^{n-2} + \dots + a_1 z + a_0} = \frac{Y(z)}{U(z)} \quad (3.57)$$

gegeben, so läßt sich die zugehörige Differenzengleichung leicht angeben. Für kausale Systeme gilt in (3.57) immer $m \leq n$.

Durch einfaches Ausmultiplizieren erhält man

$$\begin{aligned} & b_m z^m U(z) + b_{m-1} z^{m-1} U(z) + b_{m-2} z^{m-2} U(z) + \dots + b_1 z U(z) + b_0 U(z) \\ & = z^n Y(z) + a_{n-1} z^{n-1} Y(z) + a_{n-2} z^{n-2} Y(z) + \dots + a_1 z Y(z) + a_0 Y(z) . \end{aligned} \quad (3.58)$$

Nach Division durch die höchste Potenz von z ergibt sich

$$\begin{aligned} & b_m z^{-(n-m)} U(z) + b_{m-1} z^{-(n-m+1)} U(z) + b_{m-2} z^{-(n-m+2)} U(z) + \dots + b_0 z^{-n} U(z) \\ & = Y(z) + a_1 z^{-(n-1)} Y(z) + a_2 z^{-(n-2)} Y(z) + \dots + a_0 z^{-n} Y(z) . \end{aligned} \quad (3.59)$$

Nach Anwendung des Rechtstverschiebungssatzes der z -Transformation in Tabelle 3.2 erhält man aus (3.59) die Differenzgleichung

$$\begin{aligned} b_m u[k-n+m] + b_{m-1} u[k-n+m-1] + \dots + b_1 u[k-n+1] + b_0 u[k-n] \\ = y[k] + a_{n-1} y[k-1] + \dots + a_1 y[k-n+1] + a_0 y[k-n] . \end{aligned} \quad (3.60)$$

Die Gleichung (3.60) kann jetzt nach $y[k]$ aufgelöst werden und man erhält eine Rekursionsformel zur Berechnung der Ausgangsgröße für beliebige Eingangsgrößen:

$$\begin{aligned} y[k] = & -a_{n-1} y[k-1] - \dots - a_1 y[k-n+1] - a_0 y[k-n] + \\ & + b_m u[k-n+m] + \dots + b_1 u[k-n+1] + b_0 u[k-n] . \end{aligned} \quad (3.61)$$

An dieser Gleichung wird auch die Bedeutung der Kausalitätsforderung, $m \leq n$, deutlich. Ist die Bedingung $m \leq n$ nicht erfüllt, so müßten zur Berechnung von $y[k]$ auch zukünftige Eingangswerte $u[i]$ mit $i > k$ bekannt sein.

4 Pole und Nullstellen von Abtastsystemen

Die Stabilität eines kontinuierlichen Übertragungssystems kann anhand der Pole der Übertragungsfunktion überprüft werden. Die Definitionen der Stabilität linearer kontinuierlicher Systeme können unmittelbar auf das in Bild 4.1 dargestellte diskrete System übertragen werden.

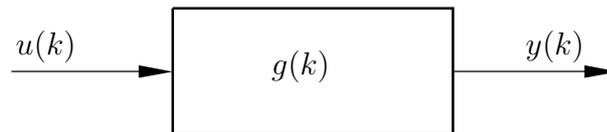


Bild 4.1: Zeitdiskretes System

Definition 4.1

- Ein autonomes zeitdiskretes System ist asymptotisch stabil, wenn es nach Auslenkung aus der Ruhelage selbsttätig dorthin zurückkehrt.
- Ein zeitdiskretes dynamisches System ist Ein-/Ausgangs-stabil (BIBO-stabil), wenn für beschränkte Eingangssignale $u(k)$ für alle t_k das Ausgangssignal $y(k)$ beschränkt ist:

$$|u(k)| \leq N < \infty \Rightarrow |y(k)| \leq M < \infty, \forall k = 0, 1, 2, \dots \quad (4.1)$$

□

Im Modul "Steuer- und Regelungstechnik" wurde bereits gezeigt, daß diese beiden Definitionen bei den hier betrachteten linearen zeitinvarianten Systemen identisch sind. Aus diesem Grund wird im weiteren nur von der Stabilität zeitdiskreter Systeme gesprochen. Aus der Definition der BIBO-Stabilität folgt mit $|u_k| < \infty$ unter Verwendung der Faltungssumme für die Ausgangsgröße

$$|y(k)| \leq \sum_{\nu=0}^{\infty} |g(k-\nu)| |u(\nu)| < \infty. \quad (4.2)$$

Daraus kann direkt diese Stabilitätsbedingung abgeleitet werden:

Satz 4.1

Ein lineares zeitinvariantes zeitdiskretes System mit der Gewichtsfolge $g(k)$ ist dann und nur dann stabil, wenn

$$\sum_{k=0}^{\infty} |g(k)| < \infty \quad (4.3)$$

gilt.

Durch Übergang in den komplexen Bereich zu der z -Transformierten $G(z)$ von $g(k)$ erhält man die folgende *notwendige* und *hinreichende* Stabilitätsbedingung in der z -Ebene:

Satz 4.2

Ein lineares zeitinvariantes zeitdiskretes System mit der Übertragungsfunktion $G(z)$ ist dann und nur dann stabil, wenn alle Pole von $G(z)$ innerhalb des Einheitskreises der z -Ebene liegen, d.h. wenn gilt

$$|z_i| < 1 \quad \text{für} \quad i = 1, 2, \dots, n \quad . \quad (4.4)$$

Der Stabilitätsbereich für die Pole der Übertragungsfunktion $G(z)$ ist in Bild 4.2 dargestellt, wobei die Pole als Wurzeln des Nennerpolynoms von $G(z)$ in der Darstellung mit positiven Potenzen von z definiert sind.

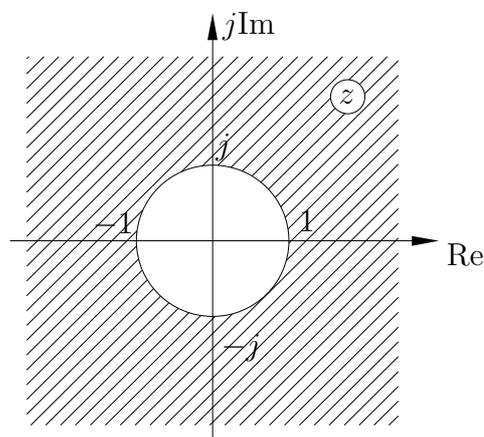


Bild 4.2: Stabilitätsgebiet in der z -Ebene

Diese Stabilitätsbedingung folgt unmittelbar aus der Analogie zwischen der s -Ebene für kontinuierliche und der z -Ebene für zeitdiskrete Systeme. Aus der Definition (3.5) der z -Transformation folgt, daß die linke s -Ebene in das Innere des Einheitskreises der z -Ebene abgebildet wird, wobei

$$|z| = e^{\sigma T} \quad (4.5)$$

und

$$\varphi = \arg z = \omega T \quad (4.6)$$

gilt.

Für die asymptotische Stabilität eines kontinuierlichen Übertragungssystem müssen alle Pole von $G(s)$ in der linken s -Halbebene liegen. Aus dem Abbildungsgesetz der z -Transformation folgt unmittelbar, daß alle Pole der z -Übertragungsfunktion im Innern des Einheitskreises der z -Ebene liegen müssen. Offensichtlich werden die Stabilitätseigenschaften eines kontinuierlichen Systems durch eine äquidistante Abtastung nicht

verändert. Allerdings kann ein instabiles kontinuierliches System mit einer aufklingende Schwingung so abgetastet werden (z.B. Abtastung in den Nulldurchgängen), daß diese Schwingung in der Folge nicht sichtbar wird.

Beispiel:

Gegeben sei ein System erster Ordnung

$$G(z) = \frac{3}{z - a}$$

mit einem Pol an der Stelle $z = a$. Erweitert man Zähler und Nenner mit z , so kann man mit Hilfe der Tabelle 3.3 (Nr. 15 für $k = t/T$) und des Rechtsverschiebungssatzes die zugehörige Gewichtsfolge sofort angeben:

$$g(k) = 3a^{k-1}, \quad k = 1, 2, \dots,$$

Für $a = 0,5$ und $a = 1,2$ sind die Folgen berechnet und in Bild 4.3 dargestellt:

k	g_1 $a = 0,5$	g_2 $a = 1,2$
0	0	0
1	3	3
2	1,5	3,6
3	0,75	4,32
4	0,375	5,184
5	0,188	6,221
6	0,094	7,465

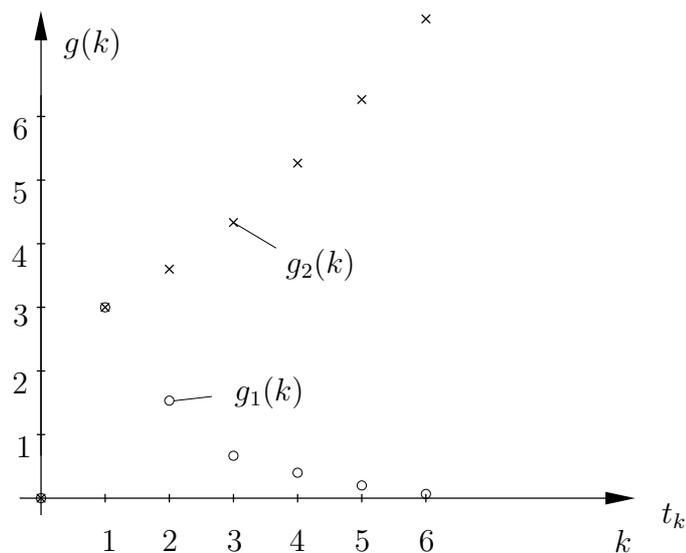


Bild 4.3: Gewichtsfolge eines stabilen und eines instabilen Systems erster Ordnung

4.1 Zusammenhang zwischen den Polen des kontinuierlichen und des Abtastsystems

Der Zusammenhang zwischen der Lage der Pole eines kontinuierlichen Systems in der s -Ebene und dem Übertragungsverhalten im Zeitbereich ist im Modul "Steuer- und Regelungstechnik" ausführlich diskutiert worden. Um die Abhängigkeit des Übertragungsverhaltens eines zeitdiskreten Systems von der Lage seiner Pole in der z -Ebene zu untersuchen, betrachtet man, was mit den Polen eines kontinuierlichen Systems beim Übergang zum zeitdiskreten System passiert. Ausgehend von der konformen (winkeltreuen) Abbildung

$$z = e^{sT} \tag{4.7}$$

mit

$$|z| = e^{\sigma T} \quad \text{und} \quad \varphi = \arg z = \omega T \tag{4.8}$$

lassen sich alle Punkte der s -Ebene eindeutig in Punkte der z -Ebene überführen. Jeder Punkt $s = \sigma + j\omega$ liefert bei der Abbildung in die z -Ebene für ein konstantes Abtastintervall T einen Betrag $|z|$, der nur vom Realteil σ abhängt. Der Winkel φ , also das Argument der komplexen Zahl z wird dann offensichtlich nur von ω bestimmt.

Betrachtet man den Ursprung der s -Ebene, so wird dieser in der z -Ebene auf den Punkt $z = e^{0T} = 1 + j0$ abgebildet. Läßt man einen Punkt in der s -Ebene auf der $j\omega$ -Achse ($\sigma = 0$) wandern, so behält die komplexe Variable den Betrag $|z| = 1$, während sich der Winkel φ im Gegenuhrzeigersinn von $-\pi$ bis π ändert, wenn ω den Bereich $-\pi/T \leq \omega \leq \pi/T$ durchläuft (vgl. Bild 4.4a). Dies entspricht einem Umlauf auf dem sogenannten *Einheitskreis* der z -Ebene. Eine weitere Änderung von ω im Bereich $\pi/T \leq \omega \leq 3\pi/T$ würde dann einen weiteren Umlauf erzeugen. Bewegt sich also ein Punkt auf der $j\omega$ -Achse von $-\infty$ bis ∞ , so wird der Einheitskreis unendlich oft durchlaufen. Das bedeutet, daß ein Punkt in z -Ebene einer unendlichen Anzahl von Punkten in der s -Ebene entspricht.

Linien mit konstantem Realteil ($\sigma = \text{const}$) in der s -Ebene werden in Kreise um den Ursprung der z -Ebene mit dem Radius $|z| = e^{\sigma T}$ abgebildet (4.4b). Linien konstanter Frequenz ($\omega = \text{const}$) in der s -Ebene werden in der z -Ebene als Strahlen abgebildet, die im Ursprung der z -Ebene beginnen und mit der reellen Achse den Winkel $\varphi = \omega T$ bilden (4.4c).

Aus diesen Zusammenhängen lassen sich unmittelbar auch die folgenden Aussagen ableiten:

- Ein kontinuierliches System verfügt über eine absolute Stabilitätsreserve σ_a (vgl. Skript SRT, Bild 3.5), wenn der Realteil aller Pole kleiner als $-\sigma_a$ ist. Ein Abtastsystem mit derselben Stabilitätsreserve hat dann Pole, die in einem Kreis mit dem Radius $e^{-\sigma_a T}$ liegen.

- Haben alle Pole eines kontinuierlichen Systems einen Imaginärteil, der kleiner als $j\omega_0$ ist, so führen diese Pole beim zeitdiskreten System auf Pole, die unterhalb einer Gerade mit dem Winkel $\omega_0 T$ liegen.

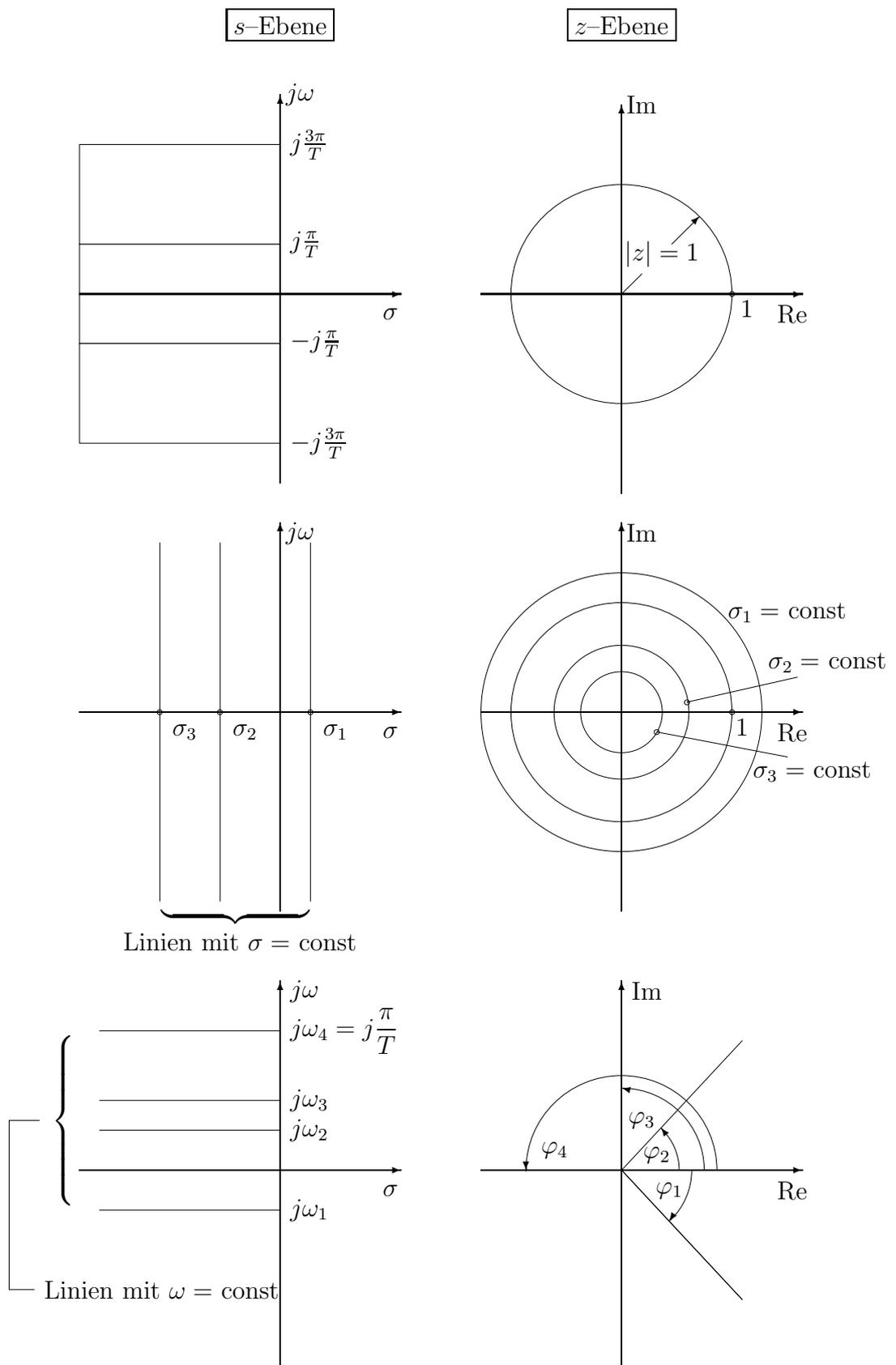


Bild 4.4: Die konforme Abbildung $z = e^{sT}$

Abschließend soll festgehalten werden, daß bei Einhaltung der Bedingung des Abstasttheorems alle Pole von $G(s)$ im sogenannten *Primärstreifen* (Streifen der Breite $2\pi/T$ in der linken s -Halbebene um die reelle Achse) liegen und die Abbildung von s nach z daher eineindeutig ist.

4.2 Zusammenhang zwischen den Nullstellen des kontinuierlichen und des Abtastsystems

Wie im vorhergehenden Abschnitt 4.1 dargestellt, existiert zwischen den Polen eines kontinuierlichen Systems und den Polen des äquivalenten zeitdiskreten Systems (Reihenschaltung eines Halteglieders 0-ter Ordnung, der kontinuierlichen Übertragungsfunktion $G(s)$ und einem idealen Abtaster) ein eindeutiger Zusammenhang. Dieser Zusammenhang zwischen den Polen in der s -Ebene und der z -Ebene wird durch die Abbildung

$$z = e^{sT} \tag{4.9}$$

mit $s = \sigma + j\omega$ und T als Abtastzeit bestimmt.

Ein solcher einfacher Zusammenhang zwischen den Nullstellen einer kontinuierlichen Übertragungsfunktion und den Nullstellen der äquivalenten z -Übertragungsfunktion existiert leider nicht. Insbesondere können minimalphasige kontinuierliche Systeme durch Abtastung nichtminimalphasig werden. Das heißt, die äquivalente z -Übertragungsfunktion hat dann Nullstellen außerhalb des Einheitskreises, obwohl die Nullstellen des kontinuierlichen Systems in der linken s -Halbebene liegen. Daher können Methoden, die bei Abtastsystemen von einer Pol-Nullstellenkürzung Gebrauch machen, bei manchen Abtastzeit hervorragend funktionieren und bei anderen Abtastzeit völlig versagen.

Die äquivalente z -Übertragungsfunktion kann sogar dann Nullstellen haben, wenn die Übertragungsfunktion des kontinuierlichen Systems keine Nullstellen hat. Dies kann man an einem einfachen Beispiel leicht zeigen.

Beispiel:

Das kontinuierliche System bestehe aus einer Reihenschaltung von zwei Integratoren mit der Übertragungsfunktion

$$G(s) = \frac{1}{s^2} . \tag{4.10}$$

Eine Berechnung der äquivalenten z -Übertragungsfunktion ist mit Hilfe der Gleichung

$$G(z) = \frac{z-1}{z} \mathcal{Z}\{h(k)\} \tag{4.11}$$

möglich, wobei $h(k)$ die abgetastete Übergangsfunktion von $G(s)$ ist.

Zunächst ist die Berechnung von $h(t)$ erforderlich:

$$\begin{aligned} h(t) &= \mathcal{L}^{-1}\{G(s) \cdot U(s)\} \quad \text{für} \quad U(s) = \frac{1}{s} \\ &= \mathcal{L}^{-1}\left\{\frac{G(s)}{s}\right\} \\ &= \mathcal{L}^{-1}\left\{\frac{1}{s^3}\right\}. \end{aligned}$$

Mit Hilfe der Korrespondenz 2 der Tabelle A2 im Skript SRT erhält man:

$$h(t) = \mathcal{L}^{-1}\left\{\frac{1}{s^3}\right\} \tag{4.12}$$

$$= \frac{1}{2}t^2. \tag{4.13}$$

Die Anwendung der z -Transformation auf die abgetastete Übergangsfunktion $h(t)$ liefert dann mittels der Korrespondenz 4 in Tabelle 3.3:

$$\mathcal{Z}\left\{\frac{1}{2}t^2\bigg|_{t=kT}\right\} = \frac{T^2 z \cdot (z + 1)}{2(z - 1)^3}. \tag{4.14}$$

Damit ergibt sich die gesuchte äquivalente z -Übertragungsfunktion des Doppelintegrators mit Hilfe der Gl. (4.11) zu:

$$G(z) = \frac{z - 1}{z} \cdot \frac{T^2 z \cdot (z + 1)}{2(z - 1)^3} \tag{4.15}$$

$$= \frac{T^2}{2} \cdot \frac{z + 1}{(z - 1)^2}. \tag{4.16}$$

Das bedeutet, obwohl die kontinuierliche Übertragungsfunktion keine Nullstellen besitzt, hat das zugehörige Abtastsystem eine Nullstelle auf dem Stabilitätsrand der z -Ebene.

□

Allgemein gilt sogar der folgende Zusammenhang, der 1984 von Aström, Hagander und Sternby bewiesen wurde: Die äquivalente z -Übertragungsfunktion der Übertragungsfunktion $G(s) = s^{-n}$ ist

$$G(z) = \frac{T^n}{n!} \frac{B_n(z)}{(z - 1)^n} \tag{4.17}$$

mit

$$B_n(z) = b_1 z^{n-1} + b_2 z^{n-2} + \dots + b_n \tag{4.18}$$

und

$$b_k^n = \sum_{l=1}^k (-1)^{k-l} l^n \binom{n+1}{k-l}, \quad k = 1, \dots, n. \tag{4.19}$$

Für $n = 1, \dots, 6$ ergeben sich folgende Polynome:

$$B_1(z) = 1$$

$$B_2(z) = z + 1$$

$$B_3(z) = z^2 + 4z + 1$$

$$B_4(z) = z^3 + 11z^2 + 11z + 1$$

$$B_5(z) = z^4 + 26z^3 + 66z^2 + 26z + 1$$

$$B_6(z) = z^5 + 57z^4 + 302z^3 + 302z^2 + 57z + 1$$

Diese Polynome haben für $n \geq 2$ immer Nullstellen auf oder außerhalb des Einheitskreises. Die Nullstellen mit einem Betrag größer als 1 für die Polynome $B_2(z), \dots, B_6(z)$ sind:

n	Nullstellen von $B_n(z)$, die außerhalb des Einheitskreises liegen
2	-1
3	-3, 732
4	-1, -9, 899
5	-2, 322, -23, 20
6	-1, -4, 542, -51, 22

Die Polynome $B_n(z)$ sind sogenannte *Euler-Frobenius-Polynome*, die auch durch die folgende Rekursionsformel berechnet werden können:

$$B_1(z) = 1 \tag{4.20}$$

$$B_{k+1}(z) = (1 + kz)B_k(z) + z(1 - z)\frac{dB_k(z)}{dz}, \quad k = 1, 2, \dots \tag{4.21}$$

Aus der Literatur ist bekannt, daß die Euler-Frobenius-Polynome folgende Eigenschaften haben:

- Die Wurzeln der Polynome $B_n(z)$ sind reell, einfach und immer negativ.
- Wenn n_i^* eine Wurzel von $B_n(z)$ ist, dann ist $\frac{1}{n_i^*}$ ebenfalls eine Nullstelle von $B_n(z)$.
- Außerdem gilt $B_n(1) = n!$.

Für beliebige rationale Übertragungsfunktionen $G(s)$ existiert kein einfacher Zusammenhang zwischen den Nullstellen des kontinuierlichen und des äquivalenten zeitdiskreten Systems. Unabhängig von der Anzahl der Nullstellen von $G(s)$ hat die z -Übertragungsfunktion in der Regel $n - 1$ Nullstellen. Allgemeingültige Aussagen können nur für die Grenzfälle $T \rightarrow 0$ und $T \rightarrow \infty$ angegeben werden.

Satz 4.3

Sei $G(s)$ eine rationale Übertragungsfunktion

$$G(s) = k \frac{(s - n_1)(s - n_2) \cdots (s - n_m)}{(s - p_1)(s - p_2) \cdots (s - p_n)} \quad (4.22)$$

in Pol–Nullstellenform mit $m < n$ und $G(z)$ die zugehörige z –Übertragungsfunktion.

Wenn die Abtastzeit T gegen Null geht, streben m Nullstellen von $G(z)$ gegen 1, so wie es $e^{n_i T}$ dann tut, und $n - m - 1$ Nullstellen von $G(z)$ gegen die Nullstellen des Euler–Frobenius–Polynoms $B_{n-m}(z)$. \square

Man kann bei den zeitdiskreten Systemen also zwei Arten von Nullstellen unterscheiden: Zum einen die sogenannten *eigentlichen Nullstellen* (intrinsic zeros), die zu den Nullstellen des kontinuierlichen Systems korrespondieren und für $T \rightarrow 0$ zu $z = 1$ wandern.

Zum anderen Nullstellen, die erst durch die Abtastung entstehen und somit im weiteren *Diskretisierungsnullstellen* genannt werden. Die Anzahl dieser Diskretisierungsnullstellen ist nur vom Differenzgrad $d = n - m$ des kontinuierlichen Systems abhängig, so daß das zeitdiskrete System in der Regel $d - 1$ Diskretisierungsnullstellen hat. Diese Nullstellen wandern für $T \rightarrow 0$ zu den Wurzeln des entsprechenden Euler–Frobenius–Polynoms (4.18).

Aus diesen Zusammenhängen und den Eigenschaften eines Euler–Frobenius–Polynoms folgt auch sofort diese weitere Aussage:

- Die z –Übertragungsfunktion eines kontinuierlichen Systems mit einem Differenzgrad $n - m > 2$ wird für entsprechend kleine Abtastzeiten T immer Diskretisierungsnullstellen außerhalb des Einheitskreises haben. Dies kann schon bei technisch relevanten Abtastzeiten auftreten, so daß nichtminimalphasige äquivalente zeitdiskrete Systeme häufig vorkommen.

Beispiel:

Betrachtet wird ein kontinuierliches System (Reihenschaltung von 3 PT₁–Systemen) mit der Übertragungsfunktion

$$G(s) = \frac{1}{(s + 1)^3} .$$

Die zugehörige z –Übertragungsfunktion ergibt sich zu

$$G(z) = \frac{b_1 z^2 + b_2 z + b_3}{(z - e^{-T})^3}$$

mit

$$\begin{aligned} b_1 &= 1 - (1 + T + T^2/2)e^{-T} \\ b_2 &= (-2 + T + T^2/2)e^{-T} + (2 + T - T^2/2)e^{-2T} \\ b_3 &= (1 - T + T^2/2)e^{-2T} - e^{-3T} . \end{aligned}$$

Diese z -Übertragungsfunktion hat eine Nullstelle außerhalb des Einheitskreises für Abtastzeiten mit $T < 1,8399$ s. Für zum Beispiel $T = 0.01$ liegen die Nullstellen bei $z = -0,2659$ und $z = -3,7042$. \square

Für den technisch weniger relevanten Fall von sehr großen Abtastzeiten gilt dieser Zusammenhang:

Satz 4.4

Sei $G(s)$ eine rationale Übertragungsfunktion mit $m < n$ und $G(0) \neq 0$ sowie $\text{Re } p_i < 0$. Dann wandern alle Nullstellen der z -Übertragungsfunktion gegen Null, wenn die Abtastzeit T unendlich groß wird. \square

Aus diesem Zusammenhang folgt sofort, daß es auch nichtminimalphasige kontinuierliche Systeme gibt, die durch Abtastung minimalphasig werden.

Beispiel:

Die kontinuierliche Übertragungsfunktion

$$G(s) = \frac{6(1-s)}{(s+2)(s+3)}$$

ist stabil und hat eine Nullstelle in der rechten s -Halbebene bei $s = 1$. Die Bedingungen des vorhergehenden Satzes sind erfüllt, so daß es eine hinreichend große Abtastzeit geben muß, für die $G(z)$ keine Nullstellen außerhalb des Einheitskreises hat.

Die Nullstelle der z -Übertragungsfunktion ergibt sich zu

$$z_1 = -\frac{8e^{-2T} - 9e^{-3T} + e^{-5T}}{1 - 9e^{-2T} + 8e^{-3T}}.$$

Für $T = 1,2485$ liegt die Nullstelle bei $z_1 = -1.0$ und für größere Abtastzeiten immer innerhalb des Einheitskreises. \square

Wie das folgende Beispiel zeigt, ist es auch möglich, daß eigentliche Nullstellen außerhalb des Einheitskreises liegen obwohl das zugehörige kontinuierliche System minimalphasig ist.

Beispiel:

Betrachtet wird eine kontinuierliche, minimalphasige Übertragungsfunktion

$$G(s) = \frac{(1+T_1s)(1+T_2s)}{(s+1)^3}$$

mit einen dreifachen Pol bei $s = -1$ und zwei Nullstellen bei $s = -1/T_1$ und $s = -1/T_2$.

Die zugehörige z -Übertragungsfunktion des Abtastsystems wird für $T_1, T_2 \neq 0$ keine Diskretisierungsnullstellen haben, da die kontinuierliche Übertragungsfunktion einen Differenzgrad von $d = 1$ hat.

Für die äquivalente z -Übertragungsfunktion kann die folgende allgemeine Form angegeben werden:

$$G(z) = \frac{b_2 z^2 + b_1 z + b_0}{(z - e^{-T})^3} \quad (4.23)$$

mit

$$b_2 = 1 - (1 + \alpha_1 T + \alpha_2 T^2) e^{-T} \quad (4.24)$$

$$b_1 = (-2 + \alpha_1 T + \alpha_2 T^2) e^{-T} + (2 + \alpha_1 T - \alpha_2 T^2) e^{-2T} \quad (4.25)$$

$$b_0 = (1 - \alpha_1 T + \alpha_2 T^2) e^{-2T} - e^{-3T} \quad (4.26)$$

$$(4.27)$$

und

$$\alpha_1 = 1 - T_1 T_2, \quad \alpha_2 = \frac{1 + T_1 T_2 - T_1 - T_2}{2}. \quad (4.28)$$

Die folgende Tabelle 4.1 listet die Werte der eigentlichen Nullstellen ($T_1 = 110$ und $T_2 = 100$) für verschiedene Abtastzeit auf.

Abtastzeit T	Nullstellen
0,1	$z_{1,2} = 0,9991 \pm 0,0003i$
0,4	$z_{1,2} = 0,9988 \pm 0,0037i$
0,5	$z_{1,2} = 1,0003 \pm 0,005i$
0,8	$z_1 = 1,0269, z_2 = 1,0026$
1,0	$z_1 = 1,0743, z_2 = 1,0016$
2,0	$z_1 = 13,9513, z_2 = 1,0009$

Tabelle 4.1: Nullstellen der z -Übertragungsfunktion für verschiedene Abtastzeiten

Die Nullstellen der kontinuierlichen Übertragungsfunktion liegen bei $s_1 = 1/T_1 = -0,0091$ und $s_2 = 1/T_2 = -0,01$ und damit in der linken s -Halbebene. Der Tabelle 4.1 kann man entnehmen, daß die eigentlichen Nullstellen der zugehörigen z -Übertragungsfunktion für Abtastzeiten $T < 0,5$ innerhalb des Einheitskreises liegen und für größere Abtastzeiten außerhalb.

4.3 Stabilitätsanalyse und w -Transformation

Bei der Stabilitätsanalyse eines zeitdiskreten Systems muß überprüft werden, ob alle Wurzeln z_i der *charakteristischen Gleichung*

$$C(z) = \alpha_n z^n + \alpha_{n-1} z^{n-1} + \dots + \alpha_1 z + \alpha_0 = 0 \quad (4.29)$$

innerhalb des Einheitskreises der z -Ebene liegen. Diese Darstellung der charakteristischen Gleichung folgt aus Gl. (3.36) durch Nullsetzen und entsprechende Darstellung des Nennerpolynoms.

Soll die Stabilität eines Systems ohne explizite Berechnung der Wurzeln der charakteristischen Gleichung überprüft werden, so ist dies bei kontinuierlichen Systemen mittels Hurwitz- und Routhkriterium möglich. Durch eine einfache Transformation läßt sich das Stabilitätsproblem zeitdiskreter Systeme nun in eines für kontinuierlichen Systeme überführen, das dann z.B. mit dem Hurwitz-Kriterium gelöst werden kann.

Durch die Transformation

$$w = \frac{z-1}{z+1}, \quad (4.30)$$

w -Transformation genannt, bildet man das Innere des Einheitskreises der z -Ebene in die linke w -Ebene ab. Diese Abbildung ist eineindeutig, so daß man durch

$$z = \frac{1+w}{1-w} \quad (4.31)$$

auch die w -Ebene in die z -Ebene abbilden kann. Damit werden bei einem stabilen zeitdiskreten System alle Wurzeln z_i der charakteristischen Gleichung (4.29) in die linke w -Ebene abgebildet. Wendet man die Transformationsbeziehung (4.31) auf die charakteristische Gleichung (4.29) an, so erhält man

$$a_n \left(\frac{1+w}{1-w} \right)^n + a_{n-1} \left(\frac{1+w}{1-w} \right)^{n-1} + \dots + a_1 \left(\frac{1+w}{1-w} \right) + a_0 = 0 \quad (4.32)$$

und nach Umformung ein Polynom n -ten Grades in w :

$$\tilde{\alpha}_n w^n + \tilde{\alpha}_{n-1} w^{n-1} + \dots + \tilde{\alpha}_1 w + \tilde{\alpha}_0 = 0. \quad (4.33)$$

Auf dieses Polynom kann dann das Hurwitz-Kriterium angewendet werden.

5 Beschreibung und Analyse im Zustandsraum

5.1 Die zeitdiskrete Zustandsraumdarstellung

Die Zustandsraumdarstellung ist bei zeitdiskreten Systemen in der gleichen Weise anwendbar wie bei kontinuierlichen Systemen. In Analogie zu den zeitkontinuierlichen Systemen kann ein zeitdiskretes System durch ein Zustandsmodell in der Form eines gekoppelten Systems von Differenzgleichungen 1. Ordnung beschrieben werden:

$$\left. \begin{aligned} \mathbf{x}(k+1) &= \mathbf{A}\mathbf{x}(k) + \mathbf{b}u(k) & ; & \quad \mathbf{x}_0 = \mathbf{x}(0) \\ y(k) &= \mathbf{c}^T\mathbf{x}(k) + d u(k) & \quad k &= 0, 1, 2, \dots \end{aligned} \right\} \quad (5.1)$$

$$\mathbf{x} \in \mathbb{R}^n \quad ; \quad u \in \mathbb{R} \quad ; \quad y \in \mathbb{R}$$

Die erste Gleichung wird wieder als Zustandsgleichung oder auch als *Zustandsdifferenzgleichung* bezeichnet, obwohl sie im engeren Sinne des Wortes nicht die Differenz aufeinanderfolgender Zustände, sondern die aufeinanderfolgenden Zustände selbst beschreibt. Die zweite Zeile ist die Ausgangsgleichung.

Das zeitdiskrete Zustandsraummodell (5.1) ist dem kontinuierlichen Modell sehr ähnlich: \mathbf{x} ist der n -dimensionale Zustandsvektor, \mathbf{A} eine $(n \times n)$ -Matrix, \mathbf{b} ein n -dimensionaler Spaltenvektor, \mathbf{c}^T ein n -dimensionaler Zeilenvektor und d ein Skalar. Allerdings unterscheiden sich \mathbf{A} und \mathbf{b} von den entsprechenden Modellparametern der kontinuierlichen Systembeschreibung (vgl. Abschnitt 5.4).

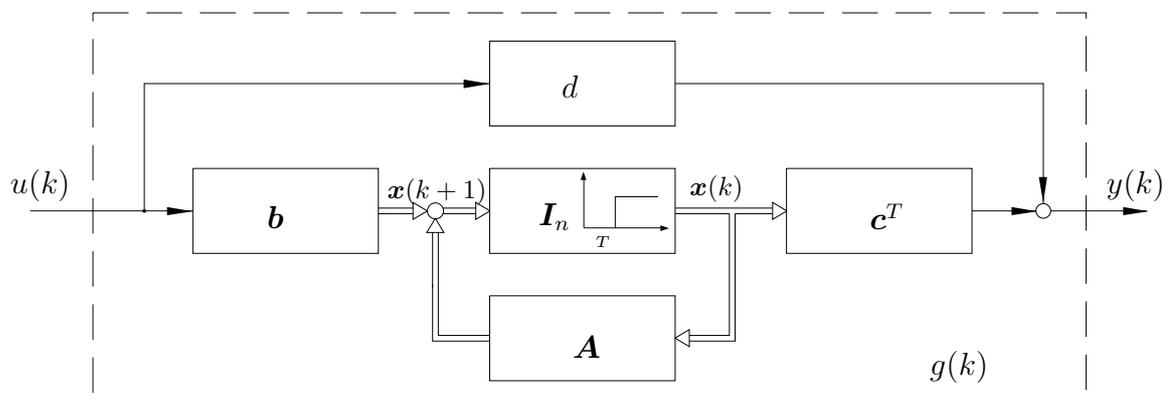


Bild 5.1: Blockschaltbild des zeitdiskreten Zustandsmodells

Das Strukturbild Bild 5.1 unterscheidet sich von dem kontinuierlicher Systeme dadurch, daß die n Integratoren durch n parallele Speicherelemente ersetzt werden, die den Zustandsvektor $\mathbf{x}(k+1)$ bis zum nächsten Taktzeitpunkt speichern und dann als aktuellen Zustand $\mathbf{x}(k)$ ausgeben. Man spricht auch von einem Laufzeitglied für den n -dimensionalen Zustandsvektor.

5.2 Ableitung des Zustandsraummodells aus der Differenzgleichung

Betrachtet wird eine Übertragungsdifferenzgleichung mit explizit angegebener Taktzeit T , wobei der Koeffizient bei der höchsten Differenz $y(k+n)$ auf 1 normiert wurde und nur der Koeffizient b_0 auf der rechten Seite von Null verschieden ist:

$$y((k+n)T) + a_{n-1}y((k+n-1)T) + \dots + a_1y((k+1)T) + a_0y(kT) = b_0u(kT), \quad k = 0, 1, 2, \dots \quad (5.2)$$

Ersetzt man kT durch k so erhält man dann:

$$y(k+n) + a_{n-1}y(k+n-1) + \dots + a_1y(k+1) + a_0y(k) = b_0u(k) \quad k = 0, 1, 2, \dots \quad (5.3)$$

Die Wahl der Zustandsvariablen fällt leicht, da die Differenzgleichung (5.3) zeigt, daß man sich die Folge $y(k), y(k+1), \dots, y(k+n-1)$ der Ausgangsgröße merken muß, um den nächsten Wert $y(k+n)$ berechnen zu können. Daher verwendet man die Zustandsvariablen

$$\left. \begin{aligned} y(k) &= x_1(k) \\ y(k+1) &= x_2(k) = x_1(k+1) \\ y(k+2) &= x_3(k) = x_2(k+1) \\ &\dots \dots \dots \dots \dots \\ y(k+n-1) &= x_n(k) = x_{n-1}(k+1) \\ y(k+n) &= x_n(k+1) \end{aligned} \right\} \quad (5.4)$$

und findet die zugehörige Normalform der Vektordifferenzgleichung zu:

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ \vdots \\ x_{n-1}(k+1) \\ x_n(k+1) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -a_0 & -a_1 & -a_2 & \dots & -a_{n-1} \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ \vdots \\ x_{n-1}(k) \\ x_n(k) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ b_0 \end{bmatrix} u(k)$$

$$y(k) = [1 \quad 0 \quad 0 \quad 0 \quad \dots \quad 0] \mathbf{x}(k) \quad . \quad (5.5)$$

Hinweis: Hierbei handelt es sich um ein System ohne Nullstellen im Endlichen, das beispielsweise einen Regelalgorithmus beschreibt. Geht man von einem kontinuierlichen – auch nullstellenfreien – System aus, dann hat das zugehörige äquivalente zeitdiskrete System normalerweise $n-1$ Nullstellen, und damit ist dann in der vorstehenden Normalform der Vektor $\mathbf{b} = [b_1, b_2, \dots, b_n]^T$ voll besetzt: $b_i \neq 0, \forall i = 1, 2, \dots, n$.

Die Systemmatrix

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1} \end{bmatrix} \quad (5.6)$$

in (5.5) hat Frobeniusform, so daß man das charakteristische Polynom des zeitdiskreten Zustandsmodells direkt ablesen kann:

$$C(\lambda) = \det(\lambda \mathbf{I} - \mathbf{A}) = \lambda^n + a_{n-1} \lambda^{n-1} + \cdots + a_1 \lambda + a_0 . \quad (5.7)$$

Die Polynomkoeffizienten a_i sind die Koeffizienten der linken Seite der Differenzgleichung (5.3).

Da das Modell zeitdiskreter Systeme dem kontinuierlicher Systeme sehr ähnlich ist, können viele Analyse- und Reglerentwurfsmethoden von den kontinuierlichen Systemen entweder direkt übernommen oder in Analogie dazu entwickelt werden. Bezüglich der strukturellen Eigenschaften wie Stabilität, Steuer- und Beobachtbarkeit ergeben sich bei diskreten Systemen keine neuen Aspekte. Es können daher die bei kontinuierlichen Systemen angestellten Überlegungen unmittelbar auf den diskreten Fall übertragen werden.

5.3 Lösung der Zustandsgleichung

Anders als im Falle der kontinuierlichen Systeme, wo die die Dynamik des Systems beschreibende *Differentialgleichung* integriert werden muß, stellen die Systemgleichungen eines zeitdiskreten Systems

$$\left. \begin{aligned} \mathbf{x}(k+1) &= \mathbf{A}\mathbf{x}(k) + \mathbf{b}u(k) \\ y(k) &= \mathbf{c}^T \mathbf{x}(k) + du(k) \end{aligned} \right\} \mathbf{x}_0 = \mathbf{x}(k)|_{k=0}; \quad k = 0, 1, 2, \dots \quad (5.8)$$

eine Rekursionsbeziehung dar, die für einen gegebenen Anfangszustand \mathbf{x}_0 und gegebene Eingangsfolge $u(k)$ nacheinander für alle interessierenden Zeitwerte t_k mit $k = 0, 1, 2, \dots$ durch Einsetzen auflösbar ist.

Aus systemtheoretischen Gründen ist es aber zweckmäßig, auch diese Gleichungen ähnlich wie bei der Differentialgleichungstheorie zu behandeln:

a) Lösung der homogenen Differenzgleichung

$$\left. \begin{aligned} \mathbf{x}(k+1) &= \mathbf{A}\mathbf{x}(k) \quad ; \quad \mathbf{x}_0 = \mathbf{x}(0) \quad ; \quad k = 0, 1, 2, \dots \\ k = 0 & \quad : \quad \mathbf{x}(1) = \mathbf{A}\mathbf{x}_0 \\ k = 1 & \quad : \quad \mathbf{x}(2) = \mathbf{A}\mathbf{x}(1) = \mathbf{A}^2 \mathbf{x}_0 \end{aligned} \right\} \quad (5.9)$$

allgemein:

$$\mathbf{x}(k) = \mathbf{A}^k \mathbf{x}_0 . \quad (5.10)$$

Ein Vergleich mit dem kontinuierlichen Fall (MA-Skript RT, S. 56) zeigt, daß sich die Übergangsmatrix (Fundamentalmatrix) $\Phi(k)$ des zeitdiskreten Systems aus

$$\Phi(k) = \mathbf{A}^k \quad (5.11)$$

berechnet.

b) Lösung der inhomogenen Gleichung für $\mathbf{x}_0 = \mathbf{0}$

$$\left. \begin{aligned} \mathbf{x}(k+1) &= \mathbf{A}\mathbf{x}(k) + \mathbf{b}u(k) \quad ; \quad k = 0, 1, 2 \\ k = 0 : \mathbf{x}(1) &= \mathbf{b}u(0) \\ k = 1 : \mathbf{x}(2) &= \mathbf{A}\mathbf{x}(1) + \mathbf{b}u(1) = \mathbf{A}\mathbf{b}u(0) + \mathbf{b}u(1) \\ k = 2 : \mathbf{x}(3) &= \mathbf{A}\mathbf{x}(2) + \mathbf{b}u(2) = \mathbf{A}^2\mathbf{b}u(0) + \mathbf{A}\mathbf{b}u(1) + \mathbf{b}u(2) \end{aligned} \right\} \quad (5.12)$$

allgemein:

$$\begin{aligned} \mathbf{x}(k) &= \sum_{i=0}^{k-1} \mathbf{A}^{k-i-1} \mathbf{b}u(i) \\ &= \sum_{i=0}^{k-1} \Phi(k-i-1) \mathbf{b}u(i) \quad k = 1, 2, \dots \end{aligned}$$

c) Gesamtlösung

Die Gesamtlösung, die zeitdiskrete Bewegungsgleichung, ergibt sich durch Überlagerung der Fälle a) und b):

$$\begin{aligned} \mathbf{x}(k) &= \Phi(k) \mathbf{x}_0 + \sum_{i=0}^{k-1} \Phi(k-i-1) \mathbf{b}u(i) \\ &= \mathbf{A}^k \mathbf{x}_0 + \sum_{i=0}^{k-1} \mathbf{A}^{(k-i-1)} \mathbf{b}u(i) \quad k = 1, 2, \dots \end{aligned}$$

Damit erhalten wir dann für den Systemausgang:

$$\left. \begin{aligned} y(k) &= \mathbf{c}^T \mathbf{x}(k) + du(k) \\ &= \underbrace{\mathbf{c}^T \left[\mathbf{A}^k \mathbf{x}_0 \right]}_{k=0,1,2,\dots} + \underbrace{\mathbf{c}^T \left[\sum_{i=0}^{k-1} \mathbf{A}^{(k-i-1)} \mathbf{b}u(i) \right]}_{k=1,2,\dots} + \underbrace{du(k)}_{k=0,1,2,\dots} \end{aligned} \right\} \quad (5.13)$$

Den Zusammenhang zwischen Zustandsmodell und Gewichtsfolge erhalten wir, wenn für $\mathbf{x}_0 = \mathbf{0}$

$$u(k) = \delta(k) = \begin{cases} 1 & \text{für } k = 0 \\ 0 & \text{für } k > 0 \end{cases}$$

eingesetzt wird:

$$g(k) = \begin{cases} d & \text{für } k = 0 \\ \mathbf{c}^T \mathbf{A}^{k-1} \mathbf{b} & \text{für } k = 1, 2, \dots \end{cases} \quad (5.14)$$

5.4 Zusammenhang zwischen der kontinuierlichen und der zeitdiskreten Zustandsraumdarstellung

Nachfolgend soll untersucht werden, wie ein gegebenes kontinuierliches System

$$\left. \begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t) \quad ; \quad \mathbf{x}_0 = \mathbf{0} \\ y(t) &= \mathbf{c}^T \mathbf{x}(t) \end{aligned} \right\} \quad (5.15)$$

unter Beachtung der Eigenschaften des Abtasters und des Halteglieders in eine äquivalente diskrete Darstellung im Zustandsraum umgewandelt werden kann, so daß die diskreten Eingangs-, Ausgangs- und Zustandsgrößen in den Abtastzeitpunkten mit den entsprechenden abgetasteten Signalen des kontinuierlichen Systems (5.15) identisch sind. Ein solcher Zusammenhang zwischen der kontinuierlichen und der diskreten Zustandsraumdarstellung kann nur dann angegeben werden, wenn der Zeitverlauf des kontinuierlichen Eingangsvektors $u(t)$ bekannt ist.

Für die weiteren Betrachtungen wird angenommen, daß die Eingangsgröße aufgrund des Halteglieders 0-ter Ordnung zwischen den Abtastzeitpunkten konstant ist:

$$u(t) = \text{const.} \quad \text{im Intervall } [t_k, t_{k+1}] . \quad (5.16)$$

Dann erhält man die Lösung der Zustandsgleichung des kontinuierlichen Systems für den Endpunkt t_{k+1} des Intervalls zu

$$\mathbf{x}(t_{k+1}) = e^{\mathbf{A}(t_{k+1}-t_k)} \mathbf{x}(t_k) + \left[\int_{t_k}^{t_{k+1}} e^{\mathbf{A}(t_{k+1}-\tau)} d\tau \right] \mathbf{b}u(t_k) . \quad (5.17)$$

Setzt man eine äquidistante Abtastung

$$t_{k+1} - t_k = T \quad , \quad \forall \quad k = 0, 1, 2, \dots , \quad (5.18)$$

voraus, so vereinfacht sich die Lösung zu

$$\mathbf{x}(k+1) = e^{\mathbf{A}T} \mathbf{x}(k) + \int_0^T e^{\mathbf{A}\tau} d\tau \mathbf{b}u(k) . \quad (5.19)$$

Mit den Abkürzungen

$$\mathbf{A}_d = e^{\mathbf{A}T} \quad ; \quad (5.20)$$

$$\mathbf{b}_d = \int_0^T e^{\mathbf{A}\tau} d\tau \mathbf{b} \quad (5.21)$$

erhalten wir für die Regelstrecke mit Halteglied am Eingang und Abtaster am Ausgang eine zeitdiskrete Ersatzbeschreibung, die das Systemverhalten zu den Abtastzeitpunkten t_k exakt beschreibt

$$\left. \begin{aligned} \mathbf{x}(k+1) &= \mathbf{A}_d \mathbf{x}(k) + \mathbf{b}_d u(k) \quad ; \quad \mathbf{x}_0 = \mathbf{x}(0) \\ y(k) &= \mathbf{c}^T \mathbf{x}(k) \quad \quad \quad k = 0, 1, 2, \dots \end{aligned} \right\} \quad (5.22)$$

Die Ausgangsgleichung erhält man dabei durch ideale Abtastung der kontinuierlichen Ausgangsgröße $y(t)$ zu den Abtastzeitpunkten $t = kT$:

$$y(kT) = \mathbf{c}^T \mathbf{x}(kT) . \quad (5.23)$$

Mit der üblichen Abkürzung $y(k)$ für $y(kT)$ zeigt diese Gleichung, daß in der kontinuierlichen und diskreten Ausgangsgleichung derselbe Vektor \mathbf{c}^T steht.

Beispiel:

Gegeben sei eine Regelstrecke $G(s) = \frac{1}{s(s+1)}$, die wie in Bild 5.2 angedeutet, über ein Halteglied angesteuert wird.

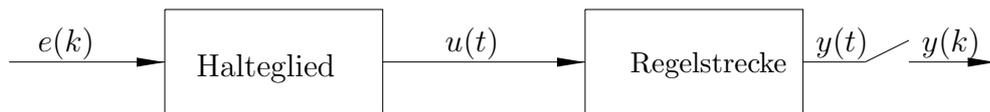


Bild 5.2: Regelstrecke des Beispiels

Mit der Partialbruchentwicklung

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

und indem man jedem der beiden Teilsysteme 1. Ordnung eine Zustandsvariable x_1 bzw. x_2 zuordnet, kann dem kontinuierlichen System mit der Übertragungsfunktion $G(s)$ dieses Zustandsmodell zugeordnet werden:

$$\begin{aligned} \begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} &= \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} \cdot u(t) \\ y(t) &= [1 \quad -1] \mathbf{x}(t) . \end{aligned}$$

Für das äquivalente System mit $e(k)$ als Eingang und $y(k)$ als Ausgang bestimmen wir:

$$\Phi(t - t_0) = e^{\mathbf{A}(t-t_0)} = \begin{bmatrix} e^{0t} & 0 \\ 0 & e^{-1t} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & e^{-t} \end{bmatrix} .$$

$$\mathbf{A}_d = \Phi(T) = e^{\mathbf{A}T} = \begin{bmatrix} 1 & 0 \\ 0 & e^{-T} \end{bmatrix}$$

Mit der Wahl $T = 0,1$ s:

$$e^{\mathbf{A}0,1} = \begin{bmatrix} 1 & 0 \\ 0 & 0,905 \end{bmatrix} .$$

$$\mathbf{M} = \int_0^T e^{\mathbf{A}\tau} d\tau = \begin{bmatrix} \int_0^{0,1} 1 d\tau & 0 \\ 0 & \int_0^{0,1} e^{-\tau} d\tau \end{bmatrix} = \begin{bmatrix} 0,1 & 0 \\ 0 & 1 - 0,905 \end{bmatrix} = \begin{bmatrix} 0,1 & 0 \\ 0 & 0,095 \end{bmatrix}$$

und $\mathbf{b} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ wird $\mathbf{M}\mathbf{b} = \begin{bmatrix} 0,1 \\ 0,095 \end{bmatrix} = \mathbf{b}_d$.

So wird das äquivalente zeitdiskrete System bestimmt zu (Bild 5.3):

$$\mathbf{x}(k+1) = \begin{bmatrix} 1 & 0 \\ 0 & 0,905 \end{bmatrix} \cdot \mathbf{x}(k) + \begin{bmatrix} 0,1 \\ 0,095 \end{bmatrix} e(k)$$

$$y(k) = [1 \quad -1] \cdot \mathbf{x}(k)$$

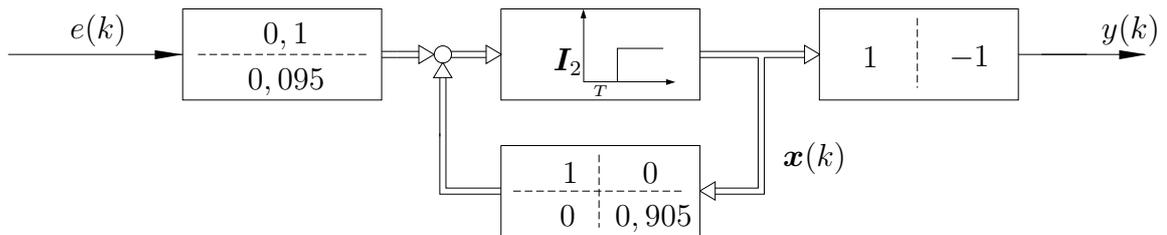


Bild 5.3: Blockschaltbild des äquivalenten zeitdiskreten Zustandsmodells des Beispiels

Dieses System soll nun mit einer Einheitsrückführung rückgekoppelt werden (Bild 5.4a).

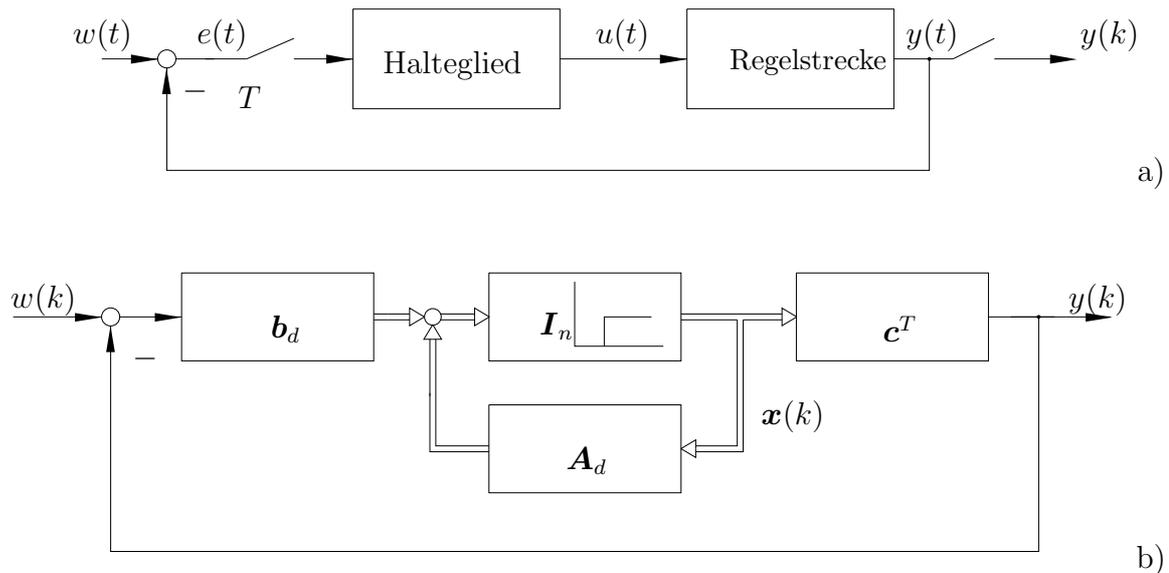


Bild 5.4: a) Übertragungsmodell des Systems mit Einheitsrückkopplung
b) zugehöriges äquivalentes Zustandsmodell

Aus Bild (5.4b) lesen wir diese Beziehung ab:

$$\mathbf{x}(k+1) = \mathbf{A}_d \mathbf{x}(k) + \mathbf{b}_d [w(k) - y(k)]$$

$$y(k) = \mathbf{c}^T \mathbf{x}(k)$$

$$\mathbf{x}(k+1) = (\mathbf{A}_d - \mathbf{b}_d \mathbf{c}^T) \mathbf{x}(k) + \mathbf{b}_d w(k)$$

$$y(k) = \mathbf{c}^T \mathbf{x}(k) \quad .$$

Mit den oben für $T = 0,1$ s berechneten Matrizen \mathbf{A}_d und $\mathbf{b}_d = \mathbf{M}\mathbf{b}$ erhalten wir die Differenzgleichung des rückgekoppelten Systems.

5.5 Steuer- und Erreichbarkeit zeitdiskreter Systeme

Die Definitionen der Steuer- und Erreichbarkeit von kontinuierlichen Systemen (siehe MA-Skript RT, Abschnitt 3.7) können ohne große Veränderungen auf zeitdiskrete Zustandsmodelle übertragen werden.

Definition 5.1 (Steuerbarkeit zeitdiskreter Systeme)

Ein zeitdiskretes System (\mathbf{A}, \mathbf{b}) heißt vollständig zustandssteuerbar, wenn es von jedem beliebigen Anfangszustand \mathbf{x}_0 durch eine endliche Eingangsfolge $u(0), \dots, u(N)$ in den Nullzustand $\mathbf{x}(N) = \mathbf{0}$ überführt werden kann. \square

Definition 5.2 (*Erreichbarkeit zeitdiskreter Systeme*)

Ein zeitdiskretes System (\mathbf{A}, \mathbf{b}) heißt vollständig erreichbar, wenn es vom Ursprung $\mathbf{x}_0 = \mathbf{0}$ durch eine endliche Eingangsfolge $u(0), \dots, u(N-1)$ in einen beliebig vorgegebenen Endzustand $\mathbf{x}(N-1) = \mathbf{x}_e$ überführt werden kann. \square

Bei zeitkontinuierlichen Systemen mußte zwischen vollständiger Erreichbarkeit und vollständiger Steuerbarkeit nicht unterschieden werden. Dies kann man unmittelbar aus den Definitionen erkennen: Wenn ein kontinuierliches System vollständig steuerbar ist, dann existiert für jeden Anfangswert \mathbf{x}_0 eine geeignete Eingangsgröße $u(t), 0 \leq t \leq t_e$, so daß

$$\mathbf{0} = \Phi(t_e)\mathbf{x}_0 + \int_0^{t_e} \Phi(t_e - \tau)\mathbf{b}u(\tau)d\tau \quad (5.24)$$

gilt und für ein vollständig erreichbares System existiert für jedes beliebiges $\mathbf{x}(t_e)$ eine Eingangsgröße $u(t), 0 \leq t \leq t_e$, die die Gleichung

$$\mathbf{x}(t_e) = \int_0^{t_e} \Phi(t_e - \tau)\mathbf{b}u(\tau)d\tau \quad (5.25)$$

erfüllt.

Ersetzt man in (5.25) die linke Seite durch $\mathbf{x}(t_e) = -\Phi(t_e)\mathbf{x}_0$, so erkennt man, daß wegen der Regularität der kontinuierlichen Transitionsmatrix $\Phi(t_e)$ die Lösbarkeit von (5.24) die von (5.25) impliziert und vice versa.

Diese Regularitätseigenschaft besitzt die zeitdiskrete Transitionsmatrix

$$\Phi(k) = \mathbf{A}^k \quad (5.26)$$

allerdings nicht. Hat die Systemmatrix \mathbf{A} des zeitdiskreten Systems einen Eigenwert bei 0, so ist die zeitdiskrete Transitionsmatrix (5.26) bereits für $k = 1$ nicht mehr invertierbar. Andererseits hat ein zeitdiskretes System mit einer Systemmatrix mit Eigenwerten bei 0 die Eigenschaft, daß die zu diesen Eigenwerten gehörenden Eigenbewegungen für beliebige Anfangsauslenkungen in wenigen Schritten mit $u(k) = 0$ in den Ursprung überführt werden können.

Beispiel:

Betrachtet wird ein zeitdiskretes Zustandsmodell mit den Matrizen

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad \text{und} \quad \mathbf{b} = \mathbf{0} .$$

Die Systemmatrix \mathbf{A} besitzt einen doppelten Eigenwert bei 0 und da $\mathbf{b} = \mathbf{0}$ ist, ist die Matrix $[\mathbf{b} \ \mathbf{A}\mathbf{b}]$ ebenfalls Null und ein kontinuierliches System mit diesen Matrizen \mathbf{A}, \mathbf{b} wäre nicht steuerbar. Die zugehörige Zustandsdifferenzgleichung lautet

$$\mathbf{x}(k+1) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \mathbf{x}(k) \quad \text{bzw.} \quad \begin{array}{l} x_1(k+1) = x_2(k) \\ x_2(k+1) = 0 \end{array} .$$

Für

$$\mathbf{x}_0 = \begin{bmatrix} x_{1_0} \\ x_{2_0} \end{bmatrix}$$

ergeben sich folgende Lösungen für die ersten Schritte:

$$\begin{aligned} k = 0 : \quad x_1(1) &= x_{1_0} \\ &x_2(1) = x_{2_0} \\ k = 1 : \quad x_1(2) &= x_2(1) = x_{2_0} \\ &x_2(2) = 0 \\ k = 2 : \quad x_1(3) &= x_2(2) = 0 \\ &x_2(3) = 0 \end{aligned}$$

Für beliebige Anfangswerte befindet sich der Zustandsvektor nach zwei Schritten im Ursprung. \square

Das bedeutet, daß jedes vollständig erreichbare System auch vollständig steuerbar ist, die Umkehrung aber nur dann gilt, wenn die Transitionsmatrix (5.26) regulär ist, d.h. es muß

$$\det \mathbf{A} \neq 0 \tag{5.27}$$

gelten.

Damit ist das Kalman–Kriterium bei zeitdiskreten Systemen nur für die Erreichbarkeit notwendig und hinreichend.

Satz 5.1 (*Erreichbarkeitskriterium für zeitdiskrete Systeme*)

Ein zeitdiskretes System (\mathbf{A}, \mathbf{b}) ist genau dann vollständig erreichbar, wenn für die Erreichbarkeitsmatrix \mathbf{Q}_E gilt:

$$\text{Rang } \mathbf{Q}_E = \text{Rang } [\mathbf{b}, \mathbf{A}\mathbf{b}, \dots, \mathbf{A}^{n-1}\mathbf{b}] = n. \tag{5.28}$$

\square

Ein Kriterium zur Überprüfung der Steuerbarkeit zeitdiskreter Systeme kann jetzt mit Hilfe des Hautus–Kriteriums angegeben werden.

Satz 5.2 (*Steuerbarkeitskriterium für zeitdiskrete Systeme*)

Ein zeitdiskretes System (\mathbf{A}, \mathbf{b}) ist genau dann vollständig steuerbar, wenn das Hautus–Kriterium

$$\text{Rang } [\lambda_i \mathbf{I} - \mathbf{A} \quad \mathbf{b}] = n \tag{5.29}$$

für alle von Null verschiedenen Eigenwerte λ_i erfüllt ist. \square

Wenn das zeitdiskrete System das äquivalente zeitdiskrete System $(\mathbf{A}_d, \mathbf{b}_d, \mathbf{c}_d)$ eines kontinuierlichen Systems ist, dann folgt aus

$$\mathbf{A}_d = e^{\mathbf{A}T}, \quad (5.30)$$

daß die Systemmatrix \mathbf{A}_d des zeitdiskreten Systems nur dann Eigenwerte bei 0 haben kann, wenn das kontinuierliche System Eigenwerte bei $-\infty$ hat. In den meisten praktisch interessierenden Fällen wird dies nicht der Fall sein, so daß man in vielen Lehrbüchern auch bei zeitdiskreten Systemen nicht zwischen Steuer- und Erreichbarkeit unterscheidet. Allerdings wird in diesen Lehrbüchern dann als Steuerbarkeit die Erreichbarkeit eines beliebigen Endzustands vom Ursprung aus definiert.

5.6 Beobachtbarkeit und Rekonstruierbarkeit zeitdiskreter Systeme

Auch die Definition der Beobachtbarkeit von kontinuierlichen Systemen (siehe MA-Skript RT, Abschnitt 3.7) kann leicht auf zeitdiskrete Zustandsmodelle übertragen werden.

Definition 5.3 (*Beobachtbarkeit zeitdiskreter Systeme*)

Ein zeitdiskretes System (\mathbf{A}, \mathbf{c}) heißt vollständig beobachtbar, wenn der Anfangszustand \mathbf{x}_0 aus dem bekannten Verlauf der endlichen Eingangsfolge $u(0), \dots, u(N)$ und der endlichen Ausgangsfolge $y(0), \dots, y(N)$ berechnet werden kann. \square

Die Erfüllung des bekannten Kalman-Kriteriums ist auch bei zeitdiskreten Systemen notwendig und hinreichend für die vollständige Beobachtbarkeit.

Satz 5.3 (*Beobachtbarkeitskriterium für zeitdiskrete Systeme*)

Ein zeitdiskretes System (\mathbf{A}, \mathbf{c}) ist genau dann vollständig beobachtbar, wenn für die Beobachtbarkeitsmatrix \mathbf{Q}_B gilt:

$$\text{Rang } \mathbf{Q}_B = \text{Rang } [\mathbf{c}, \mathbf{A}^T \mathbf{c}, \dots, (\mathbf{A}^T)^{n-1} \mathbf{c}] = n. \quad (5.31)$$

\square

Damit sind bei zeitdiskreten Systemen die Erreichbarkeit und die Beobachtbarkeit duale Eigenschaften. Die zur Steuerbarkeit duale Eigenschaft nennt man Rekonstruierbarkeit und ist wie folgt definiert:

Definition 5.4 (*Rekonstruierbarkeit zeitdiskreter Systeme*)

Ein zeitdiskretes System (\mathbf{A}, \mathbf{c}) heißt vollständig rekonstruierbar, wenn aus dem bekannten Verlauf der endlichen Eingangsfolge $u(0), \dots, u(N)$ und der endlichen Ausgangsfolge $y(0), \dots, y(N)$ der Zustand $\mathbf{x}(N)$ berechnet werden kann. \square

Zwischen der Beobachtbarkeit und der Rekonstruierbarkeit zeitdiskreter Systeme bestehen jetzt folgende Zusammenhänge:

- Jedes vollständig beobachtbare zeitdiskrete System ist auch vollständig rekonstruierbar.
- Die Umkehrung gilt jedoch nur, wenn die Systemmatrix des zeitdiskreten Systems keine Eigenwert bei Null hat.

Ein notwendiges und hinreichendes Kriterium zur Überprüfung der Rekonstruierbarkeit eines zeitdiskreten Systems kann jetzt auch wieder mit Hilfe des Hautus-Kriteriums angegeben werden:

Satz 5.4 (*Rekonstruierbarkeitskriterium für zeitdiskrete Systeme*)

Ein zeitdiskretes System (\mathbf{A}, \mathbf{c}) ist genau dann vollständig rekonstruierbar, wenn das Hautus-Kriterium

$$\text{Rang} \begin{bmatrix} \lambda_i \mathbf{I} - \mathbf{A} \\ \mathbf{c}^T \end{bmatrix} = n \quad (5.32)$$

für alle von Null verschiedenen Eigenwerte λ_i erfüllt ist. \square

5.7 Erreichbarkeit und Beobachtbarkeit des kontinuierlichen und des zeitdiskreten Systems

Das äquivalente zeitdiskrete System entsteht aus dem kontinuierlichen System in der Form, daß man ein Halteglied vor das kontinuierliche System schaltet und die Ausgangsgröße ideal abtastet. Es stellt sich jetzt die Frage, können sich die Eigenschaften der Erreichbarkeit und der Beobachtbarkeit durch Abtastung ändern?

Offensichtlich ist die Steuerbarkeit/Erreichbarkeit des kontinuierlichen Systems (\mathbf{A}, \mathbf{b}) notwendig für die Erreichbarkeit des zeitdiskreten Systems $(\mathbf{A}_d, \mathbf{b}_d)$. Wenn keine kontinuierliche Steuerfunktion $u(t)$ existiert, um einen gewünschten Endzustand in endlicher Zeit zu erreichen, so kann dieser Endzustand durch eine stückweise konstante Steuerfunktion erst recht nicht erreicht werden. Bereits von Kalman wurde der folgende Zusammenhang aufgedeckt:

Satz 5.5

Das zeitdiskrete System $(\mathbf{A}_d, \mathbf{b}_d)$, das aus dem kontinuierlichen System (\mathbf{A}, \mathbf{b}) durch Abtastung mit der Abtastzeit T entsteht, ist genau dann vollständig erreichbar,

- wenn das kontinuierliche System (\mathbf{A}, \mathbf{b}) vollständig erreichbar ist
- und wenn für zwei beliebige Eigenwerte λ_i und λ_j ($\lambda_i \neq \lambda_j$) der Matrix \mathbf{A} die Bedingung

$$e^{\lambda_i T} \neq e^{\lambda_j T} \quad (5.33)$$

erfüllt ist. \square

Die zweite Bedingung fordert, daß zwei verschiedene Eigenwerte des kontinuierlichen Systems auf zwei verschiedene Eigenwerte des zeitdiskreten Systems führen. Diese Bedingung kann nur dann verletzt werden, wenn das kontinuierliche System zwei komplexe Eigenwerte

$$\lambda_1 = \delta + j\omega_1, \quad \lambda_2 = \delta + j\omega_2 \quad (5.34)$$

mit gleichem Realteil hat und die Abtastzeit so gewählt wird, daß die Beziehung

$$e^{\lambda_1 T} = e^{\lambda_2 T} \quad (5.35)$$

gilt.

Es müßte also gelten:

$$(\omega_1 - \omega_2)T = \pm 2k\pi, \quad k = 1, 2, 3, \dots \quad (5.36)$$

Für ein konjugiert komplexes Eigenwertpaar mit $\omega_2 = -\omega_1$ ist die Erreichbarkeit des zeitdiskreten Systems dann gegeben, wenn

$$\omega_1 T \neq \pm k\pi, \quad k = 1, 2, 3, \dots \quad (5.37)$$

Setzt man für k den kleinsten Wert ein, so ergibt sich, daß die Erreichbarkeit für

$$\omega_A > 2\omega_1 \quad \text{mit} \quad \omega_A = \frac{2\pi}{T} \quad (5.38)$$

sichergestellt ist.

Diese Bedingung ist die bekannte Forderung des Abtasttheorems von Shannon, deren Einhaltung auch die Erreichbarkeit des zeitdiskreten Systems garantiert, wenn das kontinuierliche System vollständig steuerbar ist.

Hat die Systemmatrix \mathbf{A} des kontinuierlichen Systems mehrere Paare von konjugiert komplexen Eigenwerte, so stellt die Wahl der Abtastzeit

$$T < \frac{\pi}{\omega_{j,max}} \quad (5.39)$$

mit $\omega_{j,max}$ als den betraglich größten Imaginärteil aller auftretenden komplexen Eigenwerte sicher, daß die vollständige Erreichbarkeit bei der Abtastung erhalten bleibt.

Der Satz 5.5 gilt, entsprechend formuliert, auch für die Beobachtbarkeit:

Satz 5.6

Das zeitdiskrete System $(\mathbf{A}_d, \mathbf{c}_d)$, das aus dem kontinuierlichen System (\mathbf{A}, \mathbf{c}) durch Abtastung mit der Abtastzeit T entsteht, ist genau dann vollständig beobachtbar,

- wenn das kontinuierliche System (\mathbf{A}, \mathbf{c}) vollständig beobachtbar ist

- und wenn für zwei beliebige Eigenwerte λ_i und λ_j ($\lambda_i \neq \lambda_j$) der Matrix \mathbf{A} die Bedingung

$$e^{\lambda_i T} \neq e^{\lambda_j T} \quad (5.40)$$

erfüllt ist.

□

Beispiel:

Gegeben sei die Übertragungsfunktion eines schwingungsfähigen PT_2 -Systems mit einer Verstärkung von 1 und einem Dämpfungsgrad von 0:

$$G(s) = \frac{\omega_0^2}{s^2 + \omega_0^2}.$$

Ein zugehöriges Zustandsmodell ist durch

$$\begin{aligned} \mathbf{A} &= \begin{bmatrix} 0 & -\omega_0 \\ \omega_0 & 0 \end{bmatrix} & \mathbf{b} &= \begin{bmatrix} \omega_0 \\ -\omega_0 \end{bmatrix} \\ \mathbf{c}^T &= [0,5 \quad 0,5] \end{aligned}$$

gegeben.

Das zugehörige äquivalente zeitdiskrete Zustandsmodell für die Abtastzeit T ergibt sich zu:

$$\begin{aligned} \mathbf{A}_d &= \begin{bmatrix} \cos(T\omega_0) & -\sin(T\omega_0) \\ \sin(T\omega_0) & \cos(T\omega_0) \end{bmatrix} & \mathbf{b}_d &= \begin{bmatrix} b_{d1} \\ b_{d2} \end{bmatrix} \\ \mathbf{c}^T &= [0,5 \quad 0,5]. \end{aligned}$$

Die Berechnung der Erreichbarkeitsmatrix des zeitdiskreten Systems liefert:

$$\begin{aligned} \mathbf{Q}_E &= [\mathbf{b}_d \quad \mathbf{A}_d \mathbf{b}_d] \\ &= \begin{bmatrix} b_{d1} & b_{d1} \cos(T\omega_0) - b_{d2} \sin(T\omega_0) \\ b_{d2} & b_{d1} \sin(T\omega_0) + b_{d2} \cos(T\omega_0) \end{bmatrix}. \end{aligned}$$

Die Erreichbarkeit des Abtastsystems ist nur dann gegeben, wenn

$$\begin{aligned} \det \mathbf{Q}_E &= b_{d1}(b_{d1} \sin(T\omega_0) + b_{d2} \cos(T\omega_0)) - \\ &\quad - b_{d2}(b_{d1} \cos(T\omega_0) - b_{d2} \sin(T\omega_0)) \end{aligned} \quad (5.41)$$

$$= b_{d1}^2 \sin(T\omega_0) + b_{d2}^2 \sin(T\omega_0) \quad (5.42)$$

$$= (b_{d1}^2 + b_{d2}^2) \sin(T\omega_0) \quad (5.43)$$

von Null verschieden ist.

Aus (5.43) folgt sofort, daß die Erreichbarkeitsmatrix aufgrund des Sinustersms für $T\omega_0 = \pm\pi, \pm 2\pi, \dots$ singulär ist. Diese Bedingung kann durch die Wahl der Abtastzeit T zu

$$T = \frac{k\pi}{\omega_0}, \quad k = \pm 1, \pm 2, \dots$$

immer erfüllt werden.

Setzt man für k den kleinsten Wert ein, so erhält man aus

$$T = \frac{\pi}{\omega_0}$$

mit $\omega_A = \frac{2\pi}{T}$ bzw. $T = \frac{2\pi}{\omega_A}$

$$\frac{2\pi}{\omega_A} = \frac{\pi}{\omega_0}$$

die Bedingung

$$\omega_A = 2\omega_0.$$

Offensichtlich ist die Erreichbarkeit des Abtastsystems dann gewährleistet, wenn man die Abtastzeit so wählt, daß

$$\omega_A > 2\omega_0 \tag{5.44}$$

gilt. Die Bedingung (5.44) ist gerade die bekannte Bedingung des Abtasttheorems von Shannon. Wenn das kontinuierliche System also vollständig steuerbar ist, dann gewährleistet die Einhaltung des Abtasttheorems auch die vollständige Erreichbarkeit des zeitdiskreten Ersatzsystems.

Ein entsprechender Zusammenhang existiert bei Abtastsystem auch bei der Beobachtbarkeit.

6 Entwurf digitaler Regler

Für den Entwurf digitaler Regler gibt es zwei Möglichkeiten. Ist die Abtastzeit wesentlich kleiner (Faktor 5 - 10) als die maßgebenden Zeitkonstanten des Regelkreises, so kann der Regler als kontinuierlicher Regler entworfen und dann als zeitdiskreter Regler realisiert werden. Darüber hinaus gibt es Verfahren zum Entwurf zeitdiskreter Regler, die anhand eines zeitdiskreten Modells der Regelstrecke erfolgen.

6.1 Quasikontinuierliche digitale Regelung

Eine gegebene Regelungsaufgabe kann am besten mit einem kontinuierlichen Regler gelöst werden, da dieser Regler schnellstmöglich auf Regelabweichungen reagieren kann. Die Abtastung wird nur eingeführt, um das erhaltene Reglergesetz mit Hilfe eines Prozeßrechners oder eines Mikrocontrollers realisieren zu können. Wählt man dabei eine sehr kleine Abtastzeit, so stellt der digitale Regler eine gute Approximation des entworfenen kontinuierlichen Reglers dar. Das bedeutet, die Abtastzeit muß so klein sein, daß der Einfluß der Zeitdiskretisierung vernachlässigbar ist. Man entwirft also zunächst einen kontinuierlichen Regler mit Hilfe eines kontinuierlichen Modells der Regelstrecke und geht erst bei der Realisierung zu einem zeitdiskreten Regler über. Der Entwurf selbst erfolgt mit Methoden, wie sie beispielsweise in der MA-Vorlesung "Regelungstechnik" behandelt wurden.

Die Anwendung der quasikontinuierlichen Vorgehensweise hat nicht nur den Vorteil, daß der Regler mit Hilfe bekannter Methoden für kontinuierliche Systeme entworfen werden kann. Ein weiterer Vorteil ist, daß man bei der Betrachtung des kontinuierlichen Regelkreises erfährt, welche Regelgüte prinzipiell mit einer kontinuierlichen Regelung erreicht werden kann.

In diesem Abschnitt wird diese Entwurfsmöglichkeit zeitdiskreter Regler behandelt. Zunächst wird dabei untersucht, wie kontinuierliche Regelgesetze durch zeitdiskrete Regelgesetze approximiert werden können. Ein häufig auftretendes Problem bei diesem Entwurfsvorgehen besteht darin, dass der geschlossene Regelkreis zu wenig gedämpft ist, obwohl der kontinuierliche Kreis die Güteforderungen erfüllt. Der Grund dafür ist die zusätzliche Phasenverschiebung, die durch das Halteglied in den Regelkreis eingebracht wird. Diese Phasenverschiebung kann man durch Einfügen eines differenzierenden Korrekturgliedes, das die Phase wieder anhebt, ausgleichen.

Ein proportionaler Regler

$$u(t) = -K_p(y(t) - w(t)) \quad (6.1)$$

oder eine Zustandsrückführung

$$u(t) = -\mathbf{K}\mathbf{x}(t) \quad (6.2)$$

kann problemlos als zeitdiskreter Regler aufgefasst werden. Man erhält die Gleichungen für die zeitdiskrete Realisierung aus (6.1, 6.2) zu

$$u(k) = -K_p(y(k) - w(k)) \quad (6.3)$$

bzw.

$$u(k) = -\mathbf{K}\mathbf{x}(k) . \quad (6.4)$$

Diese einfache Ersetzung des kontinuierlichen durch einen zeitdiskreten Regler bedeutet jedoch nicht, daß die Eigenschaften des geschlossenen Regelkreises dabei unverändert bleiben müssen.

Der Hauptgedanke bei der quasikontinuierlichen Betrachtung besteht darin, Differentiationen und Integrationen durch Differenzenquotienten bzw. Summen zu ersetzen, wobei die aus der numerischen Mathematik bekannten Methoden verwendet werden.

Approximation der Differentiation durch den Differenzenquotienten

Treten im Regelgesetz Differentiationen auf, wie es beispielsweise beim D-Anteil eines Reglers

$$u(t) = K_D \dot{e}(t) \quad (6.5)$$

der Fall ist, so ersetzt man im zeitdiskreten Regler den Differentialquotienten

$$\dot{e}(t) = \frac{de(t)}{dt} \quad (6.6)$$

durch den Differenzenquotienten

$$\Delta e(k) = \frac{e(k) - e(k-1)}{T}, \quad (6.7)$$

wobei T die Abtastzeit ist.

Damit erhält man die Differenzengleichung eines D-Anteils zu

$$u(k) = K_D \frac{e(k) - e(k-1)}{T} . \quad (6.8)$$

Die Anwendung der z -Transformation auf diese Beziehung liefert

$$U(z) = \frac{K_D}{T} (1 - z^{-1}) E(z) \quad (6.9)$$

und hieraus folgt

$$G_D(z) = \frac{U(z)}{E(z)} = \frac{K_D}{T} (1 - z^{-1}) \quad (6.10)$$

$$= \frac{K_D}{T} \frac{z-1}{z} . \quad (6.11)$$

Interessant ist, daß der Parameter des zeitdiskreten Reglers von der Abtastzeit abhängt. Dieser Parameter muß also verändert werden, wenn man bei der Inbetriebnahme des Reglers die Abtastzeit verändert.

Die Wirkung des zeitdiskreten Reglers unterscheidet sich von der des kontinuierlichen. Der kontinuierliche Regler bringt eine positive Phasenverschiebung mit sich und wirkt deshalb bei vielen Regelstrecken stabilisierend. Demgegenüber verursacht der zeitdiskrete Regler eine Phasennacheilung, denn er hat einen Pol bei Null. Der Grund liegt in der Tatsache begründet, daß der Regler einen Takt "warten" muß, bevor er die Differenzenbildung ausführen kann.

Allerdings ist der diskrete D-Regler technisch realisierbar, da bei der Bildung des Regelgesetzes nur die tatsächlich verfügbaren Informationen verwendet werden. Der kontinuierliche D-Regler in der Form (6.5) ist demgegenüber nicht realisierbar (vgl. SRT-Skript, Abschnitt 2.6.2).

Approximation der Integration

Die zeitdiskrete Realisierung des I-Anteiles eines Reglers

$$u(t) = \frac{1}{T_I} \int_0^t e(\tau) d\tau \quad (6.12)$$

kann mit Hilfe der Verfahren der numerischen Integration berechnet werden. Verwendet man die Trapezregel

$$u(kT) = \frac{T}{T_I} \left(\frac{e(0)}{2} + \sum_{i=1}^{k-1} e(iT) + \frac{e(kT)}{2} \right), \quad (6.13)$$

so erhält man die rekursive Form des Regelgesetzes zu:

$$u(k) = u(k-1) + \frac{T}{2T_I} (e(k-1) + e(k)). \quad (6.14)$$

Im Frequenzbereich wird damit die Übertragungsfunktion

$$G_I(s) = \frac{1}{T_I s} \quad (6.15)$$

durch die z -Übertragungsfunktion

$$G_I(z) = \frac{T}{2T_I} \frac{z+1}{z-1} \quad (6.16)$$

ersetzt.

Digitaler PID–Regler

Das Zeitverhalten des analogen PID–Reglers wird durch folgende Gleichung (vgl. Gl. (4.7) im Vorlesungsskript SRT) beschrieben:

$$u_R(t) = K_R e(t) + \frac{K_R}{T_I} \int_0^t e(\tau) d\tau + K_R T_D \frac{de(t)}{dt}, \quad (6.17)$$

wobei K_R die Proportional–Verstärkung,

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

die *Nachstellzeit* und

$$T_D = \frac{K_D}{K_R}$$

die *Vorhaltezeit* ist.

Zum Aufstellen der z –Übertragungsfunktion des digitalen PID–Reglers muß man lediglich die drei Summanden einzelnen in der zuvor besprochenen Weise behandeln:

$$\begin{aligned} G_P(s) &= K_R & \rightarrow & G_P(z) = K_R \\ G_I(s) &= \frac{K_R}{T_I s} & \rightarrow & G_I(z) = K_R \frac{T}{2T_I} \frac{z+1}{z-1} \\ G_D(s) &= K_R T_D s & \rightarrow & G_D(z) = K_R \frac{T_D}{T} \frac{z-1}{z}. \end{aligned}$$

Man erhält damit für den zeitdiskreten PID–Regler im z –Bereich folgenden Zusammenhang zwischen Ein- und Ausgangsgröße:

$$U(z) = K_R \left(1 + \frac{T}{2T_I} \frac{z+1}{z-1} + \frac{T_D}{T} \frac{z-1}{z} \right) E(z). \quad (6.18)$$

Das Übertragungsverhalten des PID–Reglers wird also durch eine z –Übertragungsfunktion zweiter Ordnung beschrieben:

$$G_{\text{PID}}(z) = \frac{U(z)}{E(z)} = \frac{d_2 z^2 + d_1 z + d_0}{z(z-1)} \quad (6.19)$$

mit

$$d_2 = K_R \left(1 + \frac{T}{2T_I} + \frac{T_D}{T} \right) \quad (6.20)$$

$$d_1 = K_R \left(-1 + \frac{T}{2T_I} - \frac{2T_D}{T} \right) \quad (6.21)$$

$$d_0 = \frac{K_R T_D}{T}. \quad (6.22)$$

Für die Realisierung des Reglers wird die zu der angegebenen Übertragungsfunktion (6.19) gehörige Differenzgleichung aufgestellt. Hierzu wird (6.19) zu

$$z(z-1)U(z) = (d_2z^2 + d_1z + d_0)E(z) \quad (6.23)$$

umgeformt und beide Seiten durch z dividiert:

$$\left(1 - \frac{1}{z}\right)U(z) = \left(d_2 + d_1\frac{1}{z} + d_0\frac{1}{z^2}\right)E(z). \quad (6.24)$$

Durch Anwendung der inversen z -Transformation erhält man die gesuchte Differenzgleichung zu

$$u(k) = u(k-1) + d_2e(k) + d_1e(k-1) + d_0e(k-2). \quad (6.25)$$

In dieser Darstellung wird aus gegebenen Werten für die Regelabweichung und einem alten Wert der Stellgröße die Stellgröße zum aktuellen Zeitpunkt k berechnet. Das Entwurfsverfahren wird deshalb als *Stellungsalgorithmus* oder *Positionsalgorithmus* bezeichnet. Eine andere gebräuchliche Darstellung des PID-Reglers erhält man durch Betrachtung der Änderung der Stellgröße

$$\Delta u(k) = u(k) - u(k-1) \quad (6.26)$$

zu

$$\Delta u(k) = d_2e(k) + d_1e(k-1) + d_0e(k-2). \quad (6.27)$$

Diese Darstellung nennt man *Geschwindigkeitsalgorithmus*. Sie wird bei integralwirkenden Stellgliedern verwendet, beispielsweise, wenn ein Stellventil über einen Stellmotor angesteuert wird.

Wählt man die Abtastzeit T kleiner als $1/10$ der dominierenden Zeitkonstante des Systems, so können unmittelbar die Parameter des kontinuierlichen PID-Reglers in die Gln. (6.20) bis (6.22) eingesetzt werden, wie sie aufgrund von Einstellregeln oder Erfahrungswerte bekannt sind.

A Die Fourier–Transformation

Die Fourier–Transformation, die nach dem französischen Mathematiker Jean–Baptiste Joseph Fourier benannt ist, ist eine Transformation, die eine Funktion in ihre Sinus– und Kosinus–Bestandteile (Basisfunktionen) mit verschiedenen Frequenzen zerlegt. Entsprechend dieser physikalisch inspirierten Betrachtungsweise wird die Transformation einer Funktion in den Frequenzbereich auch als Fourieranalyse bezeichnet, die Rücktransformation als Fouriersynthese.

Alternativ kann man als Basisfunktion auch die Exponentialfunktion mit imaginärem Argument verwenden, da in diesem Fall der Realteil der Exponentialfunktion dem Kosinus und der Imaginärteil dem Sinus entspricht. Diese Form nennt man auch komplexe Fourier–Transformation, da die transformierte Funktion komplexe Werte annimmt und die zu transformierende Funktion komplexe Werte annehmen kann.

Je nach der zu transformierenden Funktion und der transformierten Funktion unterscheidet man:

- Fourier–Reihe
- Kontinuierliche Fourier–Transformation
- Diskrete Fourier–Transformation
- Schnelle Fourier–Transformation (FFT)

Als verallgemeinerte Fouriertransformation wird jede Zerlegung einer Funktion in ein System von Basisfunktionen bezeichnet. Dabei müssen die Basisfunktionen geeignet gewählt werden, so dass die Zerlegung eindeutig und umkehrbar ist.

A.1 Fourier–Reihe

Ein periodisches Signal mit der Periodendauer T_0 , z. B. ein Rechtecksignal, kann als eine unendliche Reihe harmonischer Signale dargestellt werden, entweder in dieser Form:

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} [a_n \cos(n\omega_0 t) + b_n \sin(n\omega_0 t)] \quad (\text{A.1})$$

mit

$$\left. \begin{aligned} \omega_0 &= \frac{2\pi}{T_0}; & a_n &= \frac{2}{T_0} \int_{-T_0/2}^{T_0/2} f(t) \cos(n\omega_0 t) dt \\ & & b_n &= \frac{2}{T_0} \int_{-T_0/2}^{T_0/2} f(t) \sin(n\omega_0 t) dt \end{aligned} \right\} \quad (\text{A.2})$$

oder in komplexer Notierung:

$$f(t) = \sum_{n=-\infty}^{+\infty} c_n e^{jn\omega_0 t} \quad \text{mit} \quad c_n = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} f(t) e^{-jn\omega_0 t} dt \quad (\text{A.3})$$

Die rechte Seite der Gl. (A.1) nennt man *Fourier-Reihe* von $f(t)$ mit der Grund(kreis)frequenz ω_0 . Die reellen Koeffizienten a_n und b_n heißen *Fourier-Koeffizienten* der Fourier-Reihe und können graphisch als diskrete Linienspektren dargestellt werden, Bild A.1.

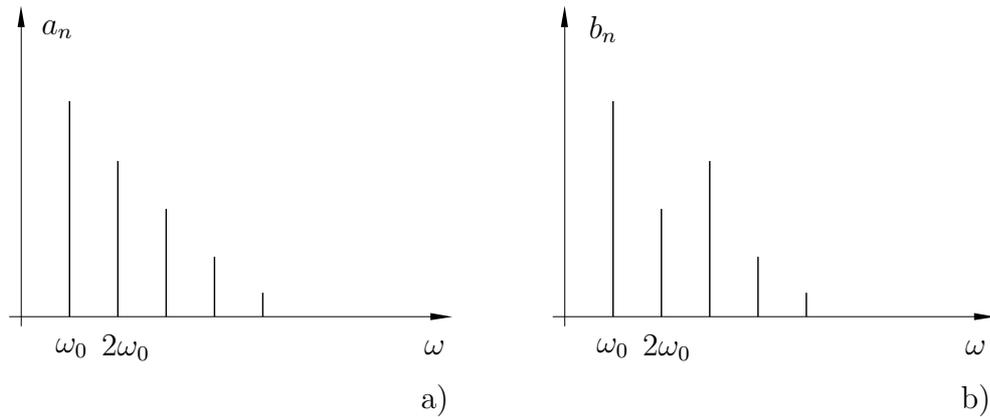


Bild A.1: Darstellung eines periodischen Signals durch Linienspektren
 a) Anteil der Cosinus-Glieder b) Anteil der Sinus-Glieder

Werden auf der rechten Seite der Gl. (A.1) die Sinus- und Cosinusschwingungen gleicher Frequenz zusammengefaßt, so erhält man folgende alternative Form der Fourier-Reihe:

$$f(t) = A_0 + \sum_{n=1}^{\infty} A_n \sin(n\omega_0 t + \varphi_n) \quad (\text{A.4})$$

mit

$$A_0 = \frac{a_0}{2} \quad (\text{A.5})$$

$$A_i = \sqrt{a_i^2 + b_i^2}, \quad i = 1, 2, \dots \quad (\text{A.6})$$

und

$$\varphi_i = \arctan \frac{a_i}{b_i}, \quad i = 1, 2, \dots \quad (\text{A.7})$$

Die Summanden für $n = 1$ in den Gleichungen (A.1) und (A.4) beschreiben die *Grundschwingung* oder erste Harmonische der Fourier-Zerlegung. Alle anderen Summanden sind Funktionen mit einem Vielfachen der Grundfrequenz ω_0 und werden als Oberwellen oder Oberschwingungen bezeichnet.

Die Form (A.3) nennt man komplexe Fourier–Reihe mit im allgemeinen komplexen Fourier–Koeffizienten c_n . Ein Vergleich mit den Koeffizienten A_i in Gl. (A.4) liefert

$$c_0 = A_0 \tag{A.8}$$

$$|c_i| = \frac{1}{2}|A_i| \quad (i \neq 0) \tag{A.9}$$

$$\arg c_i = \varphi_i \quad (i \neq 0) . \tag{A.10}$$

Der Betrag $|c_n|$ und die Phase $\arg c_i$ der komplexen Fourier–Koeffizienten lassen sich ebenfalls in der Form diskreter Spektren darstellen. Man nennt $|c_i|$ für $n = -\infty \dots \infty$ das *Amplitudenspektrum* und $\arg c_i$ das *Phasenspektrum* der periodischen Funktion $f(t)$.

Beispiele

In Bild A.2 sind ein symmetrisches periodisches Rechteck– und ein Dreiecksignal dargestellt.

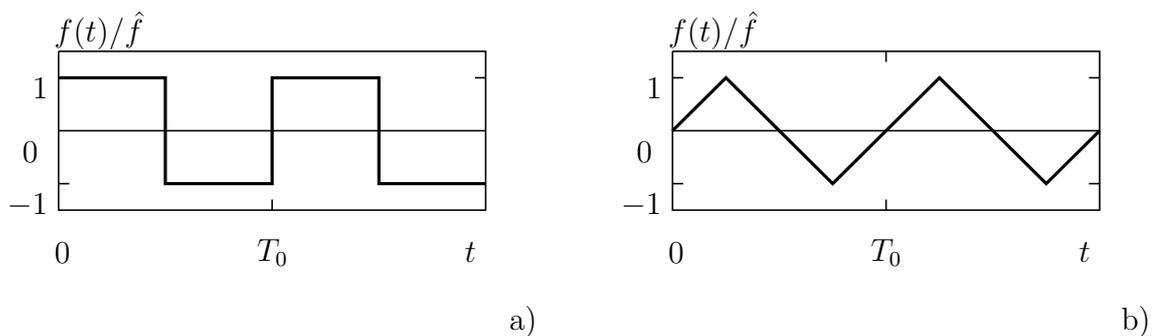


Bild A.2: Beispiel periodischer Funktionen a) Rechteckfunktion b) Dreieckfunktion

Die Fourier–Entwicklung lautet

für die Rechteckfunktion:

$$f(t) = \frac{4\hat{f}}{\pi} \left(\sin(\omega_0 t) + \frac{1}{3} \sin(3\omega_0 t) + \frac{1}{5} \sin(5\omega_0 t) + \dots \right), \tag{A.11}$$

und für die Dreieckfunktion:

$$f(t) = \frac{8\hat{f}}{\pi^2} \left(\sin(\omega_0 t) - \frac{1}{3^2} \sin(3\omega_0 t) + \frac{1}{5^2} \sin(5\omega_0 t) + \dots \right). \tag{A.12}$$

Die periodischen Funktionen setzen sich aus einer Überlagerung (Addition) unendlich vieler harmonischer Funktionen mit diskreten Frequenzen zusammen. Die zugehörigen Linienspektren zeigt Bild A.3.

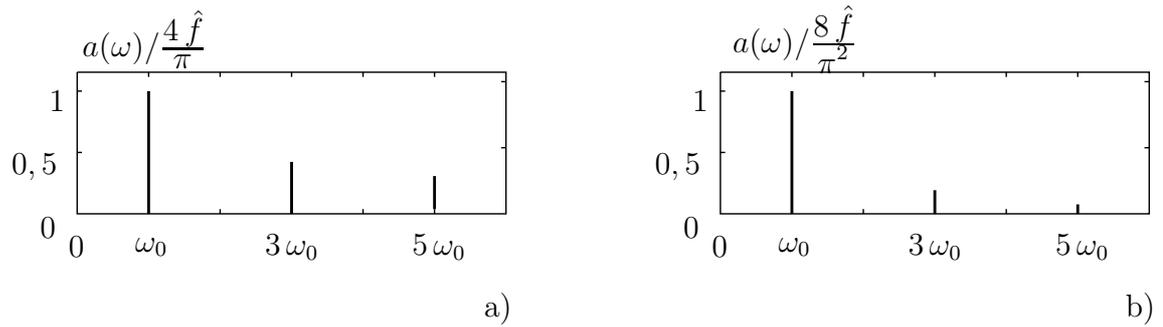


Bild A.3: Linienspektrum a) der Rechteckfunktion b) der Dreieckfunktion

Die Approximation der periodischen Funktionen durch die Grundwelle und der ersten Oberwelle ist im Bild A.4 dargestellt.

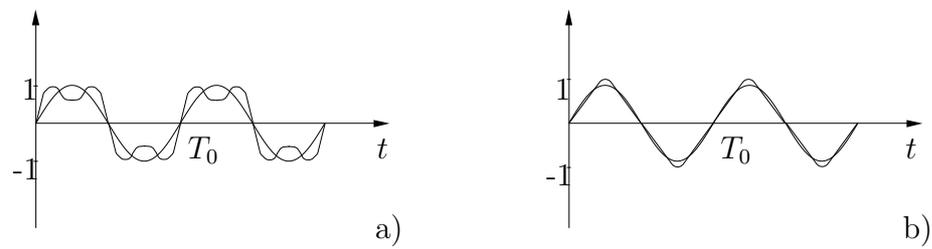


Bild A.4: Grundwelle und Summe aus Grundwelle und 1. Oberwelle a) zur Rechteckwelle b) zur Dreieckfunktion

A.2 Kontinuierliche Fourier–Transformation

Die Anwendung der Fourier–Reihe beschränkt sich auf periodische Funktionen. Wird die Periodenlänge T_0 immer weiter ausgedehnt (vgl. Bild A.5), gelangt man im Grenzfall zu den nichtperiodischen oder auch *aperiodischen* Vorgängen.

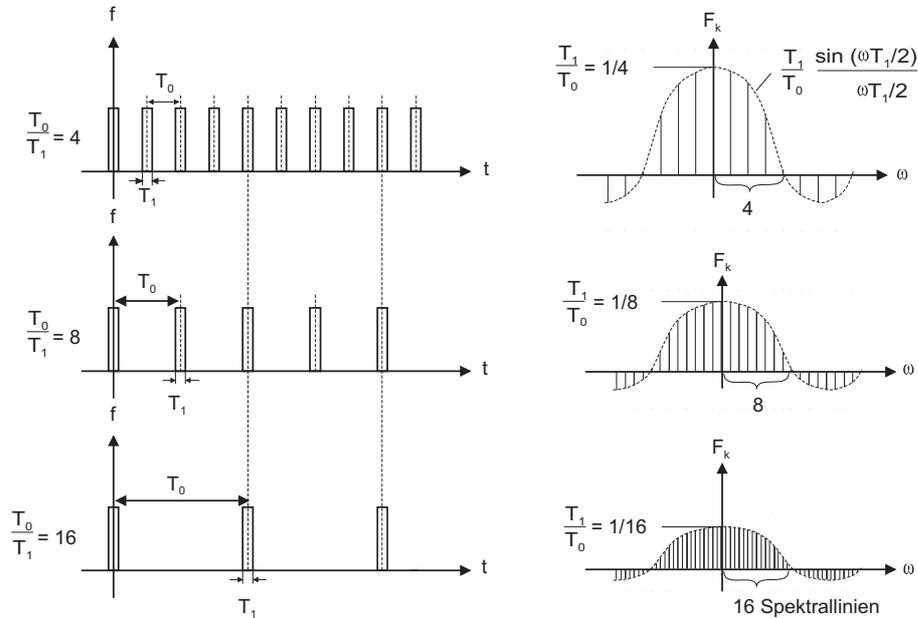


Bild A.5: Spektrum von Pulsfolgen mit verschiedenen Periodendauern

Die Grundfrequenz $\omega_0 = 2\pi/T_0$ wird demzufolge unendlich klein. Dies führt allerdings zu verschwindenden Fourier–Koeffizienten in Gleichung (A.3):

$$T_0 \rightarrow \infty \Rightarrow c_n = 0, \quad \text{für } n = -\infty \dots \infty. \quad (\text{A.13})$$

Die komplexen Fourier–Koeffizienten in (A.3) ersetzt man deshalb durch die Größen

$$F_n(T_0) = c_n T_0 = \int_{-T_0/2}^{T_0/2} f(t) e^{-jn\omega_0 t} dt, \quad (\text{A.14})$$

die als eine Amplitudendichte interpretiert werden können. Der Grenzwert der Amplitudendichte A.14 für

$$\lim_{T_0 \rightarrow \infty} F_n(T_0) = \int_{-\infty}^{\infty} f(t) e^{-j\omega t} dt = F(j\omega). \quad (\text{A.15})$$

entspricht dann der *Spektraldichte* bzw. der *Fourier–Transformierten* der aperiodischen Funktion $f(t)$.

Die *Fourier–Reihe* (A.3) geht dann in ein *Fourierintegral* über:

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} F(j\omega) e^{j\omega t} d\omega \quad (\text{A.16})$$

mit der Fouriertransformierten

$$F(j\omega) = \int_{-\infty}^{+\infty} f(t) e^{-j\omega t} dt \quad (\text{A.17})$$

von $f(t)$.

Die Fourier–Transformation (A.17) bildet eine Funktion der Zeit, $f(t)$, in eine Funktion der Frequenz, $F(j\omega)$ ab. Man schreibt symbolisch:

$$F(j\omega) = \mathcal{F}\{f(t)\}, \quad f(t) = \mathcal{F}^{-1}\{F(j\omega)\}.$$

Die Existenz einer Fourier–Transformation ist an die strenge Konvergenzbedingung

$$\int_{-\infty}^{\infty} |f(t)| dt < \infty \quad (\text{A.18})$$

gebunden. Diese Bedingung besagt, daß die Funktion $f(t)$ absolut integrierbar sein muß.

Die Fouriertransformierte oder Spektraldichte $F(j\omega)$ ist eine komplexwertige Funktion der reellen Frequenz ω mit dem Betrag $|F(j\omega)|$ und der Phase $\arg F(j\omega)$ und wird auch als *Spektrum* der Zeitfunktion $f(t)$ bezeichnet. Die graphische Darstellung von $|F(j\omega)|$ und $\arg F(j\omega)$ nennt man *Amplitudendichtespektrum* bzw. *Phasenspektrum*.

Die Fourier–Transformation zeigt, daß auch nichtperiodische Signale als Summe sinusförmiger Signale dargestellt werden können, wobei das Amplitudendichtespektrum und das Phasenspektrum kontinuierlich sind.

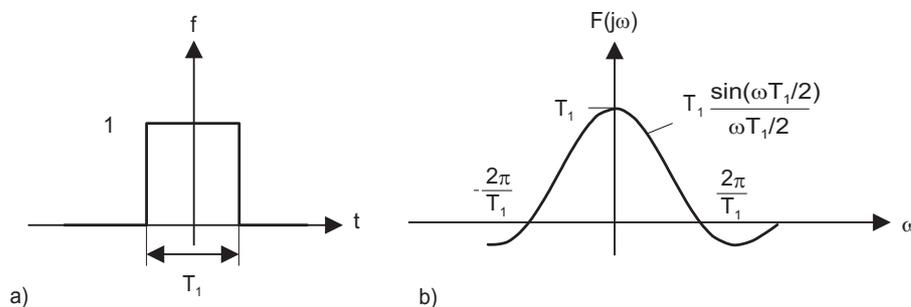


Bild A.6: Zeitverlauf (a) und Spektraldichte (b) eines einzelnen Rechtecks.

Als ein Beispiel zeigt Bild A.6 das Amplitudendichtespektrum des Rechteckimpulses, der aus der in Bild A.5 betrachteten Pulsfolge für $T_0 \rightarrow \infty$ hervorgeht.

Zur Verdeutlichung des Unterschiedes zwischen dem kontinuierlichen Amplitudendichtespektrum $|F(j\omega)|$ der Fourier–Transformation und dem diskreten Amplitudenspektrum der Fourierzerlegung kann man die Gl. (A.16) als eine Zerlegung der Funktion $f(t)$ in Elementarsignale

$$F(j\omega)e^{j\omega t}d\omega \tag{A.19}$$

deuten. Da das Frequenzintervall $d\omega$ unendlich klein ist auch die komplexe Amplitude $F(j\omega)d\omega$ dieser Elementarsignale unendlich klein. Dividiert man die komplexe Amplitude $F(j\omega)d\omega$ durch das Frequenzintervall $d\omega$, dann bekommt man die Größe $F(j\omega)$, die man daher als Amplitudendichte interpretieren kann. Demgegenüber beschreiben die Koeffizienten c_n der Fourierzerlegung die absolute Amplitude einer einzelnen Sinusschwingung.

Dieser Unterschied hat beispielsweise zur Folge, daß zwar die Fourierzerlegung einer reinen Sinusfunktion $\sin \omega_0 t$ möglich ist, nicht jedoch die Fourier–Transformation, da die Konvergenzbedingung (A.18) offensichtlich nicht erfüllt ist. Die Fourier–Transformation eines reinen Sinussignals müßte auf eine einzelne Sinusschwingung mit einer unendlich großen Amplitudendichte bei der Frequenz ω_0 führen, was einem Diracimpuls $\delta(\omega - \omega_0)$ entspricht.

A.3 Diskrete Fourier–Transformation

Die diskrete Fourier–Transformation (DFT) kann zur näherungsweise numerischen Berechnung des Spektrums kontinuierlicher Signale verwendet werden. Hierbei ist zu berücksichtigen:

- Bedingt durch den endlichen Speicherplatz im Digitalrechner können nur endlich viele Werte $f(k)$ verarbeitet werden.
- Neben der Zeit muß auch die Frequenz diskretisiert werden, da der Rechner nur diskrete Zahlenwerte verarbeiten kann.

Ausgangspunkt zur Herleitung der DFT ist die Fourier–Transformierte

$$F(j\omega) = \int_{-\infty}^{+\infty} f(t)e^{-j\omega t}dt \tag{A.20}$$

eines kontinuierlichen Signals $f(t)$. Zunächst ersetzt man in (A.20) das kontinuierliche Signal durch ein Signal, das man aus der Multiplikation des kontinuierlichen Signals $f(t)$ mit einer Folge von Dirac–Impulsen erhält:

$$f^*(t) = f(t) \sum_{n=-\infty}^{\infty} \delta(t - nT) \tag{A.21}$$

$$= \sum_{n=-\infty}^{\infty} f(nT)\delta(t - nT) . \tag{A.22}$$

Dabei kann $f^*(t)$ als kontinuierliche Darstellung eines mit $f_A = 1/T$ abgetasteten Signals angesehen werden. Die Abtastwerte $f(kT)$ treten als Gewichte der Dirac-Impulse auf.

$$F_d(j\omega) = \int_{-\infty}^{+\infty} f^*(t)e^{-j\omega t} dt \quad (\text{A.23})$$

$$= \int_{-\infty}^{+\infty} \sum_{n=-\infty}^{\infty} f(nT)\delta(t - nT)e^{-j\omega t} dt \quad (\text{A.24})$$

Vertauscht man die Reihenfolge von Summation und Integration, so ergibt sich:

$$F_d(j\omega) = \sum_{n=-\infty}^{+\infty} \int_{-\infty}^{\infty} f(nT)\delta(t - nT)e^{-j\omega t} dt . \quad (\text{A.25})$$

Da die Abtastwerte $f(nT)$ nur implizit von t abhängen, kann $f(nT)$ vor das Integralzeichen geschrieben werden:

$$F_d(j\omega) = \sum_{n=-\infty}^{+\infty} f(nT) \int_{-\infty}^{\infty} \delta(t - nT)e^{-j\omega t} dt . \quad (\text{A.26})$$

Das verbleibende Integral kann mit Hilfe der Definition des Diracimpulses und der Ausblendeigenschaft analytisch gelöst werden:

$$F_d(j\omega) = \sum_{n=-\infty}^{+\infty} f(nT)e^{-j\omega nT} \int_{-\infty}^{\infty} \delta(t - nT) dt \quad (\text{A.27})$$

$$= \sum_{n=-\infty}^{+\infty} f(nT)e^{-j\omega nT} . \quad (\text{A.28})$$

Aus der unendlichen Anzahl von Abtastwerten muß nun eine endliche Anzahl N ausgewählt werden. Dieser Signalausschnitt, auch Fensterbreite genannt, sollte wie folgt gewählt werden:

- Bei einem einmaligen, energiebegrenzten Signal die gesamte Impulsdauer.
- Bei einem periodischen Signal mindestens einige Perioden.
- Bei einem stochastischen Signal mindestens einige Perioden der tiefsten interessierenden Frequenz.

Das Spektrum des abgetasteten und gefensterterten Signals ergibt sich dann zu:

$$F_{df}(j\omega) = \sum_{n=0}^{N-1} f(nT)e^{-j\omega nT} . \quad (\text{A.29})$$

Die Funktion ist periodisch mit einer Periodenlänge T und hat nur an N Frequenzstellen linear unabhängige Funktionswerte. Die Funktion $F_{df}(j\omega)$ wird daher nur für N äquidistante Frequenzen

$$\omega = \frac{2\pi k}{TN}, \quad k = 0, 1, 2, \dots, N - 1 \quad (\text{A.30})$$

ausgewertet:

$$F_{df}\left(j\frac{2\pi k}{TN}\right) = \sum_{n=0}^{N-1} f(nT)e^{-j\frac{2\pi k}{N}n}, \quad k = 0, 1, 2, \dots, N - 1. \quad (\text{A.31})$$

Zur Vereinfachung läßt man die Kennzeichnung df und alle Argumente bis auf die Zeit- und Frequenzindizes weg:

$$F(k) = \sum_{n=0}^{N-1} f(n)e^{-j\frac{2\pi k}{N}n}, \quad k = 0, 1, 2, \dots, N - 1. \quad (\text{A.32})$$

Die Gleichung (A.32) ist die Definition der *Diskreten Fourier-Transformation* (DFT). Symbolisch schreibt man $F(k) = \text{DFT}\{f(n)\}$.

Die Definition der diskreten Fourier-Rücktransformation, auch inverse DFT (IDFT) genannt, lautet

$$f(n) = \frac{1}{N} \sum_{k=0}^{N-1} F(k)e^{j\frac{2\pi k}{N}n}, \quad n = 0, 1, 2, \dots, N - 1. \quad (\text{A.33})$$

Folgerungen aus der DFT-Definition:

- Die DFT liefert aufgrund des endlichen Meßintervalls ($N \cdot T$) eine Näherung des Spektrums an N diskreten Punkten der Frequenzachse.
- Den Abstand

$$\Delta f = \frac{1}{N}f_A, \quad \text{mit} \quad f_A = \frac{1}{T} \quad (\text{A.34})$$

zweier Frequenzpunkte nennt man Frequenzauflösung der DFT. Die Frequenzauflösung ist gleich dem Kehrwert des Meßintervalls.

- Die oberste darstellbare Signalfrequenz ergibt sich zu

$$f_{\max} = \frac{N}{2}\Delta f. \quad (\text{A.35})$$

A.3.1 Approximation der Fourier–Reihe mit Hilfe der DFT

Entsprechend Abschnitt A.1 läßt sich ein periodisches Signal mit der Grundkreisfrequenz ω_0 in die folgende Fourier–Reihe zerlegen:

$$f(t) = A_0 + \sum_{n=1}^{\infty} A_n \sin(n\omega_0 t + \varphi_n) . \quad (\text{A.36})$$

Die Amplituden A_0, A_1, \dots und die Phasenwinkel $\varphi_0, \varphi_1, \dots$ werden dann durch die DFT wie folgt approximiert:

$$A_0 \approx \frac{1}{N} |F(0)| , \quad (\text{A.37})$$

$$A_k \approx \frac{2}{N} |F(k)| , \quad k = \begin{cases} 1, 2, \dots, N/2 & N \text{ gerade} \\ 1, 2, \dots, (N-1)/2 & N \text{ ungerade} \end{cases} \quad (\text{A.38})$$

$$\varphi_k \approx \arg F(k) , \quad k = \begin{cases} 1, 2, \dots, N/2 & N \text{ gerade} \\ 1, 2, \dots, (N-1)/2 & N \text{ ungerade} \end{cases} . \quad (\text{A.39})$$

A.4 Schnelle Fourier–Transformation, FFT

Die 1965 von Cooley und Tukey entwickelte schnelle Fourier–Transformation (**F**ast **F**ourier **T**ransformation) ist der wohl wichtigste Algorithmus in der digitalen Signalverarbeitung. Die FFT liefert dieselben Ergebnisse wie die diskrete Fourier–Transformation mit einer erheblich geringeren Anzahl von Rechenoperationen. Während die DFT für N zu transformierende Werte N^2 komplexe Multiplikationen und Additionen benötigt, sind bei der FFT, sofern N eine Potenz von 2 ist, nur etwa $N \cdot \log(N)$ Operationen notwendig. Folglich spart man bei $N = 1024$ etwa 99% der Operationen ein.