

Elektrotechnische Grundlagen der Informatik

Herbert.Gruenbacher@tuwien.ac.at

Jänner 2011

Inhaltsverzeichnis

Vorwort	xvii
Einleitung	xix
I Systemtheoretische Grundlagen	1
1 Systeme	3
1.1 Lineare und nichtlineare Systeme	4
1.1.1 Lineare Systeme	4
1.1.2 Nichtlineare Systeme	5
1.1.3 Beispiele für Systeme mit linearem Verhalten	5
1.1.4 Beispiele für Systeme mit nichtlinearem Verhalten	6
1.2 Zeitunabhängige und zeitabhängige Systeme	6
1.3 Passive und verlustlose Systeme	6
1.4 Dynamische und nichtdynamische Systeme	6
1.5 Kausale und nichtkausale Systeme	7
1.6 Konzentrierte und verteilte Systeme	7
1.7 Zeitkontinuierliche und zeitdiskrete Systeme	7
1.8 Analoge und digitale Systeme	7
1.9 Stabilität	8
1.10 Zusammenfassung	8
2 Signale und Zeitfunktionen	9
2.1 Signale	9
2.1.1 Testsignale	10
2.2 Elementarfunktionen	11
2.2.1 Faltung mit Elementarfunktionen	12
2.3 Sinusfunktion	14
2.3.1 Phasenverschiebung – Zeitverschiebung	16
2.4 Die komplexe Exponentialfunktion	17
2.5 Zusammenfassung	18
3 Spektren	19
3.1 Mathematische Darstellung von zeitkontinuierlichen Signalen	19
3.2 Spektralsynthese	19
3.3 Spektralanalyse	24
3.3.1 Fourierreihen	24
3.3.2 Numerische Berechnung der Fourierkoeffizienten	30
3.3.3 Die Fouriertransformation	34
3.3.4 Eigenschaften der Fouriertransformation	37
3.3.5 Operationen mit Signalen im Zeit- und Frequenzbereich	38
3.3.6 Frequenz und Zeit – Unschärfeprinzip	40
3.3.7 Numerische Berechnung der Fouriertransformation	42
3.4 Größe von Signalen (Energie und Leistung)	43

3.5	Parseval'sches Theorem	43
3.6	Zusammenfassung	44
II Elektrotechnische Grundlagen		45
4	Elektrische Bauelemente	47
4.1	Einführung	47
4.1.1	Netzwerke, Ströme und Spannungen	48
4.1.2	Unabhängige Spannungs- und Stromquellen	49
4.1.3	Zählpfeile	50
4.2	Die Kirchhoff'schen Gleichungen	50
4.3	Ideale Bauelemente	52
4.3.1	Gesteuerte Quellen	52
4.3.2	Idealer Widerstand	52
4.3.3	Kapazität	53
4.3.4	Induktivität	55
4.3.5	Spannung und Strom an L und C bei Schaltvorgängen	56
4.3.6	Schreibweise der Netzwerkgleichungen	57
4.4	Präzisere Modelle der Bauelemente	58
4.4.1	Spannungsquelle mit Innenwiderstand	58
4.4.2	Ersatzschaltbild für den Kondensator	59
4.4.3	Ersatzschaltbild für die Spule	59
4.5	Zusammenfassung	60
5	Elektrische Netzwerke	63
5.1	Einfache Netzwerke	63
5.2	Topologische Netzwerkbeschreibung	66
5.2.1	Schleifenanalyse	67
5.2.2	Knotenanalyse	67
5.2.3	Beispiele	68
5.3	Zusammenfassung	70
6	Lösung der Netzwerkgleichungen	71
6.1	Gleichstromanalyse	72
6.2	Wechselstromanalyse	73
6.2.1	Effektivwert	73
6.2.2	Komplexe Widerstände	75
6.2.3	Leistung in Wechselstromnetzwerken	78
6.3	Allgemeine Zeitfunktion als Eingangssignal	79
6.3.1	Lösung im Frequenzbereich	79
6.4	Systemantwort	85
6.4.1	Lösung im Frequenzbereich	85
6.4.2	Lösung im Zeitbereich	93
6.4.3	Anfangsbedingungen	95
6.4.4	Eigenschaften der Systemfunktion	95
6.5	Zusammenfassung	96
III Analoge Signalverarbeitung		99
7	Verstärkerschaltungen	101
7.1	Operationsverstärker	102
7.1.1	Der unbeschaltete Operationsverstärker	103
7.1.2	Der reale Operationsverstärker	107
7.2	Gesteuerte Quellen	108
7.2.1	Spannungsgesteuerte Spannungsquellen	108

7.2.2	Stromgesteuerte Spannungsquellen	108
7.2.3	Spannungsgesteuerte Stromquellen	108
7.2.4	Stromgesteuerte Stromquellen	109
7.3	Addierer und Subtrahierer	109
7.4	Integratoren	110
7.5	Differentiatoren	110
7.6	Lösung von Differentialgleichungen	111
7.7	Zusammenfassung	113
8	Analoge Signalverarbeitung	115
8.1	Analoge Filter	116
8.1.1	Approximation eines idealen Tiefpassfilters	117
8.1.2	Realisierung von analogen Filtern	125
8.1.3	Normierung	126
8.1.4	Frequenztransformation	126
8.2	Nichtlineare analoge Signalverarbeitung	127
8.2.1	Signalgleichrichtung	127
8.2.2	Komparator	127
8.3	Zusammenfassung	129
IV	Signalabtastung	131
9	Signalabtastung	133
9.1	Vorteile digitaler Signale	134
9.2	Abtastung im Zeitbereich	134
9.3	Mathematische Darstellung des Abtastvorgangs	135
9.4	Aliasing und Folding	138
9.5	Signalrekonstruktion	139
9.6	Filterung	142
9.6.1	Interpolationsfilter	142
9.6.2	Antialiasing Filter	142
9.7	Digitalisierung	143
9.8	Die zeitdiskrete Fourier-Transformation (DTFT)	144
9.9	Abtastung im Frequenzbereich	145
9.10	Die diskrete Fourier-Transformation (DFT)	146
9.11	Systeme mit unterschiedlichen Abtastraten (Multirate Filter)	149
9.12	Zusammenfassung Fouriertransformationen	150
9.13	Zusammenfassung	151
10	AD- und DA-Umsetzer	153
10.1	Abtast-Halte-Glieder	153
10.2	AD-Umsetzer	154
10.2.1	Parallelverfahren	154
10.2.2	Wägeverfahren	155
10.2.3	Zählverfahren	155
10.3	Digital-Analog-Umsetzer	156
10.4	Zusammenfassung	158
V	Digitale Signalverarbeitung	159
11	FIR-Filter	161
11.1	Allgemeine FIR-Filter	162
11.2	Die Impulsantwort	163
11.2.1	Die Faltungssumme	164
11.3	Frequenzgang von FIR-Filtern	166

11.3.1	Frequenzgang und Überlagerungssatz	170
11.3.2	Einschaltvorgänge	170
11.4	Blockdiagramm von FIR-Filtern	171
11.4.1	Realisierung von Filtern	173
11.5	Kaskadieren von Filtern	174
11.5.1	Vom idealen zum realen Filter	175
11.6	Zusammenfassung	179
12	Die z-Transformation	181
12.1	Der z -Bereich	182
12.2	Eigenschaften der z -Transformation	183
12.3	Blockdiagramme und Systemfunktion	184
12.4	Inverse Filterung	185
12.5	Zusammenhang z -Bereich und Frequenzbereich	185
12.5.1	Pole und Nullstellen von $H(z)$	186
12.6	Zusammenfassung	189
13	IIR-Filter	191
13.1	Vom FIR- zum IIR-Filter	191
13.2	Zeitverhalten von IIR-Filtern	192
13.3	Eigenschaften von IIR-Filtern	193
13.3.1	Linearität und Zeitinvarianz	193
13.3.2	Impulsantwort	194
13.3.3	Sprungantwort	195
13.3.4	Systemfunktion von IIR-Filtern	197
13.3.5	Unendliche Folgen im z -Bereich	197
13.4	Realisierung von IIR-Filtern	198
13.4.1	Erste Direktform	198
13.4.2	Zweite Direktform	199
13.4.3	Transponierte Form	200
13.5	Pole und Nullstellen	200
13.5.1	Stabilität und Lage der Pole	201
13.6	Frequenzgang von IIR-Filtern	201
13.7	3D-Darstellung der Systemfunktion	202
13.8	Die inverse z -Transformation	203
13.8.1	Partialbruchzerlegung	205
13.9	Einschwingvorgänge und Stabilität	208
13.10	Konjugiert komplexe Pole	210
13.10.1	z -Transformierte komplexer Pole	211
13.11	Blockdiagramm der Partialbruchzerlegung	212
13.12	Darstellung von Systemen	212
13.12.1	Zeitbereich	212
13.12.2	z -Bereich	213
13.12.3	Frequenzbereich	213
13.12.4	Blockdiagramm	215
13.13	Zusammenfassung	215
VI	Anhänge	217
A	Komplexe Zahlen	219
A.1	Darstellung komplexer Zahlen	220
A.2	Die Euler'schen Formeln	222
A.3	Rechenregeln für komplexe Zahlen	224
A.3.1	Regeln für kartesische Koordinaten	224
A.3.2	Regeln für Polarkoordinaten	225

A.4	Geometrische Betrachtung	226
A.5	Achtung Phase	229
A.6	Zusammenfassung	230
B	MATLAB	233
B.1	Der Workspace	234
B.1.1	Hilfe und Orientierung	234
B.1.2	Bildschirmausgabe	234
B.1.3	Namensgebung	234
B.1.4	Zahlen	234
B.1.5	Skalare, Vektoren, Matrizen	234
B.2	Elementare Operationen	236
B.3	Elementare komplexe Operationen	236
B.4	Graphiken in Matlab	236
B.5	Programmierung	237
B.5.1	Kommentare	237
B.5.2	Scripts	237
B.5.3	Funktionen	237
B.5.4	Kontrollstrukturen	237
B.5.5	Ein- und Ausgabe	238
B.6	Kontinuierliche LTI-Systeme	238
B.6.1	Approximation analoger Filter	239
B.7	Diskrete Systeme	239
B.7.1	Approximation digitaler Filter	240
B.8	Praktischer Filterentwurf	240
B.9	Signalverarbeitung	240
B.10	Zusammenfassung	241

Tabellenverzeichnis

Abbildungsverzeichnis

1.1	System als "black box"	4
2.1	Analoge und digitale Signale	10
2.2	Sprungfunktion $\delta_{-1}(t)$	11
2.3	Darstellung des Einheitsimpulses	11
2.4	Veranschaulichung des Faltungsintegrals	13
2.5	$s(t) = A \sin(\frac{2\pi}{T_0}t - \frac{3\pi}{4})$	14
2.6	Komplexe Darstellung der Sinusfunktion	15
2.7	Phasenverschiebung Kosinus	16
2.8	Realteil der komplexen Exponentialfunktion	18
3.1	Summe von Sinusschwingungen $s(t) = s_1(t) + s_2(t) + s_3(t)$	20
3.2	Signal im Zeit- und Frequenzbereich	21
3.3	Spektrallinien bei quadratischem System	22
3.4	Spektraldarstellung einer Sägezahnfunktion	22
3.5	Zeiger mit entgegengesetzter Drehrichtung ω und $-\omega$	23
3.6	Rechteckschwingung	27
3.7	Dreieckschwingung	29
3.8	Signal in Zeigerdarstellung	31
3.9	Spektrum von $ \sin(t) $	32
3.10	Spektrum von $s(t) = t^2$	33
3.11	Frequenzsynthese der periodischen Funktion t^2 für $-1 \leq t \leq 1$	34
3.12	Aperiodisches Signal periodisch fortgesetzt	34
3.13	Rechteckimpuls im Zeit- und Frequenzbereich	36
3.14	Einheitsimpuls im Zeit- und Frequenzbereich	36
3.15	Betrags-, Phasenspektrum und Zeitfunktion von $s(t) = \delta_{-1}(t)e^{-t} \sin 10t$	37
3.16	Skalierung der Zeitachse	38
3.17	Spektrum Kosinusburst $f_0 = 3$, $\tau = 2$ und $\tau = 0.5$	41
4.1	Netzwerkanalyse	48
4.2	Idealisierte Schaltung	48
4.3	Schaltsymbole für Spannungs- und Stromquellen	49
4.4	Erzeuger- (links) und Verbraucher- (rechts) Zählpfeilsystem	50
4.5	Knotenregel	50
4.6	Zählpfeile	51
4.7	Maschenregel	51
4.8	Schaltzeichen für gesteuerte Quellen	52
4.9	Kapazität und Schaltzeichen	53
4.10	Mechanisches Analogon zur Kapazität	54
4.11	Aufladen der Kapazität	54
4.12	Strom- und Spannungsverlauf beim Laden einer Kapazität	55
4.13	Schaltungssymbol der Induktivität	55
4.14	Einschalten einer Induktivität	56
4.15	Strom- und Spannungsverlauf beim Einschalten einer Induktivität	56
4.16	Ersatzschaltbild reale Spannungsquelle	58

4.17	Belastete Quelle	58
4.18	Umwandlung Spannungs- in Stromquelle	59
4.19	Ersatzschaltbild Kondensator	59
4.20	Ladungsverteilung auf zwei Kapazitäten	60
5.1	Serienschaltung von R, L und C	64
5.2	Parallelschaltung von R, L, C	64
5.3	Spannungsteiler	65
5.4	Ersatzwiderstände im Netzwerk	65
5.5	Netzwerkgraph N	66
5.6	Spannbaum zum Graphen N	67
5.7	Baum und Glieder	67
5.8	Anwendung der Schleifenanalyse	68
5.9	Anwendung der Knotenanalyse	70
6.1	Übertragungsverhalten eines Netzwerks	71
6.2	Schleifenanalyse	72
6.3	Zeitlich veränderliche Netzwerkgrößen	73
6.4	Wechselstromleistung	74
6.5	Strom und Spannung an der Induktivität	75
6.6	Serienschaltung R und L	76
6.7	Netzwerkberechnung mit Knotenanalyse	77
6.8	Bode-Plot	78
6.9	Phasenverschiebung am Zweipol	78
6.10	RLC-Serienschwingkreis	79
6.11	Netzwerkberechnung mit Laplace-Transformation	86
6.12	LC-Tiefpassfilter	89
6.13	Pol-/Nullstellendiagramm $\frac{1}{s^3+2s^2+2s+1}$	89
6.14	PN-Diagramm von $\frac{7s+11}{s^2+4s+3}$	90
6.15	$h(t) = h_1(t) + h_2(t)$	90
6.16	Bode-Plot LC-Tiefpass $H(s) = \frac{1}{s^3+2s^2+2s+1} _{s=j\omega}$	93
6.17	Frequenzgang RC-Hochpass $\frac{R^2 C^2 s^2}{R^2 C^2 s^2 + 3RCs + 1}$	93
6.18	Zerlegung des Eingangssignals	94
6.19	Anfangsbedingungen Induktivität	95
6.20	Anfangsbedingungen Kapazität	95
6.21	Axonometrische Darstellung von $ H(s) $ in [dB]	96
7.1	Schaltungssymbole für Verstärker	101
7.2	Blockdiagramm einer Signalverarbeitungskette	102
7.3	Schaltymbol des Operationsverstärkers	103
7.4	Der invertierende OPV	104
7.5	Der nichtinvertierende OPV	105
7.6	RC-Tiefpass	105
7.7	Asymptoten an Frequenzgang	106
7.8	Frequenzgang ohne und mit Gegenkopplung	107
7.9	Stromgesteuerte Spannungsquelle	108
7.10	Spannungsgesteuerte Stromquellen	109
7.11	Summationsverstärker	109
7.12	Integrator	110
7.13	Bodediagramm des idealen Integrators	111
7.14	Differentiator	111
7.15	OPV-Schaltung zur Lösung von DGL 2. Ordnung	112
8.1	Frequenzverhalten von Filtern	116
8.2	Toleranzschema eines Filters	117
8.3	Frequenzkennlinie eines idealen TP-Filter	118

8.4 Approximation durch Potenzansatz 118

8.5 Nullstellen von (a) $F(s)F(-s) = 1 + s^8$ (b) $F(s)F(-s) = 1 - s^6$ 119

8.6 Frequenzgang Potenz- und Tschebyscheff-Filter 5. Ordnung 120

8.7 Cauerfilter 5. Ordnung (linear und doppelt logarithmisch) 121

8.8 Sinuskomponenten mit unterschiedlicher Phase 121

8.9 SZeilensignalin Falschfarbendarstellung 122

8.10 Bessel- und Tschebyscheff-Tiefpass 5. Ordnung 123

8.11 Frequenzgang und Sprungantworten Filter 5. Ordnung 124

8.12 Aktiver Tiefpass 126

8.13 Frequenztransformation 127

8.14 Diode und Ersatzschaltung (Strichlierte Linien stehen für ein genaueres Modell) 128

8.15 Gleichrichtung 128

8.16 Komparator 128

9.1 Signalabtastung 135

9.2 Periodisches Spektrum 137

9.3 Überlappen von Spektren 137

9.4 Aliasing 138

9.5 Folding 139

9.6 Drehung eines Speichenrads 139

9.7 Rekonstruktion mit Rechteckimpulsen 140

9.8 Rekonstruktion mit Dreieckimpulsen 140

9.9 Interpolation durch ideales Filter 140

9.10 Impulsantwort des idealen Filters 141

9.11 Rekonstruktion durch sinc-Pulse 142

9.12 Überlappung benachbarter Spektren 143

9.13 Blockdiagramm digitale Signalverarbeitung 143

9.14 Amplitudenquantisierung 144

9.15 DTFT von $x[n] = 0.5^n \delta_{-1}[n]$ 145

9.16 Abtastung im Frequenzbereich 146

9.17 Abtastung im Zeit- und Frequenzbereich 147

9.18 Fourierreihe (FR), Fouriertransformation (FT), Zeitdiskrete FT (DTFT), Diskrete FT (DFT) 150

10.1 Abtast- und Halteschaltung 153

10.2 Blockdiagramm 2bit-Parallelwandler 154

10.3 Blockdiagramm Wandler nach Zählverfahren 155

10.4 Blockdiagramm Dual-Slope-Wandler 156

10.5 DA-Parallelwandler 157

10.6 DA-Wandler nach dem Wägeverfahren 157

10.7 DA-Wandler nach dem Zählverfahren 157

11.1 Zeitdiskretes System 161

11.2 Mittelwertfilter 162

11.3 LTI-System 164

11.4 Impulsantwort FIR-Filter 164

11.5 Faltung 166

11.6 Frequenz- und Phasengang $b_k = [1, 2, 1]$ 168

11.7 Amplituden- und Phasengang von $B = [1, 2, 1]$ 169

11.8 Sprungfunktion 171

11.9 Eingeschaltetes Kosinussignal 172

11.10 Recheneinheiten digitaler Filter 172

11.11 Blockdiagramm Direktform FIR-Filter 172

11.12 Blockdiagramm FIR Transponierte Form 173

11.13 Kaskadieren von FIR-Filtern 174

11.14	$W_R(\hat{\omega})$ und $H_{iTP}(\hat{\omega})$ in normierter Darstellung	176
11.15	Ergebnis der Faltung $ H_{TP}(\hat{\omega}) $	177
11.16	Vergleich Tiefpass mit Dreieck- und Rechteckfenster	177
11.17	Tiefpass mit Hamming- bzw. Rechteckfenster	178
11.18	Fenstereigenschaften	178
12.1	Darstellung mit z^{-1}	183
12.2	Reihenfolge der Kaskadierung	184
12.3	Faktorisierung von Systemen	185
12.4	Einheitskreis in der komplexen Ebene	186
12.5	Betrag von $H(z) = \frac{1}{1-0.9z^{-1}}$	187
12.6	Zylinder über Einheitskreis	187
12.7	Betrag des Frequenzgangs von $H(z) = \frac{1}{1-0.9z^{-1}} _{z=j\omega}$	188
12.8	Pol-/Nullstellen-Darstellung	188
13.1	Blockdiagramm IIR-Filter	192
13.2	Eingang- & Ausgangssignal IIR-Filter	193
13.3	Sprungantwort für $a_1 = 1$ und $a_1 = -1$	196
13.4	Sprungantwort IIR-Filter	196
13.5	$y[n] = 0.7^n \delta_{-1}[n]$	198
13.6	Blockdiagramm IIR-Filter	198
13.7	IIR-Filter 2. Direktform	199
13.8	IIR-Filter 2. Direktform	199
13.9	Signalflussgraph IIR-Filter 2. Direktform	199
13.10	Transponierte Form	200
13.11	Stabiles $a_1 = -0.9$ und instabiles $a_1 = -1.1$ Verhalten	201
13.12	Grafische Ermittlung des Frequenzgangs	202
13.13	3D-Plot des Betrags eines IIR-Filters 3. Ordnung	203
13.14	Pol-/Nullstellendiagramm des IIR-Filters 3. Ordnung	204
13.15	Frequenzgang des Filters 3. Ordnung	204
13.16	Sprungantwort IIR-Filter 3. Ordnung	204
13.17	Impulsantwort $H(z) = \frac{-1+1.7z^{-1}}{1-0.4z^{-1}-0.21z^{-2}}$	207
13.18	Antwort des Systems $H(z) = \frac{1}{1+0.9z^{-1}}$ auf eingeschaltetes Kosinussignal	209
13.19	Instabiles Filter Polstelle bei $z_P = 1.1$	209
13.20	Impulsantwort IIR-Filter, Pole auf dem Einheitskreis	210
13.21	Impulsantwort	211
13.22	Impulsantwort $H(z) = \frac{1}{1+z^{-1}}$	212
13.23	Blockdiagramm Partialbruchrealisierung	213
13.24	Pol-/Nullstellendarstellung	214
13.25	Elliptischer Tiefpass 5. Ordnung	214
13.26	Blockdiagramm IIR-Filter 2. Ordnung	215
A.1	Zeiger in der komplexen Ebene	220
A.2	Zeiger in Polardarstellung	221
A.3	Polar - kartesisch	221
A.4	Zeiger im 2. Quadranten	222
A.5	Geometrische Interpretation der Euler'schen Formel	223
A.6	Rechts- und linksdrehende Zeiger	224
A.7	Zeigeraddition	226
A.8	Subtraktion von Zeigern	227
A.9	Zeigermultiplikation	227
A.10	Zeigerdivision	227
A.11	Potenzen der komplexen Zahl	228
A.12	Wurzeln von 1	228
A.13	Wurzel aus e^{j50°	229
A.14	Phasendarstellung	230

A.15 »Negativer« Betrag erzeugt $Phase = 0$ 231

Vorwort

Dieser Text ist als Unterstützung zu meinen Vorlesungen an der Technischen Universität Wien entstanden und richtet sich vor allem an Studentinnen und Studenten der Technischen Informatik.

Wegen ihrer großen Vorteile setzt man, wenn möglich, die digitale Informationsverarbeitung ein. In technischen Anwendungen kommunizieren Rechner aber häufig in Form von analogen Signalen mit Sensoren und Aktoren. Für das Verständnis des Gesamtsystems sind daher sowohl Kenntnisse im digitalen als auch im analogen Bereich unerlässlich.

Die folgenden Kapitel sollen, neben einem Überblick über die elektrotechnischen Grundlagen, ein Verständnis vermitteln, welche Aufgabenstellungen besser (oder ausschließlich) in analoger und welche besser in digitaler Technik verwirklicht werden können.

Die Grenzen zwischen analoger und digitaler Technik sind fließend:

Was heute nur analog realisiert werden kann, wird morgen vielleicht schon in digitaler Technik umsetzbar sein. Der wichtigste Einflussfaktor in dieser Entwicklung sind die technologischen Fortschritte in der Halbleitertechnik.

In der Version 2011 wurden zahlreiche Tippfehler und unklar formulierte Stellen ausgebessert. Für Überarbeitung und Korrektur lesen möchte ich mich bei Herrn Matej Pavlovic herzlich bedanken.

Fehler- und Verbesserungshinweise werden vom Autor dankbar entgegengenommen.

Herbert Grünbacher
Wien, Jänner 2011

Einleitung

Dieser Text ist in die Abschnitte Systemtheoretische Grundlagen, Elektrotechnische Grundlagen, Analoge Signalverarbeitung, Signalabtastung, Elektronik, Digitale Signalverarbeitung und einen Anhang gegliedert.

- Die *systemtheoretischen Grundlagen* umfassen neben einem Kapitel über Systeme, die Darstellung von Signalen im Zeit- und im Frequenzbereich.
- Die *elektrotechnischen Grundlagen* umfassen ein Kapitel über die elektrischen Bauelemente R, L und C auf axiomatischer Basis (ohne Feldbegriff), ein Kapitel über die systematische Erstellung der Netzwerkgleichungen, sowie ein Kapitel über die Lösung der Netzwerkgleichungen im Zeit-, Frequenz- und Bildbereich.
- Im Abschnitt *Analoge Signalverarbeitung* werden die Möglichkeiten der Signalverarbeitung mit Operationsverstärkern und analoge Filter behandelt.
- Der Abschnitt *Signalabtastung* stellt mit dem Abtasttheorem die theoretischen Zusammenhänge beim Übergang von analogen in digitale Signale dar und gibt einen Überblick über die Prinzipien von Analog-/Digital- bzw. Digital-/Analog-Wandlern.
- Der Abschnitt *Digitale Signalverarbeitung* umfasst die Kapitel Finite und Infinite Impulse Response Filter, sowie ein Kapitel über die z -Transformation.
- Im *Anhang* findet sich eine zusammenfassende Darstellung über Komplexes Rechnen, sowie einige Hinweise zur Verwendung von Matlab.

Mit diesem Text soll die Erarbeitung von Grundlagenwissen unterstützt werden und die theoretische Basis für ein Fachgebiet gelegt werden, in dem man durch Probieren nicht weiterkommt. Die Zusammenhänge werden in mathematischer Form dargestellt, um Erfahrungen in eine möglichst kompakte Form zu bringen und in dieser Form zu verbreiten: Mathematik als Kurzschrift für die Zusammenfassung von Erkenntnissen und als Anweisung für die Auswertung/Anwendung der Erkenntnisse. Durch mathematische Verfahren können viel kompliziertere Zusammenhänge erfasst werden, als es mit blosem Nachdenken möglich wäre. Um vom oft mühsamen Rechnen zu befreien, wird Matlab als Werkzeug zur Durchführung von Berechnungen und von grafischen Darstellungen eingesetzt.

Teil I

**Systemtheoretische
Grundlagen**

Kapitel 1

Systeme

Inhalt

1.1	Lineare und nichtlineare Systeme	4
1.1.1	Lineare Systeme	4
1.1.2	Nichtlineare Systeme	5
1.1.3	Beispiele für Systeme mit linearem Verhalten	5
1.1.4	Beispiele für Systeme mit nichtlinearem Verhalten	6
1.2	Zeitunabhängige und zeitabhängige Systeme	6
1.3	Passive und verlustlose Systeme	6
1.4	Dynamische und nichtdynamische Systeme	6
1.5	Kausale und nichtkausale Systeme	7
1.6	Konzentrierte und verteilte Systeme	7
1.7	Zeitkontinuierliche und zeitdiskrete Systeme	7
1.8	Analoge und digitale Systeme	7
1.9	Stabilität	8
1.10	Zusammenfassung	8

Ein System ist eine Anordnung von miteinander verbundenen Komponenten zur Realisierung einer technischen Aufgabenstellung. Ein System kann abstrakt als ein (nicht näher definierter) Operator aufgefasst werden, der Eingangsgrößen auf Ausgangsgrößen abbildet. Wird der Zusammenhang zwischen Eingang und Ausgang eines Systems mathematisch dargestellt, spricht man von Eingangs-Ausgangs-Verhalten des Systems. Ist das E/A-Verhalten – ohne die interne Struktur zu kennen – bekannt, spricht man vom System als »black box«.

Wir schreiben formal

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} = \mathcal{A} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad (1.1)$$

und beschreiben damit die Ausgangsgröße \mathbf{y} , die durch Wirkung des Systems \mathcal{A} auf die Eingangsgröße \mathbf{x} entsteht.

Ein elektrisches System, eine elektrische Schaltung, ist eine Zusammenschaltung von elektrischen Komponenten (Widerständen, Kondensatoren, Transistoren, ...). Bei einem elektrischen System werden die Systemeigenschaften durch das Verhalten der Bauelemente (z.B. das ohmsche Gesetz bei Widerständen) und durch die Verbindungen zwischen den Bauelementen (die Kirchhoff'schen Gesetze) bestimmt. Daraus entstehen Gleichungen, die Eingangs- und Ausgangsgrößen in Beziehung setzen. Diese Gleichungen stellen ein mathematisches Modell des elektrischen Systems dar.

Systeme können durch mathematische Modelle dargestellt werden. Die Systemzusam-

Ein System ist eine Anordnung von Komponenten, die Eingangsgrößen auf Ausgangsgrößen abbildet.

Die Beziehungen zwischen Eingangsgrößen und Ausgangsgrößen können durch mathematische Gleichungen beschrieben werden.

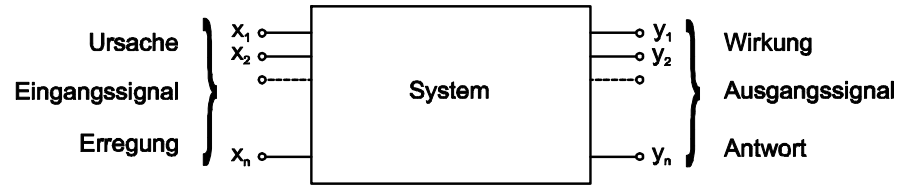


Abbildung 1.1: System als "black box"

menhänge werden in Form von linearen und nichtlinearen Gleichungen oder linearen und nichtlinearen Differenzen- oder Differentialgleichungen erfasst. Die Lösung dieser Gleichungssysteme beschreibt das Systemverhalten. Die Lösung der Systemgleichungen ist häufig schwierig oder sogar unmöglich und man muss Vereinfachungen treffen oder findet nur numerische Lösungen.

Bei kontinuierlichen Eingangsgrößen treten die Systemgleichungen in Form von Differentialgleichungen auf, bei diskreten Systemen in Form von Differenzgleichungen.

Bei der Untersuchung von Systemen treten folgende Aufgabenstellungen auf:

Systemanalyse

- Es muss ein mathematisches Modell und die Antwort des Systems auf eine gegebene Eingangsgröße gefunden werden. Diese Aufgabe nennt man Systemanalyse.

Systemsynthese

- Für eine gegebene Eingangsgröße und eine gewünschte Ausgangsgröße wird ein System gesucht. Diese Aufgabe nennt man Systemsynthese.
- Die dritte Möglichkeit, zu einem gegebenen System und einer bekannten Ausgangsgröße die Eingangsgröße zu bestimmen, hat keinen eigenen Namen. Diese Aufgabenstellung tritt in der Messtechnik auf.

Systeme können eingeteilt werden in:

1. Lineare und nichtlineare Systeme
2. Zeitunabhängige und zeitabhängige Systeme
3. Dynamische und nichtdynamische Systeme
4. Kausale und nichtkausale Systeme
5. Systeme mit konzentrierten und verteilten Parametern
6. Zeitkontinuierliche und zeitdiskrete Systeme
7. Analoge oder digitale Systeme
8. Stabile und nichtstabile Systeme

1.1 Lineare und nichtlineare Systeme

1.1.1 Lineare Systeme

Ein System ist genau dann linear, wenn der Überlagerungssatz gilt, d.h. wenn das System homogen und additiv ist.

Ein System ist homogen, wenn folgender Zusammenhang gilt:

$$\begin{array}{c} \text{Ursache} \rightarrow \text{Wirkung} \\ \downarrow \\ k \cdot \text{Ursache} \rightarrow k \cdot \text{Wirkung} \end{array}$$

die k -fache Ursache führt zur k -fachen Wirkung (für eine beliebige Konstante $k \in \mathbb{R}$). Ein System ist additiv, wenn gilt:

$$\begin{array}{c}
 \text{Ursache1} \rightarrow \text{Wirkung1} \\
 \text{Ursache2} \rightarrow \text{Wirkung2} \\
 \Downarrow \\
 \text{Ursache1} + \text{Ursache2} \rightarrow \text{Wirkung1} + \text{Wirkung2}
 \end{array}$$

Homogenität und Additivität zusammengefasst führen zum Überlagerungssatz

$$\begin{array}{c}
 k_1 \cdot \text{Ursache1} + k_2 \cdot \text{Ursache2} \rightarrow k_1 \cdot \text{Wirkung1} + k_2 \cdot \text{Wirkung2} \\
 \forall k_1, k_2 \in \mathbb{R}
 \end{array}$$

»Ursache« und »Wirkung« sind dabei als nicht näher definierte abstrakte Größen zu verstehen. Man kann sich zum Beispiel Zahlen, Funktionen (in einer oder mehreren Variablen), physikalische Größen wie Strom oder Spannung, sowie (diskrete oder kontinuierliche) Signale (Siehe Abschnitt 2.1) darunter vorstellen.

Bei linearen Systemen durchlaufen Signale das System ohne zu interagieren, d.h. ohne sich gegenseitig zu beeinflussen!

Beispiel 1 (Lineares Kopfrechnen) *Ein praktisches Beispiel für die Anwendung des Überlagerungssatz liefert das Kopfrechnen: Die »Eingangsgröße« 3023 wird über ein »System« mit der Eigenschaft $y = 4 \cdot x$ »übertragen«. 3023 wird im Kopf in die »Ursachen« $3000 + 20 + 3$ zerlegt, die »Wirkungen« auf die Ursachen durch Multiplikation mit 4 ermittelt und die Gesamtwirkung durch Addition der Einzelwirkungen ermittelt: $12000 + 80 + 12 = 12092$.*

Da bei linearen Systemen (per Definition) der Überlagerungssatz gilt, können »komplizierte« Eingangsgrößen aus einfacheren Komponenten zusammengesetzt werden. Für jede dieser Komponenten wird einzeln die Ausgangsgröße des Systems ermittelt. Die Summe der Ausgangsgrößen für alle Komponenten liefert die Ausgangsgröße für die »komplizierte« Eingangsgröße.

1.1.2 Nichtlineare Systeme

Als Beispiel eines nichtlinearen Systems betrachten wir ein einfaches quadratisches System.

Beispiel 2 *Wir zeigen das Verhalten des Systems, das durch die Funktion $y = x^2$ (die den Zusammenhang zwischen Eingang und Ausgang beschreibt) gegeben ist. Als gutes Beispiel dient wieder das Kopfrechnen. Zerlegen wir die Eingangsgröße 27 in einfachere Komponenten 20 und 7, erhalten wir die entsprechenden Ausgangsgrößen $20^2 = 400$ und $7^2 = 49$. Wenn wir diese addieren, erhalten wir $400 + 49 = 449$, was aber $\neq 27^2 = 729$ ist. Der Überlagerungssatz ist damit verletzt, das System ist also nicht linear.*

1.1.3 Beispiele für Systeme mit linearem Verhalten

- Elektrische Schaltkreise aus ohmschen Widerständen, Kondensatoren und Spulen
- Elektronische Schaltkreise wie Verstärker und Filter (innerhalb des Versorgungsspannungsbereichs)
- Mechanische Systeme aus Masse-Feder-Dämpfung
- Resonanz und Nullung
- Differentiation und Integration
- Ausbreitung von elektromagnetischen Wellen und Schallwellen in isotropen Medien
- *Alle Systeme die durch lineare Differential- oder Differenzgleichungen beschrieben werden können.*

Ein System ist linear, wenn der **Überlagerungssatz** gilt.

Bei **linearen Systemen** können die Systemantworten auf **einzelne Komponenten** der Eingangsgröße **getrennt** berechnet werden.

Reale Systeme sind nie ganz linear.

Anmerkung 3 *Linearität von realen Systemen ist nur innerhalb gewisser physikalischer Grenzen gegeben. Werden diese Grenzen überschritten, verhalten sich Systeme nichtlinear bis sie begrenzt oder zerstört werden.*

1.1.4 Beispiele für Systeme mit nichtlinearem Verhalten

- Leistung eines elektrischen Widerstands $P(t) = R * I^2(t)$
- Nichtlineare elektronische Schaltungen wie Spitzendetektoren, Quadrierer, Frequenzverdoppler, Schwellwertschalter, Komparatoren
- Nichtlineare Effekte in elektronischen Schaltkreisen wie Begrenzen, nichtlineares Verstärken, Sättigungseffekte
- Hysteresis-Effekte in magnetischen Schaltkreisen
- Multiplikation von Signalen wie bei der Frequenzmischung (Die Multiplikation mit einer Konstanten ist eine lineare Operation.)
- Digitale Logik
- Ausbreitung von elektromagnetischen Wellen und Schallwellen in anisotropen Medien
- *Alle Systeme deren Beschreibung auf nichtlineare Differential- oder Differenzgleichungen führt.*

Anmerkung 4 *In vielen nichtlinearen Systemen kann das System um den Betriebspunkt linearisiert werden und damit die nichtlineare auf eine lineare Beschreibung vereinfacht werden.*

1.2 Zeitunabhängige und zeitabhängige Systeme

Ein zeitinvariantes System ändert seine Eigenschaften mit der Zeit nicht.

Ein System ist *zeitinvariant*, wenn ein um n_0 verzögertes (verschobenes) Eingangssignal, zu einem um n_0 verzögerten Ausgangssignal führt, das System also sein Übertragungsverhalten nicht mit der Zeit ändert. Ein lineares System ist zeitunabhängig, wenn die Koeffizienten der linearen Differentialgleichungen, die das System beschreiben (bei elektrischen Netzwerken die Werte der Bauelemente), konstant sind. Man spricht von linearen und zeitinvarianten (time invariant) (LTI) Systemen.

In analogen Systemen ist Zeitinvarianz wegen der Temperaturabhängigkeit und Alterung der elektronischen Bauelemente in der Regel nur eingeschränkt gegeben, digitale Systeme haben hier einen klaren Vorteil, da ihr Übertragungsverhalten davon unbeeinflusst ist.

1.3 Passive und verlustlose Systeme

Ein System wird passiv genannt, wenn ein Eingangssignal endlicher Energie zu einem Ausgangssignal führt, dessen Energie maximal gleich der Energie des Eingangssignals ist. Wenn die Energie von Eingangs- und Ausgangssignal gleich ist, spricht man von einem verlustlosen System.

1.4 Dynamische und nichtdynamische Systeme

Dynamische Systeme ziehen auch vergangene Werte der Eingangsgröße in die "Berechnung" der Ausgangsgröße mitein.

Ein System heißt nichtdynamisch (speicherlos), wenn die Antwort des Systems $y(t)$ nur vom Wert der Eingangsgröße $x(t)$ zur Zeit t abhängt. Bei dynamischen Systemen hängt die Ausgangsgröße nicht nur vom augenblicklichen Wert der Eingangsgröße, sondern auch

von vergangenen Werten ab. Bei elektrischen Netzwerken sind Schaltungen mit Kondensatoren und Spulen dynamische Systeme (es wird elektrische und magnetische Energie in den Bauelementen gespeichert), bei digitalen Systemen kann der Speicher z.B. der Inhalt eines Registers sein.

1.5 Kausale und nichtkausale Systeme

Bei kausalen Systemen hängt die Systemantwort lediglich von gegenwärtigen und vergangenen Werten der Eingangsgröße ab, nicht jedoch von zukünftigen Werten der Erregung. Ein kausales System kann erst antworten, wenn eine Eingangsgröße anliegt. Antwortet ein System ohne Anliegen einer Eingangsgröße, so muss es Kenntnis über das zukünftige Verhalten der Eingangsgröße haben, was bei physikalischen Systemen nicht möglich ist, da die Wirkung nicht vor der Ursache eintreten kann. Jedes praktische System, das in Echtzeit arbeitet, muss daher ein kausales System sein.

Kausalität hat nur eine Bedeutung, wenn die Begriffe »vorher« und »nachher« eine Bedeutung haben. Bei der Signalverarbeitung von aufgezeichneten Daten verschwindet diese Bedeutung und Signalverarbeitungssysteme können dann auch nichtkausal sein, eine Verarbeitung in Echtzeit ist dann aber nicht möglich.

Nichtkausale Systeme haben "hellseherische Fähigkeiten".

1.6 Konzentrierte und verteilte Systeme

Bei der Beschreibung von elektrischen Systemen können wir häufig vereinfachte Bauelementbeziehungen annehmen.

So gilt zum Beispiel das ohmsche Gesetz, das den Zusammenhang zwischen Spannungsabfall und Strom durch den Widerstand beschreibt, $U = R \cdot I$.

Bei dieser Modellierung der elektrischen Zusammenhänge wird stillschweigend die Annahme gemacht, dass der Strom durch den Widerstand an jedem Punkt derselbe ist. In der Wirklichkeit ist diese Annahme nur eingeschränkt gültig, da sich elektrische Zustände nicht unendlich schnell, sondern nur mit endlicher Geschwindigkeit ausbreiten. In Wirklichkeit ist die elektrische Spannung an einem Widerstand, durch den ein Stromimpuls fließt, nicht nur eine Funktion der Zeit, sondern auch des Ortes. Bei niedrigen Frequenzen oder genauer bei Wellenlängen, die groß gegen die geometrischen Abmessungen der elektrischen Schaltung sind, kann die verteilte Natur der elektrischen Eigenschaften vernachlässigt werden.

Wenn diese Vereinfachung auf Grund hoher Frequenzen nicht mehr möglich ist, müssen die Systeme, wie z.B. Leitungen, die mit hoher Signalfrequenz gespeist werden oder Antennen, in Abhängigkeit von Ort und Zeit dargestellt werden und es entstehen bei der Modellierung (lineare oder nichtlineare) partielle Differentialgleichungen.

1.7 Zeitkontinuierliche und zeitdiskrete Systeme

Systeme deren Eingangs- und Ausgangsgrößen zeitkontinuierlich sind, nennt man zeitkontinuierliche Systeme. Wenn die Eingangsgrößen abgetastet werden und die Werte der Ausgangsgröße nur zu bestimmten Zeitpunkten bekannt sind, spricht man von zeitdiskreten Systemen (Quantisierung im Zeitbereich).

1.8 Analoge und digitale Systeme

Ein analoges System ist ein System, dessen Eingangsgröße jeden beliebigen Wert innerhalb eines kontinuierlichen Bereiches annehmen kann. Es gibt also eine unendliche Anzahl an möglichen Werten der Eingangsgröße.

Im Gegensatz dazu kann die Eingangsgröße eines digitalen Systems nur eine endliche Zahl von Werten einnehmen (Quantisierung im Amplitudenbereich). Die Begriffe analog und digital beziehen sich auf den *Wertebereich* und nicht auf den Zeitbereich. Dennoch

wird der Begriff analoges System häufig für zeit- *und* wertekontinuierliche Systeme und der Begriff digitales System für zeit- *und* wertediskrete Systeme verwendet, da diese beiden Formen die technisch häufigsten sind.

1.9 Stabilität

Ein BIBO-stabiles System produziert für eine beschränkte Eingangsgröße immer eine beschränkte Ausgangsgröße.

Ein stabiles System, das durch externe Anregung aus der Ruhelage ausgelenkt wird, kehrt nach einiger Zeit wieder in die Ruhelage zurück. Die Definition von Stabilität kann vom Eingangs-/Ausgangsverhalten ausgehen: beschränkte Eingangsgrößen (**b**ounded **i**nput) produzieren beschränkte Ausgangsgrößen (**b**ounded **o**utput). Diese Stabilitätsdefinition nennt man BIBO-Stabilität.

1.10 Zusammenfassung

Ein System ist eine Anordnung, bei der der Zusammenhang zwischen Ursache = Eingangssignal = Erregung und Wirkung = Ausgangssignal = Antwort formal mit der Beziehung $\mathbf{y} = \mathcal{A} \cdot \mathbf{x}$ beschrieben wird. Die Systemeigenschaften \mathcal{A} werden durch einen nicht näher definierten Operator beschrieben. Systeme können aus physikalischen Komponenten (Hardware) bestehen oder durch einen Algorithmus (Software) beschrieben werden.

Elektrische Schaltungen (Systeme) bestehen aus miteinander verbundenen Bauelementen (Widerständen, Kondensatoren, Spulen, ...). Die Systemeigenschaften \mathcal{A} werden durch die Beziehung zwischen Strom und Spannung an den Bauelementen, sowie durch die Gesetze der Verbindung der Bauelemente (Kirchhoff'sche Gesetze) beschrieben, wodurch wir ein mathematisches Modell des Systems erhalten. Im Gegensatz dazu werden z.B. digitale Filter durch einen Algorithmus beschrieben, der auf einem (Signal)-Prozessor ausgeführt oder in Hardware realisiert werden kann.

Systeme können nach unterschiedlichen Gesichtspunkten eingeteilt werden, von besonderer Bedeutung sind lineare, zeitinvariante (LTI) Systeme. Obwohl Systeme auf unterschiedlichste Art realisiert werden können, haben die unterliegenden mathematischen Modelle sehr viel Gemeinsamkeiten. Im weiteren Verlauf werden zeitkontinuierliche und zeitdiskrete, analoge und digitale LTI-Systeme genauer untersucht.

Kapitel 2

Signale und Zeitfunktionen

Inhalt

2.1 Signale	9
2.1.1 Testsignale	10
2.2 Elementarfunktionen	11
2.2.1 Faltung mit Elementarfunktionen	12
2.3 Sinusfunktion	14
2.3.1 Phasenverschiebung – Zeitverschiebung	16
2.4 Die komplexe Exponentialfunktion	17
2.5 Zusammenfassung	18

2.1 Signale

Ein Signal ist eine Funktion von einer oder mehreren Variablen. Signale tragen Information, daher interessieren wir uns für Signale. Beispiele für Signale sind biometrische Signale, Radarsignale, seismische Signale, Temperatursignale, Videosignale, . . . Signale spielen eine zentrale Rolle bei physikalischen Messungen, in der Unterhaltungselektronik, in der Medizintechnik, in der Telekommunikation, um nur einige Beispiele zu erwähnen. Rechnersysteme kommunizieren mit der Umwelt über Signale.

Signale werden durch Variation physikalischer Größen repräsentiert, diese Größen können verändert, gespeichert und übertragen werden. Ein Signal kann mehrere Formen einer physikalischer Darstellung annehmen. So wird beispielsweise ein akustisches Signal in Form des Schalldrucks durch ein Mikrofon in ein elektrisches Signal umgewandelt. Dieses elektrische Signal kann zur Speicherung wiederum in ein magnetisches Signal bei einer Magnetbandaufzeichnung umgewandelt werden. Es kann aber auch in Form einer binären Zahlenfolge als Muster auf einer CD-ROM abgebildet werden.

Signale sind Eingangsgrößen in signalverarbeitende Systeme, die wiederum – nach Signalverarbeitung – Ausgangssignale erzeugen.

Elektrische (oder genauer elektromagnetische Signale) haben eine besondere Bedeutung in technischen Anwendungen, da sie gegenüber anderen Signalformen, z.B. Schallsignalen, unübertreffliche Vorteile haben: Elektrische Signale breiten sich fast mit Lichtgeschwindigkeit aus und können drahtgebunden oder drahtlos übertragen werden. Sie können relativ einfach erzeugt, verarbeitet und als magnetische Signale gespeichert werden. Elektrische Signale erlauben auch bei kleinsten Energien zuverlässige Signalübertragungen (z.B. Verbindungen zu Raumsonden).

Signale sind Größen, die Information tragen. Sie werden durch Variation physikalischer Größen repräsentiert.

In praktischen Anwendungen sind elektromagnetische Signale von großer Bedeutung.

Anmerkung 5 *Signale (Nachrichten, Informationen) sind an einen physikalischen Träger gebunden, was zur Folge hat, dass Signal-Verarbeitung mit Energieverbrauch verbunden ist.*

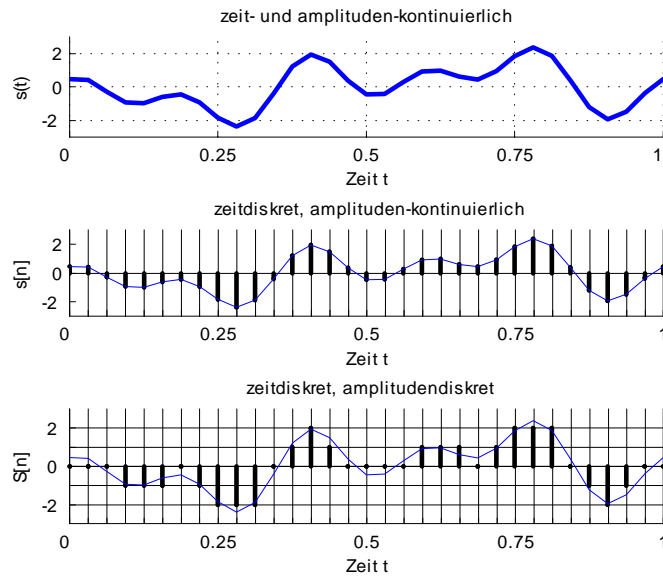


Abbildung 2.1: Analoge und digitale Signale

Nutzsignale müssen von Störsignalen getrennt werden.

Signale treten als Nutz- und Störsignale auf und eine wichtige Aufgabe der Signalverarbeitung ist es, Störsignale von Nutzsignalen zu trennen und die Störsignale zu unterdrücken. Ob ein Signal Nutz- oder Störsignal ist, hängt von der Anwendung ab. Ein »Regensignal« ist ein Nutzsignal für eine regengesteuerte Tempokontrolle, aber ein Störsignal für den Autofahrer.

Signale sind häufig Variationen einer physikalischen Größe in Abhängigkeit von der Zeit $s = f(t)$, Sprachsignale sind ein gutes Beispiel dafür. Derartige Signale sind eindimensionale kontinuierliche Signale, da sie nur von einer Variablen (im Falle des Sprachsignals von der Zeit) abhängen und da der Signalwert (innerhalb physikalischer Grenzen) beliebige Werte annehmen kann. Signale können aber auch eine Funktion des Ortes, z.B. ein Oberflächenprofil eines Werkstücks, sein. Signale können auch mehrdimensional sein, wie man am Beispiel von Bildsignalen (zweidimensionale Ortsfunktion) sehen kann. Wir beschäftigen uns im weiteren Verlauf nur mit eindimensionalen Zeitsignalen.

Die Mehrzahl der Signale ist kontinuierlich. Durch Probenentnahme in ausgewählten Zeitpunkten (Abtastung) wird aus einem kontinuierlichen Signal ein zeitdiskretes Signal. Die Amplitude (= der Signalwert) des zeitdiskreten Signals kann, wie beim kontinuierlichen Signal, beliebige Werte annehmen. Wird ein Signal in einem digitalen Rechner gespeichert oder verarbeitet, müssen die Amplitudenwerte zeitdiskreter Signale durch eine endliche Anzahl an Zahlenwerten repräsentiert werden. Aus dem zeitdiskreten Signal wird ein zeit- und amplitudendiskretes (digitales) Signal.

Zeit- und amplitudenkontinuierliche Signale bezeichnet man als analoge Signale und zeit- und amplituden-diskrete Signale als digitale Signale. Abbildung 2.1 zeigt die Zusammenhänge in graphischer Form. (Zeitkontinuierliche und amplitudendiskrete Signale haben keine technische Bedeutung.)

2.1.1 Testsignale

Systeme werden durch Eingangsgrößen (Ursache, Eingangssignal, Erregung) angeregt und man interessiert sich für die Ausgangsgrößen (Wirkung, Ausgangssignal, Antwort). Die praktisch vorkommenden Signale lassen sich im Allgemeinen nicht als Zeitfunktionen explizit angeben, da sie unregelmäßiger Natur sind (Sprache, Musik, Messwerte, ...). Außerdem wäre es unmöglich, das Verhalten von Systemen für alle praktisch vorkommenden Eingangssignale zu untersuchen. Daher verwendet man bei der Untersuchung von Systemen einfache »Testsignale«, die man so auswählt, dass sie sich gut mathema-

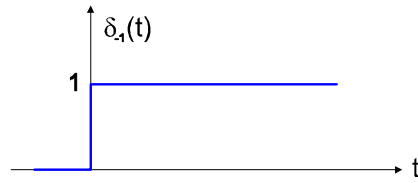


Abbildung 2.2: Sprungfunktion $\delta_{-1}(t)$

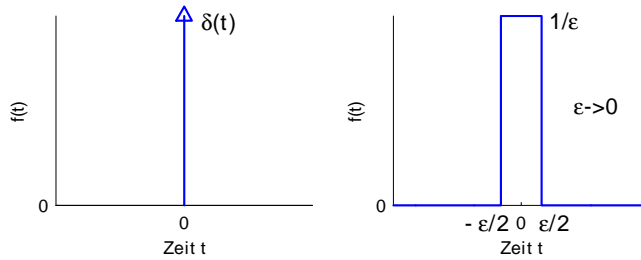


Abbildung 2.3: Darstellung des Einheitsimpulses

tisch darstellen lassen und dass sich die Antwort des Systems auf das Testsignal einfach berechnen lässt. Diese Testsignale werden außerdem so gewählt, dass sich die tatsächlichen Signale aus den Testsignalen zusammensetzen lassen. Bei linearen Systemen kann man das Eingangssignal in eine Linearkombination von Testsignalen zerlegen und die Systemantwort durch Überlagerung der Antworten auf die Testsignale berechnen. Damit sind die Systemantworten auf die Testsignale auch für praktisch vorkommende Signale anwendbar.

2.2 Elementarfunktionen

Eine wichtige Gruppe von Testsignalen sind die Elementarfunktionen. Ist die Antwort eines linearen Systems auf eine Elementarfunktion bekannt, so sind alle Eigenschaften dieses Netzwerks im Prinzip beschrieben.

Die beiden wichtigsten Elementarfunktionen sind die Sprung- und die Impulsfunktion. Die Sprungfunktion δ_{-1} ist folgendermaßen definiert:

$$\delta_{-1}(t) = \begin{cases} 0 & \text{für } t < 0 \\ 1 & \text{für } t \geq 0 \end{cases} \quad (2.1)$$

Abbildung 2.2 zeigt den Verlauf der Sprungfunktion.

Eine weitere wichtige Elementarfunktion ist der Dirac-Impuls oder kurz Impuls, auch Stoß-, Impuls- oder Deltafunktion genannt. Die Impulsfunktion δ_0 ist folgendermaßen definiert:

$$\delta_0(t) = 0 \quad \text{für } t \neq 0 \quad (2.2)$$

$$\int_{-\infty}^{\infty} \delta_0(t) dt = 1$$

Zu beachten ist, dass bei der δ_0 -Funktion lediglich das Integral – die Fläche des Signals – definiert ist. Daher kann man den Einheitsimpuls, wie in Abbildung 2.3 gezeigt, darstellen. Die Breite des Impulses geht gegen Null, die Fläche des Impulses ist 1, die Amplitude geht daher gegen ∞ .

Es können auch andere Kurvenformen (z.B. Dreieckspuls) zur Darstellung des Dirac-Impulses verwendet werden: *Die Kurvenform ist nicht festgelegt*, sondern lediglich die Tatsache, dass die Impulsdauer gegen Null geht und die Fläche 1 ist!

Die **Impulsfunktion** und die **Sprungfunktion** sind die wichtigsten Elementarfunktionen.

Die Impulsfunktion ist im Punkt Null nicht definiert. Es steht nur fest, dass das Integral über diesen Punkt 1 ist.

Anmerkung 6 Die Impulsfunktion ist keine echte Funktion und beschreibt keine eindeutige Funktion. Sie ist überall Null, mit Ausnahme der Stelle Null und auch an diesem Punkt ist sie nicht definiert. Die Bedeutung der Impulsfunktion liegt in ihrer Wirkung auf andere Funktionen, es wird nichts ausgesagt was die Impulsfunktion ist oder wie sie aussieht. Es wird lediglich die Wirkung der Impulsfunktion auf eine andere Funktion - die Signalfunktion $s(t)$ - untersucht.

Allgemein sind die Elementarfunktionen durch folgende Beziehung definiert

$$\delta_i(t) = \int_{-\infty}^t \delta_{i+1}(\tau) d\tau \quad (2.3)$$

Der Zusammenhang zwischen der Sprung- und der Deltafunktion ist daher

$$\delta_{-1}(t) = \int_{-\infty}^t \delta_0(\tau) dt = \begin{cases} 0 & \text{für } t < 0 \\ 1 & \text{für } t \geq 0 \end{cases} \quad (2.4)$$

Die Sprungfunktion ist das Integral der Impulsfunktion. Umgekehrt müsste daher gelten

$$\delta_0(t) = \frac{d}{dt} \delta_{-1}(t) \quad (2.5)$$

Wie man sieht, ist $\delta_{-1}(t)$ an der Stelle $t = 0$ nicht differenzierbar, man kann diese Beziehung daher nur formal gelten lassen. Der Impuls und seine Ableitungen sind im Rahmen der Distributionentheorie exakt definierbar. Für unsere Zwecke – es geht lediglich um die Wirkung der Impulsfunktion (und der Sprungfunktion) auf andere Funktionen – ist die Definition des Impulses ausreichend und wir verwenden auch die Bezeichnung Funktion.

Weitere Elementarfunktionen sind $\delta_1(t)$, der Doppelimpuls an der Stelle $t = 0$, die Rampenfunktion $\delta_{-2}(t)$, eine für $t \geq 0$ linear ansteigende Funktion oder $\delta_{-3}(t)$, eine Parabel für $t \geq 0$. Diese Elementarfunktionen haben aber keine Bedeutung für die Netzwerktheorie.

2.2.1 Faltung mit Elementarfunktionen

Wie erwähnt, interessiert uns lediglich die Wirkung der Impulsfunktion (und Sprungfunktion) auf andere Zeitfunktionen und hier vor allem die Faltung der Elementarfunktion mit anderen Zeitfunktionen. Die Faltung zweier Funktionen ist folgendermaßen definiert

$$g_1(t) * g_2(t) = \int_{-\infty}^{\infty} g_1(\tau) g_2(t - \tau) d\tau = \int_{-\infty}^{\infty} g_1(t - \tau) g_2(\tau) d\tau \quad (2.6)$$

Abbildung ?? veranschaulicht den Faltungsvorgang.

1. Die zu faltenden Zeitfunktionen $g_1(t)$ und $g_2(t)$ werden als Funktionen von τ dargestellt.
2. t ist die unabhängige Variable im Faltungsintegral, daher muss für die Zeitfunktionen $g_1(t)$ und $g_2(t)$ die neue Variable τ eingeführt werden.
3. Wir stellen die Funktionen zuerst für $t = 0$ dar und erhalten $g_1(-\tau)$ und $g_2(\tau)$. Die Funktion $g_1(-\tau)$ ist die entlang der vertikalen Achse gespiegelte (gefaltete) Funktion $g_1(\tau)$, daher der Name für diese Operation.
4. t nimmt die Werte von $-\infty$ bis ∞ an, was nichts anderes bedeutet, dass die Funktion $g_1(-\tau)$ entlang der τ -Achse verschoben wird $\rightarrow g_1(t - \tau)$. Zu jedem Zeitpunkt t wird das Produkt $g_1(t - \tau)g_2(\tau)$ und das Integral dieses Produkts von $-\infty$ bis ∞ liefert den Wert der Faltung zum Zeitpunkt t .

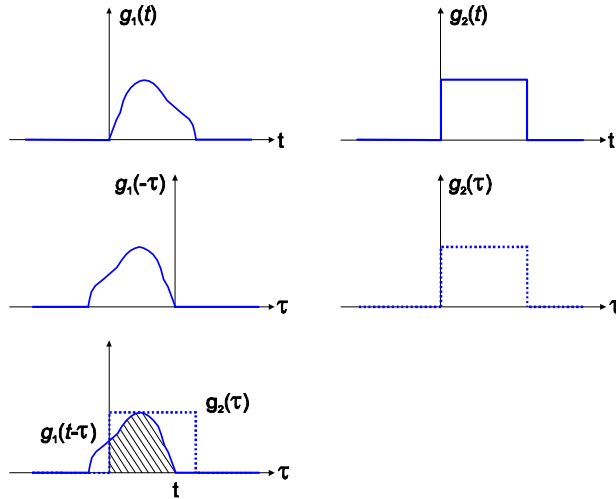


Abbildung 2.4: Veranschaulichung des Faltungsintegrals

Da wir nur Zeitfunktionen betrachten, die für $t < 0$ Null sind, kann das Integrationsintervall von 0^- bis t^+ erstreckt werden. Faltet man eine Elementarfunktion $\delta_i(t)$ mit einer Zeitfunktion $g(t)$, die für $t \geq 0$ existiert, so erhalten wir

$$g(t) * \delta_i(t) = \int_{0^-}^{t^+} g(\tau) \delta_i(t - \tau) d\tau = \int_{0^-}^{t^+} g(t - \tau) \delta_i(\tau) d\tau = g^i(t) \quad (2.7)$$

Das Ergebnis ist der i -te Differentialquotient der Zeitfunktion $g(t)$. Im Speziellen wird also

$$\text{Impuls: } i = 0 \quad g(t) * \delta_0(t) = g(t) \quad (2.8)$$

$$\text{Sprung: } i = -1 \quad g(t) * \delta_{-1}(t) = \int_{0^-}^{t^+} g(\tau) d\tau \quad (2.9)$$

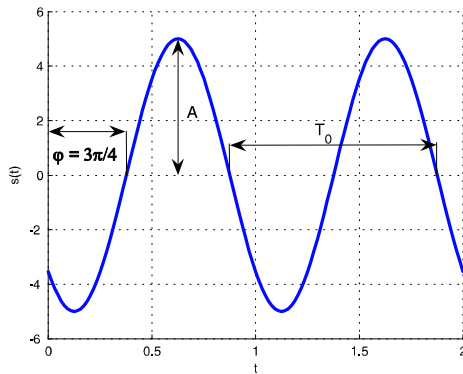
Die Impulsfunktion $\delta_0(t)$ hat die Eigenschaft einer Abtastfunktion, da die Faltung den Funktionswert an der Stelle t liefert, also der Zeitfunktion $g(t)$ an der Stelle t eine Probe entnimmt. $\delta_{-1}(t)$ legt die obere Integrationsgrenze fest, da die Faltung das Integral von $g(t)$ im Intervall 0^- bis t^+ ergibt. Damit haben wir die Möglichkeit, beliebige Zeitfunktionen aus Sprung oder Impuls (oder anderen Elementarfunktionen) zusammensetzen. Für Impuls und Sprung folgt

$$\text{Impuls: } g(t) = g(t) * \delta_0(t) = \int_{0^-}^{t^+} g(\tau) \delta_0(t - \tau) d\tau \quad (2.10)$$

$$\text{Sprung: } g(t) = \dot{g}(t) * \delta_{-1}(t) = \int_{0^-}^{t^+} \dot{g}(\tau) \delta_{-1}(t - \tau) d\tau \quad (2.11)$$

Der Punkt bedeutet die Ableitung nach der Zeit, $\dot{g}(t) = \frac{d}{dt} g(t)$

Gleichung (2.10) bedeutet die Zusammensetzung der Zeitfunktion $g(t)$ aus lauter Impulsen, die mit dem Gewicht $g(t)$ auftreten, während Gleichung (2.11) die Zusammensetzung der Zeitfunktion $g(t)$ aus lauter Sprüngen, die mit dem Gewicht $\dot{g}(t)$ auftreten, bedeutet.

Abbildung 2.5: $s(t) = A \sin\left(\frac{2\pi}{T_0}t - \frac{3\pi}{4}\right)$

2.3 Sinusfunktion

Eine weitere wichtige Aufbaufunktion ist die Sinusfunktion, da sich Signale aus Sinusschwingungen zusammensetzen und mathematisch darstellen lassen. Da wir Signale untersuchen wollen und mit ihnen rechnen müssen, fassen wir die Eigenschaften und den Umgang mit sinusförmigen Schwingungen zusammen.

$$s(t) = A \sin(\omega_0 t + \varphi) \quad (2.12)$$

$$\omega_0 = 2\pi f_0 = 2\pi \frac{1}{T_0}$$

Eigenschaften der Sinus-	
schwingung:	
Amplitude	A
Periodendauer	$T_0 = \frac{1}{f_0}$
Frequenz	$f_0 = \frac{1}{T_0}$
Kreisfreq.	$\omega_0 = 2\pi f_0$
Phase	φ

A ist die Amplitude, ω_0 die Kreisfrequenz und φ die Phase der Schwingung. f_0 ist die Frequenz der Schwingung (»Wie oft schwingt das System in einer Zeiteinheit?«) und T_0 ist die Periodendauer der Schwingung (»Wie lange dauert eine Schwingung?«). Standardmäßig wird als Zeiteinheit eine Sekunde gewählt. Die Standardeinheit für die Frequenz ist daher $\frac{1}{\text{Sekunde}} = 1\text{Hz}$ und die Standardeinheit für die Periodendauer eine Sekunde.

Anmerkung 7 Die Sinusfunktion ist eine Funktion des Winkels (in Radianen) und nicht der Zeit. Die Periode der Sinusfunktion ist immer 2π . Die Kreisfrequenz ω_0 kann man einfach als eine Konstante ansehen, die den linearen Zusammenhang zwischen Zeit und Winkel, also eine »Winkelgeschwindigkeit«, beschreibt. ω_0 gibt an, wie vielen Radianen eine Zeiteinheit (standardmäßig eine Sekunde) entspricht. Anders formuliert: um wie viele Radianen sich der Winkel ändert, wenn die Zeit um eine Zeiteinheit fortschreitet.

Abbildung 2.5 zeigt die graphische Darstellung einer Sinusfunktion.

Sinus- und Kosinus-Funktion sind verwandt, ob man die Darstellung als Sinus oder Kosinus wählt, hängt von praktischen Überlegungen ab¹. Es gelten folgende Beziehungen:

$$\text{äquivalent} \quad \sin \varphi = \cos(\varphi - \pi/2), \quad \cos \varphi = \sin(\varphi + \pi/2) \quad (2.13)$$

$$\text{gerade} \quad \cos \varphi = \cos(-\varphi) \quad (2.14)$$

$$\text{ungerade} \quad \sin \varphi = -\sin(-\varphi) \quad (2.15)$$

$$\text{periodisch} \quad \cos \varphi = \cos(\varphi + 2\pi k), \quad k \dots \text{ganzzahlig} \quad (2.16)$$

$$\text{Ableitung} \quad \frac{d}{d\varphi} \sin \varphi = \cos \varphi, \quad \frac{d}{d\varphi} \cos \varphi = -\sin \varphi \quad (2.17)$$

¹In der Elektrotechnik und Signalverarbeitung wählt man meistens die Darstellung in Form des Kosinus, da sich damit auch der Gleichanteil (Mittelwert), als Komponente der Frequenz $\omega = 0$, darstellen lässt, $A \cos(\omega t)|_{\omega=0} = A$.

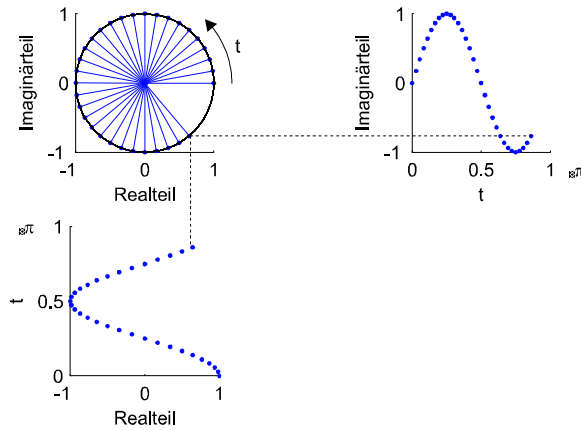


Abbildung 2.6: Komplexe Darstellung der Sinusfunktion

Weitere Zusammenhänge zwischen Sinus und Kosinus:

$$\sin^2 \varphi + \cos^2 \varphi = 1 \tag{2.18}$$

$$\sin 2\varphi = 2 \sin \varphi \cos \varphi \tag{2.19}$$

$$\cos 2\varphi = \cos^2 \varphi - \sin^2 \varphi \tag{2.20}$$

$$\sin(\alpha \pm \beta) = \sin \alpha \cos \beta \pm \cos \alpha \sin \beta \tag{2.21}$$

$$\cos(\alpha \pm \beta) = \cos \alpha \cos \beta \mp \sin \alpha \sin \beta \tag{2.22}$$

Bei der Darstellung von sinusförmigen Schwingungen sind zwei Schreibweisen gebräuchlich

$$s(t) = A \sin(\omega_0 t + \varphi) = A \sin(2\pi f_0 t + \varphi) \tag{2.23}$$

Die Darstellung unter Verwendung der Kreisfrequenz – $\sin(\omega_0 t)$ – ist weniger anschaulich als die Darstellung mit der natürlichen Frequenz – $\sin(2\pi f_0 t)$. Das Beispiel $\sin(12.57t) = \sin(2\pi 2t)$ macht das deutlich: Frequenz $f = 2\text{Hz}$ ist anschaulicher als Kreisfrequenz $\omega = 12.57\text{s}^{-1}$.

Im weiteren Verlauf werden wir häufig mit Sinus- und Kosinus-Schwingungen rechnen müssen. Da das Rechnen mit Sinus und Kosinus in reeller Darstellung aufwändig ist, setzen wir häufig die komplexe Darstellung ein, die das Rechnen wesentlich erleichtert. In der Folge werden wir sowohl die reelle als auch die komplexe Darstellung verwenden.²

Über die Euler'sche Beziehung erhalten wir den Zusammenhang:

$$\vec{s}(t) = A e^{j(\omega_0 t + \varphi)} = A \cos(\omega_0 t + \varphi) + j A \sin(\omega_0 t + \varphi) \tag{2.24}$$

Die komplexe Form (2.24) liefert eine Darstellung von Sinus *und* Kosinus, das reellwertige Signal wird durch Realteilbildung ermittelt.

$$s(t) = \text{Re}\{A e^{j(\omega_0 t + \varphi)}\} = A \cos(\omega_0 t + \varphi) \tag{2.25}$$

Die Erzeugung von Sinus und Kosinus kann man sich als Projektion eines rotierenden Zeigers vorstellen, wie in Abbildung 2.6 dargestellt. Der Kosinus wird durch Projektion auf die horizontale (reelle) Achse, der Sinus durch Projektion auf die vertikale (imaginäre) Achse erzeugt. Der Zeiger dreht sich mit der Winkelgeschwindigkeit ω_0 , die Position des Zeigers (der Winkel) zum Zeitpunkt t errechnet sich aus $\omega_0 t$ (vgl. Anmerkung 7). Hat der Zeiger eine Winkelgeschwindigkeit von 2π pro Sekunde, bedeutet das, dass sich der

Die Sinus- und die Kosinusfunktion kann man auch in **komplexer Form** anschreiben.

Die Sinus- und Kosinusfunktion kann man sich als **Projektion eines Rotierenden Zeigers** auf die vertikale und die horizontale Achse vorstellen

²Für reellwertige Signale schreiben wir $s(t)$, für komplexwertige schreiben wir $\vec{s}(t)$.

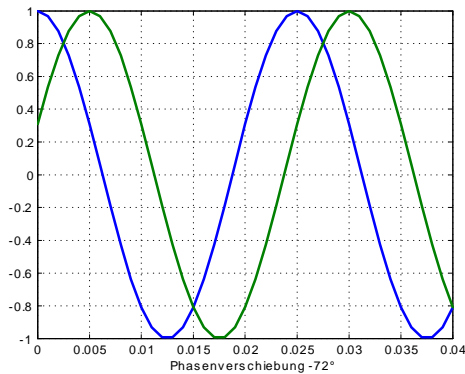


Abbildung 2.7: Phasenverschiebung Kosinus

Zeiger einmal pro Sekunde um den Kreis dreht ($360^\circ = 2\pi$). Der Winkelgeschwindigkeit von $2\pi \text{ s}^{-1}$ entspricht die Frequenz 1, es ist daher $\omega = 2\pi f$.

Bei der Rechnung mit komplexen Exponentialfunktionen gelten die im Anhang zusammengefassten Regeln der komplexen Rechnung. Die geometrische Interpretation der komplexen Rechnung liefert eine sehr anschauliche Deutung der Rechenoperationen.

2.3.1 Phasenverschiebung – Zeitverschiebung

Verschiebung im Zeitbereich

Zur Darstellung der Zeitverschiebung gehen wir vom Beispiel in Abbildung ?? aus. Eine Verschiebung nach rechts $x(t - t_1)$ erkennt man am negativen Vorzeichen, eine Verschiebung nach links $x(t + t_1)$ am positiven Vorzeichen.

Phasenverschiebung

Frequenz und Phase bestimmen die Lage der Maxima und Minima der sinusoidalen Schwingung. Die Kosinusschwingung $x(t) = A \cos(\omega_0 t + \varphi)$ hat für $\varphi = 0$ ein Maximum bei $t = 0$. Für $\varphi \neq 0$ bestimmt die Phasenverschiebung um wie viel das Maximum aus der Lage $t = 0$ verschoben ist.

Tritt das Maximum bei positivem t auf, sprechen wir von einer positiven Zeitverschiebung gegen den Kosinus mit der Phase Null.

Die Zeitverschiebung des Kosinussignals kann dadurch gefunden werden, dass das zu $t = 0$ nächstgelegene Maximum bestimmt wird. In unserem Beispiel tritt das nächstgelegene Maximum bei $t_1 = 0.005$ auf.

Der Abbildung 2.7 können wir entnehmen, dass die Periodendauer $T_0 = 0.025$ aus der wir die Frequenz ausrechnen können $f_0 = \frac{1}{T_0} = 40$. Das Maximum des Kosinus tritt auf, wenn das Argument des Kosinus Null ist, d.h. bei $\omega_0 t_1 + \varphi = 0 \Rightarrow \varphi = -\omega_0 t_1$; $\varphi = -2\pi \cdot 40 \cdot 0.005 = -1.2566 = -72^\circ$

Die Zeitverschiebung gegen den unverschobenen Kosinus $x(t) = A \cos(\omega_0 t)$ können wir über folgende Beziehung in eine Phasenverschiebung umrechnen.

$$x_0(t - t_1) = A \cos[\omega_0(t - t_1)] = A \cos[\omega_0 t - \varphi]$$

Gleichheit des obigen Ausdrucks ist nur gegeben, wenn

$$-\omega_0 t_1 = \varphi \Rightarrow t_1 = -\frac{\varphi}{\omega_0} = -\frac{\varphi}{2\pi f_0}$$

Bei positiver Zeitverschiebung ist die Phasenverschiebung negativ.

$$\varphi = -2\pi f_0 t_1 = -2\pi \frac{t_1}{T_0}$$

Das zu $t = 0$ nächstgelegene Maximum muss immer innerhalb von $|t_1| \leq \frac{T_0}{2}$ liegen, die Phase kann daher $-\pi < \varphi \leq \pi$.

Die Phase wird bestimmt, indem man das zu $t = 0$ nächstgelegene Maximum sucht und mit Hilfe von $\varphi = -2\pi \frac{t_1}{T_0}$ umrechnet.

2.4 Die komplexe Exponentialfunktion

Bei der Darstellung der Sinusfunktion haben wir uns der komplexen Schreibweise $\vec{s}(t) = Ae^{j(\omega_0 t + \varphi)}$ bedient. Diese Notation hat den entscheidenden Vorteil, dass man damit leicht Rechnen kann. Die Erregung von Netzwerken (Siehe Kapitel 6) lässt sich durch die Einführung der komplexen Exponentialfunktion weiter verallgemeinern.

Die komplexe Exponentialfunktion kann man sich wieder als einen rotierenden Zeiger in der komplexen Ebene vorstellen, dessen Länge sich aber (zusätzlich zum vorigen Fall) exponentiell mit der Zeit verändert.

Die komplexe Exponentialfunktion ist wie folgt definiert:

$$\vec{s}(t) = \vec{K} e^{st} \quad \vec{K} = Ae^{j\varphi} \quad ; \quad s = \sigma + j\omega \quad ; \quad A \geq 0 \quad (2.26)$$

\vec{K} bezeichnet man als komplexe Amplitude. Sie beinhaltet nicht nur die Information über die (Anfangs-)Länge des Zeigers A , sondern auch über seine Position (= Phase = Winkel) φ zum Zeitpunkt $t = 0$.

s nennt man komplexe Frequenz. Sie enthält einerseits die Information über die Winkelgeschwindigkeit ω des Zeigers, andererseits beschreibt die Konstante σ die Veränderung der Länge des Zeigers.

Die »physikalischen« Zeitfunktionen erhält man aus der komplexen Exponentialfunktion durch Bildung des Real- oder Imaginärteils bzw. durch Addition zweier konjugiert komplexer Funktionen.

$$\operatorname{Re} \left\{ \vec{K} e^{st} \right\} = \operatorname{Re} \left\{ Ae^{j\varphi} \cdot e^{(\sigma + j\omega)t} \right\} \quad (2.27)$$

$$= \operatorname{Re} \left\{ A \cdot e^{\sigma t} \cdot e^{j(\omega t + \varphi)} \right\} \quad (2.28)$$

$$= A \cdot e^{\sigma t} \cdot \operatorname{Re} \left\{ e^{j(\omega t + \varphi)} \right\} \quad (2.29)$$

$$= A \cdot e^{\sigma t} \cdot \cos(\omega_0 t + \varphi) \quad (2.30)$$

$$\operatorname{Im} \left\{ \vec{K} e^{st} \right\} = \dots = A \cdot e^{\sigma t} \cdot \sin(\omega_0 t + \varphi) \quad (2.31)$$

bzw.

$$\operatorname{Re} \left\{ \vec{K} e^{st} \right\} = \frac{1}{2} [\vec{s}(t) + \vec{s}^*(t)] = \frac{|\vec{K}|}{2} e^{\sigma t} [e^{j(\omega t + \varphi)} + e^{-j(\omega t + \varphi)}] \quad (2.32)$$

$$= |\vec{K}| e^{\sigma t} \cos(\omega t + \varphi) \quad (2.33)$$

Anmerkung 8 $|\vec{K}| = |A \cdot e^{j\varphi}| = A$, weil $|e^{j\varphi}| = 1$ und $A \geq 0$

Abbildung 2.8 zeigt den Realteil der komplexen Exponentialfunktion in der komplexen Ebene $[\sigma, j\omega]$.

Die Bezeichnung komplexe Frequenz für $s = \sigma + j\omega$ ist streng genommen falsch, da im physikalischen Sinn nur der Imaginärteil ω eine (Kreis)Frequenz ist. Dennoch ist diese Bezeichnung üblich, da die Verwendung der komplexen Frequenz außerordentlich zweckmäßig ist.

Die Einführung der komplexen Exponentialfunktion führt zu abstrakten Darstellungen von Signalen und in der Folge zu abstrakten Darstellungen der Eigenschaften von Systemen. Diese Darstellung hat aber entscheidende Vorteile bei der Analyse und Darstellung der Eigenschaften von Systemen.

Die **komplexe Exponentialfunktion** beschreibt einen rotierenden Zeiger, der seine Länge verändert.

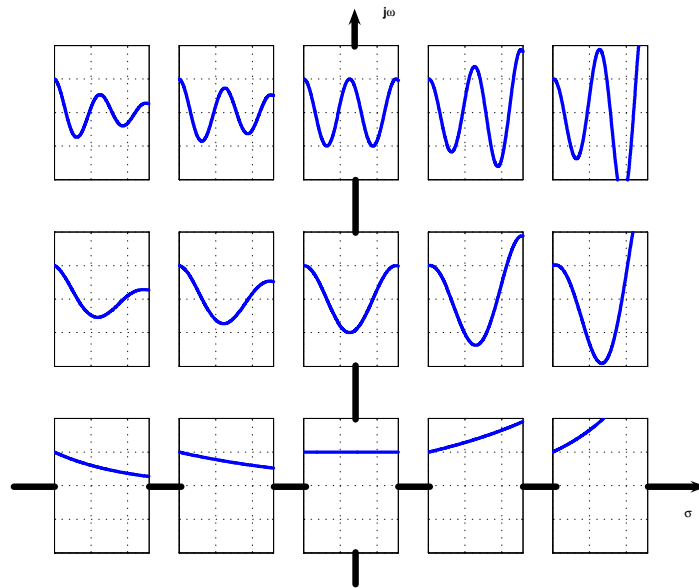


Abbildung 2.8: Realteil der komplexen Exponentialfunktion

2.5 Zusammenfassung

Um Signale (und Systeme) mathematisch erfassen zu können, setzen wir komplizierte Signale aus einfachen Aufbaufunktionen zusammen. Für diese einfachen Aufbaufunktionen ermitteln wir das Systemverhalten. Das Systemverhalten für kompliziertere Signale lässt sich bei linearen Netzwerken durch Überlagerung der Systemantworten auf die Aufbaufunktionen ermitteln.

Aufbaufunktionen sind die Impuls- und Sprungfunktion, die Sinusfunktion (reell oder komplex) und die komplexe Exponentialfunktion.

Kapitel 3

Spektren

Inhalt

3.1	Mathematische Darstellung von zeitkontinuierlichen Signalen	19
3.2	Spektralsynthese	19
3.3	Spektralanalyse	24
3.3.1	Fourierreihen	24
3.3.2	Numerische Berechnung der Fourierkoeffizienten	30
3.3.3	Die Fouriertransformation	34
3.3.4	Eigenschaften der Fouriertransformation	37
3.3.5	Operationen mit Signalen im Zeit- und Frequenzbereich	38
3.3.6	Frequenz und Zeit – Unschärfeprinzip	40
3.3.7	Numerische Berechnung der Fouriertransformation	42
3.4	Größe von Signalen (Energie und Leistung)	43
3.5	Parseval’sches Theorem	43
3.6	Zusammenfassung	44

3.1 Mathematische Darstellung von zeitkontinuierlichen Signalen

Signale haben unterschiedliches Zeitverhalten und es gibt unendlich viele Signale. Es stellt sich daher die Frage, ob es möglich ist, Signale aus einfachen Grundsignalen zusammenzusetzen und damit die Untersuchung der Eigenschaften von Signalen und in der Folge die Bearbeitung von Signalen auf eine einfachere Basis zu stellen.

Das ist tatsächlich der Fall, wir erklären die Zusammenhänge am Beispiel von akustischen Signalen. Ein akustisches Signal kann man sich aus Tönen zusammengesetzt vorstellen. Töne wiederum lassen sich mit Hilfe von Sinusschwingungen beschreiben. Beim Wählen der Telefonnummer 58801, erzeugt die Tastatur zwei Töne pro Nummer und zwar: $5 \Rightarrow 770 \text{ Hz}$ und 1336 Hz , $8 \Rightarrow 852 \text{ Hz}$ und 1336 Hz , $8 \Rightarrow 852 \text{ Hz}$ und 1336 Hz , $0 \Rightarrow 941 \text{ Hz}$ und 1336 Hz , $1 \Rightarrow 697 \text{ Hz}$ und 1209 Hz . Das Wählverfahren wird daher dual tone multifrequency (DTMF) genannt. Signale lassen sich also aus Tönen zusammensetzen (synthetisieren), elektronische Musikinstrumente machen z.B. von dieser Möglichkeit Gebrauch.

Abbildung 3.1 zeigt ein aus mehreren Sinusschwingungen zusammengesetztes Signal.

Komplizierte Signale kann man aus einfacheren Grundsignalen aufbauen und so die Untersuchung der Eigenschaften von Signalen vereinfachen.

3.2 Spektralsynthese

Komplizierte Signale lassen sich aus einfacheren Signalen aufbauen. Setzen wir Signale aus sinusoidalen Schwingungen zusammen, dann erhalten wir

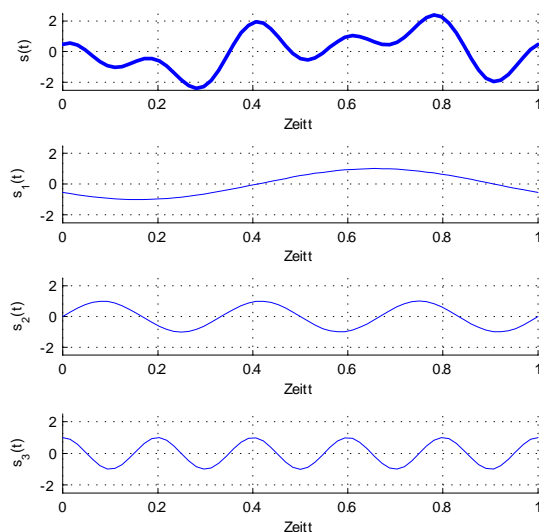


Abbildung 3.1: Summe von Sinusschwingungen $s(t) = s_1(t) + s_2(t) + s_3(t)$

$$\begin{aligned}
 s(t) &= A_0 + A_1 \cos(\omega_1 t + \varphi_1) + \dots + A_K \cos(\omega_K t + \varphi_K) = & (3.1) \\
 &= A_0 + \sum_{k=1}^N A_k \cos(\omega_k t + \varphi_k) = \sum_{k=0}^N A_k \cos(\omega_k t + \varphi_k)
 \end{aligned}$$

Benutzt man sinusoidale Schwingungen als Aufbau-signale, braucht man sich nur ihre Frequenzen, Amplituden und Phasen zu merken.

Signale mit endlicher Zeitdauer können durch eine diskrete Summe von sinusoidalen Schwingungen nicht dargestellt werden.

Da der Verlauf der Sinusschwingungen bekannt ist, braucht er nicht dargestellt zu werden. Es ist ausreichend, wenn man nur die Amplituden A_k und die Phasen φ_k über der Frequenz aufträgt. Man nennt diese Darstellung Amplituden- und Phasenspektrum. Man spricht auch von der Darstellung des Signals im Zeitbereich und im Frequenzbereich.

Die Darstellung von Signalen durch sinusoidale Schwingungen¹ setzt voraus, dass sich die Signale – ebenso wie die Sinusschwingungen – von $-\infty < t < \infty$ erstrecken. Die in der Wirklichkeit auftretenden Signale haben nur eine endliche Dauer und können nicht durch eine endliche Anzahl von Frequenzkomponenten (= Spektralkomponenten = Sinusoidale Schwingungen = Summanden in Gl. 3.1) dargestellt werden. Signale endlicher Dauer werden mathematisch mit Hilfe der Fouriertransformation dargestellt.

Abbildung 3.2 zeigt die Sicht auf ein Signal im Zeit- und im Frequenzbereich. Die Ansicht von links zeigt eine Sägezahn-schwingung im Zeitbereich $s(t)$. Die Ansicht von vorne zeigt die einzelnen Spektralkomponenten - die Darstellung im Frequenzbereich $S(\omega)$ oder $S(f)$ ². (Der Gleichanteil – die 0.te Harmonische – wurde in Abbildung 3.2 weggelassen, damit die Zeichnung übersichtlicher ist.)

Abbildung 3.4 zeigt das Betragsspektrum einer Sägezahn-schwingung. Die einzelnen Spektralkomponenten werden als Linien dargestellt, deren Länge der Amplitude der Schwingung entspricht. Man spricht daher auch vom Linienspektrum. Eine Verdopplung der Frequenz entspricht einer Oktave in der Musik.

Die Antwort eines LTI-Systems auf ein Eingangssignal kann man ermitteln, indem man das Eingangssignal in Frequenzkomponenten zerlegt und die Systemantworten auf diese Komponenten einzeln ermittelt. Die Systemantwort auf das ursprüngliche Eingangssignal setzt sich dann aus den Antworten auf die einzelnen Frequenzkomponenten zusammen. Man darf aber nicht vergessen, dass dieses Prinzip nur bei *linearen* Systemen anwendbar ist, wovon wir uns anhand des folgenden Beispiels überzeugen können.

¹Es können endlich oder unendlich viele Spektrallinien auftreten.

² $\omega = 2\pi f$, vgl. Abschnitt 2.3)

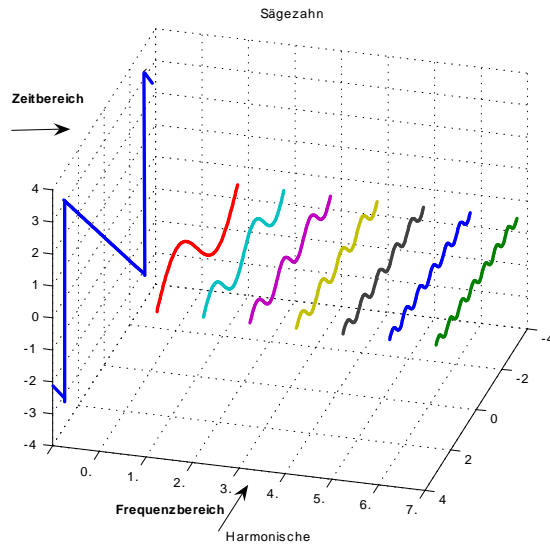


Abbildung 3.2: Signal im Zeit- und Frequenzbereich

Beispiel 9 (Nichtlineares System) Wir zeigen das Verhalten des Systems $x_A = x_E^2$, wenn zwei Eingangssignale anliegen. Die Eingangssignale nennen wir Basis-signal $A \cos(\omega t)$ und Trägersignal $B \cos(\Omega t)$

$$\begin{aligned}
 [A \cos(\omega t) + B \cos(\Omega t)]^2 &= A^2 \cos^2(\omega t) + 2A \cos(\omega t)B \cos(\Omega t) + B^2 \cos^2(\Omega t) = & (3.2) \\
 &= \frac{A^2}{2} [1 + \cos(2\omega t)] + 2AB \cos(\omega t) \cos(\Omega t) + \frac{B^2}{2} [1 + \cos(2\Omega t)] & (3.3)
 \end{aligned}$$

Wie wir sehen, durchlaufen die beiden Eingangssignale das (nichtlineare) quadratische Systems nicht ohne Interaktion. Es entsteht eine Komponente der doppelten Basisfrequenz ω , eine Komponente der doppelten Trägerfrequenz Ω und das Produkt aus Basis-signal und Trägersignal! Für dieses Produkt erhalten wir

$$\cos(\Omega t) \cos(\omega t) = \frac{1}{2} [\cos(\Omega - \omega)t + \cos(\Omega + \omega)t] \tag{3.4}$$

Es entstehen also am Ausgang des quadratischen Systems Signale mit den Frequenzen $[0, 2\omega, \Omega - \omega, \Omega + \omega, 2\Omega]$, **Signale mit den Frequenzen Ω und ω sind nicht mehr enthalten!** Abbildung 3.3 zeigt die entstehenden Spektrallinien. Bei der reellen Rechnung muss man die trigonometrischen Formeln kennen. Bei Verwendung der komplexen Rechnung benötigt man lediglich die Euler'sche Beziehung und die Regeln der Potenzrechnung, was den Rechenvorgang erheblich vereinfacht.

$$\begin{aligned}
 (\cos \omega t + 3 \cos \Omega t)^2 &= \left(\frac{1}{2}e^{j\omega t} + \frac{1}{2}e^{-j\omega t} + \frac{3}{2}e^{j\Omega t} + \frac{3}{2}e^{-j\Omega t}\right)^2 = \\
 &= \left(\frac{1}{2}e^{j\omega t} + \frac{1}{2}e^{-j\omega t}\right)^2 + 2\left(\frac{1}{2}e^{j\omega t} + \frac{1}{2}e^{-j\omega t}\right)\left(\frac{3}{2}e^{j\Omega t} + \frac{3}{2}e^{-j\Omega t}\right) + \left(\frac{3}{2}e^{j\Omega t} + \frac{3}{2}e^{-j\Omega t}\right)^2 \\
 \hline
 &= \left(\frac{1}{2}\right)^2 e^{j2\omega t} + 2\left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 e^{-j2\omega t} = 2\left(\frac{1}{2}\right)^2 + 2\left(\frac{1}{2}\right)^2 \cos 2\omega t = 2\left(\frac{1}{2}\right)^2 + \frac{1}{2} \cos 2\omega t \\
 &= 2\left(\frac{1}{2}\frac{3}{2}e^{j(\Omega+\omega)t} + \frac{1}{2}\frac{3}{2}e^{j(\Omega-\omega)t} + \frac{1}{2}\frac{3}{2}e^{j(-\Omega+\omega)t} + \frac{1}{2}\frac{3}{2}e^{j(-\Omega-\omega)t}\right) = \\
 &= 2 \times 2\frac{1}{2}\frac{3}{2} [\cos(\Omega + \omega)t + \cos(\Omega - \omega)t] = 3 [\cos(\Omega + \omega)t + \cos(\Omega - \omega)t] \\
 &= \left(\frac{3}{2}\right)^2 e^{j2\Omega t} + \left(\frac{3}{2}\right)^2 e^{-j2\Omega t} = 2\left(\frac{3}{2}\right)^2 + 2\left(\frac{3}{2}\right)^2 \cos 2\Omega t = 2\left(\frac{3}{2}\right)^2 + 4.5 \cos 2\Omega t \\
 &= 2\left(\frac{1}{2}\right)^2 + 2\left(\frac{3}{2}\right)^2 = 5
 \end{aligned}$$

Wie man leicht sieht, ist bei diesem Beispiel der Überlagerungssatz verletzt, da $[A \cos(\omega t)]^2 + [B \cos(\Omega t)]^2 \neq [A \cos(\omega t) + B \cos(\Omega t)]^2$.

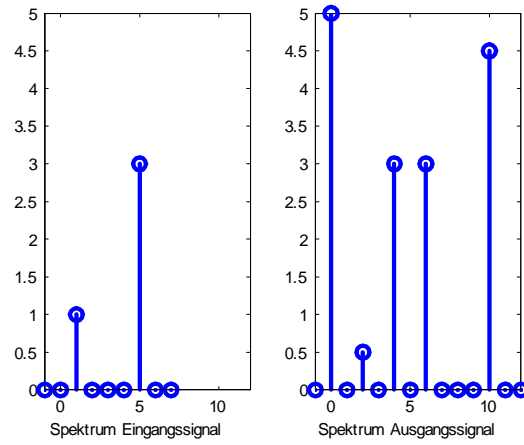


Abbildung 3.3: Spektrallinien bei quadratischem System

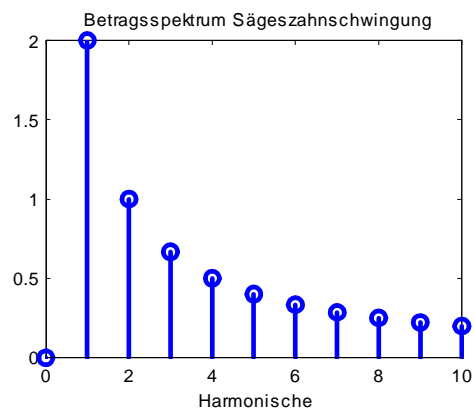


Abbildung 3.4: Spektraldarstellung einer Sägeszahnfunktion

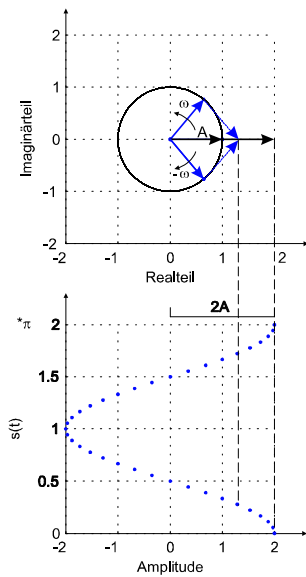


Abbildung 3.5: Zeiger mit entgegengesetzter Drehrichtung ω und $-\omega$

Anmerkung 10 Zeit- und Frequenzbereich-Darstellung sind gleichwertig und lediglich andere Betrachtensweisen für dasselbe Signal, wobei sich manche Signaleigenschaften besser im Zeitbereich, andere wiederum besser im Frequenzbereich darstellen lassen.

Aus sinusoidalen Komponenten zusammengesetzte Signale können auch in komplexer Schreibweise dargestellt werden

$$\vec{s}(t) = X_0 + \sum_{n=1}^N A_n e^{j(\omega_n t + \varphi_n)} = X_0 + \sum_{n=1}^N \vec{X}_n e^{j\omega_n t} \quad \vec{X}_n = A_n e^{j\varphi_n} \quad (3.5)$$

Die Größe $\vec{X}_n = A_n e^{j\varphi_n}$ ist die komplexe Amplitude. Die komplexe Amplitude ist eine mathematische Konstruktion, mit Hilfe der die Phasenverschiebung des Zeigers in die Amplitude verlagert wird. Damit ist eine kompaktere Schreibweise möglich, da sich der Ausdruck $e^{j(\omega_n t + \varphi_n)}$ auf $e^{j\omega_n t}$ vereinfacht. Das bedeutet aber nicht, dass sie Phase gleich Null wäre!

Der Kosinus kann mit Hilfe der inversen Euler'schen Form dargestellt werden $\cos(\varphi) = \frac{1}{2}(e^{j\varphi} + e^{-j\varphi})$. Daraus ergibt sich eine weitere Darstellung von sinusoidalen Schwingungen:

$$\begin{aligned} s(t) = A \cos(\omega t + \varphi) &= \frac{A}{2} e^{j(\omega t + \varphi)} + \frac{A}{2} e^{-j(\omega t + \varphi)} \\ &= \frac{\vec{X}}{2} e^{j\omega t} + \frac{\vec{X}^*}{2} e^{-j\omega t} \end{aligned} \quad (3.6)$$

Gleichung (3.6) kann man sich als zwei mit entgegengesetzter Drehrichtung rotierende Zeiger vorstellen, das (reellwertige) Signal $s(t)$ wird durch Addition der beiden Zeiger gebildet.

Abbildung 3.5 verdeutlicht diesen Zusammenhang. Die Summe der gegensinnig rotierenden Zeiger liegt immer auf der reellen Achse, $s(t)$ ist daher immer reellwertig.

Ein Signal, das aus mehreren sinusoidalen Komponenten zusammengesetzt ist, lässt sich wie folgt darstellen:

$$s(t) = X_0 + \sum_{n=1}^N \left\{ \frac{\vec{X}_n}{2} e^{j\omega_n t} + \frac{\vec{X}_n^*}{2} e^{-j\omega_n t} \right\} \quad (3.7)$$

Die Darstellung eines Signals nach (3.5) nennt man **einseitiges Spektrum**. Die Darstellung eines Signals nach (3.7) nennt man **zweiseitiges Spektrum**.

Die Darstellung des Spektrums unter Verwendung der inversen Euler'schen Formel nennt man zweiseitiges Spektrum (3.7), die Darstellung nach (3.5) nennt man einseitiges Spektrum. Bei der Verwendung des zweiseitigen Spektrums treten negative Frequenzen auf, die, wie sich im weiteren Verlauf zeigen wird, physikalische Bedeutung erlangen können.

Anmerkung 11 *Negative Frequenzen sind irritierend, da ja die Frequenz als Anzahl von Wiederholungen definiert und damit eine positive Zahl sein muss. Wir können negative Frequenzen leicht mit Hilfe folgender Beziehung umformen*

$$\cos(-\omega t + \varphi) = \cos(\omega t - \varphi) \quad (3.8)$$

und erhalten damit eine positive Frequenz. Gleiches gilt auch für die Beziehung

$$e^{\pm j\omega t} = \cos \omega t \pm j \sin \omega t \quad (3.9)$$

auch hier ist die Frequenz positiv.

Negative Frequenzen treten nur in der komplexen Darstellung und immer als Paar mit positiven Frequenzen auf. Wir können sie daher so wie die grafische Darstellung des zweiseitigen Spektrums verstehen (Abbildung 3.5).

3.3 Spektralanalyse

Die **Fourierzerlegung** zerlegt ein Signal in sinusoidale Schwingungen unterschiedlicher Amplitude und Phase.

Wie wir gesehen haben, lassen sich Signale aus sinusoidalen Schwingungen synthetisieren. Häufig ist aber das Signal gegeben und man möchte daraus die Spektralkomponenten berechnen. Diese Aufgabe löst die Fourierzerlegung für uns. Wir geben die Zusammenhänge zunächst für *periodische* kontinuierliche Schwingungen an, werden diese Einschränkung aber im weiteren Verlauf – bei der Fouriertransformation – aufheben.

3.3.1 Fourierreihen

Eine *periodische* Funktion³ mit der Periode $T = \frac{1}{f} = \frac{2\pi}{\omega_0}$ lässt sich durch eine Fourierreihe folgendermaßen darstellen

$$s(t) = \frac{A_0}{2} + A_1 \cos \omega_0 t + A_2 \cos 2\omega_0 t + \dots + A_k \cos k\omega_0 t + \\ + B_1 \sin \omega_0 t + B_2 \sin 2\omega_0 t + \dots + B_k \sin k\omega_0 t \quad (3.10)$$

Die Amplituden der einzelnen Schwingungen berechnet man

$$A_k = \frac{2}{T} \int_0^T s(t) \cos(k\omega_0 t) dt \quad (3.11)$$

$$B_k = \frac{2}{T} \int_0^T s(t) \sin(k\omega_0 t) dt \quad (3.12)$$

³Nicht jede periodische Funktion lässt sich in eine Fourierreihe zerlegen. Es leuchtet ein, dass für die Konvergenz einer Fourierreihe die Amplituden der Teilschwingungen endlich sein müssen. Das ist immer dann gegeben, wenn die periodische Funktion $s(t)$ absolut über eine Periode integrierbar ist, wenn also gilt: $\int_T |s(t)| dt < \infty$. Diese Bedingung ist jedoch nicht hinreichend. $s(t)$ darf darüber hinaus nur eine endliche Zahl von Maxima und Minima innerhalb einer Periode haben und darf nur eine endliche Zahl von endlichen Diskontinuitäten in einer Periode aufweisen. Diese *Dirichlet'schen Bedingungen* sind für alle praktischen Signale erfüllt.

Die Periodendauer der Schwingung ist

$$T = \frac{2\pi}{\omega_0} = \frac{1}{f_0} \quad (3.13)$$

Wir sehen, dass sich $s(t)$ im allgemeinen Fall als eine Summe von Sinus- und Kosinus-schwingungen plus einem konstanten Wert $A_0/2$ darstellen lässt. $A_0/2$ ist der Mittelwert der Funktion $s(t)$ und wird in der Elektrotechnik Gleichglied oder Gleichanteil genannt.

Es können auch abschnittsweise kontinuierliche Signale (z.B. Rechteck- oder Sägezahn-schwingung) in eine Fourierreihe zerlegt werden. In den Sprungstellen konvergiert die Reihe auf den Mittelwert des Wertes links und rechts der Unstetigkeit $[s(t_0^-) + s(t_0^+)]/2$.

Wenn die Funktion $s(t)$ gerade ist, wenn also $s(t) = s(-t)$, dann treten in der Reihenentwicklung nur gerade Anteile auf. Da die Sinusfunktion ungerade ist, werden die Amplituden $B_k = 0$. Wenn die Funktion ungerade ist, wenn also $s(t) = -s(-t)$, dann treten in der Reihenentwicklung nur ungerade Anteile auf. Da die Kosinusfunktion gerade ist, werden die Amplituden $A_k = 0$.

Die Komponenten der Fourierzerlegung bestehen aus der Grundschwingung (der Schwingung mit der Periodendauer $T = 2\pi/\omega_0$) und Oberschwingungen, die immer *ganzzahlige* Vielfache der Grundschwingung sind. Die Oberschwingungen werden auch Harmonische genannt.

In diesem Zusammenhang stellt sich die Frage, ob die Summe von sinusförmigen Schwingungen beliebiger Frequenz eine periodische Schwingung bildet und wie man bei mehreren Sinusschwingungen unterschiedlicher Frequenz die Periodendauer ermittelt. Wir nehmen als Beispiel die Schwingungen

$$\begin{aligned} s_1(t) &= 1 + 2 \sin\left(\frac{1}{2}t + \varphi_1\right) + 3 \sin\left(\frac{1}{3}t + \varphi_2\right) + 4 \sin\left(\frac{1}{5}t + \varphi_3\right) \\ s_2(t) &= 2 \sin(t + \varphi_1) + 3 \sin(\pi t + \varphi_2) \end{aligned}$$

Wir erinnern uns, dass jede Frequenzkomponente in einem periodischen Signal ein *ganzzahliges* Vielfaches der Grundfrequenz ω_0 sein muss. Daher ist das Verhältnis der (Kreis)Frequenzen zweier beliebigen Teilschwingungen m/n , wobei m und n ganze Zahlen sind. Das bedeutet, dass das Verhältnis zweier beliebiger Teilschwingungen eine rationale Zahl sein muss. In unserem Beispiel sind die Kreisfrequenzen von $s_1(t)$: $(\frac{1}{2}, \frac{1}{3}$ und $\frac{1}{5})$ ganzzahlige Vielfache von $\frac{1}{15}$, sind also Harmonische. Die Kreisfrequenzen von $s_2(t)$ stehen im Verhältnis $1/\pi$. Dieses Verhältnis ist nicht rational, $s_2(t)$ ist daher *keine* periodische Schwingung.

Neben der Darstellung von Signalen in Sinus- und Kosinuskomponenten wählt man auch die Darstellung nach Betrag und Phase. Wir rechnen um und erhalten

$$C \cos(\omega_0 t + \varphi) = \underbrace{C \cos \varphi}_{A} \cos \omega_0 t - \underbrace{C \sin \varphi}_{B} \sin \omega_0 t \quad (3.14)$$

Aus $A = C \cos \varphi$ und $B = -C \sin \varphi$ erhalten wir

$$C = \sqrt{A^2 + B^2} \quad \text{und} \quad \varphi = \arctan \frac{-B}{A} \quad (3.15)$$

Damit können wir die Fourierreihe nach Betrag und Phase darstellen

$$s(t) = C_0 + C_1 \cos(\omega_0 t + \varphi_1) + C_2 \cos(2\omega_0 t + \varphi_2) + \dots + C_k \cos(k\omega_0 t + \varphi_k) + \dots \quad (3.16)$$

wobei gilt

$$C_k = \sqrt{A_k^2 + B_k^2}, \quad \varphi_k = \arctan \left(\frac{-B_k}{A_k} \right) \quad \text{und} \quad C_0 = \frac{A_0}{2} \quad (3.17)$$

In der Reihenentwicklung einer geraden Funktion treten nur Kosinusterme, in der Reihenentwicklung einer ungeraden Funktion nur Sinusterme auf.

Eine aus sinusoidalen Komponenten zusammengesetzte Schwingung ist genau dann periodisch, wenn das Verhältnis der (Kreis)Frequenzen der Komponenten rational ist.

Die Darstellung nach Betrag und Winkel nach Gleichung (3.16) könnte man auch in Form der Sinusfunktion⁴ ausdrücken, die Kosinus-Darstellung wird aber meistens bevorzugt, da sich der Gleichanteil als Komponente der Frequenz Null darstellen lässt.

Wie man sieht, können periodische Schwingungen im Reellen durch Sinus- plus Kosinusfunktion (jeweils mit der Phase Null) oder durch eine Sinus- oder Kosinusfunktion mit allgemeiner Phase dargestellt werden.

Schließlich gibt es noch die komplexe Darstellung⁵

$$s(t) = \sum_{k=-\infty}^{+\infty} \vec{D}_k e^{jk\omega_0 t} \quad (3.18)$$

Wie man leicht sehen kann, ergeben sich die Darstellungen im Reellen durch Bildung von Real- und Imaginärteil bzw. durch Bildung von Betrag und Phase aus der komplexen Darstellung.

$$\vec{D}_k = \frac{1}{T} \int_0^T s(t) e^{-jk\omega_0 t} dt = \begin{cases} \frac{A_0}{2} & \text{für } k = 0 \\ \frac{1}{2}(A_k - jB_k) & \text{für } k > 0 \\ \frac{1}{2}(A_{-k} + jB_{-k}) & \text{für } k < 0 \end{cases} \quad (3.19)$$

Die komplexe Spektraldarstellung ist die leistungsfähigste und kompakteste aller Formen und wir werden uns daher im weiteren Verlauf häufig dieser Darstellung bedienen.

Aus der Euler'schen Beziehung können wir die Darstellung nach Betrag und Winkel $C_k \cos(\omega_0 t + \varphi_k)$ einfach in die komplexwertige Darstellung überführen

$$C_k \cos(k\omega_0 t + \varphi_k) = \frac{C_k}{2} \left[e^{j(k\omega_0 t + \varphi_k)} + e^{-j(k\omega_0 t + \varphi_k)} \right] = \quad (3.20)$$

$$= \underbrace{\left(\frac{C_k}{2} e^{j\varphi_k} \right)}_{\vec{D}_k} e^{jk\omega_0 t} + \underbrace{\left(\frac{C_k}{2} e^{-j\varphi_k} \right)}_{\vec{D}_{-k}} e^{-jk\omega_0 t} \quad (3.21)$$

und wir erhalten

$$s(t) = C_0 + \sum_{k=1}^{\infty} C_k \cos(k\omega_0 t + \varphi_k) \quad (3.22)$$

$$s(t) = D_0 + \sum_{k=1}^{\infty} \left(\vec{D}_k e^{jk\omega_0 t} + \vec{D}_{-k} e^{-jk\omega_0 t} \right) \quad (3.23)$$

Der Zusammenhang zwischen der reellen (einseitiges Spektrum) und komplexen (zwei-seitiges Spektrum) Darstellung ist

$$\vec{D}_k = \frac{1}{2} C_k e^{j\varphi_k} \quad C_k = 2 \left| \vec{D}_k \right|; \quad \varphi_k = \angle \vec{D}_k \quad (3.24)$$

Es wird auch stillschweigend angenommen, dass φ_0 gleich Null und D_0 reellwertig ist, weil $s(t)$ ein reellwertiges Signal ist.

Die kompakteste Darstellung der Fourierreihe ist die komplexe Darstellung. Es lässt sich mit ihr auch am einfachsten rechnen.

Die komplexe Darstellung der Fourierreihe entsprechend (3.18) und (3.19) ist die kompakteste Form der Darstellung. Das Rechnen mit Exponentialreihen ist wesentlich einfacher als das Rechnen mit trigonometrischen Reihen. Auch bei der Untersuchung des Verhaltens von Systemen ist die exponentielle Form geeigneter als die trigonometrische. Aus diesem Grund wird der exponentiellen Form der Vorzug gegeben.

Beispiel 12 Als Beispiel berechnen wir die Fourierkoeffizienten einer periodischen Rechteckfunktion nach Abbildung 3.6. Die Berechnung des Gleichglieds (des Mittelwerts) ist

$$^4 s(t) = c_0 + c_1 \sin(\omega_0 t + \varphi_1) + c_2 \sin(2\omega_0 t + \varphi_2) + \dots + c_k \sin(k\omega_0 t + \varphi_k) + \dots$$

wobei gilt $c_k = \sqrt{A_k^2 + B_k^2}$, $\tan \varphi_k = \frac{A_k}{B_k}$ $c_0 = \frac{A_0}{2}$

⁵Die Schreibweise \vec{D} weist in der komplexen Darstellung auf den Zeigercharakter der "komplexen" Amplitude hin.

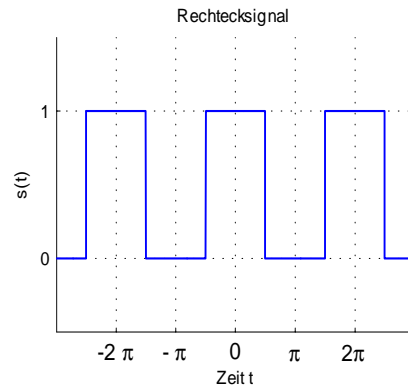


Abbildung 3.6: Rechteckschwingung

einfach: $\frac{A_0}{2} = \frac{1}{2}$. Die Länge der Periode beträgt $T = 2\pi$, die zugehörige Kreisfrequenz ist also $\omega = \frac{2\pi}{T} = 1$

$$A_k = \frac{1}{\pi} \int_{-\pi/2}^{\pi/2} \cos(kt) dt = \frac{2}{k\pi} \sin\left(\frac{k\pi}{2}\right) = \begin{cases} 0 & k \text{ gerade} \\ \frac{2}{\pi k} & k = 1, 5, 9, 13, \dots \\ -\frac{2}{\pi k} & k = 3, 7, 11, 15, \dots \end{cases}$$

$$B_k = \frac{1}{\pi} \int_{-\pi/2}^{\pi/2} \sin(kt) dt = 0$$

Beim Integrieren einer periodischen Zeitfunktion über eine Periode spielt keine Rolle, über welches Zeitintervall integriert wird, solange die Breite des Intervalls der Periodendauer entspricht. Außerdem braucht man nicht über einen Bereich zu integrieren, wo der Integrand sicher Null ist. So kommen die Integrationsgrenzen $-\frac{\pi}{2}$ und $\frac{\pi}{2}$ zustande, obwohl man nach Einsetzen in die Formeln (3.11) bzw. (3.12) vielleicht 0 und 2π erwarten würde.

Für die Reihenentwicklung ergibt sich daher:

$$s(t) = \frac{1}{2} + \frac{2}{\pi} \left(\cos t - \frac{1}{3} \cos 3t + \frac{1}{5} \cos 5t - \frac{1}{7} \cos 7t + \dots \right)$$

$s(t)$ ist eine gerade Funktion, es treten daher in der Fourierreihe nur gerade Aufbaufunktionen, also Kosinusfunktionen auf.

Anmerkung 13 In der Schreibweise $C \cos(\omega t + \varphi)$ ist die Amplitude definitionsgemäß positiv und in der Reihenentwicklung treten daher auch nur positive Amplituden auf. Die auftretenden negativen Vorzeichen der Reihenentwicklung des vorherigen Beispiels kann man leicht durch folgende Identität beseitigen

$$-\cos \alpha = \cos(\alpha - \pi) \tag{3.25}$$

$$\Downarrow$$

$$-\frac{1}{3} \cos 3t = \frac{1}{3} \cos(3t - \pi) \tag{3.26}$$

um die korrekte Darstellung der Reihe zu finden. In vielen Fällen ist man aber lediglich an der Amplitude der Spektrallinie interessiert und akzeptiert daher, dass der Koeffizient negative Werte annimmt.

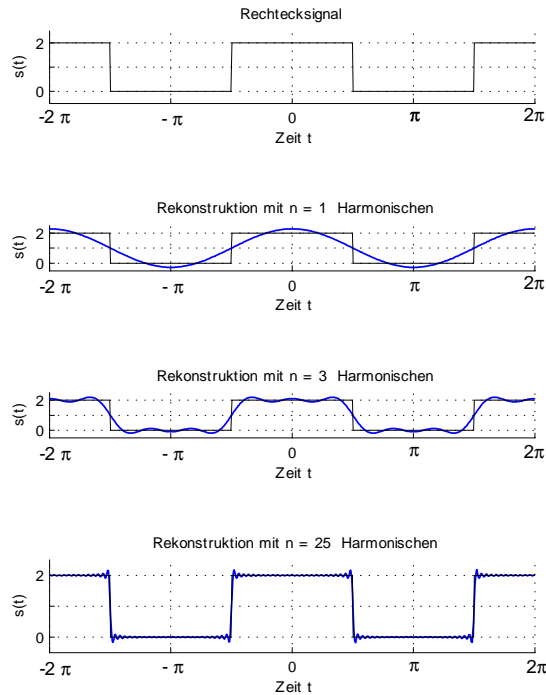
Beispiel 14 Abbildung 14 zeigt die Fourierentwicklung, wenn die Reihe bei $n = 1, 3$ und 25 abgeschnitten wird. Man beobachtet, dass auch für großes n ein Überschwingen an den

Je mehr Elemente der Fourierreihe man summiert, desto besser deckt sich die Summe mit der ursprünglichen Funktion.

Unstetigkeitsstellen auftritt. Die Rechteckschwingung – ein Signal mit »Ecken« – wird durch Kosinusfunktionen – Signale ohne »Ecken« – nachgebildet. Es überrascht daher nicht, dass die Reihendarstellung nicht exakt ist. Die Näherung ist aber für $n \rightarrow \infty$ so gut, dass die Differenz der beiden Darstellungen die Energie Null hat.

$$\int_0^T [f(t) - s_n(t)]^2 dt \rightarrow 0 \text{ für } n \rightarrow \infty.$$

Das Überschwingen verschwindet aber nicht, sondern erreicht für große n einen Wert von 9%. Diese Erscheinung wird Gibbs'sches Phänomen genannt.



Fourierzerlegung für verschiedene n

Beispiel 15 Als nächstes Beispiel betrachten wir die periodische Impulsfunktion mit der Periodendauer T

$$s(t) = \begin{cases} A & 0 < t \leq \tau \\ 0 & \tau < t \leq T \end{cases} \quad (3.27)$$

Wir erhalten die Fourierzerlegung

$$s(t) = A \frac{\tau}{T} + \sum_{k=1}^{\infty} \left[A \frac{\sin(k\omega_0\tau)}{k\pi} \cos(k\omega_0 t) + 2A \frac{\sin^2(k\omega_0\tau/2)}{k\pi} \cos(k\omega_0 t) \right] \quad (3.28)$$

$$\omega_0 = \frac{2\pi}{T} \quad (3.29)$$

Beim Tastverhältnis 1 : 1 (Rechteckschwingung) verschwindet die 2., 4., 6., ... Oberschwingung, beim Verhältnis 1 : 4 verschwindet die 4., 8., 12., ..., beim Verhältnis 1 : 5 die 5., 10., 15., ..., bei 1 : 10 die 10., 20., 30., ... Oberschwingung.

Lässt man $\tau \rightarrow 0$ gehen, dann werden aus den Rechteckimpulsen Diracimpulse $\delta_0(t)$ ($A\tau \equiv 1$) und wir erhalten

$$s_0(t) = \frac{1}{T} + \frac{2}{T} \sum_{k=1}^{\infty} \cos(k\omega_0 t) = \frac{1}{T} \sum_{k=-\infty}^{\infty} e^{jk\omega_0 t} \quad (3.30)$$

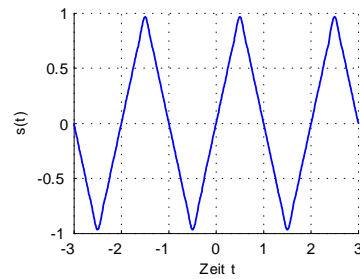


Abbildung 3.7: Dreieckschwingung

Das Signal $s_0(t)$, das aus Diracimpulsen im Abstand T besteht, wird Kammfunktion genannt. Das Spektrum einer Kammfunktion ist wieder eine Kammfunktion und enthält alle ganzzahligen Vielfachen von ω_0 von 0 bis ∞ !

Beispiel 16 Als weiteres Beispiel geben wir die Fourierentwicklung für eine Dreieckschwingung nach Abbildung 3.7 an. In diesem Beispiel ist die Länge der Periode $T = 2$ und daher $\omega = \frac{2\pi}{2} = \pi$. Der Mittelwert von $s(t) = 0$. Die Funktion ist ungerade, es entfallen daher die Kosinusanteile. Das Signal lässt sich darstellen

$$s(t) = \begin{cases} 2t & |t| \leq \frac{1}{2} \\ 2(1-t) & \frac{1}{2} < t \leq \frac{3}{2} \end{cases}$$

$$\begin{aligned} b_k &= \int_{-1/2}^{1/2} 2t \sin(k\pi t) dt + \int_{1/2}^{3/2} 2(1-t) \sin(k\pi t) dt = \\ &= \frac{8}{k^2\pi^2} \sin\left(\frac{k\pi}{2}\right) = \begin{cases} 0 & k \text{ gerade} \\ \frac{8}{k^2\pi^2} & k = 1, 5, 9, 13, \dots \\ -\frac{8}{k^2\pi^2} & k = 3, 7, 11, 15, \dots \end{cases} \end{aligned}$$

$$s(t) = \frac{8}{\pi^2} \left(\sin \pi t - \frac{1}{9} \sin 3\pi t + \frac{1}{25} \sin 5\pi t - \frac{1}{49} \sin 7\pi t + \dots \right)$$

Der Vergleich der Amplitudenspektren der Rechteck- und der Dreieckschwingung zeigt, dass die Amplituden der Oberschwingungen der Dreieckschwingung schneller abklingen als die der Rechteckschwingung. Die Rechteckschwingung hat einen höheren Oberschwingungsanteil, braucht also mehr Frequenzbandbreite, da mehr Schwingungen erforderlich sind, um den steilen Anstieg der Flanken »nachzuzeichnen«.

Anmerkung 17 Rechteckschwingungen treten in digitalen elektronischen Schaltkreisen auf und haben einen hohen Oberschwingungsanteil. Bei hohen Taktfrequenzen können diese Oberschwingungen abstrahlen und elektromagnetische Interferenzen erzeugen, die zu Störungen benachbarter Geräte führen können.

Um zu einer zusätzlichen Sicht der Zusammenhänge zwischen Zeit- und Frequenzbereich zu gelangen, betrachten wir die Fourierreihe aus einem anderen Blickwinkel. Wir gehen dazu von der Darstellung eines Signals mit der Periodendauer $T = 2\pi/\omega_0$ in komplexer Schreibweise aus (wir wissen, dass sich das Signal $s(t)$ durch die folgende Summe darstellen lässt):

$$s(t) = \sum_{k=-\infty}^{+\infty} \vec{D}_k e^{jk\omega_0 t} \quad (3.31)$$

Die einzigen Unbekannten sind hier die Fourierkoeffizienten \vec{D}_k . Um eine von ihnen (die m -te) auszurechnen, multiplizieren wir (3.31) mit $e^{-jm\omega_0 t}$

$$s(t)e^{-jm\omega_0 t} = \left(\sum_{k=-\infty}^{+\infty} \vec{D}_k e^{jk\omega_0 t} \right) e^{-jm\omega_0 t} = \sum_{k=-\infty}^{+\infty} \vec{D}_k e^{j(k-m)\omega_0 t} \quad (3.32)$$

Integrieren wir nun über die Periodendauer T der periodischen Schwingung, so erhalten wir

$$\int_T s(t)e^{-jm\omega_0 t} dt = \int_T \left(\sum_{k=-\infty}^{+\infty} \vec{D}_k e^{j(k-m)\omega_0 t} \right) dt \quad (3.33)$$

Da das Integrieren linear ist, können wir das Integral in die Summe schreiben und die Konstanten \vec{D}_k vor das Integral ziehen. Wir erhalten

$$\int_T s(t)e^{-jm\omega_0 t} dt = \sum_{k=-\infty}^{+\infty} \vec{D}_k \int_T e^{j(k-m)\omega_0 t} dt \quad (3.34)$$

Der »Trick« hinter dem Ganzen ist, dass wegen

$$\int_T e^{j(k-m)\omega_0 t} dt = \begin{cases} T & \text{für } k = m \\ 0 & \text{für } k \neq m \end{cases} \quad (3.35)$$

alle Summanden bis auf einen (den m -ten) Null werden. Die ganze rechte Seite von (3.34) reduziert sich also auf $\vec{D}_m T$. Wir erhalten so für den m -ten Fourierkoeffizienten

$$\vec{D}_m = \frac{1}{T} \int_T s(t)e^{-jm\omega_0 t} dt, \quad m = 0, \pm 1, \pm 2, \dots \quad (3.36)$$

Um zu einer physikalischen Interpretation der Zusammenhänge zu kommen, gehen wir von Abbildung 3.8 aus, die einige Zeiger (D_1, D_{-2} und D_{+n}) von $s(t)$ zeigt, die mit unterschiedlichen Winkelgeschwindigkeiten $k\omega_0$ rotieren. Die Summe aller rechts- und linksdrehenden Zeiger ergibt die Funktion $s(t)$.

Wenn man einen rotierenden Zeiger »anhalten« möchte, rotiert man das ganze Diagramm mit der selben Geschwindigkeit in die Gegenrichtung.

Möchte man \vec{D}_k einer Frequenzkomponente bestimmen, dreht man das ganze Zeigerdiagramm mit $k\omega_0$ in Gegenrichtung, multipliziert also mit $e^{-jk\omega_0 t}$. Dadurch »steht« die betrachtete Komponente, während sich alle anderen Komponenten mit Vielfachen von ω_0 drehen. Bildet man nun den Mittelwert über die Periode T , so ergibt sich \vec{D}_k für diese Komponente. Für die anderen – sich drehenden Komponenten – ist der Mittelwert Null, da der Mittelwert einer Kreisfunktion über Vielfache ihrer Periode Null ist.

3.3.2 Numerische Berechnung der Fourierkoeffizienten

Häufig liegen Signale nicht in analytischer Form vor oder lassen sich nur schwer analytisch darstellen. In diesen Fällen kann man die Fourierkoeffizienten durch numerische Berechnung des Integrals (3.37) ermitteln.

$$\vec{D}_k = \frac{1}{T} \int_T s(t)e^{-jk\omega_0 t} dt, \quad k = 0, \pm 1, \pm 2, \dots \quad (3.37)$$

Zur Berechnung des Integrals (der Fläche) zerlegen wir die Fläche in rechteckige Streifen im Abstand T_s . Bei der Periodendauer T erhalten wir dadurch $N = T/T_s$ Streifen, die summiert werden müssen.

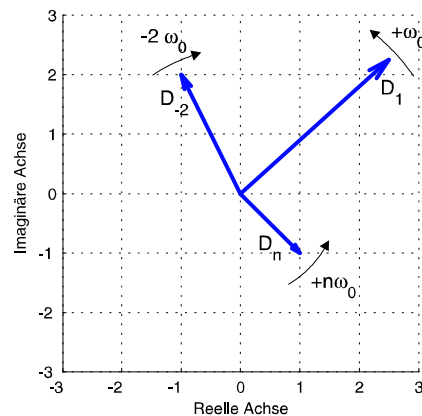


Abbildung 3.8: Signal in Zeigerdarstellung

$$\vec{D}_k = \frac{1}{T} \int s(t) e^{-jk\omega_0 t} dt, \quad k = 0, \pm 1, \pm 2, \dots = \tag{3.38}$$

$$= \lim_{T_s \rightarrow 0} \frac{1}{T} \sum_{n=0}^{N-1} s(nT_s) e^{-jk\omega_0 nT_s} \tag{3.39}$$

Mit den Abkürzungen $\Omega_0 = \omega_0 T_s$ und $N = T/T_s$ erhalten wir

$$\vec{D}_k = \lim_{T_s \rightarrow 0} \frac{1}{N} \sum_{n=0}^{N-1} s(nT_s) e^{-jk\Omega_0 n} \tag{3.40}$$

Für praktische Berechnungen können wir T_s beliebig klein, aber nicht Null machen, was bedeutet, dass bei der numerischen Berechnung des Integrals immer ein Fehler auftreten wird. Wenn wir diesen Fehler ignorieren, erhalten wir für die Fourierkoeffizienten

$$\vec{D}_k = \frac{1}{N} \sum_{n=0}^{N-1} s(nT_s) e^{-jk\Omega_0 n} \tag{3.41}$$

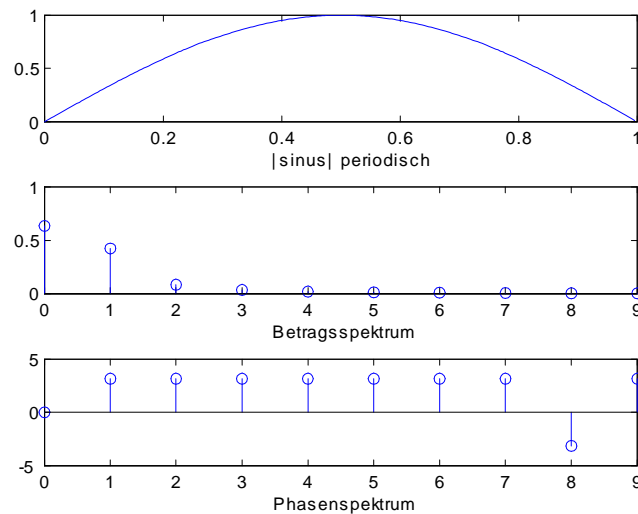
Je größer N gewählt wird, d.h. je schmaler die rechteckigen Streifen werden, desto kleiner ist der Fehler der numerischen Integration

Durch Einsetzen sehen wir, dass $\vec{D}_{k+N} = \vec{D}_k$, was bedeutet, dass das Spektrum \vec{D}_k sich mit der Periodendauer N wiederholt! Die Beziehung nach Gleichung (3.41) nennt man diskrete Fouriertransformation (DFT), die noch genauer behandelt wird. Die DFT liefert die Koeffizienten \vec{D}_k für $n \geq 0$ bis $n = N/2$. Die Koeffizienten oberhalb von $N/2$ sind wegen der Periodizität des Spektrums $\vec{D}_{k+N} = \vec{D}_k$ die Koeffizienten der negativen Frequenzen. Für $N = 64$ ist daher $\vec{D}_{33} = \vec{D}_{-31}$, $\vec{D}_{34} = \vec{D}_{-30}$, ..., $\vec{D}_{63} = \vec{D}_{-1}$.

Ein besonders effizienter Algorithmus zur Berechnung der Gleichung (3.41) ist die Fast Fourier Transformation (FFT), für die N eine Potenz von Zwei – $N = 2^m$, $m \dots$ ganzzahlig – sein muss.

Bei der **DFT** werden die Fourierkoeffizienten numerisch berechnet.

Anmerkung 18 Die Verwendung der Begriffe *DFT* und *FFT* führt oft zu Verwirrung. Während die *DFT* ganz abstrakt die »numerische Berechnung der Fourierkoeffizienten« bezeichnet, steht die *FFT* für einen konkreten Algorithmus, der vorschreibt, wie die Koeffizienten berechnet werden. Die *FFT* kann man also als eine Implementierung der *DFT* ansehen. In der Umgangssprache bezeichnen diese zwei Begriffe meistens eine und die selbe Sache.

Abbildung 3.9: Spektrum von $|\sin(t)|$

Beispiel 19 Als Beispiel berechnen wir die Fourierkoeffizienten der gleichgerichteten Sinusschwingung $s(t) = |\sin \omega_0 t|$ mit Hilfe von Matlab. Wir berechnen die Werte von $s(t)$ beginnend bei $t = 0$, den letzten Wert entnehmen wir an der Stelle $t = T - T_s$. (Der letzte Wert ist nicht an der Stelle T , da dieser Wert identisch mit dem Wert an der Stelle $t = 0$ ist, die neue Periode beginnt an der Stelle $t = T$.)

```

N=256;
t=linspace (0,1,(N+1));
t=t([1:N]);
s=abs(sin(pi*t));
subplot (311), plot(t,s), xlabel '|sinus| periodisch'
d=fft(s)/N; % komplexes Spektrum
dMag=abs(d); dPhase=angle(d);
M=10; % Fourierkoeffizienten von M Spektrallinien
d0=dMag(1); dM=2*dMag(2:M); % Umrechnen auf einseitiges Spektrum
cMag=[d0,dM]; cPhase=dPhase(1:M);
subplot (312), stem((0:(M-1)),cMag), xlabel 'Betragsspektrum';
subplot (313), stem((0:(M-1)),cPhase), xlabel 'Phasenspektrum';N=256;

```

Abbildung 3.9 zeigt die Zeit- und Frequenzbereichdarstellung. Aus einer Formelsammlung erhalten wir die Fourierzerlegung der gleichgerichteten Sinusschwingung ($T = \pi$, $\omega = 2\pi/T = 2$)

$$s(t) = \frac{2}{\pi} - \frac{4}{\pi} \left(\frac{\cos 2t}{1.3} + \frac{\cos 4t}{3.5} + \frac{\cos 6t}{5.7} + \dots \right) = \quad (3.42)$$

$$= 0.6366 - 0.4244 \cos 2t - 0.0849 \cos 4t - 0.0364 \cos 6t + \dots \quad (3.43)$$

Vergleich der analytischen und numerischen Lösung

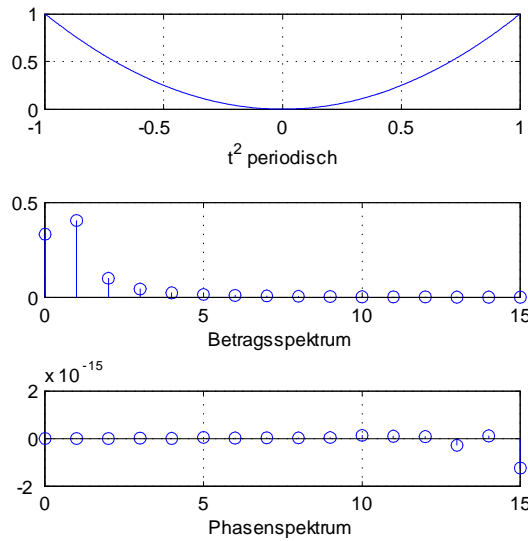


Abbildung 3.10: Spektrum von $s(t) = t^2$

	<i>Betrag (FFT)</i>	<i>Phase (FFT)</i>	<i>Betrag (Reihe)</i>	<i>Phase (Reihe)</i>
	0.6366	0	0.6366	0
	0.4244	3.1416	0.4244	π
	0.0849	3.1416	0.0849	π
	0.0364	3.1416	0.0364	π
Beispiel 20	0.0202	3.1416	0.0202	π
	0.0129	3.1416	0.0129	π
	0.0089	3.1416	0.0089	π
	0.0065	3.1416	0.0065	π
	0.0050	-3.1416	0.0050	$\pi \triangleq -\pi^6$
	0.0040	3.1416	0.0039	π

Beispiel 21 Als weiteres Beispiel berechnen wir das Spektrum der periodischen Parabelfunktion nach Abbildung 3.10 auf numerischem Weg. Durch analytische Berechnung erhalten wir

$$s(t) = \frac{1}{3} - \frac{4}{\pi^2} \left(\cos \pi t - \frac{1}{2^2} \cos 2\pi t + \frac{1}{3^2} \cos 3\pi t - \frac{1}{4^2} \cos 4\pi t + \dots \right) = \tag{3.44}$$

$$= \frac{1}{3} + \frac{4}{\pi^2} \left[\cos(\pi t - \pi) + \frac{1}{4} \cos 2\pi t + \frac{1}{9} \cos(3\pi t - \pi) + \frac{1}{16} \cos 4\pi t + \dots \right] \tag{3.45}$$

$$= \frac{1}{3} + \sum_{k=1}^{\infty} (-1)^k \frac{4}{k^2 \pi^2} \cos(k\pi t) \tag{3.46}$$

Wir sehen einen scheinbaren Widerspruch zwischen der numerischen Berechnung, die für die Phase der Spektrallinien Null ($< 10^{-15}$) liefert, während bei der analytischen Berechnung die Phase der Spektrallinien abwechselnd $-\pi$ und Null ist.

Werten wir die analytische Berechnung aus und summieren die ersten 16 Frequenzkomponenten, dann erhalten wir Abbildung 3.11. Abbildung 3.11 (a) liefert die Synthese nach (3.44).

Abbildung 3.11 (b) stellt die periodische Funktion $s(t) = (1 - t)^2$ dar und unterscheidet sich von Abbildung (a) lediglich durch die Zeitverschiebung. Das Spektrum der periodischen Funktion $(1 - t)^2$ ist $s(t) = \frac{1}{3} + \sum_{k=1}^{\infty} \frac{4}{k^2 \pi^2} \cos(k\pi t)$, die Kosinusfunktionen haben die Phase Null. Diese Lösung haben wir auch bei der numerischen Berechnung gefunden.

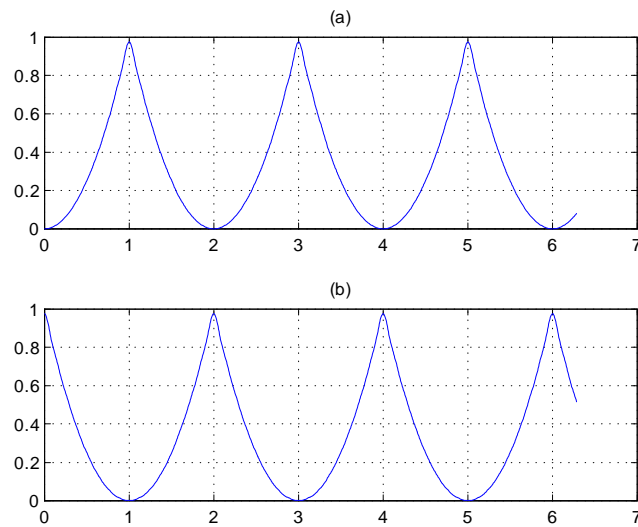
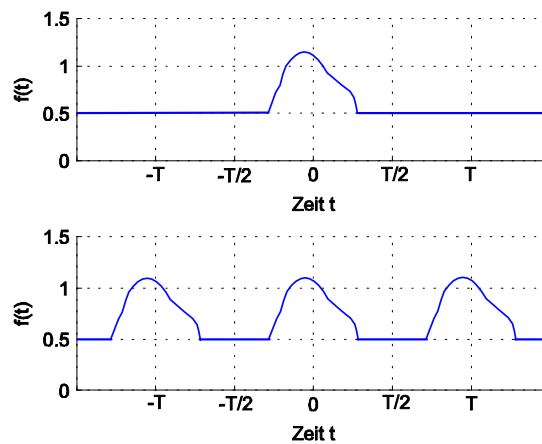
Abbildung 3.11: Frequenzsynthese der periodischen Funktion t^2 für $-1 \leq t \leq 1$ 

Abbildung 3.12: Aperiodisches Signal periodisch fortgesetzt

3.3.3 Die Fouriertransformation

Wir haben die Darstellung im Frequenzbereich durch Entwicklung einer periodischen Funktion in eine Fourierreihe eingeführt. Periodische Funktionen sind theoretische Konstrukte, die bei tatsächlichen Signalen nicht auftreten können, da jedes praktische Signal endliche Dauer hat.

Zur mathematischen Darstellung des endlichen (= aperiodischen) Signals $f(t)$ erzeugen wir ein neues Signal $f_T(t)$, indem wir $f(t)$ in Intervallen von T wiederholen und damit die aperiodische Funktion periodisch machen. Dieses periodische Signal $f_T(t)$ können wir in eine Fourierreihe zerlegen. Lassen wir nun die Periodendauer gegen unendlich gehen ($T \rightarrow \infty$) erhalten wir daraus unser endliches, aperiodisches Signal

$$\lim_{T \rightarrow \infty} f_T(t) = f(t) \quad (3.47)$$

Abbildung 3.12 stellt diesen Zusammenhang graphisch dar.

Wenn wir die Periodendauer T verdoppeln, wird die Grundfrequenz ω_0 der Reihenentwicklung halbiert, es entstehen daher doppelt so viele Spektrallinien wie vorher. Gleichzeitig wird Einhüllende des Spektralfunktion $\frac{1}{T}F(\omega)$ entsprechend kleiner. Weiteres Ver-

Beim Berechnen der **Fouriertransformation** stellen wir uns eine aperiodische Funktion als eine periodische Funktion mit unendlich großer Periode vor.

doppeln führt wiederum zu einer höheren Zahl von Spektrallinien mit niedriger Amplitude, wobei die Form der Einhüllenden $F(\omega)$ gleich bleibt. Schließlich wird das Spektrum so dicht, dass der Abstand zwischen den Spektrallinien infinitesimal klein wird und die Amplitude der entsprechenden Spektrallinie infinitesimal Null wird. Wir erhalten nichts von allem und haben doch etwas!

Führen wir diesen Grenzübergang durch, dann erhalten wir nach einer Reihe von Rechenschritten

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) e^{j\omega t} d\omega \quad (3.48)$$

Gleichung (3.48) nennt man Fourierintegral. Ein aperiodisches Signal lässt sich mit dem Fourierintegral (im Gegensatz zur Fourierreihe beim periodischen Signal) darstellen. Zum Unterschied von der Fourierreihe, bei der diskrete Spektrallinien \vec{D}_k auftreten

$$s(t) = \sum_{k=-\infty}^{+\infty} \vec{D}_k e^{jk\omega_0 t},$$

tritt bei der Fouriertransformation ein kontinuierliches⁷ Spektrum $F(\omega)$ auf, wobei $F(\omega)$ in der Regel komplex ist⁸. $F(\omega)$ ist eine Spektraldichtefunktion, wird aber meistens kurz als Spektrum bezeichnet.

Wir fassen die Darstellung aperiodischer Signale im Zeit- und Frequenzbereich zusammen

$$F(\omega) = \int_{-\infty}^{\infty} f(t) e^{-j\omega t} dt \quad (3.49)$$

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) e^{j\omega t} d\omega \quad (3.50)$$

(3.49) wird direkte, (3.50) wird inverse Fouriertransformation genannt. Die Funktion $F(\omega)$ nennt man die Fouriertransformierte von $f(t)$.

Bei aperiodischen Signalen treten keine Spektrallinien auf, sondern eine Spektraldichtefunktion, d.h. die »Spektrallinien« sind in infinitesimalem Abstand und das Spektrum erstreckt sich immer von $-\infty$ bis ∞ .

Beispiel 22 Als Beispiel der Anwendung der Fouriertransformation berechnen wir das Spektrum eines rechteckigen Impulses nach Abbildung 3.13.

$$\begin{aligned} F(\omega) &= \int_{-\tau/2}^{\tau/2} e^{-j\omega t} dt = -\frac{1}{j\omega} \left(e^{-j\omega\tau/2} - e^{j\omega\tau/2} \right) = \frac{2 \sin\left(\frac{\omega\tau}{2}\right)}{\omega} \\ &= \tau \frac{\sin\left(\frac{\omega\tau}{2}\right)}{\left(\frac{\omega\tau}{2}\right)} = \tau \operatorname{sinc}\left(\frac{\omega\tau}{2}\right) \end{aligned}$$

Diese Funktion wird *sinc-Funktion*⁹ genannt. Die *Si-Funktion* erstreckt sich von $-\infty < \omega < \infty$. Das Spektrum des Rechtecksignals der Abbildung 3.13 ist reell. Der *Si-Funktion* werden wir noch öfter begegnen.

⁷Kontinuierliche Spektren sind aus der Physik bekannt, weißes Licht setzt sich z.B. aus den Spektralfarben des Regenbogens zusammen.

⁸Weist jedem ω eine komplexe Zahl zu.

⁹In der deutschsprachigen Literatur wird diese Funktion *Si-Funktion* oder *Spaltfunktion* genannt. Wir verwenden wie in Matlab die Bezeichnung *sinc-Funktion*.

Bei der Fouriertransformation einer aperiodischen Funktion werden »**Spektrallinien unendlich dicht**« aneinandergedrückt.»

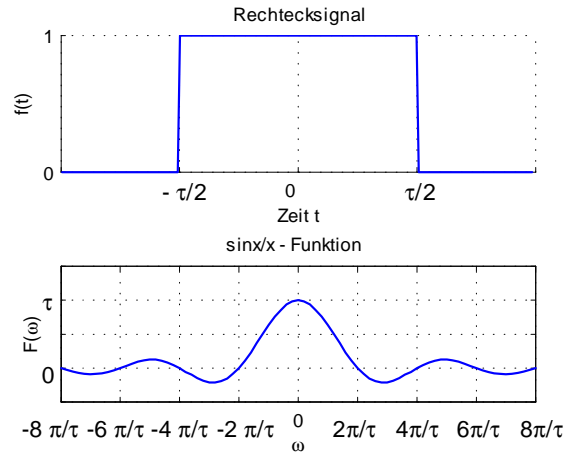


Abbildung 3.13: Rechteckimpuls im Zeit- und Frequenzbereich

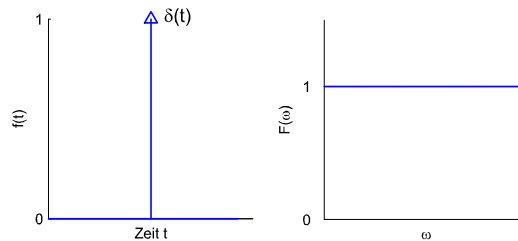


Abbildung 3.14: Einheitsimpuls im Zeit- und Frequenzbereich

Beispiel 23 Wir berechnen die Fouriertransformierte des Einheitsimpulses¹⁰:

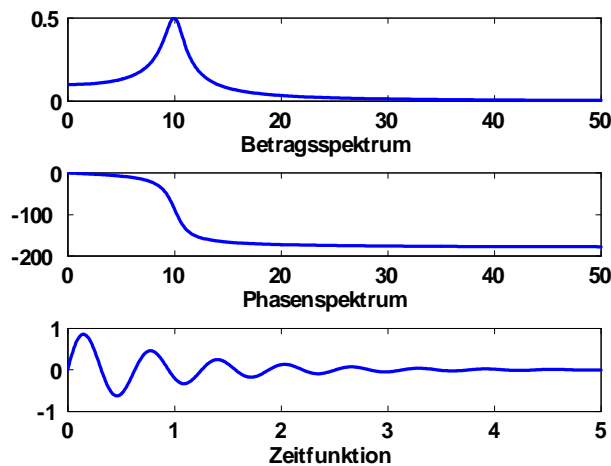
$$F\{\delta_0(t)\} = F(\omega) = \int_{-\infty}^{\infty} \delta(t) \underbrace{e^{-j\omega t}}_{=1 \text{ für } \omega=0} dt = 1$$

$$\delta_0(t) \iff 1$$

Abbildung 3.14 stellt den Einheitsimpuls im Zeit- und im Frequenzbereich dar. Die Impulsfunktion ist keine echte Funktion und beschreibt keine eindeutige Funktion. Wir sehen an diesem Beispiel, dass lediglich die Wirkung der Impulsfunktion auf eine andere Funktion – $e^{-j\omega t}$ – untersucht wird.

Anmerkung 24 Das Spektrum der Impulsfunktion enthält alle Frequenzen von $-\infty$ bis ∞ , deren Amplitude ist für alle Frequenzen gleich ist. Im Impuls »stecken« also Frequenzen. Diese Eigenschaft macht die Impulsfunktion für die Untersuchung von Systemen aus theoretischer Sicht so bedeutend. Legt man an den Eingang eines Systems einen Impuls an, so wird dieses System gleichzeitig mit allen Frequenzen angeregt. Aus messtechnischer Sicht ist der Impuls weniger geeignet, er müsste die Amplitude ∞ haben, was technisch nicht möglich ist. Für messtechnische Zwecke ist die Sprungfunktion $\delta_{-1}(t)$ besser geeignet.

¹⁰Siehe Abschnitt 2.2: Elementarfunktionen

Abbildung 3.15: Betrags-, Phasenspektrum und Zeitfunktion von $s(t) = \delta_{-1}(t)e^{-t} \sin 10t$

3.3.4 Eigenschaften der Fouriertransformation

Einige Transformationsbeziehungen

	$s(t)$	$S(\omega)$
1	$\delta_{-1}(t)e^{-at}$	$\frac{1}{a+j\omega} \quad a > 0$
2	$\delta_0(t)$	1
3	1	$2\pi\delta_0(\omega)$
4	$e^{j\omega t}$	$2\pi\delta_0(\omega - \omega_0)$
5	$\cos \omega_0 t$	$\pi[\delta_0(\omega - \omega_0) + \delta_0(\omega + \omega_0)]$
6	$\sin \omega_0 t$	$j\pi[\delta_0(\omega + \omega_0) - \delta_0(\omega - \omega_0)]$
7	$\delta_{-1}(t)$	$\pi\delta_0(\omega) + \frac{1}{j\omega}$
8	$\delta_{-1}(t) \cos \omega_0 t$	$\frac{\pi}{2}[\delta_0(\omega - \omega_0) + \delta_0(\omega + \omega_0)] + \frac{j\omega}{\omega_0^2 - \omega^2}$
9	$\delta_{-1}(t) \sin \omega_0 t$	$\frac{\pi}{2j}[\delta_0(\omega - \omega_0) - \delta_0(\omega + \omega_0)] + \frac{\omega}{\omega_0^2 - \omega^2}$
10	$\delta_{-1}(t)e^{-at} \cos \omega_0 t$	$\frac{a+j\omega}{(a+j\omega)^2 + \omega_0^2} \quad a > 0$
11	$\delta_{-1}(t)e^{-at} \sin \omega_0 t$	$\frac{\omega_0}{(a+j\omega)^2 + \omega_0^2} \quad a > 0$

Die Transformationsbeziehung (2) zeigt, dass das Spektrum des Dirac-Impulses konstant ist, d.h. alle Frequenzen treten mit gleicher Amplitude auf.

Die Transformationsbeziehung (3) zeigt, dass der Dirac-Impuls im Frequenzbereich ($\delta_0(\omega)$ ist ein »Impuls« bei $\omega = 0$) einem konstanten Signal im Zeitbereich entspricht, das sich von $-\infty$ bis ∞ erstreckt. Aus (2) und (3) kann man auch gut die Symmetrie der Fouriertransformation erkennen.

(5) und (6) zeigt die Spektrallinien des Sinus bzw. Kosinus bei den positiven und negativen Frequenzen ω_0 . Konsequenterweise liefert (4) nur die Spektrallinie bei ω_0 . Bekanntlich ist $\cos(\omega_0 t) = \frac{1}{2}(e^{j\omega_0 t} + e^{-j\omega_0 t})$, wir erhalten das Spektrum entsprechend durch Verwendung von (4).

Abbildung 3.15 zeigt Zeit- und Frequenzdarstellung der eingeschalteten, abklingenden Sinusfunktion. Die Fouriertransformation ist symmetrisch¹¹, der Betrag von $F(\omega)$ ist eine

¹¹Unter der Annahme reeller Zeitfunktionen, die bei technischen Signalen immer erfüllt ist.

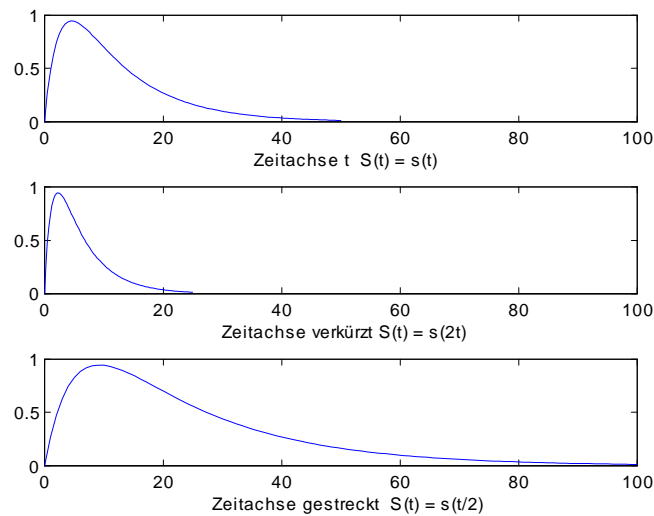


Abbildung 3.16: Skalierung der Zeitachse

gerade Funktion, die Phase von $F(\omega)$ ist eine ungerade Funktion.

3.3.5 Operationen mit Signalen im Zeit- und Frequenzbereich

Zeitskalierung

Abbildung 3.16 zeigt ein Signal mit normaler Zeitachse $s(t)$, mit verkürzter Zeitachse $s(2t)$ und mit gedehnter Zeitachse $s(t/2)$.

Wird das Signal $s(t)$ auf Band aufgezeichnet und dann wiedergegeben, dann entspricht $s(2t)$ der doppelten Abspielgeschwindigkeit, $s(t/2)$ der halben Abspielgeschwindigkeit. Allgemein schreiben wir

$$s(kt) \quad \text{bzw.} \quad s\left(\frac{t}{k}\right) \quad (3.51)$$

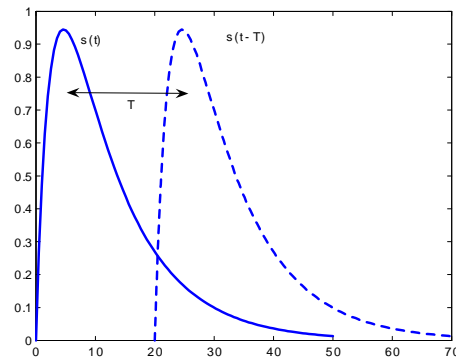
Im Frequenzbereich erhalten wir

$$s(kt) \Leftrightarrow \frac{1}{|k|} S\left(\frac{\omega}{k}\right) \quad (3.52)$$

Abspielen eines Signals in der halben Zeit entspricht der Verdoppelung der Frequenz, bzw. Abspielen in der doppelten Zeit entspricht der Halbierung der Frequenz.

Verschiebung im Zeitbereich

Abbildung 3.3.5 zeigt das Signal $s(t)$ und das nach rechts verschobene – verzögerte – Signal $s(t - T)$. Positives T bewirkt eine Verschiebung nach rechts (zeitlich nach vorne), negatives T bewirkt eine Verschiebung nach links (zeitlich zurück).



Verschiebung im Zeitbereich

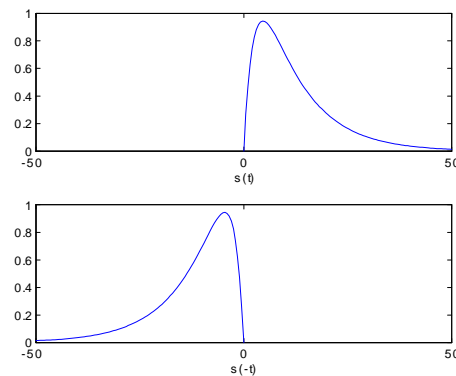
Im Frequenzbereich erhalten wir

$$s(t - T) \Leftrightarrow S(\omega)e^{-j\omega T} \quad (3.53)$$

Die Verschiebung im Zeitbereich ändert das Amplitudenspektrum nicht, das Phasenspektrum wird um den Betrag $-\omega T$ verschoben. Beachten Sie das die Frequenzkomponenten proportional zur Phase (linear mit der Phase) verschoben werden. Diese Eigenschaft hat Bedeutung bei der Signalfilterung (Filter mit linearer Phase), die später behandelt wird.

Zeitspiegelung

Abbildung 3.3.5 zeigt das Signal $s(t)$ mit gespiegelter Zeitachse $s(-t)$ bzw. mit Skalierung $k = -1$.



Zeitumkehr

Verschiebung im Frequenzbereich

Auch bei der Verschiebung im Zeitbereich kann man die Symmetrie der Fouriertransformation gut erkennen.

$$s(t)e^{j\omega_0 t} \Leftrightarrow S(\omega - \omega_0) \quad (3.54)$$

$e^{j\omega_0 t}$ ist keine reelle Zeitfunktion und kann daher nicht erzeugt werden. Die Verschiebung im Frequenzbereich wird daher praktisch durch Multiplikation mit einer sinusoidalen Funktion erzeugt

$$s(t) \cos \omega_0 t = \frac{1}{2} [s(t)e^{j\omega_0 t} + s(t)e^{-j\omega_0 t}] \quad (3.55)$$

$$s(t) \cos \omega_0 t \Leftrightarrow \frac{1}{2} [S(\omega - \omega_0) + S(\omega + \omega_0)] \quad (3.56)$$

Anmerkung 25 Die Beziehung $s(t) \cos \omega_0 t$ tritt in der Nachrichtentechnik als Amplitudenmodulation auf. Ein Signal im Basisband (z.B. Sprachsignal beim Kurzwellenfunk) wird in einen anderen Frequenzbereich (ω_0) verschoben. Durch Wahl unterschiedlicher Frequenzen ω_0 , den sogenannten Trägerfrequenzen, ist es möglich mehrere Basisbänder nebeneinander im Frequenzbereich unterzubringen (Frequenzmultiplex). Beachten Sie, dass durch die Verschiebung um ω_0 , die negativen Frequenzen des Basisbandsignals zu positiven Frequenzen werden und auch physikalisch auftreten, man spricht vom unteren und oberen Seitenband.

Faltung im Zeit- und Frequenzbereich

Wie wir bereits gesehen haben, ist die Faltung folgendermaßen definiert:

$$f(t) * g(t) = \int_{-\infty}^{\infty} f(\tau)g(t - \tau)d\tau = \int_{-\infty}^{\infty} f(t - \tau)g(\tau)d\tau$$

Im Frequenzbereich erhalten wir

$$f(t) * g(t) \Leftrightarrow F(\omega)G(\omega) \quad (3.57)$$

$$f(t)g(t) \Leftrightarrow \frac{1}{2\pi}F(\omega) * G(\omega) \quad (3.58)$$

Aus der Faltung im Zeitbereich wird die Multiplikation im Frequenzbereich, aus der Faltung im Frequenzbereich wird die Multiplikation im Zeitbereich. Aus diesem Zusammenhang wird auch klar, dass die Bandbreite B des Produkts von zwei Signalen $f(t)$ und $g(t)$, gleich der Summe der Bandbreiten $B_g + B_f$ ist.

3.3.6 Frequenz und Zeit – Unschärfeprinzip

Es gehört zu den wichtigsten Erkenntnissen der Physik, dass bestimmte Größen nicht unabhängig voneinander gemessen werden können. Dazu gehört auch die Messung von Zeit und Frequenz:

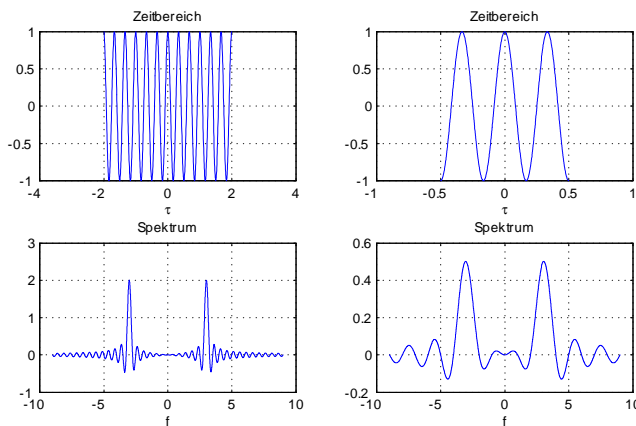
Je kürzer die Zeitdauer Δt eines Signals, desto breiter wird sein Frequenzband Δf , schmaler das Frequenzband Δf eines Signals, desto länger muss die Zeitdauer Δt des Signals sein.

- Ein langer Sinuston (ein paar Sekunden) klingt wie erwartet.
- Kürzt man die Zeitdauer des Sinustons und erzeugt Sinus-Bursts, geht der Klang von einem Sinuston in ein Knattern über.
- Je kürzer die Zeitdauer des Bursts, desto breiter wird sein Spektrum.

Wir müssen aber noch festlegen, was wir unter Bandbreite verstehen.

Die Bandbreite des Signals geht zwar gegen ∞ , der Amplitudenverlauf geht aber sehr schnell gegen Null, sodass dieser Teil des Frequenzbandes vernachlässigbar ist. Für Messzwecke können wir als Bandbreite z. B. als die halbe mittlere Breite des Hauptmaximums bezeichnen.

- Wenn wir die Zeitdauer Δt halbieren, so verdoppelt sich die Bandbreite Δf . Δt und Δf verhalten sich also umgekehrt proportional $\Delta t = k \times \Delta f$ oder $\Delta f \times \Delta t = k$.

Abbildung 3.17: Spektrum Kosinusburst $f_0 = 3$, $\tau = 2$ und $\tau = 0.5$

Zur Berechnung von k untersuchen wir eine zeitlich begrenzte Kosinusfunktion mit der Amplitude $A = 1$

$$f(t) = \begin{cases} \cos \omega_0 t & -\tau \leq t \leq \tau \\ 0 & |t| > \tau \end{cases} \quad (3.59)$$

Wir wissen von der Fouriertransformation, dass bei zeitbegrenzten Signalen ein Dichtespektrum und keine diskreten Spektrallinien auftreten. Es kann daher die in Gleichung (3.59) auftretende Kreisfrequenz ω_0 nicht als diskrete Spektrallinie auftreten!

Wir berechnen die Dichtefunktion von (3.59). Da $f(t)$ eine gerade Funktion ist, können wir $F(\omega) = \int_{-\infty}^{\infty} f(t)e^{-j\omega t} dt \Rightarrow \int_{-\infty}^{\infty} f(t) \cos(\omega t) dt$ vereinfachen und erhalten

$$F(\omega) = \int_{-\infty}^{\infty} f(t) \cos(\omega t) dt = \int_{-\tau}^{\tau} \cos(\omega_0 t) \cos(\omega t) dt = \quad (3.60)$$

$$= \frac{1}{2} \int_{-\tau}^{\tau} \{ \cos [(\omega_0 - \omega)t] + \cos [(\omega_0 + \omega)t] \} dt = \quad (3.61)$$

$$= \frac{1}{2} \left[\frac{\sin [(\omega_0 - \omega)t]}{\omega_0 - \omega} + \frac{\sin [(\omega_0 + \omega)t]}{\omega_0 + \omega} \right]_{-\tau}^{\tau} = \quad (3.62)$$

$$= \left[\frac{\sin [(\omega_0 - \omega)\tau]}{\omega_0 - \omega} + \frac{\sin [(\omega_0 + \omega)\tau]}{\omega_0 + \omega} \right] \quad (3.63)$$

$f(t)$ und $F(\omega)$ sind in Abbildung 3.17 dargestellt und zeigen Kosinusbursts der Kreisfrequenz $\omega_0 = 3$, einmal mit der Länge $\tau = 2$ und einmal mit der Länge $\tau = 0.5$. Die Nulldurchgänge des Spektrums errechnen sich aus den Nulldurchgängen der Sinus-Funktion $\sin[(\omega - \omega_0)\tau] = 0 \Rightarrow (\omega - \omega_0)\tau = n\pi$, für $n = (0), 1, 2, \dots$. An der Stelle $\omega \rightarrow \omega_0$ wird $F(\omega_0) = \tau$, die Nullstellen von $F(\omega)$ liegen, von ω_0 ausgehend, in den Abständen π/τ .

Wir erkennen deutlich, dass die Spektraldichte im Bereich von $\omega_0 = 3$ größer wird. Je länger die Zeitdauer τ des Kosinusbursts, desto näher rücken die die Nullstellen zusammen und desto besser entspricht der Burst der kontinuierlichen Kosinusschwingung. Im Grenzfall $\tau \rightarrow \infty \Rightarrow \Delta\omega \rightarrow 0$ erhalten wir die kontinuierliche Kosinusschwingung mit dem Spektrum $F(\omega) \rightarrow \delta_0(\omega - \omega_0)$, also eine Spektrallinie.

Wir können aus Abbildung 3.17 erkennen, dass das Spektrum, d.h. die »Frequenz« einer zeitlich begrenzten Kosinusfunktion – *unschärf* – ist. Bezeichnen wir die Breite der Hauptkeule der Funktion in Abbildung 3.17 als Spektrum, so erhalten wir $\Delta\omega = 2\frac{\pi}{\tau}$, bzw. $\Delta f = \frac{1}{\tau}$. Die Zeitdauer der Kosinusschwingung beträgt $\Delta t = 2\tau$. Damit erhalten wir für den Zusammenhang zwischen Frequenz und Zeitdauer die Beziehung $\Delta f \times \Delta t = 2$

Nimmt man einen Sinus-Burst mit Gauß-Funktion als Einhüllende und misst man Δt auf halber Amplitudenhöhe im Zeitbereich, Δf auf halber Amplitudenhöhe im Frequenzbereich (das Spektrum hat ebenfalls eine Gauß-Verteilung), dann wird $\Delta t \times \Delta f = 1$

Unschärfeprinzip in der Nachrichtentechnik
 $\Delta t \times \Delta f \geq 1$

Das Unschärfeprinzip in der Nachrichtentechnik wird daher

$$\Delta f \times \Delta t \geq 1 \quad (3.64)$$

Anmerkung 26 Je kürzer die Zeitdauer Δt eines Signals desto breiter Δf wird sein Spektrum; umgekehrt gilt, je schmaler das Spektrum, desto länger die Zeitdauer des Signals. Zeitdauer und Spektrum sind untrennbar verbunden, Zeit und Frequenz können nicht unabhängig voneinander gemessen werden.

3.3.7 Numerische Berechnung der Fouriertransformation

Wenn sich die Transformation $F(\omega) = \int_{-\infty}^{\infty} f(t)e^{-j\omega t} dt$ nicht analytisch berechnen lässt, muss das Spektrum numerisch berechnet werden. Da unsere Zeitsignale sich nur von 0 bis T erstrecken (außerhalb dieses Intervalls werden sie als Null definiert), können wir die Integrationsgrenzen anpassen.

$$F(\omega) = \int_0^T f(t)e^{-j\omega t} dt \quad (3.65)$$

$$F(\omega) \approx \sum_{n=0}^{N-1} f(n\Delta t)e^{-j\omega n\Delta t} \Delta t \quad t \rightarrow n\Delta t, \quad \Delta t = \frac{T}{N} \quad (3.66)$$

Δt (bzw. die Zahl der Werte N) muss so gewählt werden, dass die gewünschte Genauigkeit bei der numerischen Integration erreicht wird.

Wir tasten das Signal $f(t)$ mit $f_{\text{abtaast}} = \frac{1}{\Delta t}$ ab und nehmen an, dass die Spektralkomponenten des Signals oberhalb der halben Abtastfrequenz $|\omega| \geq \omega_{\text{max}} = \frac{1}{2}\omega_{\text{abtaast}}$ vernachlässigbar sind. Das Zeitsignal setzt sich dann aus seinen spektralen Komponenten wie folgt zusammen

$$f(n\Delta t) = \frac{1}{2\pi} \int_{-\omega_{\text{max}}}^{\omega_{\text{max}}} F(j\omega)e^{j\omega n\Delta t} d\omega \quad (3.67)$$

Unter Verwendung der DFT erhalten wir

$$F(k\Delta\omega) \approx \Delta t \sum_{n=0}^{N-1} f(n\Delta t)e^{-jk\Delta\omega n\Delta t} \quad (3.68)$$

$$\Delta\omega = \frac{1}{N} 2\pi \frac{1}{\Delta t} \quad (3.69)$$

$$F(n\Delta\omega) \approx \Delta t \sum_{n=0}^{N-1} f(n\Delta t)e^{-j\frac{2\pi}{N}kn} \quad (3.70)$$

Die Fouriertransformierte einer nichtperiodischen Funktion $f(t)$ kann also numerisch mit Hilfe der DFT berechnet werden

$$F(k\Delta\omega) = \begin{cases} \frac{1}{f_{\text{abtaast}}} F[k] & k < N/2 \\ 0 & k \geq N/2 \end{cases} \quad (3.71)$$

wenn das Signal mit $f_{\text{abtaast}} = \frac{1}{\Delta t} = \frac{T}{N}$ abgetastet wird und die Spektralkomponenten bei Frequenzen größer als $\frac{N}{2T}$ vernachlässigbar sind.

3.4 Größe von Signalen (Energie und Leistung)

Um die »Größe« eines Signals zu erfassen, wird man sowohl die Amplitude als auch die Dauer berücksichtigen müssen. Das zunächst naheliegende Maß der Fläche des Signals ist ungeeignet, da sich positive und negative Signalflächen aufheben können. Die Verwendung des Betrages ist aus mathematischer Sicht nicht sehr praktisch, daher definiert man die Signalenergie

$$E_s = \int_{-\infty}^{\infty} s^2(t) dt \quad \text{bzw.} \quad E_s = \int_{-\infty}^{\infty} |s(t)|^2 dt \quad (3.72)$$

für reelle und komplexe Signale. Diese Definition liefert nur dann ein brauchbares Maß, wenn die Energie endlich ist. Das ist dann gegeben, wenn die Signalamplitude gegen Null geht, wenn $|t| \rightarrow \infty$. Ist das nicht der Fall, dann liefert die Signalleistung,

$$P_s = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} s^2(t) dt \quad \text{bzw.} \quad P_s = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} |s(t)|^2 dt \quad (3.73)$$

der quadratische Mittelwert, ein brauchbares Maß. Die Wurzel daraus ist der Effektivwert (root mean square - rms).

Anmerkung 27 Es muss darauf hingewiesen werden, dass die Begriffe Signalenergie und Signalleistung zwar ein Maß für die Energie darstellen aber nicht die tatsächliche Signalenergie sind. Erst wenn $s(t)$ der Amplitudenverlauf eines elektrischen Stromes oder einer elektrischen Spannung an einem ohmschen Widerstand sind, kann man im physikalischen Sinn von einer Energie sprechen. Konsequenterweise sind auch die Dimensionen von E_s und P_s nicht Joule und Watt, wir verwenden diese Begriffe lediglich zur Messung der Größe eines Signals.

3.5 Parseval'sches Theorem

Wir wissen, dass die Fourierentwicklung eines periodischen Signals durch die Beziehung

$$s(t) = C_0 + \sum_{k=1}^N C_k \sin(\omega_k t + \varphi_k) \quad (3.74)$$

dargestellt werden kann. Selbstverständlich muss die Leistung des gesamten Signals $s(t)$ gleich der Summe der Leistungen seiner spektralen Komponenten sein. Zur Berechnung der Leistung eines periodischen Signals berechnet man die Energie pro Periode und dividiert sie durch die Periodendauer. Es ergibt sich für die Leistung einer (der k -ten) Spektralkomponente (zur Vereinfachung wird die Phase weggelassen)

$$\begin{aligned} P_k &= \frac{1}{T} \int_T (C_k \sin \omega_k t)^2 dt = \frac{C_k^2}{2T} \int_T (1 - \sin 2\omega_k t) dt = \\ &= \frac{C_k^2}{2T} T - \frac{C_k^2}{2T} \int_T \sin(2\omega_k t) dt = \frac{C_k^2}{2} \end{aligned} \quad (3.75)$$

Die Leistung eines sinusförmigen Signals ist unabhängig von der Frequenz und beträgt $C_k^2/2$. Wir erhalten für die Leistung eines beliebigen periodischen Signals mit den Fourierkoeffizienten C_k

$$P_s = C_0^2 + \frac{1}{2} \sum_{k=1}^{\infty} C_k^2 \quad (3.76)$$

Die **Signalenergie** ist die Fläche unter dem quadrierten Betrag des Signals.

Die **Signalleistung** eines Zeitsignals ist die mittlere Signalenergie pro Zeit.

Die Leistung eines sinusförmigen Signals ist gleich dem halben Quadrat seiner Amplitude. Sie hängt nicht von der Frequenz ab.

3.6 Zusammenfassung

In diesem Kapitel haben wir gezeigt, dass periodische Signale durch Überlagerung von sinusförmigen Funktionen dargestellt werden können. Wir können ein Signal als Zeitfunktion oder als Summe von sinusförmigen Komponenten betrachten und gelangen dadurch zur Darstellung im Zeit- und im Frequenzbereich.

Periodische Signale sind immer aus ganzzahligen Vielfachen der Grundschiwingung aufgebaut, im Frequenzbereich führt das zu Linienspektren. Signale, die aus sinusförmigen Schwingungen zusammengesetzt sind, deren Frequenzverhältnis nicht rational ist, sind keine periodischen Signale, haben aber ebenfalls ein Linienspektrum. Die Breite des Spektrums hängt vom Signal ab, Linienspektren können endliche Breite haben oder können sich von $-\infty$ bis ∞ erstrecken.

Signale sind in der Regel nicht periodisch, sondern haben endliche Dauer. Für die Berechnung der Spektralkomponenten aperiodischer Signale wird die Fourierreihe zur Fouriertransformation. Aus dem Linienspektrum der periodischen Signale wird das kontinuierliche Spektrum der aperiodischen Signale. Das Spektrum eines aperiodischen Signals erstreckt sich immer von $-\infty$ bis ∞ .

Der Aufbau der Fourierreihe und der Fouriertransformation ist sehr ähnlich.

$$\begin{aligned}
 s(t) &= \sum_{k=-\infty}^{+\infty} D_k e^{jk\omega_0 t} & \Leftrightarrow & \quad D_k = \frac{1}{T} \int_T s(t) e^{-jk\omega_0 t} dt \\
 f(t) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) e^{j\omega t} d\omega & \Leftrightarrow & \quad F(\omega) = \int_{-\infty}^{\infty} f(t) e^{-j\omega t} dt
 \end{aligned}$$

Bei der Fourierreihe wird das Zeitsignal als Summe von diskreten komplexen Exponentialfunktionen $e^{jk\omega_0 t}$ mit den (komplexen) Amplituden D_k dargestellt. Bei der Fouriertransformation wird das Zeitsignal als Summe von unendlich vielen (Integral) komplexen Exponentialfunktionen $e^{j\omega t}$ mit der (komplexen) Amplitudendichte $F(\omega)$ dargestellt.

Die Amplitude der Frequenzkomponenten D_k des periodischen Signals errechnet man aus dem Zeitsignal $s(t)$ durch »Anhalten« des Zeigerdiagramms, indem man es mit der Frequenz der gesuchten Komponente in die entgegengesetzte Richtung $e^{-jk\omega_0 t}$ dreht und den Mittelwert (Integration) bildet. In ähnlicher Weise erhält man die »Frequenzkomponenten« $F(\omega)$ des aperiodischen Signals aus dem Zeitsignal $f(t)$ durch »Anhalten« des Zeigerdiagramms, indem man es mit der Frequenz der gesuchten Komponente in die entgegengesetzte Richtung $e^{-j\omega t}$ dreht und den Mittelwert (Integration) bildet. Zum Unterschied vom periodischen Signal liegen die »Spektrallinien« ∞ dicht nebeneinander, man erhält daher die Frequenzdichte $F(\omega)$.

Teil II

Elektrotechnische Grundlagen

Kapitel 4

Elektrische Bauelemente

Inhalt

4.1	Einführung	47
4.1.1	Netzwerke, Ströme und Spannungen	48
4.1.2	Unabhängige Spannungs- und Stromquellen	49
4.1.3	Zählpfeile	50
4.2	Die Kirchhoff'schen Gleichungen	50
4.3	Ideale Bauelemente	52
4.3.1	Gesteuerte Quellen	52
4.3.2	Idealer Widerstand	52
4.3.3	Kapazität	53
4.3.4	Induktivität	55
4.3.5	Spannung und Strom an L und C bei Schaltvorgängen	56
4.3.6	Schreibweise der Netzwerkgleichungen	57
4.4	Präzisere Modelle der Bauelemente	58
4.4.1	Spannungsquelle mit Innenwiderstand	58
4.4.2	Ersatzschaltbild für den Kondensator	59
4.4.3	Ersatzschaltbild für die Spule	59
4.5	Zusammenfassung	60

4.1 Einführung

Signale werden durch Variation physikalischer Größen repräsentiert, diese Größen können verändert, gespeichert und übertragen werden. Elektrische und magnetische Größen sind besonders gut veränderbar, übertragbar und speicherbar und haben daher eine überraschende technische Bedeutung.

Elektrische und magnetische Vorgänge treten in Form von beweglichen Ladungen sowie von elektrischen und magnetischen Feldern auf. Für eine genaue Analyse müssten diese physikalischen Erscheinungen genau nach Ort und Zeit erfasst werden, was in den meisten Fällen aber zu schwierig und häufig nicht durchführbar ist. Die Berechnung der Feldgleichungen ist meistens auch nicht erforderlich, da nur bestimmte Aspekte des Feldes und integrale Größen wie Spannung und Strom interessieren. Daher wurde eine eigene Theorie entwickelt, die auf diesen integralen Größen aufbaut. Das war übrigens auch der in der physikalischen Forschung eingeschlagene Weg. Man abstrahiert Beobachtungen und Messungen in Form von Modellen, je genauer die tatsächlichen Vorgänge nachgebildet werden sollen, desto komplizierter sind die Modelle. Wegen der besseren Übersichtlichkeit und der einfacheren Berechenbarkeit ist man daran interessiert möglichst einfache Modelle zu finden, die nur die wesentlichen Zusammenhänge beschreiben. Die idealen

Statt den Verlauf elektromagnetischer Felder direkt zu berechnen, werden **Modelle** eingeführt, die sich nur auf bestimmte Aspekte der Felder und integrale Größen wie Spannung oder Strom **beschränken**.

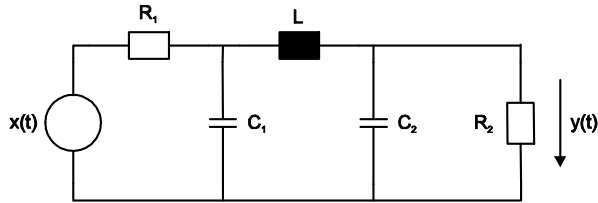


Abbildung 4.1: Netzwerkanalyse

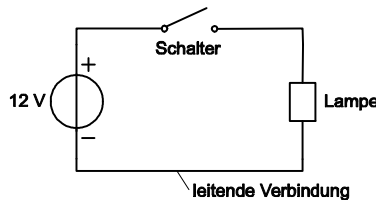


Abbildung 4.2: Idealisierte Schaltung

elektrischen Bauelemente und die mit ihnen gebildeten elektrischen Netzwerke sind solche Modelle und bilden die Basis für die *Netzwerktheorie*.

Bei der Untersuchung von Netzwerken gibt es zwei Arten von Aufgabenstellungen:

Netzwerkanalyse

- Gegeben ist eine Anordnung von miteinander verbundenen Bauelementen, das Netzwerk N , an das ein Eingangssignal $x(t)$ angelegt wird. Gesucht ist das Ausgangssignal $y(t)$. Diese Aufgabe nennt man Netzwerkanalyse. Abbildung 4.1 zeigt ein Beispiel.

Netzwerksynthese

- Gegeben ist das Eingangssignal $x(t)$ und ein gewünschtes Ausgangssignal $y(t)$, gesucht ist das Netzwerk N . Diese Aufgabe nennt man Netzwerksynthese.
- Schließlich gibt es noch die Aufgabe zu einem gegebenen Netzwerk N und dem Ausgangssignal $y(t)$, das Eingangssignal zu ermitteln. Diese Aufgabe hat keinen eignen Namen und tritt vor allem in der Messtechnik auf, wobei das Messgerät durch das »Netzwerk« N modelliert wird.

4.1.1 Netzwerke, Ströme und Spannungen

Das Beispiel in Abbildung 4.2 soll in die Begriffswelt einführen.

Wir haben eine Batterie, die über Schalter und Kabel mit einer Lampe verbunden ist. In der Batterie finden komplizierte elektrochemische Vorgänge statt, wir können aber die Spannung an den Klemmen der Batterie leicht messen. Wird der Schalter geschlossen, fließt elektrische Ladung (Elektronen) durch die Kabel, die Lampe wird warm und leuchtet. Die Kabel bestehen aus einem isolierten sehr guten elektrischen Leiter (Kupfer). Die Elektronen können gut im Leiter fließen (dem Ladungsfluss wird nur ein sehr kleiner Widerstand entgegengesetzt), nicht aber in der Isolierung. Der Leuchtfaden der Glühlampe besteht aus einem sehr dünnen wendelförmig aufgerollten Wolframdraht. Der Wolframdraht setzt dem Ladungsfluss einen großen Widerstand entgegen (die Elektronen kollidieren mit den Wolframatomen), der Draht wird heiß und beginnt zu glühen. (Wolfram kann hohen Temperaturen standhalten, das ist der Grund dafür, dass der Wolframdraht glüht und erst nach längerer Zeit »durchbrennt«.) Die dabei von der Quelle an die Lampe abgegebene Leistung (in Form von Wärme und Licht) ist gleich dem Produkt des elektrischen Stroms (Ladungen/Zeiteinheit) und der Spannung an der Batterie. Die tatsächlichen Zusammenhänge in der Batterie, beim Stromfluss durch den Leiter und in der Lampe sind sehr kompliziert. Wir können aber die wesentlichen Zusammenhänge durch eine einfache elektrische Schaltung, bestehend aus Spannungsquelle, Schalter

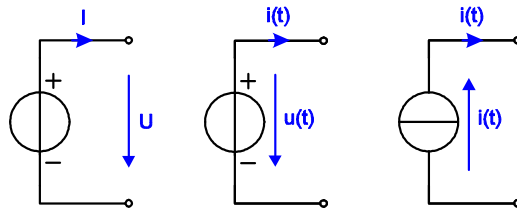


Abbildung 4.3: Schaltsymbole für Spannungs- und Stromquellen

und Widerstand (Lampe) beschreiben. Die Drähte werden als Linien gezeichnet, die die Lampe (widerstandsfrei) mit der Spannungsquelle verbinden.

Elektrische Schaltkreise verhalten sich ähnlich wie Wasserleitungssysteme: Die Batterie entspricht der Pumpe, die die Ladung – das Wasser – in Bewegung setzt. Die Leiter – meistens Kupferdrähte – durch die der elektrische Strom fließt, entsprechen den (reibungsfreien) Rohren, durch die das Wasser fließt. Die elektrische Spannung entspricht dem Druck der Pumpe. Die Lampe (der elektrische Widerstand) entspricht einer Einengung in einem Wasserrohr in der durch Turbulenzen Energie in Wärme umgewandelt wird. Der Strom ist ein Maß für den Fluss von Ladung *durch* den Querschnitt der Netzwerkelements, während die Spannung *über* den Enden (zwischen den Enden) des Netzwerkelements gemessen wird. Die Leistung ist das Produkt der Flussgröße Strom mit der Abfallgröße Spannung.

Analogie Schaltkreis - Wasserleitungssystem:

- Draht Rohr
- Ladung Wasser
- Quelle Pumpe
- Widerstand Rohreinengung
- Spannung Wasserdruck
- Strom Durchfluss

4.1.2 Unabhängige Spannungs- und Stromquellen

Zur Aufrechterhaltung des Stromes in einer Schaltung müssen Quellen vorhanden sein, die die im Stromkreis fließenden Elektronen und Energie in das Netzwerk liefern. Genau genommen handelt es sich bei den Quellen nicht um Energieerzeuger sondern um Energieumwandler. In einem Akkumulator wird beispielsweise chemische Energie in elektrische Energie umgewandelt, diese elektrische Energie wird im ohmschen Widerstand wiederum in Wärmeenergie umgewandelt.

Eine ideale Gleichspannungsquelle ist ein Schaltelement, dessen Quellspannung unabhängig von der angeschlossenen Last (vom durchfließenden Strom) ist. Die ideale Spannungsquelle ist nur ein Modell einer Spannungsquelle und entspricht nicht der physikalischen Wirklichkeit. Wenn der an die Quelle angeschlossene Widerstand kleiner und kleiner wird, steigt der Strom immer mehr an und damit die Leistungsabgabe der Quelle. Jede reale Quelle hat eine Grenze bei ihrer Stromlieferfähigkeit. Akkumulatoren oder elektronische Netzgeräte kommen aber im Betriebsbereich einer idealen Quelle sehr nahe.

Die Spannung einer **idealen Spannungsquelle ist unabhängig vom durchfließenden Strom**. Solche Quellen gibt es in der Wirklichkeit nicht.

Die elektrische Spannung wird in Volt (V), der Strom in Ampere (A) und der elektrische Widerstand in Ohm (Ω) gemessen. Eine ideale Gleichspannungsquelle hat folgende Eigenschaften:

Größe	Einheit
Spannung (U)	Volt (V)
Strom (I)	Ampere (A)
Widerstand (R) ..	Ohm (Ω)

- Die Ausgangsspannung ist unabhängig vom angeschlossenen Netzwerk.
- Im Falle eines angeschlossenen ohmschen Widerstands stellt sich der Strom $I = U_{Quelle}/R$ ein, für $R \rightarrow 0$ (Kurzschluss) wird $I \rightarrow \infty$.

Eine ideale Gleichstromquelle hat die Eigenschaften:

- Der Ausgangsstrom ist unabhängig vom angeschlossenen Netzwerk.
- Im Falle eines angeschlossenen ohmschen Widerstands stellt sich die Spannung $U = I_{Quelle}R$ ein, für $R \rightarrow \infty$ (Leerlauf) wird $U \rightarrow \infty$.

Für Spannungs- und Stromquellen werden die Symbole der Abbildung 4.3 verwendet. Spannung und Strom von idealen Quellen kann auch zeitlich veränderlich sein.

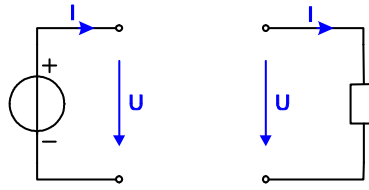


Abbildung 4.4: Erzeuger- (links) und Verbraucher- (rechts) Zählfeilsystem

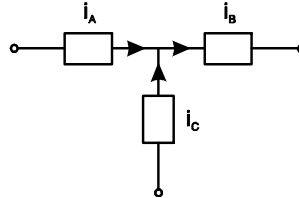


Abbildung 4.5: Knotenregel

4.1.3 Zählfeile

Zählfeile geben die Richtungen für Strom und Spannung **im Modell** an, unabhängig von der wirklichen physikalischen Richtung.

Die Quellen sind die Ursache für den Stromfluss und die Spannungsabfälle im Netzwerk. Bei komplizierten Netzwerken ist die Richtung von Strom und Spannung nicht von vornherein klar, sondern muss erst durch die Netzwerkanalyse ermittelt werden. In Schaltungen werden Strom und Spannung durch Zählfeile¹ (beliebige) Richtungen gegeben, die Netzwerkanalyse liefert die tatsächliche Richtung. (Falls sie mit der Richtung des Zählfeils nicht übereinstimmt, ist der Wert der entsprechenden Größe negativ.)

Größe **Einheit**
Leistung (P) ... Watt (W)

Ein durch einen Verbraucher fließender Strom I erzeugt einen Spannungsabfall U , Strom und Spannung zeigen in dieselbe Richtung, die im Verbraucher umgesetzte Leistung ist $P = U \cdot I$. Die Leistung wird in Watt (W) gemessen. In unserem mechanischen Analogon bedeutet das, dass Wasser der Schwerkraft folgt und daher bergab fließt. Man spricht in diesem Fall vom Verbraucher-Zählfeilsystem.

In Quellen wird der aus der positiven Klemme herausfließende Strom positiv gezählt, Strom und Spannung haben daher die entgegengesetzte Richtung, man spricht vom Erzeugerfeilsystem. In unserem mechanischen Analogon bedeutet das, dass Wasser in der Pumpe bergauf fließt. Die Leistung des Erzeugers $P = -U \cdot I$ ist negativ und damit an den physikalischen Hintergrund angepasst, dass der Generator Energie liefert, der Verbraucher Energie aufnimmt und der gesamte Energieumsatz Null sein muss. Abbildung 4.4 stellt die Zusammenhänge dar.

4.2 Die Kirchhoff'schen Gleichungen

Ein Knoten in einem elektrischen Netzwerk ist ein Punkt, an dem zwei oder mehr Netzwerkelemente verbunden sind. Die *Kirchhoff'sche Knotenregel* besagt, dass die Summe aller Ströme in einen Netzwerkknoten Null sein muss.

Die Kirchhoff'sche Knotenregel:

Die Summe aller in einen Knoten fließenden Ströme ist gleich Null.

$$\sum I = 0 \quad (4.1)$$

Die Zählrichtung der Ströme ist entsprechend Abbildung 4.5 frei wählbar, $I_a - I_b + I_c = 0$, in den Knoten fließende Ströme sind positiv, aus dem Knoten fließende negativ².

¹Strom und Spannung sind skalare Größen, in Schaltungen werden ihnen durch Zählfeile Richtungen zugeordnet. Diese Pfeile geben lediglich die Richtung des Stromflusses durch das Bauelement bzw. des Spannungsabfalls am Bauelement an und sind *keine* Vektoren.

²Ebenso könnten aus dem Knoten fließenden Ströme positiv, in den Knoten fließende Ströme negativ gezählt werden. Eine einmal festgelegte Vereinbarung muss aber innerhalb des gesamten Netzwerks einheitlich angewendet werden!

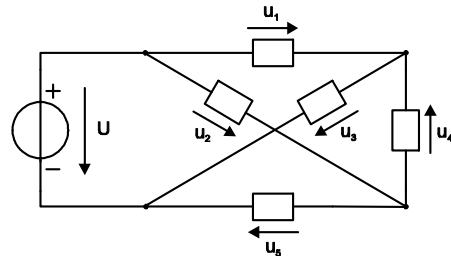


Abbildung 4.6: Zählpfeile

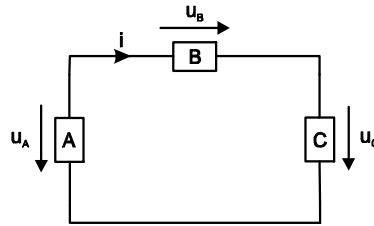


Abbildung 4.7: Maschenregel

Zu einem physikalischen Verständnis gelangen wir, in dem wir mikroskopisch vom elektrischen Strom als Fluss von elektrischen Ladungen ausgehen. Ladung pro Zeiteinheit ist gleich dem Strom. Die elektrische Ladung³ wird in Coulomb (C) gemessen.

$$I = \frac{Q}{t} \tag{4.2}$$

Die Ladungen bleiben erhalten und können sich im Knoten nicht anhäufen, daher muss die Summe der zufließenden Ladungen gleich der Summe der abfließenden Ladungen sein.

Beispiel 28 *Interesshalber wollen wir berechnen was passieren würde, wenn sich Ladungen im Knoten anhäuft. Wir nehmen an, dass sich eine Ladungsmenge von 1 C in einem Knoten anhäuft. Da das Netzwerk nach außen ungeladen bleiben muss, nehmen wir an, dass sich an einem anderen Knoten die Ladung -1 C anhäuft. Wir wollen annehmen, dass die Knoten sich in einem Abstand von einem Meter befinden. Die Kraftwirkung zwischen diesen Ladungen wäre nach dem Coulomb'schen Gesetz*

$$F = \frac{1}{4\pi\epsilon_0} \frac{Q_1 Q_2}{r^2} = \frac{1}{4\pi \times 8.854 \times 10^{-12}} = 8.89 \times 10^9 \text{ N}$$

Das entspricht dem Gewicht von rund 300.000 LKW-Zügen!

Eine Masche in einem elektrischen Netzwerk ist ein geschlossener Pfad, der ausgehend von einem Knoten des Netzwerks über mehrere Bauelemente zum Ausgangsknoten zurückkehrt. Die Kirchhoff'sche Maschenregel besagt, dass die Summe aller Spannungen in einer Masche Null sein muss.

$$\sum U = 0 \tag{4.3}$$

Die Zählrichtung der Spannungen ist entsprechend Abbildung 4.6 beliebig wählbar, $-U + u_1 - u_4 + u_5 = 0$.

Die Kirchhoff'sche Maschenregel folgt aus dem Satz der Erhaltung von der Energie.

Abbildung 4.7 zeigt drei in Serie geschaltete Bauelemente A,B,C, durch die der Strom i fließt. Die in den Bauelementen umgesetzte Leistung ist gegeben durch $p_a = -u_a i$, $p_b = u_b i$, $p_c = u_c i$. Die Summe der Leistungen im Netzwerk muss nach dem Energieerhaltungssatz Null sein, $p_a + p_b + p_c = 0$ und wir erhalten $-u_a i + u_b i + u_c i = 0$, woraus

³Im elektrischen Leiter stehen die Ladungen in Form von Elektronen zu Verfügung. Die Ladung eines Elektrons beträgt $-1.602 \times 10^{-19} \text{ C}$.

Größe	Einheit
Ladung (Q)	Coulomb (C)

Die Kirchhoff'sche Maschenregel:
Die Summe aller Spannungen in einer Masche ist gleich Null.

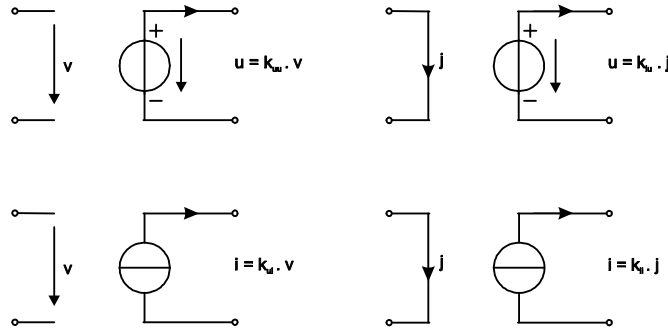


Abbildung 4.8: Schaltzeichen für gesteuerte Quellen

folgt, dass $-u_a + u_b + u_c = 0$, was aber nichts anderes als die Summe der Spannungen in der Masche ist. Es ist zu beachten, dass bei der Aufstellung der Kirchhoff'schen Maschengleichungen die Zählpfeilsysteme für Erzeuger und Verbraucher eingehalten werden, bei Verbrauchern zeigen Strom und Spannung in dieselbe Richtung, bei Erzeugern in die entgegengesetzte Richtung. Die Erzeuger liefern also eine »negative Leistung«.

4.3 Ideale Bauelemente

Bauelemente mit zwei Anschlusspunkten (z.B. R, L, C, U, I) bezeichnet man in der Elektrotechnik als Zweipole, Bauelemente mit vier Anschlusspunkten (z.B. gesteuerte Quellen) nennt man Vierpole.

4.3.1 Gesteuerte Quellen

Neben den unabhängigen Spannungs- und Stromquellen nach Abbildung 4.3 gibt es noch gesteuerte Spannungs- und Stromquellen. Abbildung 4.8 zeigt die Schaltsymbole der gesteuerten Quellen, wobei wir spannungs- und stromgesteuerte Spannungsquellen und spannungs- und stromgesteuerte Stromquellen unterscheiden müssen. Die gesteuerten Quellen haben vor allem im Bereich der Verstärkertechnik und analogen Signalverarbeitung Bedeutung und wir werden in diesen Kapiteln noch darauf zurückkommen.

4.3.2 Idealer Widerstand

Ohm'sches Gesetz:

Der Spannungsabfall (U) am Widerstand (R) ist direkt proportional zum durchfließenden Strom (I).

$$U = R \cdot I$$

Bei idealen Widerständen ist der Spannungsabfall (= die entstandene Spannung) am Widerstand proportional zum Strom durch den Widerstand.

$$U = R \cdot I \quad (4.4)$$

Diesen Zusammenhang nennt man Ohm'sches Gesetz, die Proportionalitätskonstante R – der ohmsche Widerstand – wird in Ohm (Ω) gemessen. Der ohmsche Widerstand eines zylindrischen Leiters ist abhängig von der Länge L und dem Querschnitt A des Leiters, sowie dem Material des Leiters, ausgedrückt durch die Materialkonstante ρ . Je länger der Leiter desto größer der Widerstand (daher lange Zuleitungen vermeiden), je kleiner der Querschnitt desto größer ist der Widerstand (daher bei Starthilfekabeln keinen Klingeldraht verwenden).

$$R = \rho \frac{L}{A} \quad (4.5)$$

Der **Widerstand** eines homogenen Zylinders hängt von der **Länge**, dem **Querschnitt** und dem verwendeten **Material** ab.

Den Proportionalitätsfaktor (die Materialkonstante) nennt man spezifischer Wider-

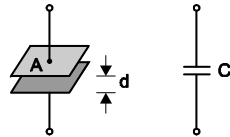


Abbildung 4.9: Kapazität und Schaltzeichen

stand, er beträgt bei Kupfer (nach Silber dem besten elektrischen Leiter) $\rho_{Cu} = 1.72 \times 10^{-8} \Omega\text{m}$, bei Glas $\rho_{Glas} = 1 \times 10^{12} \Omega\text{m}$.⁴

Dieses mathematische Modell des idealen Widerstands berücksichtigt die geometrischen Abmessungen L, A und die Materialkonstante ρ . Der Einfluss der Temperatur des Leiters auf den elektrischen Widerstand wird vernachlässigt.

Anmerkung 29 *Mathematische Modelle der Physik sind nicht beweisbar im Sinne eines mathematischen Beweises. Die Überprüfung kann nur durch das Experiment erfolgen!*

Die im Widerstand umgesetzte Leistung ist

$$P = U \cdot I \quad \text{bzw.} \quad p(t) = u(t)i(t) \tag{4.6}$$

Unter der Annahme eines konstanten Widerstandes R erhalten wir nach Einsetzen in das Ohm'schen Gesetz

$$P = \frac{U^2}{R} \quad \text{bzw.} \quad p(t) = \frac{u^2(t)}{R} \tag{4.7}$$

Anmerkung 30 *Betrachtet man $u(t)$ in (4.7) als die physikalische Ausprägung eines Signals (was in technischen Anwendungen meistens der Fall ist), so sieht man, dass die Definition der Signalleistung in Abschnitt 3.4 als proportional zum Quadrat der Amplitude vernünftig ist.*

P, U, I sind zeitlich unveränderliche Größen (Gleichspannung und Gleichstrom), p, u, i sind zeitlich veränderliche Größen.⁵

4.3.3 Kapazität

Während Widerstände Energie nur in Wärme umsetzen, können Kondensatoren (und Spulen) elektrische Energie speichern. Kondensatoren sind technische Bauelemente, ideale Kondensatoren bezeichnen wir als Kapazitäten.

Die Abbildung 4.9 zeigt eine Kapazität und ihr Schaltungssymbol. Elektrisch leitende Platten der Fläche A sind im Abstand d durch eine nichtleitendes Material voneinander isoliert. Die Kapazität dieser Anordnung beträgt

$$C = \epsilon\epsilon_0 \frac{A}{d} \tag{4.8}$$

sie wird größer, wenn die Fläche größer wird oder wenn der Abstand kleiner wird. Die Proportionalitätskonstante ϵ nennt man Dielektrizitätskonstante⁶, die materialabhängig und dimensionslos ist. $\epsilon_0 = 8.85 \cdot 10^{-12} \frac{\text{As}}{\text{Vm}}$ ist die Influenzkonstante, die eine Naturkonstante ist. Die Kapazität wird in Farad (F) gemessen.

Wie Abbildung 4.10 zeigt, sammeln sich Ladungen an den gegenüberliegenden Platten der Kapazität an, wenn Strom fließt. Ein Flüssigkeitsäquivalent der Kapazität wäre

Größe **Einheit**
Kapazität (C) ... Farad (F)

Eine **Kapazität** speichert Energie im elektrischen Feld.

⁴Der Widerstandsunterschied zwischen dem Leiter Kupfer und dem Nichtleiter Glas beträgt 20 Zehnerpotenzen, die technologisch beherrscht werden können: Einer der Gründe für den besonderen Stellenwert der Elektrotechnik in der Informationstechnik.

⁵In der Elektrotechnik werden unveränderliche Größen in der Regel mit Großbuchstaben, veränderliche Größe mit Kleinbuchstaben bezeichnet.

⁶Die Dielektrizitätskonstante ist Vakuum (1), Glas (5-10) und Wasser (81), überstreicht also im Vergleich zum spezifischen Widerstand nur einen sehr kleinen Bereich.

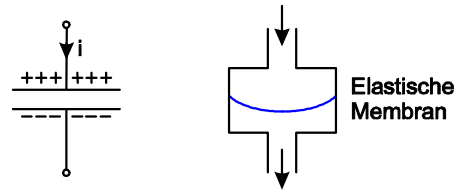


Abbildung 4.10: Mechanisches Analogon zur Kapazität

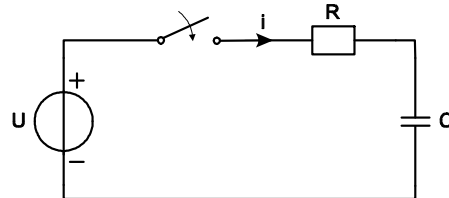


Abbildung 4.11: Aufladen der Kapazität

ein Behälter mit einer Membran. Die Flüssigkeit strömt in den Behälter und lenkt die Membran aus. Wenn der Druck der Membran dem Druck des Wassers entspricht, hört der Fluss auf. Nimmt man den Druck des Wassers weg, dann wird die Membran das Wasser aus dem Behälter drücken, bis sie in der Ausgangslage ist. Die Membran »speichert« Energie und gibt sie wieder ab.

Ähnlich verhält sich die Kapazität. Es fließt Strom, der Kondensator lädt sich auf⁷. Wenn der »Druck« an der Kapazität soweit angestiegen ist, dass er dem »Wasserdruck« entspricht, ist der Kondensator auf die am Bauelement angelegte Spannung aufgeladen und der Stromfluss hört auf. Verbindet man die Platten der aufgeladenen Kapazität mit einem Widerstand, fließt die Ladung von den Platten über den Widerstand ab und die Kapazität entlädt sich zum Ausgangszustand (Spannung Null). Die in der Kapazität gespeicherte Energie wird an den Widerstand abgegeben.

Die im Kondensator gespeicherte Ladung ist proportional zur Spannung, auf die der Kondensator aufgeladen ist. Die Kapazität C ist die Proportionalitätskonstante.

$$Q = C \cdot U \quad (4.9)$$

Der elektrische Strom ist definiert als Ladung pro Zeit⁸, der Zusammenhang zwischen Strom und Spannung an der Kapazität ist daher

$$i_c = \frac{dq}{dt} = C \frac{du}{dt} \quad (4.10)$$

Die in einer Kapazität gespeicherte elektrische Energie ist

$$W_c = \frac{1}{2} CU^2 = \frac{1}{2} QU \quad (4.11)$$

Beispiel 31 Als Beispiel berechnen wir den Aufladevorgang einer Kapazität wie in Abbildung 4.11 dargestellt. Wir nehmen an, dass die Kapazität ungeladen ist. Zum Zeitpunkt $t = 0$ schließen wir den Schalter. Zur Berechnung des Stroms wenden wir die Kirchhoff'sche Knotenregel auf den durch R und C gebildeten Knoten an und erhalten aus

$$i_R - i_C = 0. \quad \frac{u_R(t)}{R} - C \frac{du_c(t)}{dt} = 0$$

Für die Spannung u_R am Widerstand können wir aufgrund der Maschenregel $u_R(t) =$

⁷Der Stromfluss durch den Kondensator, insbesondere durch das Vakuum, lässt sich nur mit dem Verschiebestrom der Maxwell'schen Gleichungen erklären.

⁸Die elektrische Ladung ist nur sehr schwer messbar, man verwendet daher im SI-Einheitensystem Länge, Zeit, Masse und elektrischen Strom als Grundgrößen.

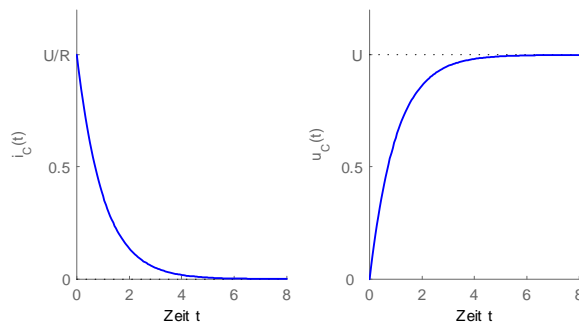


Abbildung 4.12: Strom- und Spannungsverlauf beim Laden einer Kapazität

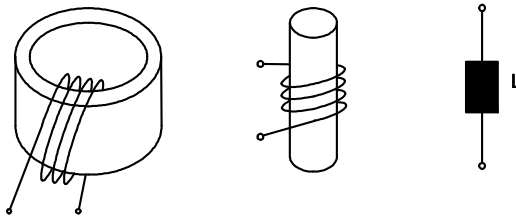


Abbildung 4.13: Schaltungssymbol der Induktivität

$U - u_c(t)$ einsetzen und erhalten nach Umformung

$$C \frac{du_c(t)}{dt} + \frac{u_c(t) - U}{R} = 0 \implies RC \frac{du_c(t)}{dt} + u_c(t) = U \tag{4.12}$$

Gleichung (4.12) ist eine Differentialgleichung 1. Ordnung mit der Lösung

$$u_c(t) = U(1 - e^{-\frac{t}{RC}}) \tag{4.13}$$

Den zeitlichen Verlauf von Spannung und Strom an der Kapazität zeigt Abbildung 4.12.

4.3.4 Induktivität

Spulen werden durch Aufwickeln von leitenden Drähten wie in Abbildung 4.13 gezeigt hergestellt.

Der durch den Draht fließende Strom erzeugt ein magnetisches Feld das mit der Spule verbunden ist. Die ideale Spule nennen wir Induktivität. Die Induktivität wird in Henry (H) gemessen. Bei der Induktivität besteht zwischen Strom und Spannung der Zusammenhang

$$u_L = L \cdot \frac{di_L}{dt} \tag{4.14}$$

Der Spannungsabfall an der Induktivität ist proportional der Stromänderung. Der Proportionalitätsfaktor L kann über – in der Regel komplizierte – Berechnungen des magnetischen Feldes bzw. magnetischen Flusses berechnet werden.

Die Flüssigkeitsanalogie für die Induktivität ist die Trägheit der Flüssigkeitssäule in einer reibungslosen Leitung konstanten Durchmessers. Der Druckunterschied zwischen den Enden der Leitung entspricht der Spannung, die Durchflussmenge dem Strom, Zu- oder Abnahme (Beschleunigung) der Durchflussmenge entsprechen der Stromänderung. Eine Druckdifferenz zwischen den Enden der Leitung existiert nur dann, wenn die Durchflussmenge zu- oder abnimmt.

Die in einer Spule gespeichert Energie ist

$$W_L = \frac{1}{2}LI^2 \tag{4.15}$$

Größe **Einheit**
Induktivität (L) Henry (H)

Eine **Induktivität** speichert Energie im magnetischen Feld.

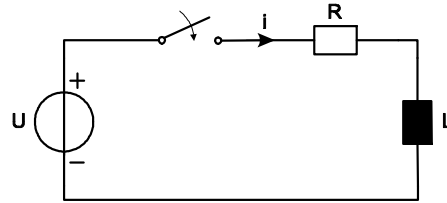


Abbildung 4.14: Einschalten einer Induktivität

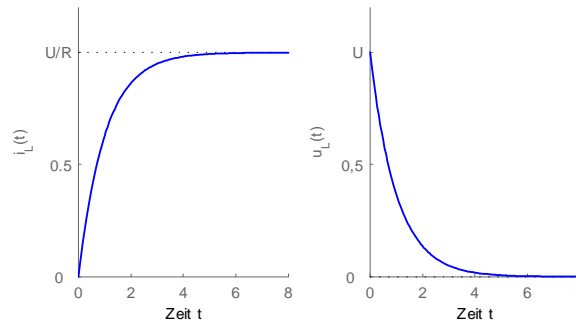


Abbildung 4.15: Strom- und Spannungsverlauf beim Einschalten einer Induktivität

Beispiel 32 Als Beispiel berechnen wir den Einschaltstrom einer Induktivität nach Abbildung 4.14. Wir wenden die Kirchhoff'sche Maschenregel an und erhalten aus $-U + u_R + u_L = 0$

$$-U + i(t)R + L \frac{di_L(t)}{dt} = 0 \quad \Rightarrow \quad \frac{L}{R} \frac{di_L(t)}{dt} + i(t) = \frac{U}{R} \quad (4.16)$$

Die Differentialgleichung (4.16) hat die Form von (4.12) und wir erhalten als Lösung

$$i_L = \frac{U}{R} (1 - e^{-\frac{R}{L}t}) \quad (4.17)$$

Den zeitlichen Verlauf von Spannung und Strom an der Induktivität zeigt Abbildung 4.15

4.3.5 Spannung und Strom an L und C bei Schaltvorgängen

Aus den Abbildungen 4.12 und 4.15 können wir folgendes Verhalten erkennen:

Die Energie (und damit auch die **Spannung**) einer **Kapazität** kann sich **nicht sprunghaft ändern**.

- Beim Einschalten der Kapazität ist die Spannung an der Kapazität unmittelbar nach dem Einschalten Null, da sich die Kapazität erst aufladen muss (Das elektrische Feld der Kapazität kann sich nicht sprunghaft ändern, sondern muss sich erst aufbauen.): Die Energie (des elektrischen Felds) von C kann nicht sprunghaft zunehmen. Unmittelbar nach dem Einschalten verhält sich C daher wie ein Kurzschluss und der Einschaltstrom zum Zeitpunkt t_{0+} ist $i_{0+} = U/R$. Wenn C geladen ist, dann fließt kein Strom und die Spannung an C wird U . Die Energie der Kapazität steckt im elektrischen Feld.

Die Energie (und damit auch der **Strom**) einer **Induktivität** kann sich **nicht sprunghaft ändern**.

- Beim Einschalten der Induktivität ist der Strom unmittelbar nach dem Einschalten Null, da sich die Induktivität erst »aufladen« muss (Das magnetische Feld der Induktivität muss sich erst aufbauen.): Die Energie (des magnetischen Felds) von L kann nicht sprunghaft zunehmen. Unmittelbar nach dem Einschalten verhält sich L daher wie eine Unterbrechung. Die Spannung an L zum Zeitpunkt t_{0+} ist $u_{L0+} = U$. Wenn L »geladen« ist, wird die Spannung an L Null und es fließt der Strom U/R . Die Energie der Induktivität steckt im magnetischen Feld.

4.3.6 Schreibweise der Netzwerkgleichungen

Für die auftretenden Differentialquotienten führen wir folgende Schreibweise ein

$$\frac{d^i}{dt} = s^i \quad \text{mit} \quad s = \sigma + j\omega \quad (4.18)$$

Damit erhalten wir folgende Beziehungen, wobei wir für die Variablen der Funktionen von s Großbuchstaben verwenden

$$u(t) = L \frac{di(t)}{dt} \Rightarrow U(s) = LsI(s) \Rightarrow I(s) = \frac{U(s)}{Ls} \quad (4.19)$$

$$i(t) = C \frac{du_c(t)}{dt} \Rightarrow I(s) = CsU(s) \Rightarrow U(s) = \frac{I(s)}{Cs} \quad (4.20)$$

Während bei zeitlich veränderlichen Größen zur Darstellung der Netzwerkvariablen Kleinbuchstaben verwendet werden, werden für die Darstellung in s Großbuchstaben verwendet. Den Parameter s werden wir noch ausführlich erörtern. Vorläufig lässt er sich folgendermaßen erklären:

- Als Operator für Differentiation und Integration im Sinne der Gleichung (4.18). Eine in s geschriebene Gleichung kann jederzeit in eine Differentialgleichung umgewandelt werden.
- Als unabhängige Variable (komplexe Frequenz). Die in s geschriebene algebraische Gleichung stellt die Laplace-Transformierte⁹ der entsprechenden Differentialgleichungen für energielosen Anfangszustand dar. Zum Auffinden der Zeitfunktion ist die Rücktransformation in den Zeitbereich (inverse Laplace-Transformation) erforderlich.
- Als Frequenzvariable für den Erregeranteil der Netzwerkantwort bei der Erregung mit der Zeitfunktion e^{st} . Für $s = j\omega$ wird der Erregeranteil bei stabilen Netzwerken zum stationären Anteil und die Gleichungen entsprechen vollständig den Beziehungen der komplexen Wechselstromrechnung.

Bei Netzwerken die aus R, L, C , wie in den Abbildungen 4.11 und 4.14, aufgebaut sind, entstehen (lineare) Differentialgleichungen (mit konstanten Koeffizienten) für die gesuchten Ströme und Spannungen. Die Lösung dieser Differentialgleichungen ergibt den zeitlichen Verlauf von Strömen und Spannungen wie beispielsweise in (4.13) und (4.17) gezeigt.

Die direkte Lösung der Differentialgleichungen – *Lösung im Zeitbereich* genannt – wird für die Berechnung linearer elektrischer Netzwerke (linearer Systeme) kaum angewendet, da sie zu umständlich und aufwändig ist¹⁰. Man bedient sich zur Lösung der Differentialgleichungen der Laplace-Transformation. Die Laplace-Transformation stellt einen Übergang vom *Zeitbereich* (Originalbereich in der Mathematik) in den *Frequenzbereich* (Bildbereich in der Mathematik) dar, man spricht daher von der *Lösung im Frequenzbereich*. Neben der Lösung der Differentialgleichungen bietet die Frequenzbereichsdarstellung noch weitere vorteilhafte Möglichkeiten zur Darstellung von Systemeigenschaften, wie zum Beispiel bei der Untersuchung von Stabilitätseigenschaften eines Netzwerks.

Selbst bei einfachen Erregungsfunktionen wie Impuls oder Sprung löst man die Differentialgleichungen selten direkt, sondern benutzt die Laplace-Transformation. Die Laplace-Transformation ist die wichtigste Methode zur Berechnung von Differentialgleichungen die bei Netzwerken (bzw. ganz allgemein bei LTI-Systemen) auftreten.

Die bei der Modellierung elektrischer Netzwerke entstandenen Differentialgleichungen können mit Hilfe der **Laplace-Transformation** einfacher gelöst werden.

⁹Zum Unterschied von Funktionen, die einer Größe x eine zweite Größe y in eindeutiger Weise zuordnen, sind Transformationen mathematische Operationen die aus einer gegebenen Funktion (Originalfunktion) eine neue Funktion (Bildfunktion) erzeugen. Der Vorteil der Transformationen ist die einfachere Lösbarkeit eines Problems im Bildbereich. Z.B. werden Differentialgleichungen durch die Laplace-Transformation zu algebraischen Gleichungen, die im Bildbereich einfach gelöst werden. Die Lösung im Originalbereich wird durch Rücktransformation aus dem Bildbereich gefunden.

¹⁰Eine Ausnahme stellt die Beschreibung von Systemen im Zustandsraum dar.

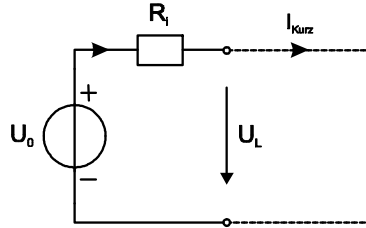


Abbildung 4.16: Ersatzschaltbild reale Spannungsquelle

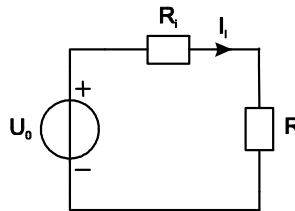


Abbildung 4.17: Belastete Quelle

4.4 Präzisere Modelle der Bauelemente

4.4.1 Spannungsquelle mit Innenwiderstand

Reale Quellen haben einen endlichen **Innenwiderstand**.

Legt man eine Last an eine reale Spannungsquelle an, dann bleibt die Spannung der Quelle bei zunehmendem Strom nicht konstant, sondern nimmt ab. Ein Ersatzschaltbild, das das Verhalten der Quelle besser modelliert, ist in Abbildung 4.16 gezeigt.

Der maximale Strom der Spannungsquelle nach 4.16 beträgt $I_{kurz} = U_0/R_i$. R_i nennt man den Innenwiderstand der Quelle.

Beispiel 33 Wir fragen uns, bei welchem Lastwiderstand die Quelle die maximale Leistung abgibt. Der Strom im Schaltkreis in Abbildung 4.17 ist $I = U_0/(R_i + R_l)$, die Leistungsaufnahme der Last ist $P_l = I_l^2 R$. Wir setzen ein und erhalten

$$P_l = \frac{U_0^2 R_l}{(R_i + R_l)^2} \quad (4.21)$$

Die Leistung ist eine Funktion des Lastwiderstands R_l . Das Maximum erhalten wir durch Nullsetzen der ersten Ableitung

$$\frac{dP_l}{dR_l} = \frac{U_0^2 (R_i + R_l)^2 - 2U_0^2 R_l (R_i + R_l)}{(R_i + R_l)^4} = 0 \quad (4.22)$$

Wir lösen diese Gleichung und erhalten

$$R_i = R_l \quad (4.23)$$

Die Quelle gibt die höchste Leistung $P_{l\max} = U_0^2/4R_i$ ab, wenn sie mit ihrem Innenwiderstand¹¹ abgeschlossen ist.

Bei der Netzwerkanalyse ist es gelegentlich vorteilhaft, Spannungsquellen in Stromquellen umzuwandeln, wie Abbildung 4.18 zeigt.

(Die Abbildung 4.18 habe ich neu gezeichnet. Ich habe den Lastwiderstand und Pfeile für U dazugezeichnet.)

Der Zusammenhang zwischen Spannung bzw. Strom der Quellen ist

$$I_0 = \frac{U_0}{R_i} \quad (4.24)$$

¹¹ Wenn der Innenwiderstand der Quelle komplex ist, muss für Anpassung mit dem konjugiert komplexen Widerstand $Z_l = Z_i^*$ abgeschlossen werden.

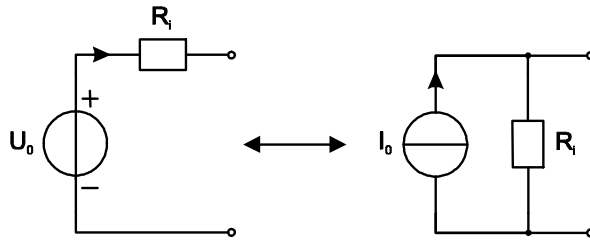


Abbildung 4.18: Umwandlung Spannungs- in Stromquelle

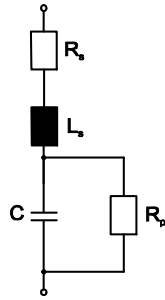


Abbildung 4.19: Ersatzschaltbild Kondensator

I_0 ist der Kurzschlussstrom der Spannungsquelle, U_0 ist die Leerlaufspannung der Spannungsquelle.

Die beiden Modelle einer Quelle sind äquivalent und können daher beliebig vertauscht werden. Sie haben nach außen immer das gleiche Verhalten. Die Ausgangsspannung U ist (unabhängig vom gewählten Modell) eine Funktion des angeschlossenen Lastwiderstandes R_l (U_0, I_0 und R_i sind Konstanten)

Die **Spannungsquelle** und die **Stromquelle** (Abb. 4.18) sind **äquivalente Modelle**.

$$U = U_0 \cdot \frac{R_l}{R_i + R_l} = I_0 \cdot \frac{R_i R_l}{R_i + R_l} \tag{4.25}$$

Man sieht, dass bei $R_l = 0$ (Kurzschluss) die Spannung U ebenfalls Null wird und bei $R_l \rightarrow \infty$ (Leerlauf) $U \rightarrow U_0$.¹²

4.4.2 Ersatzschaltbild für den Kondensator

Das einfache Schaltbild der Kapazität nach Abbildung 4.9 ist in vielen Fällen nicht ausreichend und man verwendet für den Kondensator kompliziertere Ersatzschaltbilder wie in Abbildung 4.19 gezeigt.

In diesem Ersatzschaltbild modelliert R_s den Zuleitungswiderstand des Kondensators. Da jeder Strom ein Magnetfeld führt, müssen wir auch dieses Feld durch eine Induktivität L_s modellieren. Schließlich hat jedes Dielektrikum einen endlichen ohmschen Widerstand, der durch R_p modelliert wird.

4.4.3 Ersatzschaltbild für die Spule

Das einfache Schaltbild der Induktivität nach Abbildung 4.13 ist in den meisten Fällen für Spulen nicht ausreichend und man verwendet komplizierter Ersatzschaltbilder wie in Abbildung 4.4.3 gezeigt.

Induktivitäten werden durch Aufwickeln von Draht gebildet, wie in Abbildung 4.13 schematisch gezeigt. Dieser Draht hat immer einen ohmschen Widerstand, der durch R_s modelliert wird. Zwischen den Wicklungen der Induktivität entsteht ein elektrisches Feld, das durch C_p modelliert wird. Die Wicklungen der Induktivität sind häufig auf einem

¹²Die Anwendung der Regel von de L'Hospital ist hier nötig.

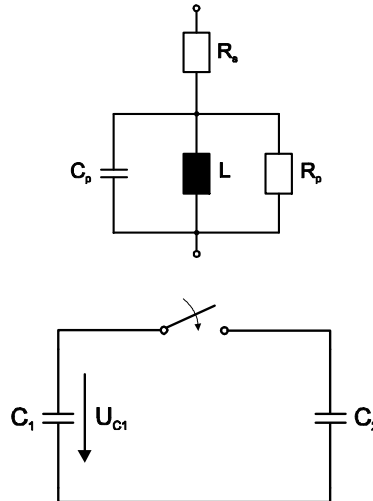


Abbildung 4.20: Ladungsverteilung auf zwei Kapazitäten

Reale Kondensatoren und Spulen werden durch kompliziertere Netzwerke von idealen Bauelementen modelliert.

Ferrit-Stab oder -Ring aufgebracht. Durch Induktion fließen im Stab oder Ring Ströme, die zu Verlusten führen, die durch R_p modelliert werden.

Beispiel 34 Bei der Beschreibung elektrischer Netzwerke mit Hilfe von idealen Bauelementen muss man immer in Erinnerung behalten, dass es sich um vereinfachte Modelle handelt, die nur die wesentlichen Zusammenhänge beschreiben. Das bekannte Beispiel in Abbildung 4.20 soll das verdeutlichen. Die Kapazität C_1 sei auf die Spannung U_{C_1} aufgeladen, ihre Energie ist daher $W_{C_1} = \frac{1}{2}C_1U_1^2$. Schließen wir den Schalter, so findet ein Ausgleich der Ladungen zwischen den beiden Kapazitäten statt, die Gesamtladung muss aber erhalten bleiben, $Q_{ges} = Q_1 + Q_2$. Nach dem Schließen des Schalters liegen die beiden Kapazitäten parallel und die Gesamtkapazität beträgt daher $C_{ges} = C_1 + C_2$. An den Kapazitäten stellt sich die Spannung $U = Q_{ges}/C_{ges}$ ein. Der Einfachheit halber nehmen wir an, dass die beiden Kapazitäten gleich groß sind, $C_1 = C_2 = C$. Wir erhalten dann für die Energie vor dem Schließen des Schalters $W_{vor} = \frac{1}{2}CU_1^2$. Nach dem Schließen des Schalters wird die Spannung an den Kapazitäten $U_{nach} = U_1/2$, die Energie an einer Kapazität also $W_{C_1} = \frac{1}{2}C_1(U_1/2)^2 = \frac{1}{2}C_1 \frac{U_1^2}{4}$, die Energie an beiden Kapazitäten ist daher doppelt so groß, also

$$W_{C_1+C_2} = \frac{1}{2}C_1 \frac{U_1^2}{2} = \frac{W_{vor}}{2} \quad (4.26)$$

Die Hälfte der Energie ist nach dem Schließen des Kondensators verschwunden! Das für diese Überlegungen verwendete Modell führt offensichtlich zu einem Widerspruch. Man müsste daher die Leitung zwischen den Kapazitäten und den Schalter genauer modellieren, um zu einer Übereinstimmung zwischen Rechnung und physikalischer Wirklichkeit zu kommen.

4.5 Zusammenfassung

Die genaue Analyse elektrischer Netzwerke ist nur über die Berechnung der mit dem Netzwerk verbundenen elektrischen und magnetischen Felder möglich. Diese Feldberechnung ist in der Regel nicht durchführbar, aber auch nicht notwendig. Man kann sich für die Untersuchung von Netzwerken auf die integralen Feldgrößen Spannung und Strom beschränken, die elektrischen Felder werden durch die Kapazitäten, die magnetischen Felder durch die Induktivitäten repräsentiert. Ströme und Spannungen im Netzwerk ergeben sich durch den Knoten- und Maschensatz, sowie durch die Bauelementbeziehungen. Damit gelangt man zu einer eigenständigen Theorie der elektrischen Netzwerke.

Es darf nie vergessen werden, dass die elektrischen Netzwerke nur eine Annäherung an die physikalische Wirklichkeit darstellen. Durch entsprechend aufwändige Ersatzschaltungen kann aber das Verhalten eines Netzwerks in der gewünschten Genauigkeit beschrieben werden.

Kapitel 5

Elektrische Netzwerke

Inhalt

5.1	Einfache Netzwerke	63
5.2	Topologische Netzwerkbeschreibung	66
5.2.1	Schleifenanalyse	67
5.2.2	Knotenanalyse	67
5.2.3	Beispiele	68
5.3	Zusammenfassung	70

Ströme und Spannungen in elektrischen Netzwerken berechnet man mit Hilfe der Kirchhoff'schen Gleichungen und der Bauelementbeziehungen. Bei einfachen Netzwerken kann das mit Handrechnung geschehen, für Netzwerke mit vielen Komponenten wird man sich entsprechender Netzwerkanalyseprogramme bedienen.

5.1 Einfache Netzwerke

Serienschaltungen von R , L und C nach Abbildung 5.1 lassen sich einfach über den Strom durch die Serienschaltung analysieren. Mehrere in Serie geschaltete Widerstände, Induktivitäten oder Kapazitäten lassen sich dabei durch ein äquivalentes Bauelement (das Ersatzbauelement) modellieren.

Ein Strom i durch die Serienschaltung von Widerständen erzeugt an jedem Widerstand einen Spannungsabfall

Serienschaltung:

$$R_{ers} = \sum R$$

$$L_{ers} = \sum L$$

$$\frac{1}{C_{ers}} = \sum \frac{1}{C}$$

$$iR_1 + iR_2 + \dots + iR_n = u \quad (5.1)$$

$$i(R_1 + R_2 + \dots + R_n) = u$$

$$R_{ers} = R_1 + R_2 + \dots + R_n$$

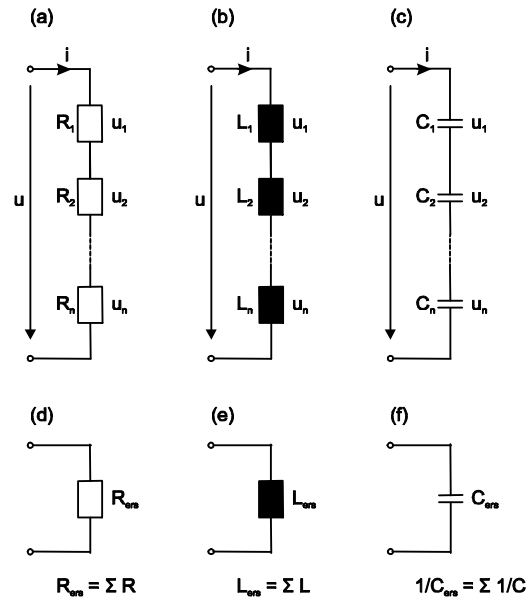
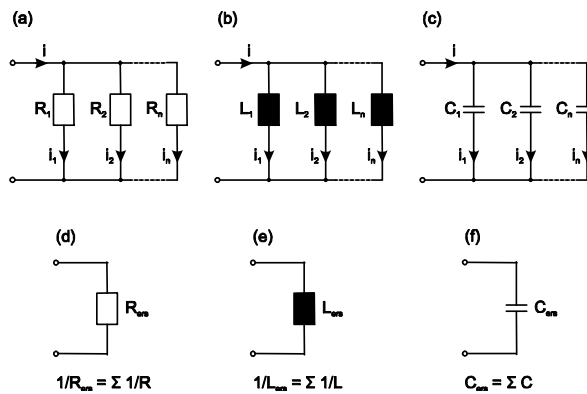
Der Ersatzwiderstand R_{ers} ist gleich der Summe aller Einzelwiderstände. Der gleiche Ansatz liefert für die Ersatzinduktivität einer Serienschaltung L_{ers} ist die Summe aller Einzelinduktivitäten.

$$L_1 \frac{di}{dt} + L_2 \frac{di}{dt} + \dots + L_n \frac{di}{dt} = u \quad (5.2)$$

$$(L_1 + L_2 + \dots + L_n) \frac{di}{dt} = u$$

$$L_{ers} = L_1 + L_2 + \dots + L_n$$

Bei Kapazitäten gilt $i = Cdu/dt$, daraus folgt $u = \frac{1}{C} \int idt$. Für die Serienschaltung von Kapazitäten erhalten wir

Abbildung 5.1: Serienschaltung von R , L und C Abbildung 5.2: Parallelschaltung von R , L , C

$$u = \frac{1}{C_1} \int idt + \frac{1}{C_2} \int idt + \dots + \frac{1}{C_n} \int idt \quad \Rightarrow \quad \frac{1}{C_{ers}} = \frac{1}{C_1} + \frac{1}{C_2} + \dots + \frac{1}{C_n} \quad (5.3)$$

Der Kehrwert der Ersatzkapazität der Serienschaltung von Kapazitäten ist gleich der Summe der Kehrwerte der Einzelkapazitäten.

Parallelschaltung:

$$\frac{1}{R_{ers}} = \sum \frac{1}{R}$$

$$\frac{1}{L_{ers}} = \sum \frac{1}{L}$$

$$C_{ers} = \sum C$$

Parallelschaltungen von R, L und C nach Abbildung 5.2 lassen sich einfach über die Spannung an der Parallelschaltung analysieren. Mehrere parallel geschaltete Widerstände, Induktivitäten oder Kapazitäten lassen sich dabei durch ein äquivalentes Bauelement (das Ersatzbauelement) modellieren.

Eine Spannung u an der Parallelschaltung von Widerständen bewirkt einen Strom

$$i = i_1 + i_2 + \dots + i_n = \frac{U}{R_1} + \frac{U}{R_2} + \dots + \frac{U}{R_n} \quad \Rightarrow \quad \frac{1}{R_{ers}} = \frac{1}{R_1} + \frac{1}{R_2} + \dots + \frac{1}{R_n} \quad (5.4)$$

Der Kehrwert des Ersatzwiderstands der Parallelschaltung von Widerständen ist gleich der Summe der Kehrwerte der Einzelwiderstände. Bei Induktivitäten liefert der gleiche Ansatz, dass der Kehrwert der Ersatzinduktivität der Parallelschaltung von Induktivitäten ist gleich der Summe der Kehrwerte der Einzelinduktivitäten ist.

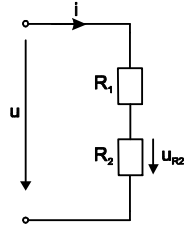


Abbildung 5.3: Spannungsteiler

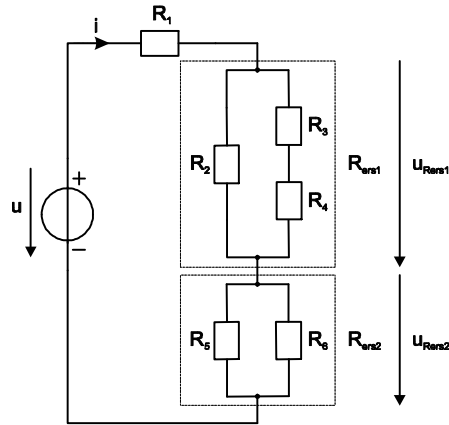


Abbildung 5.4: Ersatzwiderstände im Netzwerk

Bei Kapazitäten ist

$$i = i_1 + i_2 + \dots + i_n = C_1 \frac{du}{dt} + C_2 \frac{du}{dt} + \dots + C_n \frac{du}{dt} \Rightarrow C_{ers} = C_1 + C_2 + \dots + C_n \quad (5.5)$$

Die Ersatzkapazität bei Parallelschaltung ist daher die Summe der Einzelkapazitäten.

Beispiel 35 Als Beispiel berechnen wir einen Spannungsteiler nach Abbildung 5.3. Der Strom durch den Spannungsteiler ist $i = u/R_{ers} = u/(R_1 + R_2)$, dieser Strom bewirkt am Widerstand R_2 den Spannungsabfall $u_2 = R_2 i$, die Spannung u_2 ist daher

$$u_2 = \frac{R_2}{R_1 + R_2} u \quad (5.6)$$

Beispiel 36 Bei komplizierteren Widerstandsnetzwerken können wir Ströme und Spannungen häufig durch Zusammenfassen von Serien- und Parallelschaltungen berechnen. An Hand von Abbildung 5.4 wird die Vorgangsweise bei der Ermittlung aller Ströme und Spannungen im Netzwerk beschrieben. Für die Berechnung von i müssen wir den Ersatzwiderstand der Schaltung ermitteln, $i = u/R_{ers}$.

$$R_{ers} = R_1 + R_{ers1} + R_{ers2} \quad (5.7)$$

$$\frac{1}{R_{ers1}} = \frac{1}{R_2} + \frac{1}{R_3 + R_4} \quad (5.8)$$

$$\frac{1}{R_{ers2}} = \frac{1}{R_5} + \frac{1}{R_6} \quad (5.9)$$

Wenn wir i berechnet haben, können wir die Spannung $u_{R1} = iR_1$ berechnen. Als nächstes berechnen die Spannung $u_{ers1} = iR_{ers1}$, damit haben wir die Spannung $u_{R2} = u_{ers1}$ ermittelt. Der Strom durch R_2 ist $i_{R2} = u_{R2}/R_2$. Die Spannung an den Widerständen R_3 und R_4 ermitteln wir über die Spannungsteilerregel. $u_{R3} = \frac{R_3}{R_3+R_4} u_{R2}$. Daraus

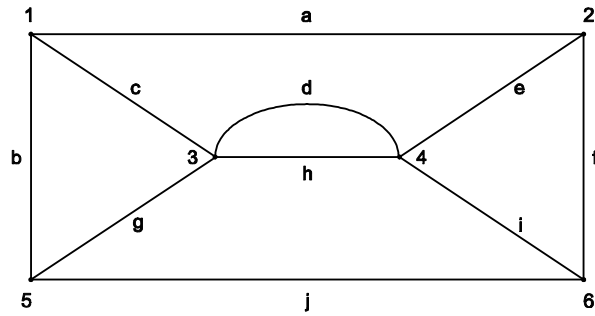


Abbildung 5.5: Netzwerkgraph N

berechnen wir $u_{R_4} = u_{R_2} - u_{R_3}$. Der Strom durch R_3 und R_4 ist gleich und beträgt $i_{R_4} = i_{R_3} = u_{R_3}/R_3$. Schließlich erhalten wir $u_{ers2} = u - u_{R_1} - u_{ers1} = u_{R_5} = u_{R_6}$, woraus wir die Ströme $i_{R_5} = u_{R_5}/R_5$ und $i_{R_6} = u_{R_6}/R_6$ berechnen können. Damit haben wir alle Ströme und Spannungen dieser Schaltung ermittelt.

Selbst bei Widerstandsnetzwerken ist dieses Verfahren aufwändig. Wenn ein Netzwerk Speicherelemente enthält und die Netzwerkgleichungen dadurch zu Differentialgleichungen werden, wird es völlig unübersichtlich und man bedient sich daher leistungsfähigerer Verfahren.

5.2 Topologische Netzwerkbeschreibung

Bei komplizierten Netzwerken ist nicht ohne weiteres zu erkennen, wie die Kirchhoff'schen Gleichungen anzuwenden sind, um die interessierenden Netzwerkgrößen mit einem möglichst geringen Rechenaufwand zu ermitteln. Aus der Struktur des Netzwerks lässt sich aber der Lösungsansatz systematisch ableiten.

Bei der systematischen Analyse von komplizierteren Netzwerken werden diese als **Grafen** dargestellt.

Dazu stellen wir das Netzwerk in Form eines *Grafen* dar. Dabei ersetzt man die Zweige des Netzwerks durch Kanten und Netzwerkknotten werden zu Knoten des Grafen. Man erhält einen Grafen mit z Kanten (Netzwerkzweige) und k Knoten. Der Graf in Abbildung 5.5 hat 6 Knoten und 10 Kanten. *Schleifen* sind beliebige geschlossene Wege im Grafen, eine *Masche* ist eine einfache Schleife, die von keinen anderen Kanten durchkreuzt wird. Durch Entfernen von Kanten entsteht ein *Teilgraf* des Grafen. Ein für die Netzwerkberechnung wichtiger *Teilgraf* ist der *Spannbaum*. Ein Spannbaum ist ein Teilgraf der alle Knoten des ursprünglichen Grafen enthält. Die Knoten des Baums sind alle miteinander verbunden, jedoch so, dass keine geschlossenen Schleifen auftreten. Im Allgemeinen gibt es mehrere Spann bäume¹ zu einem Grafen. Die entfernten Kanten des Grafen nennt man *Glieder*. Abbildung 5.6 zeigt einen Spannbaum zum Grafen der Abbildung 5.5, die Glieder sind strichliert dargestellt.

Zur vollständigen Beschreibung eines Netzwerks müssen alle Spannungen und Ströme berechnet werden, bei einem Netzwerk mit z Kanten treten daher $2z$ Unbekannte auf. Da für die Bauelemente die Beziehung zwischen Strom und Spannung bekannt ist, müssen nur z Gleichungen mit Hilfe der Kirchhoff'schen Gleichungen erstellt werden. Die Anzahl der Netzwerkvariablen ist meistens größer als erforderlich, beispielsweise ist in Abbildung 5.4 $u_{R_5} = u_{R_6}$. Es ist daher wichtig, die Unbekannten so zu wählen, dass man mit möglichst wenigen, unabhängigen Gleichungen auskommt. Während bei einfachen Netzwerken das Aufstellen unabhängiger Gleichungen einfach ist, kommt man bei komplizierten Netzwerken nur mit einer gewissen Systematik weiter.

¹Die Wahl des Baums erfolgt nach Gesichtspunkten der Zweckmäßigkeit. Bei der Schleifenanalyse ist es z. B. zweckmäßig die Spannungsquellen in unabhängige Kanten (Netzwerkzweige) zu legen. Dadurch kommt jede Quelle nur in einer Schleife vor und ist im Gleichungssystem nur in einer Gleichung vorhanden.

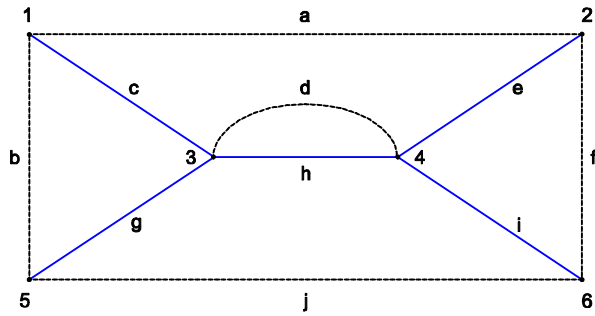


Abbildung 5.6: Spannbaum zum Graphen N

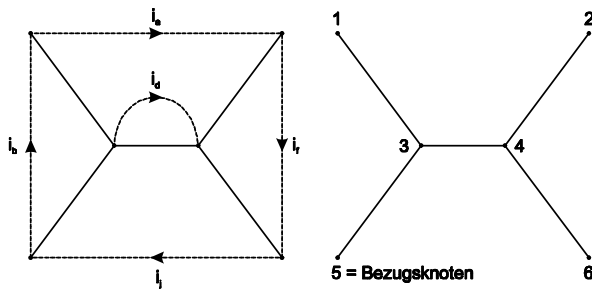


Abbildung 5.7: Baum und Glieder

5.2.1 Schleifenanalyse

Dazu gehen wir von einem Netzwerk aus, das in Spannbaum und Glieder zerlegt ist. In Abbildung 5.7 ist jedem Glied des Grafen ein Strom zugeordnet, der sich über den Baum schließt und eine eindeutige Schleife bildet. Die in den Gliedern eines Netzwerks fließenden Schleifenströme sind ein geeigneter Satz von Unbekannten zur vollständigen Berechnung des Netzwerks. Die Zahl der Unbekannten ist gleich der Zahl der Glieder $g = z - b = z - k + 1$, da die Zahl der Kanten eines Baums $b = k - 1$ ist. Kennt man die Schleifenströme, so kann man die Zweigströme aus Linearkombinationen der Schleifenströme ermitteln. Das zugehörige Berechnungsverfahren nennt man *Schleifenanalyse*².

Wir haben bisher angenommen, dass Netzwerke nur aus R, L, C und (ungesteuerten) Spannungsquellen bestehen. Kommen im Netzwerk Konstantstromquellen, gesteuerte Quellen und Übertrager vor, müssen diese für die Schleifenanalyse in Ersatzschaltungen umgewandelt werden.

5.2.2 Knotenanalyse

Kennt man die Knotenspannungen eines Netzwerks gegenüber einem Bezugsknoten (die Spannungen u_1, u_3, u_2, u_4 , und u_6 für den Bezugsknoten (5) in Abbildung 5.7), so sind alle Zweigspannungen als Differenz zweier Knotenspannungen angebar.

Bei der Knotenanalyse ist der Referenzknoten frei wählbar. Er wird meistens so gewählt, dass er möglichst viele direkte Verbindungen zu anderen Knoten hat³. Das ist in der Regel der Masseknoten.

Wir haben bisher angenommen, dass Netzwerke nur aus R, L, C und (ungesteuerten) Stromquellen bestehen. Kommen im Netzwerk Konstantspannungsquellen, gesteu-

Bei der **Schleifenanalyse** werden die Netzwerkgrößen über Schleifenströme ermittelt.

Bei der **Knotenanalyse** werden die Netzwerkgrößen über Knotenspannungen ermittelt.

²Bei ebenen Netzwerken ist die Suche nach einem Spannbaum nicht erforderlich, es kann auch das Maschenverfahren angewendet werden. Als Netzwerk-Unbekannte werden die Maschenströme in allen Maschen des Netzwerk genommen. In unserem Beispiel sind die Maschen durch die Bauelementzweige $(u, R_1, C_1), (C_1, C_2, C_3), (C_2, L)$ und (C_3, R_2) gegeben.

³Es gibt Fälle, in denen es keinen Referenzknoten gibt, von dem aus alle anderen Knoten direkt erreichbar sind. In diesem Fällen muss das Knotenpotenzialverfahren verallgemeinert werden (Schnittmengenanalyse).

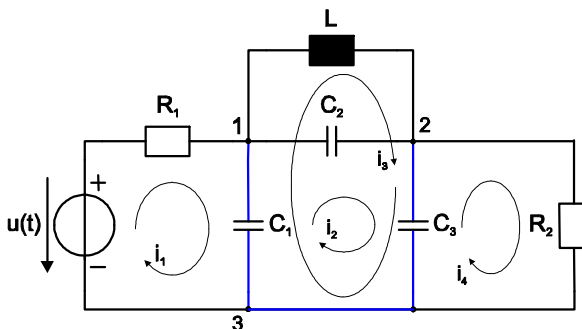


Abbildung 5.8: Anwendung der Schleifenanalyse

erte Quellen und Übertrager vor, müssen diese für die Knotenpotenzialanalyse in Ersatzschaltungen umgewandelt werden.

Die Knotenspannungen sind also ein geeigneter Satz von Unbekannten zur vollständigen Berechnung des Netzwerks. Das zugehörige Berechnungsverfahren wird *Knotenanalyse* genannt.

5.2.3 Beispiele

Schleifenanalyse

Wir berechnen die Netzwerkgrößen der Abbildung 5.8 mit der Schleifenanalyse.

1. Wir wählen einen Spannbaum bestehend aus den Kanten mit den Bauelementen C_1 und C_2 .
2. Wir zeichnen die Schleifenströme ein. Die Schleifenströme bilden einen geschlossenen Pfad, fließen durch die Bauelemente der Glieder und schließen sich durch Bauelemente im Baum, z.B. $L_{\text{Glieder}} \leftrightarrow C_{3\text{Baum}} \leftrightarrow C_{1\text{Baum}}$ oder $C_{2\text{Glieder}} \leftrightarrow C_{3\text{Baum}} \leftrightarrow C_{1\text{Baum}}$. Die Richtung der Schleifenströme kann beliebig gewählt werden.
3. Wir machen einen vollständigen Umlauf in Richtung eines jeden Schleifenstroms und erstellen die Netzwerkgleichungen $\sum u = 0$. Die Spannungsabfälle an den (passiven) Bauelementen werden mit Hilfe der Schleifenströme ausgedrückt, wobei zu beachten ist, dass durch ein Bauelement mehrere Schleifenströme fließen können.

Wir erhalten

$$\text{(Schleife 1): } -u + i_1 R_1 + \frac{1}{sC_1} (i_1 - i_2 - i_3) = 0 \quad (5.10)$$

$$\text{(Schleife 2): } \frac{1}{sC_2} i_2 + \frac{1}{sC_3} (i_2 + i_3 - i_4) + \frac{1}{sC_1} (i_2 + i_3 - i_1) = 0 \quad (5.11)$$

$$\text{(Schleife 3): } sLi_3 + \frac{1}{sC_3} (i_3 + i_2 - i_4) + \frac{1}{sC_1} (i_3 + i_2 - i_1) = 0 \quad (5.12)$$

$$\text{(Schleife 4): } R_2 i_4 + \frac{1}{sC_3} (i_4 - i_2 - i_3) = 0 \quad (5.13)$$

Nach Umordnen und Darstellung in Matrixschreibweise erhalten wir

$$\begin{pmatrix} R_1 + \frac{1}{sC_1} & -\frac{1}{sC_1} & -\frac{1}{sC_1} & 0 \\ -\frac{1}{sC_1} & \frac{1}{sC_1} + \frac{1}{sC_2} + \frac{1}{sC_3} & \frac{1}{sC_1} + \frac{1}{sC_3} & -\frac{1}{sC_3} \\ -\frac{1}{sC_1} & \frac{1}{sC_2} + \frac{1}{sC_3} & sL + \frac{1}{sC_1} + \frac{1}{sC_3} & -\frac{1}{sC_3} \\ 0 & -\frac{1}{sC_3} & -\frac{1}{sC_3} & R_2 + \frac{1}{sC_3} \end{pmatrix} \cdot \begin{pmatrix} i_1 \\ i_2 \\ i_3 \\ i_4 \end{pmatrix} = \begin{pmatrix} u \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad (5.14)$$

Wenn wir dieses Gleichungssystem⁴ lösen, erhalten wir die Schleifenströme, aus denen sich durch Überlagerung die Zweigströme berechnen lassen. Die Zweigspannungen erhalten wir über die Bauelementebeziehungen aus den Zweigströmen. Lösen wir das Gleichungssystem für i_4 , erhalten wir

$$i_4 = -u(1 + s^2 C_2 L) / [(R_1 + R_2 + s(L + C_1 R_1 R_2 + C_3 R_1 R_2) + s^2(C_1 L R_1 + C_2 L R_1 + C_2 L R_2 + C_3 L R_2) + s^3(C_1 C_2 L R_1 R_2 + C_1 C_3 L R_1 R_2 + C_2 C_3 L R_1 R_2)]$$

Knotenanalyse

Wir berechnen nun die Netzwerkgrößen der Abbildung ?? mit der Knotenanalyse.

1. Wir wählen den Knoten (3) als Referenzknoten⁵.
2. Die Knotenspannungen u_1 und u_2 sind von den Knoten 1 und 2 zum Knoten 3 gerichtet und in dieser Richtung positiv⁶.
3. Wir erstellen die Netzwerkgleichungen $\sum i = 0$ für alle Knoten des Netzwerks, wobei die in die Knoten fließenden Ströme positiv, die aus den Knoten fließenden negativ gezählt werden.⁷

$$\text{Knoten 1: } i_{R_1} + i_{C_1} + i_{C_2} + i_L = 0 \quad (5.15)$$

$$\text{Knoten 2: } i_{R_2} + i_{C_3} - i_{C_2} - i_L = 0 \quad (5.16)$$

Wir drücken die Zweigströme durch die Knotenspannungen aus

$$\begin{aligned} i_{R_1} &= \frac{u_{R_1}}{R_1} = \frac{u - u_1}{R_1} \\ i_{C_1} &= sC_1 u_C = -sC_1 u_1 \\ i_{C_2} &= sC_2 (u_2 - u_1) \\ i_{C_3} &= -sC_3 u_2 \\ i_L &= \frac{1}{sL} (u_2 - u_1) \\ i_{R_2} &= -\frac{u_2}{R_2} \end{aligned}$$

und erhalten

$$\frac{(u - u_1)}{R_1} - sC_1 u_1 + sC_2 (u_2 - u_1) + \frac{1}{sL} (u_2 - u_1) = 0 \quad (5.17)$$

$$-\frac{u_2}{R_2} - sC_3 u_2 - sC_2 (u_2 - u_1) - \frac{1}{sL} (u_2 - u_1) = 0 \quad (5.18)$$

Nach Umordnen wird daraus

$$\underbrace{\begin{pmatrix} \frac{-1}{R_1} - sC_1 - sC_2 - \frac{1}{sL} & sC_2 + \frac{1}{sL} \\ sC_2 + \frac{1}{sL} & -\frac{1}{R_2} - sC_3 - sC_2 - \frac{1}{sL} \end{pmatrix}}_{\mathcal{Y}} \cdot \underbrace{\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}}_{\mathcal{U}} = \underbrace{\begin{pmatrix} -\frac{u}{R_1} \\ 0 \end{pmatrix}}_{\mathcal{I}} \quad (5.19)$$

Wenn wir dieses Gleichungssystem (im s -Bereich) lösen, erhalten wir die Knotenspannungen (in Bezug auf den Referenzknoten). Die Zweigspannungen lassen sich aus den Knotenspannungen berechnen, die Zweigströme können über die Bauelementebeziehungen aus den Zweigspannungen berechnet werden.

Wir lösen das Gleichungssystem und erhalten für u_2

⁴Die Gleichungen sind im s -Bereich erstellt und können mit Hilfe der Laplace-Transformation gelöst werden. Alternativ könnte man mit der Beziehung $s = d/dt$ ein System von Differentialgleichungen erhalten, das im Zeitbereich gelöst werden kann.

⁵Es könnte jeder Knoten des Netzwerks den Referenzknoten bilden, es ist allerdings praktisch den Referenzknoten auf das Potential Null (Erde) zu legen und damit alle Spannungen auf dieses Potential zu beziehen.

⁶Auch hier könnte man, wie bei den Schleifenströmen, eine beliebige Richtung wählen. Ein einheitliches Vorgehen, z.B. alle Schleifen rechtsdrehend und alle Knotenspannungen positiv zum Bezugsknoten, erhöht die Übersichtlichkeit bei der Erstellung der Netzwerkgleichungen.

⁷Die Stromrichtungen können ganz willkürlich gewählt werden. Es ist nur darauf aufzupassen, dass durch einen Zweig der Strom nur in einer Richtung fließen kann (also immer von einem Knoten weg in einen anderen Knoten hinein).

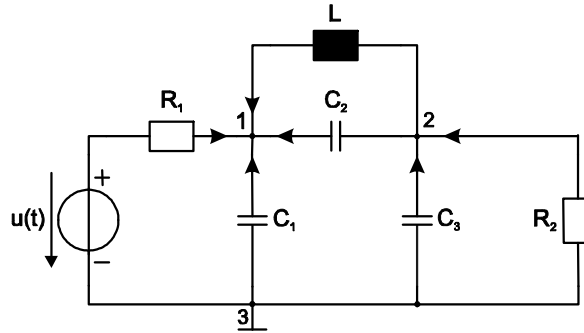


Abbildung 5.9: Anwendung der Knotenanalyse

$$u_2 = R_2(1 + C_2 L s^2)u / (R_2 + Ls + C_2 L R_2 s^2 + C_3 L R_2 s^2 + R_1((1 + C_3 R_2 s)(1 + C_2 L s^2) + C_1 s(R_2 + Ls + C_2 L R_2 s^2 + C_3 L R_2 s^2)))$$

Formt man das Nennerpolynom um, erhält man die Lösung wie bei der Schleifenanalyse. Das Zählerpolynom der Schleifenanalyse ist

$$i_4 = -u(1 + s^2 C_2 L). \text{ Daraus folgt für } u_2 = -i_4 R_2 = R_2(1 + C_2 L s^2)u.$$

Schleifen- und Knotenanalyse liefern selbstverständlich dasselbe Ergebnis für die Netzwerkgrößen. Die Entscheidung, ob die Schleifenanalyse oder die Knotenanalyse verwendet wird, ist eine Frage der Zweckmäßigkeit.

Netzwerkanalyseprogramme, wie z.B. SPICE (Simulation Program with Integrated Circuit Emphasis), erlauben die (graphische) Eingabe von elektrischen Schaltungen, die Netzwerkgleichungen werden automatisch erstellt und der Benutzer kann entscheiden, welche Netzwerkuntersuchungen (Gleichstrom-, Wechselstrom-, transiente Analyse, ...) durchgeführt werden sollen.

5.3 Zusammenfassung

Bei der Berechnung der Zweigspannungen und -ströme in einem elektrischen Netzwerk mit z Zweigen sind $2z$ Netzwerkgrößen zu ermitteln. Da in jedem Zweig Strom und Spannung über die Bauelementbeziehung verknüpft ist, sind letztlich nur z Netzwerkgrößen zu ermitteln.

Bei einfachen Netzwerken kann die Erstellung der Netzwerkgleichungen per Hand erfolgen, bei umfangreichen Netzwerken ist ein methodisches Vorgehen angeraten, um linear unabhängige Netzwerkgleichungen zu finden. Sind im Netzwerk Speicherelemente enthalten, dann sind die entstehenden Netzwerkgleichungen Integro-Differentialgleichungen. Wir haben uns bei unseren Beispielen für die Erstellung der Netzwerkgleichungen der Operatorenrechnung bedient, die entstehenden Netzwerkgleichungen werden dadurch algebraische Gleichungen, die die Laplace-Transformierte der Netzwerk-Differentialgleichungen⁸ sind. Die Lösung der Netzwerkgleichungen erfolgt durch Umkehrung (inverse) der Laplace-Transformation mit Hilfe von Tabellenverfahren, wie im Kapitel 6 über die Lösung der Netzwerkgleichung gezeigt.

⁸Für energielosen Anfangszustand des Netzwerks.

Kapitel 6

Lösung der Netzwerkgleichungen

Inhalt

6.1	Gleichstromanalyse	72
6.2	Wechselstromanalyse	73
6.2.1	Effektivwert	73
6.2.2	Komplexe Widerstände	75
6.2.3	Leistung in Wechselstromnetzwerken	78
6.3	Allgemeine Zeitfunktion als Eingangssignal	79
6.3.1	Lösung im Frequenzbereich	79
6.4	Systemantwort	85
6.4.1	Lösung im Frequenzbereich	85
6.4.2	Lösung im Zeitbereich	93
6.4.3	Anfangsbedingungen	95
6.4.4	Eigenschaften der Systemfunktion	95
6.5	Zusammenfassung	96

Mit Hilfe topologischer Methoden können wir die Kirchhoff'schen Gleichungen systematisch erstellen und erhalten ein System von unabhängigen Gleichungen, die das Netzwerk beschreiben. Aus diesen Gleichungen werden die Schleifenströme bzw. Knotenspannungen berechnet, aus denen sich alle Zweigströme und -spannungen ermitteln lassen.

An das Netzwerk wird ein (oder mehrere) Eingangssignal angelegt und ein (oder mehrere) Ausgangssignal ermittelt. In vielen Fällen interessieren nicht die Zweigspannungen und -ströme innerhalb des Netzwerks, sondern das Übertragungsverhalten des Netzwerkes, d.h. das Verhältnis von Ausgangs- zu Eingangssignal. Wir beschränken uns im weiteren auf Systeme mit einem Eingang und einem Ausgang, wie in Abbildung 6.1 gezeigt.

Wir unterscheiden drei Typen von Eingangsgrößen.



Abbildung 6.1: Übertragungsverhalten eines Netzwerkes

Häufig ist in einem Netzwerk nur das **Übertragungsverhalten** interessant, d.h. das Verhältnis von Eingangs- zu Ausgangssignal.

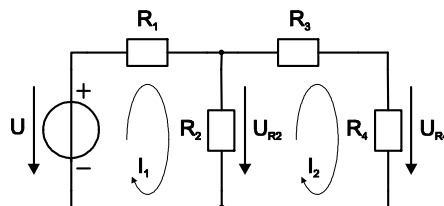


Abbildung 6.2: Schleifenanalyse

Das Netzwerkverhalten kann für konstante Erregersignale (**Gleichstromanalyse**), sinusförmige Erregersignale (**Wechselstromanalyse**) oder beliebige Zeitsignale untersucht werden.

- Der einfachste Fall liegt vor, wenn das »Eingangssignal« eine Gleichspannung oder ein -strom ist. Hier können wir eigentlich nicht von Signalen sprechen, da die Eingangsgröße unveränderlich ist und keine Information trägt. Die Analyse elektrischer Netzwerke für Gleichgrößen nennt man *Gleichstromanalyse*.
- Wir haben im Kapitel über Signale gesehen, dass man sich allgemeine Signale aus sinusförmigen Grundsignalen zusammengesetzt denken kann. Das Verhalten von Systemen für sinusförmige Eingangssignale ist Thema der *Wechselstromanalyse*.
- Schließlich bleibt noch das Verhalten von Systemen auf *allgemeine zeitabhängige Signale* zu untersuchen. Die Lösung der Netzwerkgleichungen für diesen Fall kann durch Lösung im Zeit- oder im Frequenzbereich erfolgen.

6.1 Gleichstromanalyse

Bei der Gleichstromanalyse kann man Kapazitäten durch offene und Induktivitäten durch geschlossene Verbindungen ersetzen.

Die Gleichstromanalyse ist der einfachste Fall der Netzwerkuntersuchungen. Bei der Gleichstromanalyse treten nur (konstante oder gesteuerte) Spannungs- und Stromquellen und Widerstände auf. In einem Gleichstromnetzwerk verhalten sich Kapazitäten wie offene Schaltkreise, Induktivitäten wie kurzgeschlossene Schaltkreise. Nach Erstellung der Netzwerkgleichungen mit Hilfe der Kirchhoff'schen Sätze erhalten wir ein System gewöhnlicher Gleichungen das gelöst werden muss.

Beispiel 37 Wir berechnen das Netzwerk in Abbildung¹ 6.2 mit Hilfe der Schleifenanalyse. Dieses einfache Beispiel könnte man auch per Hand rechnen, wir lösen es aber auf systematische Art mit Hilfe der Maschenströme I_1 und I_2 .

$$-U + I_1 R_1 + I_1 R_2 - I_2 R_2 = 0 \quad (6.1)$$

$$I_2 R_2 - I_1 R_1 + I_2 R_3 + I_2 R_4 = 0 \quad (6.2)$$

$$\underbrace{\begin{pmatrix} R_1 + R_2 & -R_2 \\ -R_1 & R_2 + R_3 + R_4 \end{pmatrix}}_{\mathbf{R}} \cdot \underbrace{\begin{pmatrix} I_1 \\ I_2 \end{pmatrix}}_{\mathbf{I}} = \underbrace{\begin{pmatrix} U \\ 0 \end{pmatrix}}_{\mathbf{U}} \quad (6.3)$$

Dieses Gleichungssystem mit zwei Unbekannten kann gelöst werden und wir erhalten für die Schleifenströme

$$I_1 = \frac{U}{R_1 + R_2 - \frac{R_1 R_2}{R_2 + R_3 + R_4}} = \frac{(R_2 + R_3 + R_4)U}{R_2^2 + R_1 R_3 + R_2 R_3 + R_1 R_4 + R_2 R_4} \quad (6.4)$$

$$I_2 = \frac{R_1}{R_2 + R_3 + R_4} I_1 = \frac{R_1 U}{R_2^2 + R_1 R_3 + R_2 R_3 + R_1 R_4 + R_2 R_4} \quad (6.5)$$

Setzen wir für die Zahlenwerte $R_1 = 1\Omega$, $R_2 = 2\Omega$, $R_3 = 3\Omega$, $R_4 = 4\Omega$, $U = 10V$ ein, so erhalten wir

¹Für die Bezeichnung von Gleichgrößen verwendet man in der Elektrotechnik in der Regel Großbuchstaben.

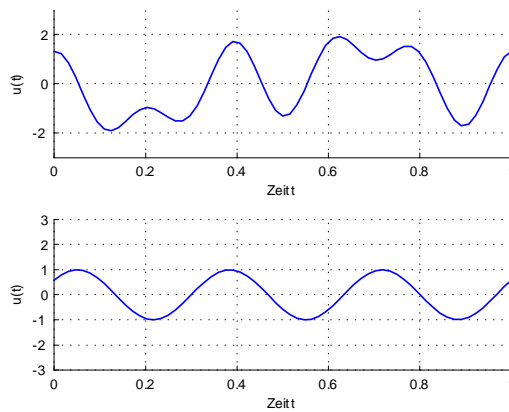


Abbildung 6.3: Zeitlich veränderliche Netzwerkgrößen

$$\underbrace{\begin{pmatrix} 3 & -2 \\ -1 & 9 \end{pmatrix}}_{\mathbf{R}} \cdot \underbrace{\begin{pmatrix} I_1 \\ I_2 \end{pmatrix}}_{\mathbf{I}} = \underbrace{\begin{pmatrix} 10 \\ 0 \end{pmatrix}}_{\mathbf{u}}$$

Wir lösen die Matrixgleichung durch Linksmultiplikation mit \mathbf{R}^{-1} und erhalten $\mathbf{I}=\mathbf{R}^{-1}\mathbf{U}$

$$\begin{pmatrix} I_1 \\ I_2 \end{pmatrix} = \begin{pmatrix} 0.36 & 0.08 \\ 0.04 & 0.12 \end{pmatrix} \cdot \begin{pmatrix} 10 \\ 0 \end{pmatrix} = \begin{pmatrix} 3.60 \\ 0.40 \end{pmatrix}$$

Spannungen und Ströme im Netzwerk lassen sich über die Schleifenströme ermitteln. Die Spannung $U_{R_4} = I_2 R_4$ lässt sich direkt aus dem Schleifenstrom I_2 ermitteln. Die Spannung U_{R_2} müsste aus der Summe der beiden Schleifenströme berechnet werden $U_{R_2} = (I_1 - I_2)R_2$.

6.2 Wechselstromanalyse

Wechselspannungen und -ströme sind Spannungen und Ströme, die sich zeitlich ändern, wie in Abbildung 6.3 gezeigt. Im engeren Sinn verstehen wir unter Wechselgrößen Größen mit sinusförmiger Zeitabhängigkeit $u(t) = \hat{U} \sin(\omega t + \varphi)$ bzw. $i(t) = \hat{I} \sin(\omega t + \varphi)$.

6.2.1 Effektivwert

Die Amplituden der Wechselgrößen (\hat{U}, \hat{I}) nennt man Scheitelwert. Die Leistung in elektrischen Netzwerken ist durch das Produkt aus Strom mal Spannung definiert. Liegt an einem ohmschen Widerstand R eine Wechselspannung $u(t)$ an, so ist die im Widerstand verbrauchte Leistung

$$p(t) = u(t) \cdot i(t) = u(t) \cdot \frac{u(t)}{R} = \frac{u^2(t)}{R} \tag{6.6}$$

So wie die Spannung, ist auch die Leistung zeitabhängig, wie Abbildung 6.4 zeigt. Abbildung 6.4 zeigt, dass die Frequenz der Leistung gleich der doppelten Frequenz der Spannung² ist. Zur Berechnung der Wärmewirkung des Wechselstroms interessiert der Mittelwert P_{mittel} der verbrauchten Leistung. Wir berechnen die mittlere Leistung durch Integration und Mittelung über die Periodendauer

²Die Helligkeitsvariation einer Glühlampe in unserem Stromnetz, die wir wegen der Trägheit des Auges nicht sehen, aber mit entsprechenden Messgeräten feststellen können, erfolgen daher mit 100 Hz.

Wechselspannung oder **-strom** im engeren Sinne heißt Spannung oder Strom mit **sinusförmigem Zeitverlauf**.

Der **Effektivwert** einer Wechselspannung ist der Wert einer Gleichspannung, die am ohmschen Widerstand die **gleiche Leistung** verbraucht. Der Effektivwert ist **von der Kurvenform abhängig!**

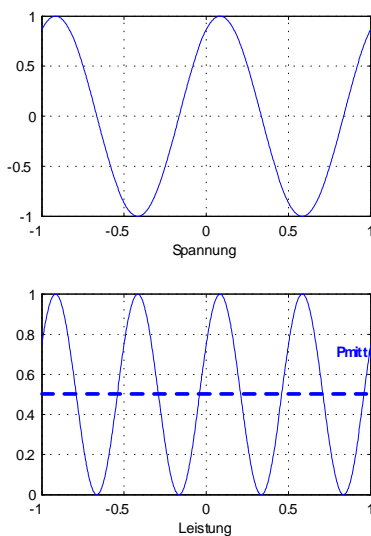


Abbildung 6.4: Wechselstromleistung

$$P_{\text{mittel}} = \frac{1}{T} \int_0^T \frac{u^2(t)}{R} dt \quad (6.7)$$

Die Größe

$$U_{\text{eff}} = \sqrt{\frac{1}{T} \int_0^T u^2(t) dt} \quad (6.8)$$

nennt man den Effektivwert der Spannung. Die mittlere Leistung ist daher

$$P_{\text{mittel}} = \frac{U_{\text{eff}}^2}{R} \quad (6.9)$$

Die Effektivspannung ist jene Spannung, die eine Gleichspannungsquelle haben müsste, um die gleiche Leistung an einem ohmschen Widerstand zu verbrauchen wie die Wechselspannungsquelle. Aus Gleichung (6.8) sehen wir, dass die Effektivspannung von der Kurvenform der Wechselspannung $u(t)$ abhängt!

Beispiel 38 Als Beispiel berechnen wir den Effektivwert der sinusförmigen Wechselspannung $u(t) = \hat{U} \sin \omega t$, \hat{U} ist die Amplitude der Sinusschwingung

$$U_{\text{eff}} = \sqrt{\frac{1}{T} \int_0^T \hat{U}^2 \sin^2(\omega t) dt} \quad (6.10)$$

Unter Verwendung der Beziehung $\sin^2 \alpha = \frac{1}{2}(1 - \cos 2\alpha)$ erhalten wir

$$U_{\text{eff}} = \sqrt{\frac{\hat{U}^2}{2T} \int_0^T [1 - \cos(2\omega t)] dt} \quad (6.11)$$

und erhalten

$$U_{\text{eff}} = \sqrt{\frac{\hat{U}^2}{2T} \left[t - \frac{1}{2\omega} \sin(2\omega t) \right] \Big|_0^T} = \hat{U} \sqrt{\frac{1}{2}} \quad (6.12)$$

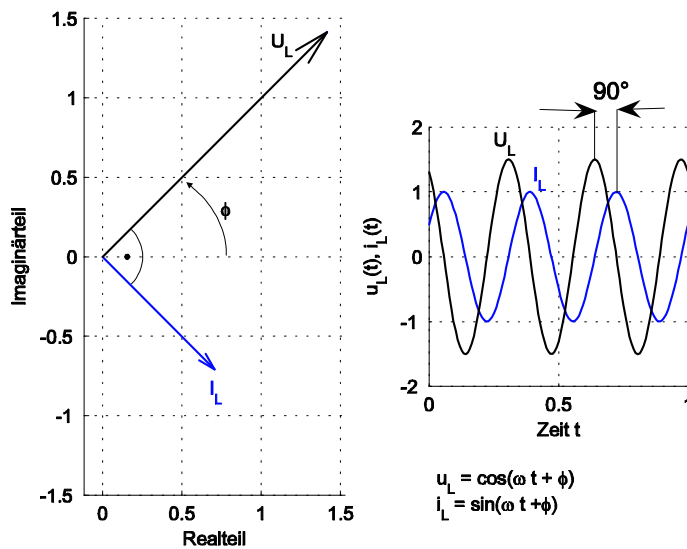


Abbildung 6.5: Strom und Spannung an der Induktivität

Die Wechselspannung unseres Stromnetzes beträgt 230 Volt und ist in Effektivwert angegeben, die Scheitelspannung beträgt daher 325 Volt. Die Umrechnung $\hat{U} = U_{eff}\sqrt{2}$ gilt nur für sinusförmige Wechselspannungen, da der Effektivwert von der Kurvenform der Zeitfunktion abhängt! Bei sinusförmigen Größen bezeichnet man die Amplituden nur mit U, I , ob Effektiv- oder Scheitelwert gemeint ist, geht aus dem Zusammenhang hervor. Wenn es sich um leistungsbezogene Größen handelt wird in der Regel der Effektivwert verwendet, bei Signalen meistens der Scheitelwert.

6.2.2 Komplexe Widerstände

Im Abschnitt über Signale haben wir die komplexe Rechnung zur Darstellung von Sinus- und Kosinusfunktion eingeführt und damit eine wesentliche Vereinfachung der Rechnung erreicht. Durch die Einführung der komplexen Rechnung wird die Netzwerkanalyse für sinusförmige Größen sehr einfach und praktisch gleich wie die Gleichstromanalyse: Bei der Gleichstromanalyse haben wir ohmsche Widerstände, die nur reelle Werte annehmen können. Bei der Wechselstromanalyse treten zusätzlich zu den reellen ohmschen Widerständen noch die imaginären Widerstände von L und C auf. Abgesehen davon dass Widerstände reell, imaginär oder komplex sein können, erfolgt die Berechnung von Wechselstromnetzwerken gleich wie die Berechnung von Gleichstromnetzwerken.

Die Wechselstromanalyse erfolgt wie die Gleichstromanalyse mit dem Unterschied, dass **Widerstände** im Allgemeinen **komplexe Werte** annehmen.

Induktivität

Bei der Induktivität und sinusförmigen Größen haben wir folgenden Zusammenhang

$$i_L = I \sin(\omega t + \varphi) \quad u_L = L \frac{di_L(t)}{dt} \quad (6.13)$$

und erhalten daraus durch Einsetzen

$$u_L = \omega LI \cos(\omega t + \varphi) \quad (6.14)$$

Der Zusammenhang zwischen Strom und Spannung als Funktion der Zeit in reeller und in komplexer Darstellung ist in Abbildung 6.5 gezeigt.

Wie wir aus Abbildung 6.5 erkennen können, beträgt die Phasenverschiebung zwischen Strom und Spannung $\pi/2 = 90^\circ$, man sagt, der Strom eilt der Spannung um 90° nach.

Bei einer Induktivität eilt der Strom der Spannung um $\pi/2$ nach.

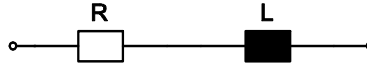


Abbildung 6.6: Serienschaltung R und L

Sinusförmige Größen können auch mit Hilfe der komplexen Exponentialfunktion³ als rotierende Zeiger dargestellt werden. Bei einer Sinusschwingung ändert sich die Länge des Zeigers nicht, daher $\sigma = 0$ und $s = j\omega$. Unter Verwendung der komplexen Schreibweise erhalten wir $i_L(t) = I \sin(\omega t + \varphi) \Rightarrow I e^{j(\omega t + \varphi)}$. Wir setzen in (6.13) ein und erhalten

$$u_L = L \frac{di_L(t)}{dt} \Rightarrow \vec{u}_L = L \frac{d(I e^{j(\omega t + \varphi)})}{dt} = j\omega L I e^{j(\omega t + \varphi)} \quad (6.15)$$

\vec{u}_L wird dadurch zu einer komplexen Größe, die zunächst keine physikalische Bedeutung hat, mit der aber gut gerechnet werden kann. Durch die Differentiation bleibt die Funktion $e^{j\omega t}$ erhalten⁴! Wir können daher auf das Anschreiben der komplexen Exponentialfunktion verzichten und schreiben kurz⁵

$$\vec{U}_L = j\omega L \cdot \vec{I}_L = sL \cdot \vec{I}_L \Big|_{s=j\omega} \quad (6.16)$$

Den komplexen Widerstand nennt man **Impedanz**. Die Impedanz ist **von der Frequenz abhängig**.

Die Impedanz einer Induktivität ist

$$\vec{Z}_L = j\omega L$$

Bei einer Kapazität eilt die Spannung dem Strom um $\pi/2$ nach. Die Impedanz einer Kapazität ist

$$\vec{Z}_C = \frac{1}{j\omega C}$$

Bei hohen Frequenzen haben Kapazitäten einen kleinen "Widerstand", Induktivitäten einen großen. Für niedrige Frequenzen ist es umgekehrt.

Der komplexe Widerstand $\vec{Z} = \frac{\vec{U}}{\vec{I}}$ (vgl. $R = \frac{U}{I}$) wird Impedanz genannt. Das »ohmsche« Gesetz für die Induktivität lautet nach Einführung der Impedanz

$$\vec{U}_L = \vec{Z}_L \vec{I}_L \quad \text{für} \quad \vec{Z}_L = j\omega L = sL \quad (6.17)$$

Die Impedanz der Induktivität ist rein imaginär⁶, der ohmsche Widerstand rein reell. Die Serienschaltung eines ohmschen Widerstands und einer Induktivität nach Abbildung 6.6 hat die Ersatzimpedanz $\vec{Z} = R + j\omega L$ und ist eine komplexe Größe.

Kapazität

Für die Kapazität erhalten wir in komplexer Darstellung

$$\vec{U}_C = \vec{Z}_C \vec{I}_C \quad \text{für} \quad \vec{Z}_C = 1/j\omega C = -j \frac{1}{\omega C} = \frac{1}{sC} \quad (6.18)$$

Schaltkreisanalyse mit komplexen Widerständen

Bei der Analyse von Schaltkreisen für sinusförmige Wechselspannungen und -ströme können wir die Kirchhoff'schen Gesetze in der gewohnten Form anwenden:

$$\sum u = 0 \quad \text{und} \quad \sum i = 0.$$

- Wir ersetzen die Zeitbeschreibung der sinusoidalen Netzwerkgrößen durch die komplexe Exponentialfunktion. Die komplexen Exponentialfunktionen können wir uns als rotierende Zeiger in der komplexen Ebene vorstellen, die Drehgeschwindigkeit der Zeiger entspricht der Kreisfrequenz ω der Netzwerkgröße. Im weiteren Verlauf der Berechnungen interessieren wir uns nur mehr für den Betrag und die Phase der Zeiger, da wir ja wissen, dass die Zeiger sinusoidale Größen repräsentieren und die sinusoidale Kurvenform in linearen Netzwerken erhalten bleibt. Wir können das Netzwerk immer aber nur für *eine* Frequenz untersuchen. Wenn mehrere Quellen in einem Netzwerk auftreten, müssen alle dieselbe Frequenz haben!
- Induktivitäten und Kapazitäten haben die Impedanzen $\vec{Z}_L = j\omega L$ bzw. $\vec{Z}_C =$

³Siehe Abschnitt 2.4 über die komplexe Exponentialfunktion.

⁴Die (komplexe) Exponentialfunktion ist die einzige Funktion auf die das zutrifft! Wir verwenden $e^{j\omega t}$ zur Darstellung der Sinusfunktion.

⁵Strom, Spannung und Widerstand (Impedanz) sind in der Wechselstromrechnung komplexe Größen. Zur Unterscheidung verwenden wir daher $\vec{U}, \vec{I}, \vec{Z}$.

⁶Rein imaginäre Impedanzen werden auch Reaktanzen genannt.

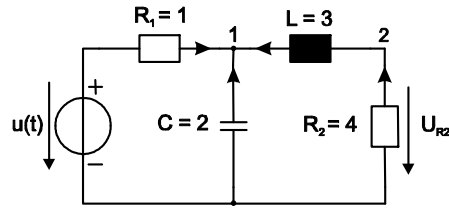


Abbildung 6.7: Netzwerkberechnung mit Knotenanalyse

$1/j\omega C$, ohmsche Widerstände die Impedanz $Z_R = R$. Impedanzen sind frequenzabhängige Widerstände, ohmsche Widerstände hängen nicht von der Frequenz ab. Induktivitäten haben bei niedrigen Frequenzen einen kleinen Widerstand (bei $\omega = 0$ verhalten sie sich wie ein Kurzschluss), der Widerstand steigt frequenzproportional an $Z_L \sim \omega$. Kapazitäten haben bei niedrigen Frequenzen einen großen Widerstand (bei $\omega = 0$ verhalten sie sich wie ein Leerlauf, d.h. ein offener Schaltkreis), bei steigender Frequenz sinkt der Widerstand verkehrt proportional zur Frequenz $Z_C \sim \frac{1}{\omega}$.

- Wir erstellen die Netzwerkgleichungen unter Verwendung der bekannten Verfahren, wobei in den Netzwerkmatrizen in der Regel komplexe Zahlen auftreten werden.

Beispiel 39 Das Beispiel in Abbildung 6.7 zeigt die Vorgangsweise, wobei wir in diesem Fall die Knotenanalyse anwenden wollen.

$$\frac{\vec{U} - \vec{U}_1}{R_1} - j\omega C \vec{U}_1 + \frac{1}{j\omega L} (\vec{U}_2 - \vec{U}_1) = 0 \quad (6.19)$$

$$-\frac{1}{j\omega L} (\vec{U}_2 - \vec{U}_1) - \vec{U}_2 \frac{1}{R_2} = 0 \quad (6.20)$$

$$\begin{pmatrix} -\frac{1}{R_1} - j\omega C - \frac{1}{j\omega L} & \frac{1}{j\omega L} \\ \frac{1}{j\omega L} & -\frac{1}{R_2} - \frac{1}{j\omega L} \end{pmatrix} \cdot \begin{pmatrix} \vec{U}_1 \\ \vec{U}_2 \end{pmatrix} = \begin{pmatrix} -\frac{\vec{U}}{R_1} \\ 0 \end{pmatrix} \quad (6.21)$$

Wir lösen die Netzwerkgleichungen und erhalten

$$\vec{U}_2 = \frac{R_2 \vec{U}}{R_1 + R_2 + Ls + CR_1 R_2 s + CLR_1 s^2} \Big|_{s=j\omega} \quad (6.22)$$

In vielen Fällen ist bei der Untersuchung von Netzwerken lediglich das Verhältnis Ausgangsgröße zu Eingangsgröße gefragt, z.B. U_2/U . Dieses Verhältnis ist eine komplexe Zahl, da die Netzwerkgleichungen komplex sind. Man stellt dieses Verhältnis meistens nach Betrag (in logarithmischem Maßstab) und Phase dar und nennt diese Darstellung Bode-Plot. Wenn wir (6.22) grafisch darstellen wollen, müssen wir Zahlenwerte für die Bauelemente einsetzen und die Funktion für die interessierenden Frequenzen berechnen.

Für die gegebenen Bauelementewerte ist $\frac{\vec{U}_2}{\vec{U}} = \frac{4}{6s^2 + 11s + 5}$.

Mit Hilfe von Matlab erhalten wir `bode(tf(ZP,NP))`, wie in 6.8 dargestellt.

Wir »überprüfen« die grafische Darstellung und berechnen die Knotenspannung \vec{U}_2 für $R_1 = 1$, $C = 2$, $L = 3$, $R_2 = 4$, $U = 10$ und für die Kreisfrequenz $\omega = 1$ durch Einsetzen in (6.21)

$$\begin{pmatrix} -1 - j2 - \frac{1}{j3} & \frac{1}{j3} \\ \frac{1}{j3} & -\frac{1}{4} - \frac{1}{j3} \end{pmatrix} \cdot \begin{pmatrix} \vec{U}_1 \\ \vec{U}_2 \end{pmatrix} = \begin{pmatrix} -10 \\ 0 \end{pmatrix} \quad (6.23)$$

$$\begin{pmatrix} -1 - j1.667 & -j0.333 \\ -j0.333 & -0.25 + j0.333 \end{pmatrix} \cdot \begin{pmatrix} \vec{U}_1 \\ \vec{U}_2 \end{pmatrix} = \begin{pmatrix} -10 \\ 0 \end{pmatrix} \quad (6.24)$$

Wir multiplizieren von links mit der inversen Matrix und erhalten

$$\begin{pmatrix} \vec{U}_1 \\ \vec{U}_2 \end{pmatrix} = \begin{pmatrix} -0.2377 + j0.3852 & 0.0328 + j0.3607 \\ 0.0328 + j0.3607 & -1.2459 - j1.7049 \end{pmatrix} \cdot \begin{pmatrix} -10 \\ 0 \end{pmatrix} = \begin{pmatrix} 2.3770 - j3.8525 \\ -0.3279 - j3.6066 \end{pmatrix}$$

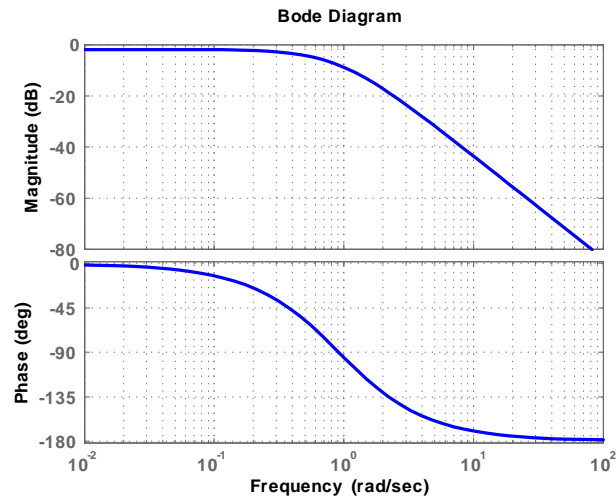


Abbildung 6.8: Bode-Plot

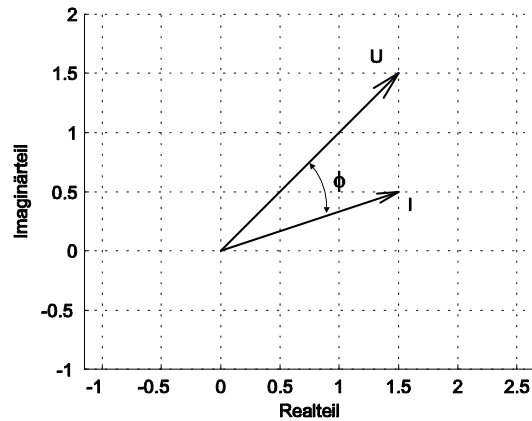


Abbildung 6.9: Phasenverschiebung am Zweipol

Die gesuchte Spannung $\vec{U}_{R_2} = \vec{U}_2 = 3.6214e^{-j1.66152} = 3.6214e^{-j95.19446^\circ}$. Die Amplitude der Ausgangsspannung hat von 10V auf 3.6214V abgenommen, die Phasenverschiebung beträgt -95° . Das Verhältnis von Ausgangs- zu Eingangsspannung ist 0.36214 oder im Dezibelmaß $20\log_{10}(0.36214) = -8.8\text{dB}$.

6.2.3 Leistung in Wechselstromnetzwerken

Liegt an einem Zweipol mit der Impedanz $\vec{Z} = R + jX$ eine Spannung \vec{U} und fließt ein Strom \vec{I} , der, wie in Abbildung 6.9 gezeigt, gegen die Spannung eine Phasenverschiebung von ϕ hat, so ist die vom Zweipol aufgenommene Wirkleistung $P = U_{eff}I_{eff} \cos \phi = I_{eff}^2 R$, das ist jene Leistung, die im ohmschen Anteil des Zweipols in Wärme umgesetzt wird. Die Größe $P_B = U_{eff}I_{eff} \sin \phi = I_{eff}^2 X$ bezeichnet man als (effektive) Blindleistung. Als Scheinleistung bezeichnet man die Größe $P_S = U_{eff} \cdot I_{eff} = \sqrt{P^2 + P_B^2}$.

In elektrischen Netzen trägt nur die Wirkleistung zur Energieumsetzung bei. Die Scheinleistung steckt im elektrischen und magnetischen Feld der Komponenten und es wird lediglich Energie zwischen dem elektrischen und dem magnetische Feld verschoben, die keinen Beitrag zur Wirkleistung leistet. Beim Verschieben der Energie zwischen elektrischem und magnetischen Feld fließen jedoch Ströme über Leitungen, deren ohmscher

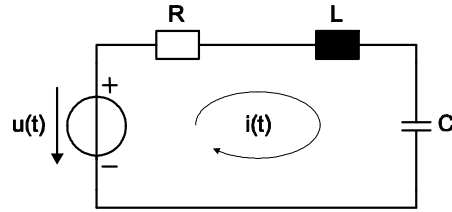


Abbildung 6.10: RLC-Serienschwingkreis

Widerstand zwar klein, aber nicht Null ist⁷. Der Blindstrom erzeugt durch die endlichen Widerstände der Leitungen Wärmeverluste. Die Energieversorgungsunternehmen verrechnen daher industriellen Abnehmern auch die Blindleistung, die zwar keine echte Leistung ist, aber auf den Leitungen ohmsche Verluste erzeugt.

Bei Verwendung der komplexen Rechnung ermittelt man die komplexe Leistung⁸ mit der Beziehung $\vec{P} = P + jP_B = \vec{U}^* \vec{I} = U e^{-j\omega t} I e^{j(\omega t + \phi)} = UI e^{j\phi}$.

6.3 Allgemeine Zeitfunktion als Eingangssignal

Wir erstellen die Netzwerkgleichungen für das Beispiel nach Abbildung 6.10 mit Hilfe des Schleifenstroms $i(t)$ und erhalten

$$L \frac{di(t)}{dt} + Ri(t) + \frac{1}{C} \int_0^T i(t) dt + u_C(0) = u(t) \quad (6.25)$$

Wir differenzieren Gleichung (6.25) und erhalten

$$L \frac{d^2 i(t)}{dt^2} + R \frac{di(t)}{dt} + \frac{1}{C} i(t) = \frac{du(t)}{dt} \quad (6.26)$$

$$\frac{d^2 i(t)}{dt^2} + \frac{R}{L} \frac{di(t)}{dt} + \frac{1}{LC} i(t) = \frac{1}{L} \frac{du(t)}{dt} \quad (6.27)$$

Mit den Abkürzungen $\alpha = \frac{R}{2L}$ und $\omega_0 = \frac{1}{\sqrt{LC}}$ wird daraus

$$\frac{d^2 i(t)}{dt^2} + 2\alpha \frac{di(t)}{dt} + \omega_0^2 i(t) = f(t) \quad (6.28)$$

Die Differentialgleichung (6.28) nennt man Schwingungsgleichung. $f(t)$ ist die Erregungsfunktion. Ähnlich den mechanischen Schwingungen in einem Masse-Federsystem, treten in einem LC-Netzwerk elektrische Schwingungen auf. Die Differentialgleichung (6.28) ist für die gegebene Erregungsfunktion zu lösen, wobei die Anfangsbedingungen (Anfangsspannung an der Kapazität und Anfangsstrom der Induktivität) berücksichtigt werden müssen.

Die direkte Lösung von Netzwerk-Differentialgleichungen werden wir nicht behandeln, da sie umständlich und aufwändig ist und daher für lineare Netzwerke selten angewendet wird.

6.3.1 Lösung im Frequenzbereich

Die Lösung der Netzwerkgleichungen kann vorteilhaft mit Hilfe der Laplace-Transformation durchgeführt werden. Da die Laplace-Transformation einen Übergang der Darstellung vom Zeitbereich in den Frequenzbereich darstellt, spricht man von der Lösung im Frequenzbereich.

Die Bezeichnung Frequenzbereich ist nicht ganz zutreffend, da bei der Laplace-Trans-

⁷Nur in unserer idealisierten Darstellung hat die Verbindung zwischen den Komponenten den Widerstand Null. Einen Zweipol mit rein imaginärer Impedanz gibt es also in der Wirklichkeit nicht.

⁸Die Größe $\vec{P} = \vec{U} \vec{I}$ ist nicht zweckmäßig, da sie nicht zur Definition der Wirk- und Blindleistung führt.

Um Netzwerkgleichungen einfacher lösen zu können, werden sie mittels der **Laplace-Transformation** in den s -Bereich transformiert.

formation das Netzwerk in einen abstrakten mathematischen Bereich (die s -Ebene) abgebildet wird. Für technische Frequenzen gilt $(\sigma = 0) \rightarrow s = j\omega$, wie wir in der Wechselstromrechnung bereits gesehen haben. Die technischen sinusoidalen Schwingungen liegen in der s -Ebene bei $\sigma = 0$, also entlang der imaginären Achse.

Die Laplace-Transformation ist eine Methode zur Lösung von Differentialgleichungen, die sehr vorteilhaft auf elektrische Netzwerke bzw. ganz allgemein auf LTI-Systeme angewendet werden kann. Fasst man den Parameter s der Netzwerkgleichungen als komplexen Frequenzparameter⁹ auf, so hat man bereits die Laplace-transformierten Differentialgleichungen des Netzwerks im energielosen¹⁰ Anfangszustand.

Die Beziehung

$$F(s) = \int_{0^-}^{\infty} f(t)e^{-st} dt \quad (6.29)$$

nennt man (einseitige) Laplace-Transformation. Da wir uns nur für Vorgänge interessieren, die zu einem bestimmten Zeitpunkt beginnen, den wir (willkürlich) bei $t = 0$ festlegen, ist die untere Integrationsgrenze der Gleichung (6.29) 0^- . (Wir definieren einfach $f(t) = 0$ für $t < 0$.) Die Laplace-Transformation bildet über die Beziehung (6.29) eine Funktion vom Zeitbereich $f(t)$ in den Frequenzbereich (s -Bereich) $F(s)$ ab.

Durch die Laplace-Transformation werden Differentialgleichungen zu algebraischen Gleichungen.

Durch diese Abbildung werden aus den Differentialgleichungen des Zeitbereichs gebrochen rationale Funktionen im s -Bereich. Da wir die Laplace-Transformation bereits auf Bauteilebene durchführen (wir ersetzen $u_L(t) = L \frac{di_L(t)}{dt}$ durch $U_L(s) = sLI_L(s)$ und $i_C(t) = C \frac{du_C(t)}{dt}$ durch $I_C(s) = sCU_C(s)$), erstellen wir unsere Netzwerkgleichungen bereits im s -Bereich und erhalten daher keine Differentialgleichungen wie z.B. in (6.28). Ermittelt man die Netzwerkgleichungen in Abbildung 6.10 direkt im s -Bereich (mit Hilfe der Operatorenrechnung), dann wird

$$-U(s) + RI(s) + sLI(s) + \frac{1}{sC}I(s) = 0 \quad (6.30)$$

Nach Umformung erhalten wir daraus unter Verwendung der in (6.28) verwendeten Abkürzungen

$$I(s) = \frac{\frac{1}{L}s}{s^2 + \frac{R}{L}s + \frac{1}{LC}} \cdot U(s) = \underbrace{\frac{\frac{1}{L}s}{s^2 + 2\alpha s + \omega_o^2}}_{\text{Netzwerk}} \cdot \underbrace{U(s)}_{\text{Erregung}} \quad (6.31)$$

Wie man leicht sehen kann, ist die Erstellung der Netzwerkgleichungen im s -Bereich einfacher durchzuführen als die Erstellung der Differentialgleichungen. Die Netzwerkgleichung ist im s -Bereich eine gebrochen rationale Funktion mit dem Zählerpolynom $P(s) = \frac{1}{L}s$ und dem Nennerpolynom $Q(s) = s^2 + 2\alpha s + \omega_o$. Was wir zur Beschreibung des Systemverhaltens im s -Bereich noch ergänzen müssen, ist die Laplace-Transformierte der Erregung.

Schließlich muss vom s -Bereich in den Zeitbereich zurücktransformiert werden, da uns ja letztlich das Zeitverhalten der Schaltung interessiert. Diese Vorgangsweise mag kompliziert erscheinen, ist aber eine sehr elegante und effiziente Methode zur Lösung von Differentialgleichungen. Darüber hinaus liefert die Analyse der gebrochen rationalen Funktionen weitere Einsichten in die Eigenschaften eines Netzwerks, wie wir im weiteren Verlauf noch sehen werden.

Für die Rücktransformation in den Zeitbereich werden meistens vorberechnete Transformationspaare aus Korrespondenztabelle verwendet.

Die inverse Laplace-Transformation – die Rücktransformation vom s -Bereich in den Zeitbereich – ist durch folgende Beziehung gegeben

$$f(t) = \frac{1}{2\pi j} \int_{\sigma-j\infty}^{\sigma+j\infty} F(s)e^{st} ds \quad (6.32)$$

⁹ s wird meistens kurz komplexe Frequenz genannt.

¹⁰ Anfangsbedingungen wie geladene Kapazitäten oder stromdurchflossene Induktivitäten müssen getrennt berücksichtigt werden, was aber keine Einschränkung des Verfahrens darstellt.

die aber ohne praktische Bedeutung ist. Für die Rücktransformation vom s -Bereich in den Zeitbereich zerlegt man die Netzwerkgleichungen in einfache Ausdrücke, deren Transformationsbeziehung $f(t) \iff F(s)$ man aus den Korrespondenztabelle entnehmen kann. Die Zerlegung der gebrochenen rationalen Funktionen in einfache Ausdrücke wird mit Hilfe der Partialbruchzerlegung durchgeführt.

Rechnen mit der Laplace-Transformation

Operationen

Operation	$f(t)$	$F(s)$
Linearität	$\sum_i c_i f_i(t)$	$\sum_i c_i F_i(s)$
Differentiation	$\frac{df(t)}{dt}$	$sF(s) - f(0^-)$
Integration	$\int_{0^-}^{t^+} f(\tau) d\tau$	$\frac{F(s)}{s}$
Verschiebung im Zeitbereich	$f(t - a)$	$e^{-as} F(s)$
Verschiebung im Frequenzbereich	$e^{at} f(t)$	$F(s - a)$
Maßstabsänderung $a > 0$	$f\left(\frac{t}{a}\right)$	$aF(as)$
Faltung $\int_{0^-}^{t^+} f_1(\tau) f_2(t - \tau) d\tau$	$f_1(t) * f_2(t)$	$F_1(s) F_2(s)$

Korrespondenzen

	$f(t)$	$F(s)$
0	$\delta_i(t)$	s^i
1	$\delta_0(t)$	1
2	$\delta_{-1}(t)$	$\frac{1}{s}$
3	$\delta_{-1}(t)t$	$\frac{1}{s^2}$
4	$\delta_{-1}(t)e^{-at}$	$\frac{1}{s+a}$
5	$\delta_{-1}(t)te^{-at}$	$\frac{1}{(s+a)^2}$
6	$\delta_{-1}(t)\sin\omega t$	$\frac{\omega}{s^2+\omega^2}$
7	$\delta_{-1}(t)\cos\omega t$	$\frac{s}{s^2+\omega^2}$
8	$\delta_{-1}(t)e^{-at}\sin\omega t$	$\frac{\omega}{(s+a)^2+\omega^2}$
9	$\delta_{-1}(t)e^{-at}\cos\omega t$	$\frac{s+a}{(s+a)^2+\omega^2}$
10	$\delta_{-1}(t)re^{-at}\cos(\omega t + \varphi)$	$\frac{0.5re^{j\varphi}}{s+(a-j\omega)} + \frac{0.5re^{-j\varphi}}{s+(a+j\omega)}$
11	$\delta_{-1}(t)re^{-at}\cos(\omega t + \varphi)$	$\frac{As+B}{s^2+2as+c} = \frac{As+B}{(s+a)^2+\omega^2}$
(a)	$r = \sqrt{\frac{A^2c+B^2-2ABa}{\omega^2}}$	
(b)	$\varphi = \arctan \frac{Aa-B}{A\omega}$	
(c)	$\omega = \sqrt{c-a^2}$	

Das »Einschalten« eines Signals wird durch die Multiplikation mit der Funktion $\delta_{-1}(t)$ modelliert.

Wir betrachten nur solche Zeitfunktionen, die für $t < 0$ verschwinden oder anders ausgedrückt zum Zeitpunkt $t = 0$ »eingeschaltet« werden. Die mathematische Darstellung des Einschaltens erfolgt durch Multiplikation mit $\delta_{-1}(t)$. In dieser Tabelle finden wir auch wichtige Erregungsfunktionen, die in Gleichung (6.31) noch gefehlt haben, z.B. der Einheitssprung, der Impuls oder die (eingeschaltete) Sinusoidalfunktion.

Partialbruchzerlegung

Bei der Erstellung von Netzwerkgleichungen im s -Bereich treten gebrochen rationale Funktionen auf, die Zähler- und Nennerpolynome hohen Grades aufweisen können. In unseren Korrespondenztabelle finden wir aber nur Polynome $F(s)$ mit maximal 2. Grad. Um die Korrespondenztabelle verwenden zu können müssen daher die Netzwerkfunktionen in eine Summe von Polynomen (maximal zweiten Grades) zerlegt werden. Die Zerlegung können wir mit Hilfe der Partialbruchzerlegung durchführen.

Wenn bei der **Partialbruchzerlegung** der Grad des Zählerpolynoms größer als der Grad des Nennerpolynoms ist, muss **Polynomdivision** durchgeführt werden.

Bei der Partialbruchzerlegung muss m , der Grad des Zählerpolynoms $P(s)$, kleiner als n , der Grad des Nennerpolynoms¹¹ $Q(s)$, sein. Wenn das nicht der Fall ist, muss zuerst durch Division $m < n$ erreicht werden.

¹¹Wir sehen in den Korrespondenztabelle, dass der Zählergrad stets kleiner als der Nennergrad ist.

Beispiel 40

$$F(s) = \frac{s^2 + 7s + 11}{s^2 + 4s + 5} \quad (6.33)$$

$$(s^2 + 7s + 11) : (s^2 + 4s + 5) = 1 \quad (6.34)$$

$$\frac{-(s^2 + 4s + 5)}{(0s^2 + 3s + 6)} : \quad (6.35)$$

$$(0s^2 + 3s + 6) : \quad (6.36)$$

$$F(s) = \frac{s^2 + 7s + 11}{s^2 + 4s + 5} = 1 + \frac{3s + 6}{s^2 + 4s + 5} \quad (6.37)$$

Wir transformieren die Netzwerkgleichung aus dem s -Bereich in den Zeitbereich und verwenden dafür folgende Entsprechungen aus der Korrespondenztabelle

$$(1) \quad \delta_0(t) \Leftrightarrow 1 \quad (6.38)$$

$$(11) \quad \delta_{-1}(t)re^{-at} \cos(\omega t + \varphi) \Leftrightarrow \frac{As + B}{s^2 + 2as + c} \quad (6.39)$$

Die Zeitfunktion $f(t)$ wird nach Einsetzen in (11 a,b,c)

$$f(t) = \delta_0(t) + \delta_{-1}(t)3e^{-2t} \cos(t) \quad (6.40)$$

Alternativ können wir umformen und erhalten

$$F(s) = 1 + \frac{3s + 6}{s^2 + 4s + 5} = 1 + 3 \frac{s + 2}{(s + 2)^2 + 1} \quad (6.41)$$

und daraus unter Verwendung der Korrespondenz (9) dasselbe Ergebnis.

Beispiel 41 Die Polynomdivision kann mit Hilfe der Matlabfunktion `deconv` durchgeführt werden.

$$P = [1 \ 7 \ 11]; \quad Q = [1 \ 4 \ 3]$$

$$[Q, R] = \text{deconv}(P, Q)$$

$$Q = 1 \quad , \quad R = [0 \ 3 \ 8]$$

Bei der Partialbruchzerlegung müssen zuerst die Nullstellen des Nennerpolynoms berechnet werden und das Nennerpolynom in die Wurzelarstellung gebracht werden. Dann müssen die Zählerkoeffizienten (die Residuen¹²) ermittelt werden.

$$\text{Wir zerlegen } F(s) = \frac{7s+11}{s^2+4s+3}$$

$$1. \text{ Berechnen der Nullstellen von } s^2 + 4s + 3 = 0 \Rightarrow s_{0,1} = -3, \quad s_{02} = -1$$

$$2. F(s) = \frac{7s+11}{(s+3)(s+1)} = \frac{R_1}{s+3} + \frac{R_2}{s+1}$$

Die Zählerkoeffizienten erhalten wir durch Ausmultiplizieren und Koeffizientenvergleich

$$7s + 11 = R_1(s + 1) + R_2(s + 3) = s(R_1 + R_2) + (R_1 + 3R_2) \quad (6.42)$$

$$7 = R_1 + R_2 \quad ; \quad 11 = R_1 + 3R_2 \quad (6.43)$$

$$R_1 = 5 \quad ; \quad R_2 = 2 \quad (6.44)$$

$$F(s) = \frac{7s + 11}{s^2 + 4s + 3} = \frac{5}{s + 3} + \frac{2}{s + 1} \quad (6.45)$$

Unter Verwendung der Beziehung (4) aus der Korrespondenztabelle erhalten wir nach Rücktransformation

$$f(t) = \delta_{-1}(t)5e^{-3t} + \delta_{-1}(t)2e^{-t} \quad (6.46)$$

¹²Bei der direkten Berechnung der inversen Laplace-Transformation wird das Integral mit Hilfe des Residuensatzes ausgewertet. Die Zählerkoeffizienten der Partialbruchzerlegung nennt man Residuen.

Beispiel 42 Die Partialbruchzerlegung kann mit Hilfe der Matlabfunktion `residue` durchgeführt werden. In diesem Fall ist bei $m \geq n$ keine explizite Polynomdivision erforderlich, `residue` liefert als Ergebnis die Residuen (R), Pole (P) und Quotienten (K). Für das Beispiel $F(s) = \frac{7s+11}{s^2+4s+3}$ erhalten wir

$$P = [7 \ 11]; \quad Q = [1 \ 4 \ 3]$$

$$[R, P, K] = \text{residue}(P, Q) \implies R' = [5 \ 2] \quad P' = [-3 \ -1] \quad K = []$$

Die **Residuen-Methode** ist bei der Partialbruchzerlegung schneller und einfacher als der Koeffizientenvergleich.

Residuen-Methode Die Berechnung der Residuen durch Koeffizientenvergleich ist aufwändig, eine schnellere Lösung für die händische Berechnung der Zählerkoeffizienten finden wir über den folgenden Zusammenhang

$$F(s) = \frac{b_m s^m + b_{m-1} s^{m-1} + \dots + b_1 s + b_0}{s^n + a_{n-1} s^{n-1} + \dots + a_1 s + a_0} = \frac{P(s)}{Q(s)} \quad (6.47)$$

$$= \frac{R_1}{s - \lambda_1} + \frac{R_2}{s - \lambda_2} + \dots + \frac{R_n}{s - \lambda_n} \quad (6.48)$$

Multiplizieren wir beide Seiten der Gleichung mit $s - \lambda_1$ und werten an der Stelle $s = \lambda_1$ aus, dann wird

$$(s - \lambda_1) F(s) |_{s=\lambda_1} = R_1 + \frac{R_2 (s - \lambda_1)}{s - \lambda_2} + \dots + \frac{R_n (s - \lambda_1)}{s - \lambda_n} |_{s=\lambda_1} \quad (6.49)$$

Auf der rechten Seite verschwinden alle Terme bis auf R_1 , da $(s - \lambda_1) |_{s=\lambda_1} = 0$, wir erhalten daher

$$R_1 = (s - \lambda_1) F(s) |_{s=\lambda_1} \quad (6.50)$$

Diesen Vorgang können wir für auch für alle anderen Residuen durchführen und erhalten

$$R_k = (s - \lambda_k) F(s) |_{s=\lambda_k} \quad k = 1, 2, \dots, n \quad (6.51)$$

Beispiel 43 Wir berechnen die Residuen der Funktion

$$\frac{As + B}{s^2 + 2as + c} = \frac{As + B}{(s + a)^2 + \omega^2} = \frac{R_1}{s + a + j\omega} + \frac{R_2}{s + a - j\omega} \quad (6.52)$$

$$s^2 + 2as + a^2 + \omega^2 \rightarrow \omega = \sqrt{c - a^2} \quad (6.53)$$

$$R_1 = \frac{(As + B) \boxed{(s + a + j\omega)}}{\boxed{(s + a + j\omega)} (s + a - j\omega)} |_{s=-(a+j\omega)} \quad (6.54)$$

$$\frac{A[-(a + j\omega)] + B}{-(a + j\omega) + a - j\omega} = \frac{A\omega + j(B - Aa)}{2\omega} \quad (6.55)$$

$$|R_1| = \frac{1}{2} \sqrt{\frac{A^2 c - 2ABa + B^2}{\omega^2}} \quad (6.56)$$

$$\varphi = \arctan\left(\frac{Aa - B}{A\omega}\right) \quad (6.57)$$

$$R_2 = R_1^* \quad (6.58)$$

$$\frac{As + B}{s^2 + 2as + c} = \frac{0.5 |R_1| e^{j\varphi}}{s + (a - j\omega)} + \frac{0.5 |R_1| e^{-j\varphi}}{s + (a + j\omega)} \quad (6.59)$$

Mehrfache Nullstellen Wir haben bisher nur Netzwerkfunktionen mit einfachen Nullstellen des Nennerpolynoms betrachtet. Treten mehrfache Nullstellen im Nennerpolynom auf, dann wird die Partialbruchentwicklung

$$F(s) = \frac{P(s)}{(s-\alpha)^n (s-\lambda_1)(s-\lambda_2)\dots(s-\lambda_m)} \quad (6.60)$$

$$= \frac{A_0}{(s-\alpha)^n} + \frac{A_1}{(s-\alpha)^{n-1}} + \dots + \frac{A_{n-1}}{(s-\alpha)} \quad (6.61)$$

$$+ \frac{R_1}{(s-\lambda_1)} + \frac{R_2}{(s-\lambda_2)} + \dots + \frac{R_m}{(s-\lambda_m)} \quad (6.62)$$

Die Koeffizienten R_1, R_2, \dots, R_m können wie bisher berechnet werden. Die Koeffizienten A_0, A_1, \dots, A_{n-1} erhält man

$$A_0 = (s-\alpha)^n F(s) |_{s=\alpha} \quad (6.63)$$

$$A_1 = \frac{d}{ds} [(s-\alpha)^n F(s)] |_{s=\alpha} \quad (6.64)$$

$$A_i = \frac{1}{i!} \frac{d^i}{ds^i} [(s-\alpha)^n F(s)] |_{s=\alpha} \quad (6.65)$$

Beispiel 44

$$F(s) = \frac{1}{s(s+2)^2} = \frac{R}{s} + \frac{A_0}{(s+2)^2} + \frac{A_1}{(s+2)} \quad (6.66)$$

$$R = sF(s) |_{s=0} = \frac{1}{4} \quad (6.67)$$

$$A_0 = (s+2)^2 \frac{1}{s(s+2)^2} \Big|_{s=-2} = -\frac{1}{2} \quad (6.68)$$

$$A_1 = \frac{d}{ds} \left[(s+2)^2 \frac{1}{s(s+2)^2} \right] \Big|_{s=-2} = -\frac{1}{s^2} \Big|_{s=-2} = -\frac{1}{4} \quad (6.69)$$

$$F(s) = \frac{1/4}{s} - \frac{1/2}{(s+2)^2} - \frac{1/4}{(s+2)} \quad (6.70)$$

Oder durch Berechnung mit Matlab¹³ $P = [1]$, $Q = [1, 4, 4, 0]$.

$[R, P, K] = \text{residue}[P, Q]$, $R = [-0.25 \ -0.5 \ 0.25]$; $P = [-2 \ -2 \ 0]$ $K = []$

$$\frac{R(j)}{(s-p_j)} + \frac{R(j+1)}{(s-p_j)^2} + \dots + \frac{R(j+n-1)}{(s-p_j)^n} \quad (6.71)$$

$$F(s) = -\frac{0.25}{(s+2)} - \frac{0.5}{(s+2)^2} + \frac{0.25}{(s)} \quad (6.72)$$

6.4 Systemantwort

Unter Systemantwort verstehen wir die Antwort eines Systems auf ein gegebenes Eingangssignal.

6.4.1 Lösung im Frequenzbereich

Das Verhältnis von Antwort (Ausgangssignal) zu Erregung (Eingangssignal) im s -Bereich nennt man Systemfunktion (Netzwerkfunktion, Übertragungsfunktion). Bei LTI-Netzwerken ist die Systemfunktion stets eine gebrochen rationale Funktion, die sich in folgender Form darstellen lässt

¹³Die Partialbruchzerlegung von gebrochen rationalen Funktionen ist ein schlecht konditioniertes Problem. Kleine Änderungen in den Koeffizienten oder Rundungsfehler bewirken große Änderungen bei den Polen und Residuen!

Die **Systemfunktion** eines LTI-Netzwerkes (das Verhältnis von Eingangssignal zu Ausgangssignal im s -Bereich) ist immer eine gebrochen rationale komplexwertige Funktion.

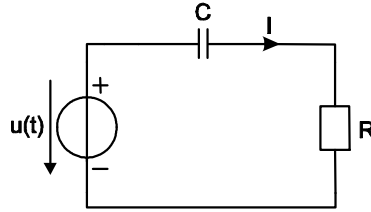


Abbildung 6.11: Netzwerkberechnung mit Laplace-Transformation

$$H(s) = \frac{B(s)}{A(s)} = \frac{b_m s^m + b_{m-1} s^{m-1} + \dots + b_1 s + b_0}{a_n s^n + a_{n-1} s^{n-1} + \dots + a_1 s + a_0} \quad (6.73)$$

Da s komplex ist, ist $H(s)$ eine Funktion komplexer Variablen und liefert daher ein komplexes Ergebnis. Die Koeffizienten b_k , a_k sind durch die Bauelemente des Netzwerks bestimmt und daher stets reell.

Lösung mit Hilfe der Laplace-Transformation

Die Netzwerkgleichungen können sehr vorteilhaft mit der Laplace-Transformation gelöst werden.

Beispiel 45 Das einfache Beispiel nach Abbildung 6.11 zeigt die Vorgangsweise. Das RC-Netzwerk wird von der Eingangsspannung $\vec{u}_E = U_E e^{s_E t}$, $s_E = \sigma_E + j\omega_E$ angeregt. Wie man sieht ist die Eingangsspannung für $\omega_E \neq 0$ eine komplexe Zeitfunktion, was physikalisch sinnlos ist. Wir rechnen dennoch mit komplexen Größen und finden im weiteren Verlauf der Rechnung physikalisch sinnvolle Interpretationen. Der Strom in dieser Schaltung ist

$$I(s) = \frac{U(s)}{R + \frac{1}{Cs}} = \frac{s}{s + \frac{1}{RC}} \frac{1}{R} U(s) \quad (6.74)$$

Die Laplace-Transformierte der Eingangsspannung $u_E = U_E e^{s_E t}$ wird laut Transformationstabelle (4) $U(s) = \frac{U_E}{s - s_E}$. Für den Strom $I(s)$ erhalten wir dann

$$I(s) = \frac{s}{(s - s_N)(s - s_E)} \frac{1}{R} U_E, \quad s_N = -\frac{1}{RC} \quad (6.75)$$

Wir zerlegen diesen Ausdruck in Partialbrüche und erhalten

$$I(s) = \frac{U_E}{R} \left(\frac{R_1}{s - s_N} + \frac{R_2}{s - s_E} \right) \quad (6.76)$$

$$I(s) = \frac{U_E}{R} \left(\underbrace{\frac{\frac{s_N}{s_N - s_E}}{s - s_N}}_{\text{Einschwing-Anteil}} + \underbrace{\frac{\frac{s_E}{s_E - s_N}}{s + s_E}}_{\text{Erreger-Anteil}} \right) \quad (6.77)$$

Gleichung (6.77) besteht aus zwei Teilen: Dem Einschwinganteil der vom Netzwerk stammt und dem Erregeranteil, der von der Spannungsquelle stammt.

Anmerkung 46 Die Unterteilung in Einschwing- und Erregeranteil darf nicht zur Annahme verleiten, die beiden Anteile ließen sich unabhängig voneinander berechnen und man müsse bei einem gegebenen Netzwerk für eine andere Erregung lediglich den Erregeranteil neu berechnen. System und Erregung wirken über die Partialbruchkoeffizienten auf beide Anteile ein!

Unter Verwendung der Transformationstabelle (4) erhalten wir durch Rücktransformation

$$i(t) = \frac{U_E}{R} \left(\frac{s_N}{s_N - s_E} e^{s_N t} + \frac{s_E}{s_E - s_N} e^{s_E t} \right) \quad (6.78)$$

Die Eingangsfunktion $u_E = U_E e^{s_E t}$ hat keine physikalische Bedeutung.

Wenn aber $s_E = \sigma_E$ und damit s_E reell wird, dann ist $u_E = U_E e^{\sigma_E t}$ eine zum Zeitpunkt $t = 0$ eingeschaltete, abklingende Exponentialfunktion mit der Antwort

$$i(t) = \frac{U_E}{R} \left(\frac{1}{1 + RC\sigma_E} e^{-\frac{1}{RC}t} + \frac{\sigma_E}{\sigma_E + \frac{1}{RC}} e^{\sigma_E t} \right) \quad (6.79)$$

Der zweite technisch interessante Fall liegt vor, wenn die Eingangsspannung sinusförmig ist, d.h. $\sigma_E = 0$. Dann wird $\vec{u}_E = U_E e^{j\omega_E t}$, die zum Zeitpunkt $t = 0$ eingeschaltete komplexe Exponentialfunktion und $\vec{i}(t)$ wird.

$$i(t) = \frac{U_E}{R} \left(\underbrace{\frac{1}{1 + jRC\omega_E} e^{-\frac{1}{RC}t}}_{\text{abklingender Anteil}} + \underbrace{\frac{j\omega_E RC}{1 + j\omega_E RC} e^{j\omega_E t}}_{\text{Wechselstromanteil}} \right) \quad (6.80)$$

Wie wir sehen, liefert Gleichung (6.80) als Ergebnis einen Strom mit komplexen Werten, was physikalisch nicht sinnvoll ist. Wir müssen daher bei der Auswertung von (6.80) den Realteil bilden.

Gleichung (6.80) besteht aus zwei Teilen, die abklingender Anteil und Wechselstromanteil genannt werden. Der abklingende Anteil verschwindet für $t \rightarrow \infty$ da $e^{-\frac{1}{RC}t} \rightarrow 0$ und es bleibt der Wechselstromanteil $e^{j\omega_E t}$ (mit der komplexen Amplitude $\frac{j\omega_E RC}{1 + j\omega_E RC}$) übrig.

Dieses Ergebnis hätte man auch direkt, mit Hilfe der Wechselstromrechnung, erhalten.

Ein interessanter Fall liegt vor, wenn Eingangsfunktion und Netzwerkfunktion (abgesehen von einem konstanten Faktor) dieselbe Laplace-Transformierte haben (Resonanz-Fall).

$$I(s) = \frac{s}{(s - s_N)^2} \frac{U_E}{R} \quad (6.81)$$

Die Nullstelle im Nenner tritt doppelt auf, die Partialbruchzerlegung liefert

$$I(s) = \frac{s}{(s - s_N)^2} \frac{U_E}{R} = \frac{U_E}{R} \left(\frac{1}{s - s_N} + \frac{s_N}{(s - s_N)^2} \right) \quad (6.82)$$

Nach Rücktransformation in den Zeitbereich erhalten wir

$$i(t) = \frac{U_E}{R} (1 + s_N t) e^{s_N t} = \frac{U_E}{R} \left(1 - \frac{t}{RC} \right) e^{-\frac{t}{RC}} \quad (6.83)$$

Wie man aus Gleichung (6.83) erkennen kann, steigt die Amplitude des Stroms $i(t)$ linear – proportional zu $t e^{s_N t}$ – an (Resonanzeffekt). Die Exponentialfunktion klingt aber schneller ab, sodass sich der Amplitudenanstieg nicht auswirken kann. (Die Gültigkeit der Lösung von Gleichung (6.83) ist auf die Zeit $t \geq 0$ eingeschränkt, man müsste daher mit $\delta_{-1}(t)$ multiplizieren. Da man den Gültigkeitsbereich kennt, wird die Multiplikation mit $\delta_{-1}(t)$ häufig weggelassen.)

Die folgende Tabelle fasst die Systemfunktion einfacher Schaltungen zusammen:

		$H(s) = \frac{U_a(s)}{U_e(s)}$ (6)			HP	$\frac{LCs^2}{LCs^2+1}$
(1)		$\frac{1}{RCs+1}$ (7)			TP	$\frac{1}{LCs^2+R}$
(2)		$\frac{R}{Ls+R}$ (8)			TP	$\frac{1}{R^2C^2s^2+}$
(3)		$\frac{RCs}{RCs+1}$ (9)			HP	$\frac{R^2C}{R^2C^2s^2+}$
(4)		$\frac{Ls}{Ls+R}$ (10)			BP	$\frac{RCs}{LCs^2+R}$
(5)		$\frac{1}{LCs^2+1}$ (11)			BSp	$\frac{LCs^2-}{LCs^2+R}$

TP...Tiefpass, HP...Hochpass, BP...Bandpass, BSp...Bandsperr

Null- und Polstellen

Die Nullstellen des Zählerpolynoms von $H(s)$ heißen **Nullstellen** der Systemfunktion, die Nullstellen des Nennerpolynoms heißen **Polstellen**.

Wie wir gesehen haben, treten bei der Analyse von elektrischen Netzwerken (LTI-Systemen) gebrochen rationale Funktionen auf. Eine besondere Bedeutung für das Systemverhalten hat die Lage der Nullstellen des Zähler- und Nennerpolynoms dieser Funktionen. Zur Unterscheidung bezeichnet man die Nullstellen des Zählerpolynoms als Nullstellen, da Netzwerkfunktion Null wird. Die Nullstellen des Nennerpolynoms bezeichnet man als Polstellen. In den Polstellen wird die Netzwerkfunktion ∞ . Wird die Lage der Pole und Nullstellen in der komplexen Ebene dargestellt, spricht man vom Pol-/Nullstellen-Diagramm (PN-Diagramm). Zur Unterscheidung werden Nullstellen als \circ und Pole als \times gezeichnet.

Die Systemfunktion ist bis auf einen konstanten Faktor durch Angabe der Pol- und Nullstellen eindeutig bestimmt.

Anmerkung 47 Durch Angabe der Pol- und Nullstellen ist die Systemfunktion eines LTI-Systems bis auf eine reelle Konstante vollständig bestimmt. Der konstante Faktor muss getrennt ermittelt werden. Beschreibt die Systemfunktion z.B. das Verhältnis von Ausgangs- zu Eingangsspannung mit $H(s) = V \frac{U_a}{U_e}$, dann nennt man V die Spannungsverstärkung - oder abschwächung.

Beispiel 48 Wir berechnen die Systemfunktion des in Abbildung 6.12 gezeigten Netzwerks. Durch Knotenanalyse erhalten wir die die Knotenspannungen

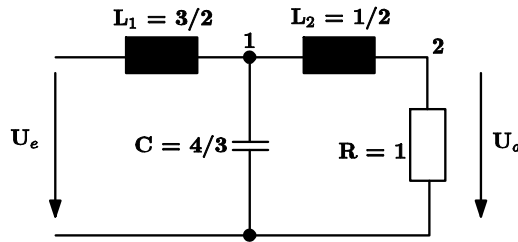
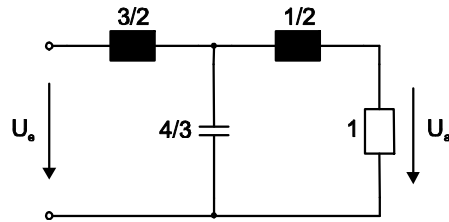


Abbildung 6.12: LC-Tiefpassfilter

Abbildung 6.13: Pol-/Nullstellendiagramm $\frac{1}{s^3+2s^2+2s+1}$

$$U_1 = \frac{(R + L_2 s) U_e}{R + s(L_1 + L_2) + s^2 C L_1 R + s^2 C L_1 L_2} \quad (6.84)$$

$$U_2 = \frac{R U_e}{R + s(L_1 + L_2) + s^2 C L_1 R + s^2 C L_1 L_2} \quad (6.85)$$

Die Ausgangsspannung ist $U_a = U_2$, die Systemfunktion ist $\frac{U_a}{U_e}$. Nach Einsetzen der konkreten Werte für L_1, L_2, C und R und Vereinfachen des entstandenen Ausdrucks erhalten wir als Ergebnis der Netzwerkanalyse

$$H(s) = \frac{U_a(s)}{U_e(s)} = \frac{1}{s^3 + 2s^2 + 2s + 1} \quad (6.86)$$

Da die Systemfunktion $H(s) = \frac{U_a(s)}{U_e(s)}$ als das Verhältnis von Eingangssignal und Ausgangssignal (im s -Bereich) definiert ist, erhält man die Laplace-Transformierte des Ausgangssignals einfach durch Multiplikation der Laplace-Transformierten des Eingangssignals mit der Systemfunktion.

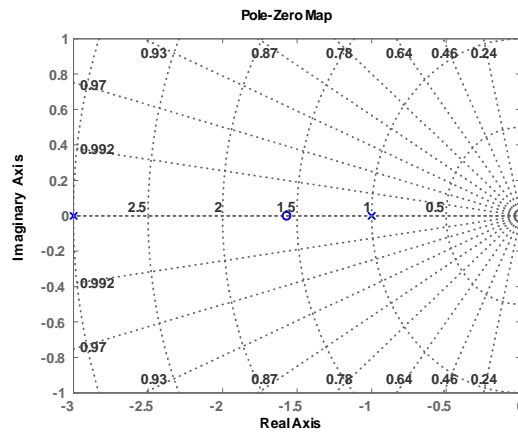
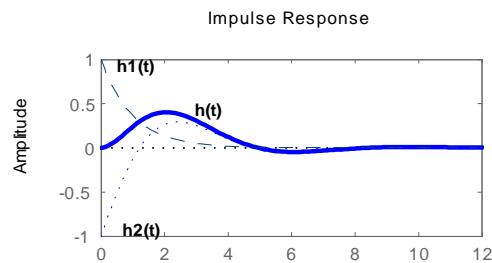
$$U_a(s) = U_e(s) \cdot \frac{U_a(s)}{U_e(s)} = U_e(s) \cdot H(s) \quad (6.87)$$

Beispiel 49 Wir berechnen (mit Matlab) die Nullstellen von $s^3 + 2s^2 + 2s + 1$ und stellen das Nennerpolynom in Produktform dar
`roots([1 2 2 1])` $\Rightarrow (-1.0), (-0.5+0.8660i), (-0.5-0.8660i)$
 $\Rightarrow (s + 1)(s + 0.5 + j0.866)(s + 0.5 - j0.866) = (s + 1)(s - e^{-j120^\circ})(s - e^{j120^\circ})$. Das zugehörige PN-Diagramm ist in Abbildung 6.13 gezeigt.

Beispiel 50 Für die Darstellung des PN-Diagramms stellt Matlab die Funktion `pzmap` zur Verfügung. Für das Beispiel $F(s) = \frac{7s+11}{s^2+4s+3}$ erhalten wir
`P = [7 11]; Q = [1 4 3]`
`sys = tf(P,Q); pzmap(sys)`
eine Nullstelle bei $-\frac{11}{7}$ und Polstellen bei $-1, -3$.

Das Nennerpolynom der Systemfunktion $H(s)$ nennt man charakteristisches Polynom. Die Wurzeln des charakteristischen Polynoms (die Pole) werden auch Eigenfrequenzen

Die Wurzeln des Nennerpolynoms (= charakteristischen Polynoms) einer Systemfunktion (die Pole) stellen **Eigenfrequenzen** des zugehörigen Systems dar.

Abbildung 6.14: PN-Diagramm von $\frac{7s+11}{s^2+4s+3}$ Abbildung 6.15: $h(t) = h_1(t) + h_2(t)$

des Netzwerks genannt. Um zu einem besseren Verständnis der Eigenfrequenzen von Netzwerken zu gelangen, gehen wir vom Netzwerk mit der Systemfunktion (6.86) aus und transformieren in den Zeitbereich.

Die Laplace-Transformierte des Dirac-Impulses ist $\delta_0(t) \Leftrightarrow 1$. Wird also das Netzwerk mit einem Dirac-Impuls angeregt, ist Die Laplace-Transformierte des Ausgangssignals $1 \cdot H(s)$. Die Rücktransformation der Systemfunktion in den Zeitbereich liefert daher die Impulsantwort. Die Eigenschwingungen sind die Reaktionen des Systems auf Anregung mit dem Dirac-Impuls.

Die **Impulsantwort** eines Systems setzt sich aus dessen Eigenschwingungen zusammen.

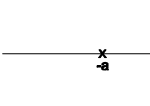
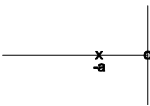
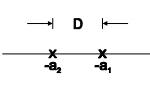
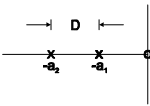
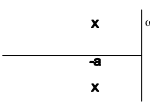
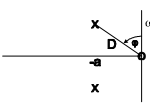
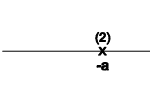
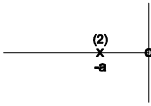
$$H(s) = \frac{1}{s^3 + 2s^2 + 2s + 1} = \frac{1}{s + 1} - \frac{s}{s^2 + s + 1} \quad (6.88)$$

$$h(t) = \underbrace{\delta_{-1}(t)e^{-t}}_{h_1(t)} - \underbrace{\delta_{-1}(t) \frac{2}{\sqrt{3}} e^{-\frac{t}{2}} \cos\left(\frac{\sqrt{3}}{2}t + 30^\circ\right)}_{h_2(t)} \quad (6.89)$$

Wie man in Abbildung 6.15 sieht, sind die Eigenschwingungen unseres Beispiels die abklingende Exponentialfunktion e^{-t} und die abklingende Kosinusfunktion $\frac{2}{\sqrt{3}} e^{-\frac{t}{2}} \cos\left(\frac{\sqrt{3}}{2}t + 30^\circ\right)$, die den Polen bei -1 und $e^{\pm j120^\circ}$ entsprechen.

Anmerkung 51 Ein Netzwerk nennt man stabil, wenn die Eigenfrequenzen abklingen. Das ist immer dann der Fall, wenn die Pole in der linken offenen Halbebene des s -Bereichs liegen. Liegen die Pole in der offenen rechten Halbebene, dann klingen die Schwingungen an, was physikalisch nicht möglich ist und wir haben ein instabiles Netzwerk vorliegen. Pole auf der imaginären Achse stellen ein Netzwerk mit ungedämpften Schwingungen dar (Oszillator). (Vgl. Abbildung 2.8 im Abschnitt 2.4)

Die folgende Tabelle stellt die PN-Diagramme einfacher Systeme und die zugehörigen Impulsantworten dar.

PN-Diagramm	$F(s)$	$f(t)^{14}$
	$\frac{1}{s+a}$	e^{-at}
	$\frac{s}{s+a}$	$\delta_0(t) - ae^{-at}$
	$\frac{1}{(s+a_1)(s+a_2)}$	$\frac{1}{D} (e^{-a_1 t} - e^{-a_2 t})$
	$\frac{s}{(s+a_1)(s+a_2)}$	$\frac{1}{D} (-a_1 e^{-a_1 t} + a_2 e^{-a_2 t})$
	$\frac{1}{(s+a+j\omega)(s+a-j\omega)} = \frac{1}{(s+a)^2 + \omega^2}$	$\frac{1}{\omega} e^{-at} \sin \omega t$
	$\frac{s}{(s+a+j\omega)(s+a-j\omega)} = \frac{s}{(s+a)^2 + \omega^2}$	$\frac{D}{\omega} e^{-at} \cos(\omega t + \varphi)$
	$\frac{1}{(s+a)^2}$	te^{-at}
	$\frac{s}{(s+a)^2}$	$(1-at)e^{-at}$

Systemantwort auf eine beliebige Erregung

Die Systemantwort berechnet man aus $A_g(s) = H(s)G(s)$. Hat man $H(s)$ durch Netzwerkanalyse ermittelt, ist die Laplace-Transformierte der Eingangsfunktion zu ermitteln. Bei der Untersuchung des Übertragungsverhaltens von Netzwerken beschränkt man sich bei der Eingangsfunktion meistens auf sinusförmige Eingangsgrößen (aus denen beliebige Eingangssignale aufgebaut werden können) und auf die Impuls- und Sprungfunktion. Für diese Eingangssignale können wir die Laplace-Transformierte den Transformationstabellen entnehmen. Bei der Ermittlung der Systemantwort auf beliebige Erregung gehen wir wie folgt vor.

1. Die Erregung wird vom Zeitbereich in den Frequenzbereich transformiert: $g(t) \Rightarrow G(s)$
2. Die Systemfunktion wird mit der Erregung multipliziert $A_g(s) = H(s)G(s)$

¹⁴Zur Erinnerung: Die Zeitfunktion $f(t)$ ist nur für $t \geq 0$ definiert. Der kürzeren Schreibweise willen, ist die Multiplikation mit $\delta_{-1}(t)$ weggelassen.

3. Die Antwort wird durch Rücktransformation ermittelt $a_g(t) \Leftarrow A_g(s)$

Mit Hilfe der Partialbruchzerlegung zerlegt man die Systemantwort $A_g(s)$ in einfache Ausdrücke, für die man in Transformationstabellen die entsprechende Zeitfunktion nachschlagen kann.

Anstelle der Lösung der Netzwerkgleichungen mit Hilfe von Differentialgleichungen gelingt es durch die Operatorenrechnung (Laplace-Transformation) die Aufgabenstellung auf die Lösung algebraischer Gleichungen zu vereinfachen. Da die Eigenschaften des Systems ausschließlich im Frequenzbereich (s -Bereich) formuliert sind, spricht man auch von der Lösung im Frequenzbereich.

Stationäre Lösung (Wechselstromrechnung)

Ein wichtiger Sonderfall (Wechselstromrechnung) ist die Erregung eines Netzwerks mit einer stationären Schwingung der Kreisfrequenz ω_i

$$\vec{g}(t) = \vec{K} e^{j\omega_i t} \Rightarrow G(s) = \frac{\vec{K}}{s - j\omega_i} \quad (6.90)$$

$$A_g(s) = H(s) \frac{\vec{K}}{s - j\omega_i} \quad (6.91)$$

Bei stabilen Systemen klingt der Einschwinganteil ab und es bleibt nur der Erregeranteil übrig.

Die Systemantwort $A_g(s)$ besteht aus einem Einschwing- und einem Erregeranteil. Bei stabilen System klingt der Einschwinganteil ab und es bleibt nur der stationäre Erregeranteil übrig. Der stationäre Anteil hat einen einfachen Pol bei $s_\infty = j\omega_i$ mit dem Residuum \vec{R}_{stat}

$$A_{g_{stat}}(s) = \frac{\vec{R}_{stat}}{s - j\omega_i} \quad \vec{R}_{stat} = H(j\omega_i) \vec{K} \quad (6.92)$$

$$\vec{a}_{g_{stat}}(t) = \vec{R}_{stat} e^{j\omega_i t} = H(j\omega_i) \vec{K} e^{j\omega_i t} \quad (6.93)$$

$\vec{a}_{g_{stat}}$ ist die Antwort auf die Erregung $\vec{K} e^{j\omega_i t}$, die (komplexe) Eingangsamplitude \vec{K} wird zur (komplexen) Ausgangsamplitude $H(j\omega_i) \vec{K}$ verändert. Für die reelle Zeitfunktion $a_{g_{stat}}$ ist der Realteil zu bilden $a_{g_{stat}}(t) = \text{Re} \{ \vec{a}_{g_{stat}}(t) \}$.

Wird ein stabiles System durch eine **sinusoidale Schwingung** ange-regt, entsteht am Ausgang (nach dem Abklingen des Einschwinganteils) wieder eine sinusoidale Schwin-gung mit der **gleichen Frequenz**, die die sich von der Erregerschwin-gung nur in der Amplitu-de und Phasenlage unter-scheidet.

Bei Erregung eines stabilen Systems durch eine stationäre Schwingung der Frequenz ω ist die Antwort (nach Abklingen der Einschwingzeit) ebenfalls eine stationäre Schwingung der Frequenz ω , deren komplexe Amplitude aus der komplexen Amplitude der Erregung durch Multiplikation mit $H(j\omega)$ folgt. Die Funktion $H(j\omega)$ nennt man die Frequenzant-wort des Systems.

Im Frequenzbereich ist die Betrachtung der Kurvenform nicht erforderlich, da die Kurvenform bekannt ist, es interessiert lediglich die komplexe Amplitude. Die komplexe Amplitude nach Betrag und Winkel als Funktionen von ω dargestellt, nennt man Amplituden- und Phasengang oder Bodediagramm.

Beispiel 52 Für die numerische Berechnung des Frequenzgangs stellt Matlab die Funk-tion `bode` zur Verfügung. Für das Netzwerk (6.86) $\mathbf{B} = [1]$; $\mathbf{A} = [1, 2, 2, 1]$ erhalten wir unter Verwendung von `bode(B,A)` das Bodediagramm in Abbildung 6.16.

Anmerkung 53 Der Frequenzgang von LTI-Systemen interessiert meistens über einen großen Frequenz- und Amplitudenbereich. Daher verwendet man häufig eine logarithmi-sche Darstellung für beide Achsen.

Das Verhältnis von Eingangs- zu Ausgangsgröße wird immer als Leistungsverhältnis und in Dezibel [dB] ausgedrückt.

$$a[\text{Bel}] = \log \frac{P_{aus}}{P_{ein}}; \quad a[\text{dB}] = 10 \log \frac{P_{aus}}{P_{ein}} \quad \text{bei Leistungen} \quad (6.94)$$

$$a[\text{dB}] = 10 \log \left(\frac{U_{aus}}{U_{ein}} \right)^2 = 20 \log \frac{U_{aus}}{U_{ein}} \quad \text{bei Spannungen} \quad (6.95)$$

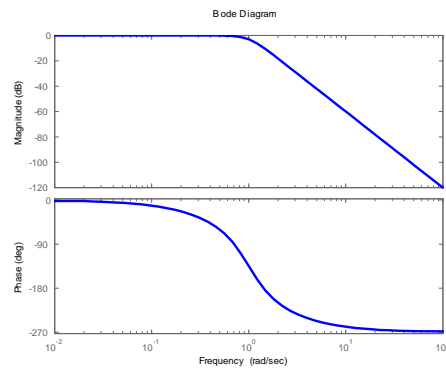


Abbildung 6.16: Bode-Plot LC-Tiefpass $H(s) = \frac{1}{s^3 + 2s^2 + 2s + 1} \Big|_{s=j\omega}$

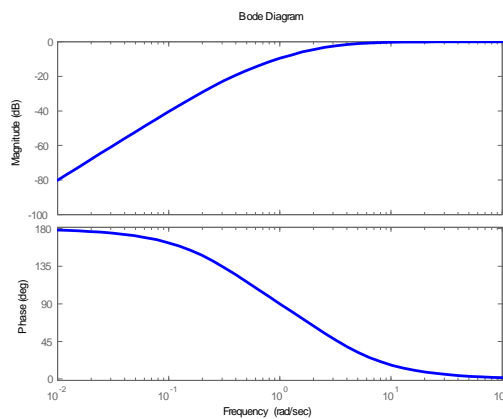


Abbildung 6.17: Frequenzgang RC-Hochpass $\frac{R^2 C^2 s^2}{R^2 C^2 s^2 + 3RCs + 1}$

Beispiel 54 Matlab stellt zur Umrechnung in Dezibel die Funktion $\mathbf{dB}(\mathbf{x})$ zur Verfügung

$$\mathbf{x} = [0.01, 0.1, \frac{1}{\sqrt{2}}, 1, \sqrt{2}, 10, 100]$$

$$\mathbf{dB}(\mathbf{x}) \Rightarrow [-40 \ -20 \ -3 \ 0 \ 3 \ 20 \ 40]$$

% Das Verhältnis $\sqrt{2} \triangleq 3\text{dB}$ entspricht der doppelten Leistung.

Beispiel 55 Für das einfache RC-Netzwerk (9) $H(s) = \frac{R^2 C^2 s^2}{R^2 C^2 s^2 + 3RCs + 1}$ mit $RC = 1$ wird

$B=[1 \ 0 \ 0]$; $A=[1 \ 3 \ 1]$ und wir erhalten das Bodediagramm in Abbildung 6.17

6.4.2 Lösung im Zeitbereich

Die direkte Berechnung der Systemantwort im Zeitbereich durch Lösung der Differentialgleichung ist für komplizierte Erregungsfunktionen zu umständlich. Für einfache Zeitfunktionen wie die Impulsfunktion $\delta_0(t)$ kann man die Systemantwort aber leicht berechnen.

Im Kapitel über die Elementarfunktionen haben wir gesehen, dass sich beliebige Eingangssignale durch Abtastung mit der Impulsfunktion darstellen lassen.

$$g(t) * \delta_0(t) = \int_{-0}^t g(\tau) \delta_0(\tau - t) d\tau = g(t) \tag{6.96}$$

Wie wir wissen, gilt bei LTI-Systemen der Überlagerungssatz. Denken wir uns das

Das **Ausgangssignal** kann man sich als **Überlagerung von Systemantworten auf Impulse**, aus denen sich das Eingangssignal zusammensetzt, vorstellen.

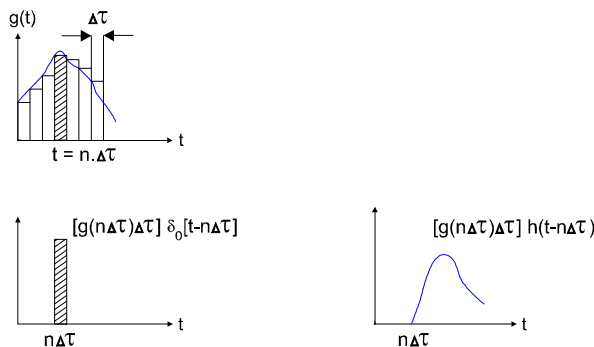


Abbildung 6.18: Zerlegung des Eingangssignals

Eingangssignal aus lauter Impulsen zusammengesetzt, so können wir für jeden dieser Impulse die Impulsantwort bestimmen, die Überlagerung dieser Impulsantworten liefert dann die Systemantwort auf das Eingangssignal. Um die Zusammenhänge verständlich zu machen, stellen wir das Eingangssignal als Summe von Rechteckimpulsen, wie in Abbildung 6.18 gezeigt dar.

Um den Übergang zu Dirac-Impulsen herzustellen, gehen wir wie bei der Definition der Dirac-Impulse vor: Der Impuls hat eine Breite $\Delta\tau$, die Höhe $1/\Delta\tau$, die Fläche 1 und $\Delta\tau \rightarrow 0$. Die »Stärke« des Impuls entspricht seiner Fläche, also $g(n\Delta\tau)\Delta\tau$, der Impuls wird also $[g(n\Delta\tau)\Delta\tau] \delta_0(t - n\Delta\tau)$.

Sie Systemantwort auf den Impuls $\delta_0(t)$ ist $h(t)$ und wir erhalten

$$\delta_0(t) \Rightarrow h(t) \quad (6.97)$$

$$\delta_0(t - n\Delta\tau) \Rightarrow h(t - n\Delta\tau) \quad (6.98)$$

$$\underbrace{[g(n\Delta\tau)\Delta\tau] \delta_0(t - n\Delta\tau)}_{\text{Eingang Einzelpuls}} \Rightarrow \underbrace{[g(n\Delta\tau)\Delta\tau] h(t - n\Delta\tau)}_{\text{Ausgang Einzelpuls}} \quad (6.99)$$

Das Eingangssignal müssen wir aus einer Summe von einzelnen Dirac-Impulsen zusammensetzen und erhalten

$$\lim_{\Delta\tau \rightarrow 0} \sum_{n=-\infty}^{\infty} \underbrace{g(n\Delta\tau)\delta_0(t - n\Delta\tau)\Delta\tau}_{\text{Eingangssignal}} \Rightarrow \lim_{\Delta\tau \rightarrow 0} \sum_{n=-\infty}^{\infty} \underbrace{g(n\Delta\tau)h(t - n\Delta\tau)\Delta\tau}_{\text{Ausgangssignal}} \quad (6.100)$$

$$\int_{-\infty}^{\infty} g(\tau)\delta_0(t - \tau)d\tau \Rightarrow \int_{-\infty}^{\infty} g(\tau)h(t - \tau)d\tau \quad (6.101)$$

Da wir nur Zeitfunktionen für $t \geq 0$ betrachten können wir die Integration von -0 bis t erstrecken und erhalten

$$a_g(t) = g(t) * h(t) = \int_{-0}^t h(\tau)g(t - \tau)d\tau \quad (6.102)$$

Die Systemantwort ergibt sich durch die **Faltung des Eingangssignals mit der Impulsantwort** des Systems.

Die Antwort eines Systems auf das Eingangssignal $g(t)$ berechnet man durch Faltung des Eingangssignals mit der Impulsantwort. Gleichung (6.102) nennt man Superpositionsintegral und spricht von der Lösung der Systemantwort im Zeitbereich.

Die Impulsantwort $h(t)$ kann man durch direkte Lösung der Differentialgleichung des LTI-Systems bestimmen oder zweckmäßiger durch Berechnung der Impulsantwort durch inverse Laplace-Transformation der Systemfunktion $H(s)$.

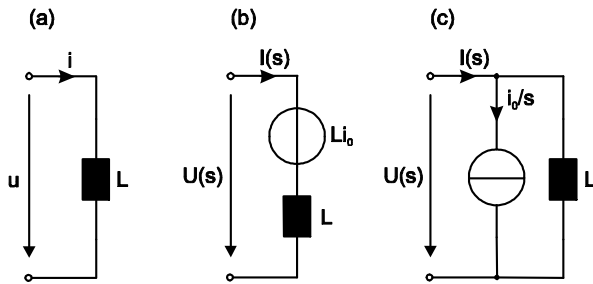


Abbildung 6.19: Anfangsbedingungen Induktivität

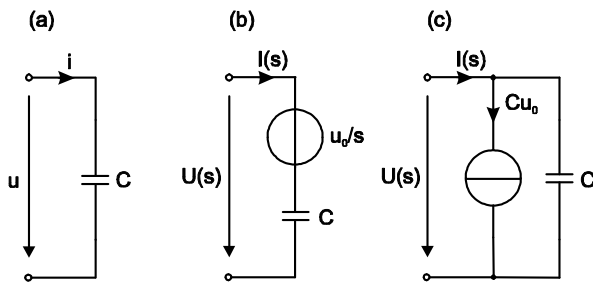


Abbildung 6.20: Anfangsbedingungen Kapazität

6.4.3 Anfangsbedingungen

Alle bisherigen Untersuchungen treffen nur auf Systeme mit energielosem Anfangszustand zu. Das bedeutet, dass zum Zeitpunkt $t = 0$ alle Kapazitäten ungeladen und alle Induktivitäten stromlos sein müssen. Wenn diese Bedingung nicht zutrifft, müssen die Anfangsbedingungen durch zusätzliche Erregungen berücksichtigt werden. Der Zustand eines Netzwerkes zu einem gegebenen Zeitpunkt t ist durch die Ströme in den Spulen und die Spannungen an den Kapazitäten vollständig bestimmt. Sein zukünftiges Verhalten kann aus diesem Zustand und den ab diesem Zeitpunkt wirksamen Erregungen berechnet werden.

Für die Induktivität mit Anfangsstrom erhalten wir das Ersatzschaltbild in Abbildung 6.19.

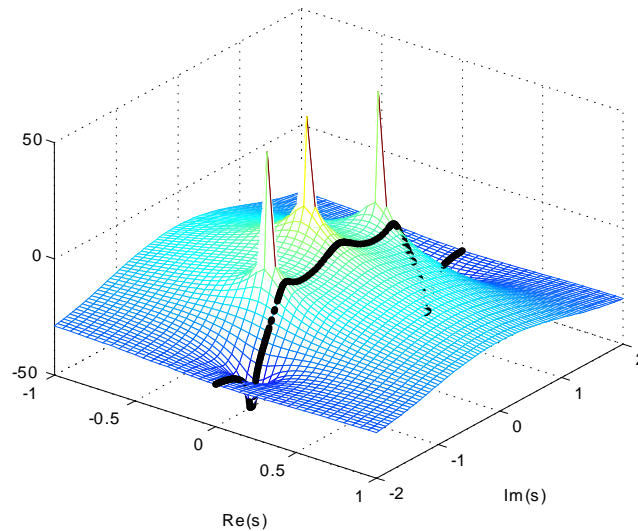
Für die Kapazität mit Anfangsladung erhalten wir das Ersatzschaltbild in Abbildung 6.20.

Wenn das Netzwerk nicht im energielosen Anfangszustand ist, müssen die Netzwerkgleichungen mit den Ersatzschaltungen in den Abbildungen 6.19 und 6.20 in gewohnter Form erstellt werden. Die Berechnung des Netzwerkes wird dadurch zwar umständlicher, jedoch prinzipiell nicht schwieriger. Durch die zusätzlichen Erregungen, die die Anfangsbedingungen berücksichtigen, verliert die Systemfunktion als Verhältnis von *Antwort/Erregung* ihren ursprünglichen Sinn, da das Netzwerk jetzt an mehreren Stellen erregt wird.

6.4.4 Eigenschaften der Systemfunktion

Die Systemfunktion ist eine Funktion der komplexen Variablen $s = \sigma + j\omega$, die für alle s definiert ist. Bei der Berechnung der Systemantwort für beliebige Erregungen ist die Systemfunktion und die Laplace-Transformierte der Erregung erforderlich. Für den Sonderfall $s = j\omega$, liegen die Werte der Systemfunktion über der imaginären Achse, die technischen Frequenzen liegen also auf der imaginären Achse. Die inverse Laplace-Transformation der Systemfunktion ist die Impulsantwort des Systems. Sowohl Systemfunktion, als auch PN-Diagramm (bis auf einen konstanten Faktor), als auch Impulsantwort liefern eine vollständige Beschreibung des Übertragungsverhalten eines Systems.

Zum Zeitpunkt $t = 0$ geladene Kapazitäten oder stromdurchflossene Induktivitäten werden durch zusätzliche Strom- oder Spannungsquellen modelliert.

Abbildung 6.21: Axonometrische Darstellung von $|H(s)|$ in [dB]

Die Systemfunktion von elektrischen Netzwerken hat nur reelle Koeffizienten, die durch die Bauelementewerte bestimmt sind. Die Pole und Nullstellen der Systemfunktion können daher nur entweder reell oder konjugiert komplex sein, wobei die Zahl der Pole und Nullstellen der Ordnung der Nenner- und Zählerpolynome entspricht. Die konjugiert komplexe Systemfunktion ist gleich der Systemfunktion für konjugiert komplexes Argument.

$$H^*(s) = H(s^*); \quad H^*(j\omega) = H(-j\omega) \quad (6.103)$$

Für $s = j\omega$ ist der Realteil der Systemfunktion eine gerade, der Imaginärteil eine ungerade Funktion. Der Betrag der Systemfunktion für $s = j\omega$ ist eine gerade Funktion, die Phase eine ungerade Funktion.

Trägt man den Betrag von $|H(s)|$ über der komplexen Ebene auf, so entsteht ein »Funktionsgebirge«. Der Schnitt des Gebirges mit der Ebene durch die imaginäre Achse (die dicke Linie in Abbildung 6.21) ist der Amplitudengang der Funktion. Abbildung 6.21 zeigt drei Pole und zwei Nullstellen in axonometrischer Darstellung¹⁵. *Die axonometrische Darstellung dient dem besseren Verständnis der Systemfunktion und ist sehr aufwändig. Daher werden in der Regel nur die Pol- und Nullstellen der Systemfunktion dargestellt (PN-Diagramm).*

6.5 Zusammenfassung

Bei der Lösung der Netzwerkgleichungen für *Gleichstromnetzwerke*, die mit Hilfe der Schleifen- oder Knotenanalyse erstellt werden, muss ein System von linearen Gleichungen gelöst werden und man erhält die gewünschten Netzwerkgrößen.

Nach der Einführung der komplexen Widerstände für L und C kann die Netzwerkanalyse in gleicher Form wie bei der Gleichstromanalyse durchgeführt werden. Anstelle der reellen Widerstände der Gleichstromanalyse treten die komplexen Widerstände der *Wechselstromanalyse*. Die gewünschten Netzwerkgrößen können durch Lösung eines Systems von linearen Gleichungen mit komplexen Koeffizienten gefunden werden. Komplexe Widerstände können nur für sinusförmige Wechselgrößen definiert werden, diese Form der Netzwerkberechnung ist daher nur für sinusförmige Erregungen anwendbar.

¹⁵Der Betrag ist logarithmisch dargestellt. Die Nullstellen liegen daher bei $-\infty$.

Für Erregungen mit *beliebigem Zeitverlauf* liefert die Netzwerkanalyse ein System von Integro-Differentialgleichungen, die nach den Regeln der Lösung von Differentialgleichungen zu lösen sind. Derartige Verfahren werden in der Netzwerkanalyse allerdings kaum eingesetzt, da durch die auf der Basis der Laplace-Transformation eingeführte Operatorenrechnung ein wesentlich leistungsfähigeres Verfahren zur Lösung der Differentialgleichungen zur Verfügung steht, das darüber hinaus noch weitere Aussagen über die Eigenschaften des Netzwerks ermöglicht. Die erregenden Zeitfunktionen sind auf den Bereich $t \geq 0$ eingeschränkt, was in technischen Anwendungen sinnvoll ist, da Zeitfunktionen »eingeschaltet« werden. Mit Hilfe der Laplace-Transformation stellen wir das Netzwerk im s -Bereich (Frequenzbereich) dar, aus den linearen Differentialgleichungen werden dadurch algebraische Gleichungen. Die Rücktransformation aus dem s -Bereich in den Zeitbereich wird mit Hilfe bekannter Korrespondenzen zwischen Zeit- und s -Bereich durchgeführt. Dazu wird die Systemfunktion mit Hilfe der Partialbruchzerlegung in einfache gebrochen rationale Funktionen zerlegt, für die man Korrespondenzen in Transformationstabellen findet. Einschränkend ist festzuhalten, dass mit dieser Methode nur Zeitfunktionen erfassbar sind, deren Laplace-Transformierte, so wie die Systemfunktion, eine gebrochen rationale Funktion ist. Diese Einschränkung ist für die wichtigsten Zeitfunktionen Sinus, Impuls und Sprung erfüllt. Eine Ausnahme ist die nichtrationale Funktion e^{-as} , die über den Verschiebungssatz $f(t-a) \Leftrightarrow e^{as}F(s)$ berücksichtigt werden kann.

Die Laplace-Transformation ist zunächst nur für energielosen Anfangszustand anwendbar, Anfangsbedingungen können aber durch entsprechende Ersatzschaltungen für L und C berücksichtigt werden. Die Wechselstromanalyse ist als Sonderfall $s = j\omega$ im s -Bereich enthalten und beschreibt das Netzwerk, wenn der Einschwingvorgang, durch das »Einschalten« des Zeitsignals, abgeklungen ist. (Die Zeitsignale sind nur für $t \geq 0$ definiert, während bei der Wechselstromrechnung sinusförmige Zeitsignale verwendet werden, die für $-\infty < t < \infty$ definiert sind.) Bei stabilen Systemen ist der Einschwingvorgang praktisch in endlicher Zeit abgeschlossen (exponentielles Abklingen).

Wenn die Pole der Systemfunktion in der offenen linken komplexen Ebene liegen, ist das System stabil, da es nur abklingende Eigenschwingungen hat. Nullstellen auf der imaginären Achse sind Nullstellen der Systemfunktion, Erregungen mit einer Frequenz, die diesen Nullstellen entspricht, werden vom System unterdrückt.

Den Frequenzgang (das Bode-Diagramm) eines Systems erhält man durch Berechnen des Betrags und der Phase der Systemfunktion für $s = j\omega$.

Für die Lösung im Zeitbereich muss die Impulsantwort des Systems bekannt sein. Die Antwort des Systems auf eine beliebige Erregung wird dann durch Faltung (Superpositionsintegral) der Impulsantwort mit der Erregung gefunden. Die Impulsantwort kann entweder durch direkte Lösung der Netzwerk-Differentialgleichungen gefunden werden oder durch inverse Laplace-Transformation der Systemfunktion.

Teil III

Analoge Signalverarbeitung

Kapitel 7

Verstärkerschaltungen

Inhalt

7.1	Operationsverstärker	102
7.1.1	Der unbeschaltete Operationsverstärker	103
7.1.2	Der reale Operationsverstärker	107
7.2	Gesteuerte Quellen	108
7.2.1	Spannungsgesteuerte Spannungsquellen	108
7.2.2	Stromgesteuerte Spannungsquellen	108
7.2.3	Spannungsgesteuerte Stromquellen	108
7.2.4	Stromgesteuerte Stromquellen	109
7.3	Addierer und Subtrahierer	109
7.4	Integratoren	110
7.5	Differentiatoren	110
7.6	Lösung von Differentialgleichungen	111
7.7	Zusammenfassung	113

Verstärker sind Schaltelemente die Signale kleiner Eingangsamplitude in Signale größerer Ausgangsamplitude umsetzen, um nachfolgende Einheiten ansteuern zu können. Zum Beispiel muss das Signal eines Mikrofons vom μV -Bereich in den Volt-Bereich verstärkt werden, um es über einen Lautsprecher wiedergeben zu können. Verstärker werden aus elektronischen Bauelementen (Transistoren, Widerständen, Kondensatoren) aufgebaut und sind meistens monolithisch integriert. Der Aufbau von Verstärkern aus Einzeltransistoren in diskreter Schaltungstechnik wird heute nur in Sonderfällen, für die keine integrierte Lösung vorliegt, durchgeführt. Abbildung 7.1 zeigt das Schaltungssymbol eines Verstärkers.

Verstärker setzen Signale kleiner Amplitude in Signale großer Amplitude um.

Die heutige Halbleitertechnologie stellt sehr leistungsfähige und kostengünstige Schaltkreise zur Verfügung und die Verarbeitung von Signalen wird daher wegen der technischen Vorteile zunehmend digital mit Mikroprozessoren, Signalprozessoren oder dedizierter digitaler Hardware durchgeführt. Die zu verarbeitenden Signale stammen aber in der Regel von analogen Quellen (magnetische Sensoren, lichtempfindliche Bauelemente, Thermoelemente, Empfangsantennen) und müssen vor der Digitalisierung verstärkt werden. Nach der digitalen Verarbeitung müssen die Signale wieder in analoge Form gebracht werden und für die "Wiedergabe" an Aktoren (Elektromotoren, Anzeigeeinrichtungen, Heizelemente, Sendeantennen, ...) verstärkt werden.



Abbildung 7.1: Schaltungssymbole für Verstärker

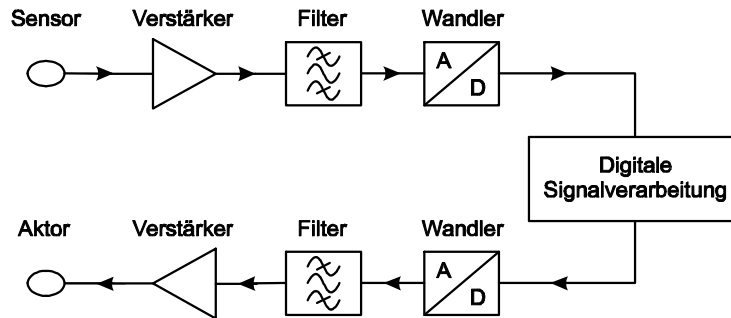


Abbildung 7.2: Blockdiagramm einer Signalverarbeitungskette

Abbildung 7.2 zeigt den typischen Aufbau einer Signalverarbeitungskette.

Obwohl die Schaltungstechnik sehr ähnlich ist, werden die Verstärker für ihre Aufgabe optimiert. Ein (Kleinsignal-)Verstärker für ein Mikrofonsignal ist anders dimensioniert als ein (Leistungs-)Verstärker zur Ansteuerung eines Lautsprechers für ein Popkonzert.

Verstärker werden für verschiedene Leistungen und Frequenzbereiche dimensioniert.

Ein weiteres Unterscheidungsmerkmal ist der Frequenzbereich (untere und obere Grenzfrequenz) für den der Verstärker ausgelegt ist. Bei der unteren Grenzfrequenz unterscheidet man zwischen Gleich- und Wechselspannungsverstärkern. Bei der oberen Grenzfrequenz spricht man von Nieder- und Hochfrequenzverstärkern, wobei die Grenze fließend ist und bei etwa 1 MHz liegt. Verstärker arbeiten mit schmalen und breiten Frequenzbändern und können auch nach der Anwendung in Audio- Video-, Zwischenfrequenz-, Radiofrequenz- und Mikrowellenverstärker eingeteilt werden. Meistens werden hohe Verstärkungsfaktoren verlangt, die durch mehrstufigen Aufbau der Verstärker realisiert werden. Der prinzipielle Schaltungsaufbau der Stufen ist fast identisch, das wesentliche Unterscheidungsmerkmal ist die Kopplung der einzelnen Stufen: Direkte (galvanische) Kopplung bei Gleichspannungsverstärkern, kapazitive Kopplung bei Wechselspannungsverstärkern oder selektive Kopplung bei Schmalbandverstärkern.

Eine besondere Stellung nehmen die Operationsverstärker ein, die universell einsetzbare Gleichspannungsverstärker sind. Für Standardaufgaben kommen fast ausschließlich Operationsverstärker zu Einsatz.

7.1 Operationsverstärker

Operationsverstärker sind universell einsetzbare mehrstufige Gleichspannungsverstärker mit sehr hoher Verstärkung.

Operationsverstärker (OPV) sind mehrstufige Gleichspannungsverstärker mit sehr hoher Verstärkung der Eingangsspannung U_d in Abbildung 7.1. Die (Differenz)verstärkung von Standard-OPVs liegt im Bereich von $10^4 - 10^5$ und erreicht Werte bis 10^7 für Präzisionsverstärker. Es besteht kein prinzipieller schaltungstechnischer Unterschied zwischen normalen Verstärkern und Operationsverstärkern. Der Unterschied zwischen den beiden Typen rührt daher, dass die Eigenschaften des normalen Verstärkers durch seinen inneren Aufbau gegeben sind, während die Eigenschaften des Operationsverstärkers durch die externe Beschaltung festgelegt sind. Das ist möglich, da die Verstärkung der OPVs sehr hohe Werte annimmt.

Die Eigenschaften des OPV sind durch seine **äußere Beschaltung** festgelegt.

Operationsverstärker wurden in der Vergangenheit in Analogrechnern verwendet. Analogrechner wurden eingesetzt, um mathematische Berechnungen (z. B. lösen von Differentialgleichungen) durchzuführen. Dabei mussten Rechenoperationen wie Addition und Integration durchgeführt werden, daher die Bezeichnung Operationsverstärker. Analogrechner haben heute keine Bedeutung, da durch die Digitaltechnik wesentlich genauere und leistungsfähigere Verfahren zur Lösung von Differentialgleichungen zur Verfügung stehen.

Operationsverstärker werden praktisch ausschließlich als monolithisch integrierte Schaltungen angeboten und unterscheiden sich in Größe und Preis kaum von Einzeltransistoren. Es gibt Dutzende von OPV-Typen, die entweder für allgemeine Aufgaben (Universaltypen) oder für spezielle Aufgaben (hohe Genauigkeit, geringes Rauschen, hohe

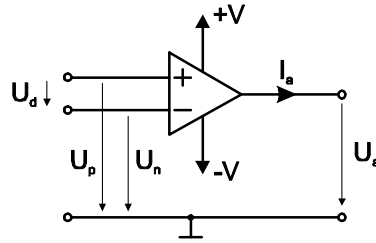


Abbildung 7.3: Schaltsymbol des Operationsverstärkers

Ausgangsspannung, hohe Bandbreite, gute Schalteigenschaften) entwickelt sind.

Die Innenschaltung von Operationsverstärkern wird dem Anwender nicht bekanntgegeben und steht lediglich als Prinzipschaltung oder als Rechnermodell zur Verfügung. Die Berechnung der grundlegenden Eigenschaften der OPV-Schaltungen ist so einfach, dass sie mit der Hand durchgeführt werden kann. Detailliertere Untersuchungen werden mit entsprechenden OPV-Modellen und einem Netzwerksimulator (z. B. SPICE) durchgeführt.

Man unterscheidet vier Typen von OPVs:

1. Die Eingangsspannung U_d steuert die Ausgangsspannung: $U_a = V_d U_d$. Diesen Typ nennt man normalen OPV.
2. Die Eingangsspannung U_d steuert den Ausgangsstrom: $I_a = S_d U_d$. Diesen Typ nennt man Transkonduktanz-Verstärker.
3. Der Eingangsstrom I_{in} steuert die Ausgangsspannung: $U_a = Z I_{in}$. Diesen Typ nennt man Transimpedanz-Verstärker.
4. Der Eingangsstrom I_{in} steuert den Ausgangsstrom: $I_a = V_i I_{in}$. Diesen Typ nennt man Strom-Verstärker.

Wir werden uns im weiteren Verlauf auf den normalen Operationsverstärker beschränken.

7.1.1 Der unbeschaltete Operationsverstärker

Abbildung 7.3 zeigt das Schaltungssymbol eines normalen idealen Operationsverstärkers.

Der nichtinvertierende P-Eingang wird mit einem +, der invertierende N-Eingang wird mit einem – gekennzeichnet. Der OPV wird mit zwei Betriebsspannungen, die eine gegen Masse positiv, die andere gegen Masse negativ, versorgt. Damit sind Eingangs- und Ausgangsruhepotentiale von 0 V möglich. Die OPVs selbst besitzen keinen Masseanschluss, obwohl Eingangs- und Ausgangsspannungen darauf bezogen werden.

Die Differenzverstärkung des idealen OPV ist sehr hoch, im theoretischen Grenzfall ist sie $V \rightarrow \infty$. Der Eingangswiderstand (Widerstand zwischen Eingang und Masse) des Verstärkers beim P- und N-Eingang ist theoretisch unendlich $R \rightarrow \infty$. Wird also eine Spannung (bezogen auf die Masse) an einen Eingang gelegt, fließt kein Strom in oder aus dem OPV.

OPVs werden nur im beschalteten Zustand eingesetzt, im unbeschalteten Zustand reichen schon sehr kleine Spannungen aus, um den Operationsverstärker bis an die positive oder negative Versorgungsspannung auszusteuern.

Die Verstärkung des idealen OPVs ist ∞ . Der Eingangswiderstand ist ebenfalls ∞ .

OPVs werden nur im beschalteten Zustand eingesetzt.

Der invertierende Operationsverstärker

Abbildung 7.4 zeigt den OPV in invertierender Beschaltung.

Zur Berechnung des Verhältnisses U_a/U_e wenden wir die Kirchhoff'sche Knotenregel auf den Knoten (1) an

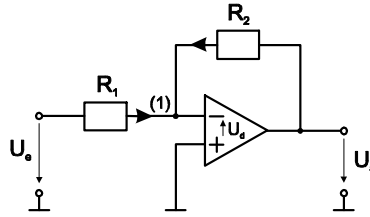


Abbildung 7.4: Der invertierende OPV

$$I_{R_1} + I_{R_2} = 0 \rightarrow \frac{U_{R_1}}{R_1} + \frac{U_{R_2}}{R_2} = 0 \quad (7.1)$$

In den negativen Eingang des Verstärkers fließt kein Strom, da der Eingangswiderstand des idealen OPV unendlich ist. Daher ist

$$I_{R_1} + I_{R_2} = 0 \rightarrow \frac{U_{R_1}}{R_1} + \frac{U_{R_2}}{R_2} = 0 \quad (7.2)$$

$$\frac{U_e + U_d}{R_1} + \frac{U_a + U_d}{R_2} = 0 \quad (7.3)$$

$$U_a = VU_d \quad (7.4)$$

Wir setzen $U_d = U_a/V$ ein und erhalten

$$-\frac{U_e}{U_a} = \frac{R_1}{R_2} \left(1 + \frac{1}{V}\right) + \frac{1}{V} \quad (7.5)$$

Die Größe V bezeichnet man als Leerlaufverstärkung (open loop gain) des Operationsverstärkers. Das Verhältnis von Ausgangsspannung zu Eingangsspannung $A = U_a/U_e$ nennt man Spannungsverstärkung (closed loop gain).

Typische Werte für die Leerlaufverstärkung sind $V = 10^5$ und für die Spannungsverstärkung $A = 10^2$. Für große Werte von V können wir $1/V$ vernachlässigen und erhalten

Die Spannungsverstärkung A des gegengekoppelten OPV ist negativ, praktisch unabhängig von V (solange V groß genug ist) und hängt nur von der externen Beschaltung ab.

Anmerkung 56 Wegen der hohen Verstärkung von OPVs ist die Differenzspannung U_d sehr klein und man sagt, dass der negative Eingang des OPV virtuell auf Masse liegt.

Der Eingangswiderstand des invertierenden OPV nach Abbildung 7.4 ist gleich R_1 , da der invertierende Eingang virtuell auf Masse liegt. Der Ausgangswiderstand ist sehr gering, das Verhalten der Ausgangsspannung ist nahe dem Verhalten einer idealen Spannungsquelle.

Der nichtinvertierende Verstärker

Die Beschaltung des nichtinvertierenden OPV ist in Abbildung 7.5 gezeigt.

Die Untersuchung der Schaltung liefert folgende Zusammenhänge

$$U_n = \frac{R_1}{R_1 + R_2} U_a \quad (7.6)$$

$$U_a = VU_d = V(U_e - U_n) \quad (7.7)$$

Nach Umformung wird daraus

$$A = \frac{U_a}{U_e} = \frac{V}{1 + \alpha V} \quad \alpha = \frac{R_1}{R_1 + R_2} \quad (7.8)$$

Die Spannungsverstärkung des **invertierenden OPV** ist negativ und hängt praktisch nur von dem Verhältnis der angeschlossenen Widerstände ab.

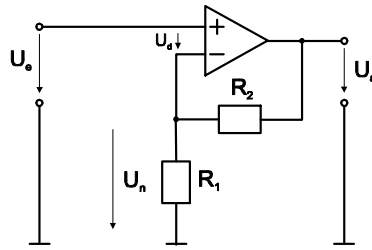


Abbildung 7.5: Der nichtinvertierende OPV

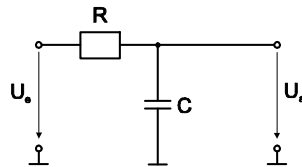


Abbildung 7.6: RC-Tiefpass

Für den Ausdruck $\frac{U_e}{U_a} = \frac{1}{A} = \frac{1}{V} + \frac{R_1}{R_1 + R_2}$ sehen wir, dass für hohe Werte der Verstärkung $1/V$ vernachlässigt werden kann und wir erhalten

$$\frac{1}{A} \doteq \frac{R_1}{R_1 + R_2} \tag{7.9}$$

$$A \doteq 1 + \frac{R_2}{R_1} \tag{7.10}$$

Eingangsspannung und Ausgangsspannung haben dasselbe Vorzeichen, die Eingangsspannung wird bei dieser Beschaltung des OPV nicht invertiert.

Das Eingangssignal der Schaltung in Abbildung 7.5 liegt direkt am nichtinvertierenden Eingang des OPV, der Eingangswiderstand ist daher hochohmig. Der Ausgangswiderstand ist, wie beim invertierenden OPV, sehr niederohmig.

Im Gegensatz zum invertierenden OPV bei dem die Eingangsspannungen U_n und U_p praktisch Null sind, tritt beim nichtinvertierenden OPV eine Gleichtaktaussteuerung auf, da $U_n \doteq U_p = U_e$. Ideale OPVs verstärken lediglich die Differenzspannung U_d und zwar unabhängig von der Gleichtaktspannung. Bei realen OPVs ist das nicht der Fall, man misst eine Gleichtaktverstärkung, die typisch in der Größenordnung von 1 liegt und damit mehrere Zehnerpotenzen kleiner als die Differenzspannungsverstärkung ist. Als Kenngröße für OPVs wird die Gleichtaktunterdrückung $G = \frac{U_{\text{differenz}}}{U_{\text{gleich}}}$ angegeben.

Frequenzverhalten des Operationsverstärkers

Die Differenzverstärkung eines realen Operationsverstärkers nimmt bei höheren Frequenzen ab. Die Ursache dafür sind Kapazitäten der Transistoren der Verstärkerschaltungen und Schaltungswiderstände, die Tiefpässe bilden. In erster Näherung kann das Frequenzverhalten des OPV durch einen RC-Tiefpass, wie in Abbildung 7.6 gezeigt, modelliert werden.

Die Netzwerkanalyse der Schaltung Abbildung 7.6 liefert unter Verwendung der Abkürzung $f_g = 1/2\pi RC$

$$\frac{\vec{U}_a}{\vec{U}_e} = \frac{1}{1 + j(f/f_g)} = \vec{H}(f) \tag{7.11}$$

Daraus berechnen wir den Betrag der Systemfunktion

Beim **nichtinvertierenden OPV** haben Eingangsspannung und Ausgangsspannung das selbe Vorzeichen. Die Verstärkung wird durch die externe Beschaltung bestimmt.

Das **Frequenzverhalten des OPV** kann annähernd durch einen **RC-Tiefpass** modelliert werden.

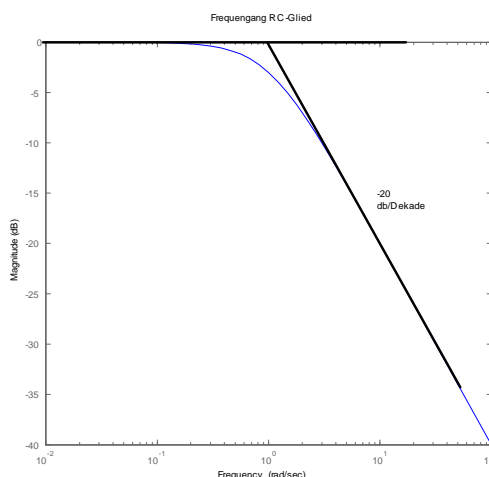


Abbildung 7.7: Asymptoten an Frequenzgang

$$|H(f)|_{dB} = 20 \log |H(f)| = 20 \log \left(\frac{1}{\sqrt{1 + (f/f_g)^2}} \right) \quad (7.12)$$

$$= 20 \log(1) - 20 \log \sqrt{1 + (f/f_g)^2} \quad (7.13)$$

$$= -10 \log \left[1 + \left(\frac{f}{f_g} \right)^2 \right] \quad (7.14)$$

Gleichung (7.14) liefert für kleine Werte $f \ll f_g$ den Wert Null und kann daher in Abbildung 7.7 als Gerade durch den Ursprung dargestellt werden. Für große Werte $f \gg f_g$ erhalten wir $|H(f)|_{dB} = -20 \log(f/f_g)$, eine Gerade mit der Steigung -20 dB/Dekade (in logarithmischer Darstellung), wie in Abbildung 7.7 gezeigt. Den Schnittpunkt beider Geraden bezeichnen wir als Grenzfrequenz. Bei f_g ist das Verhältnis von Ausgangsspannung zu Eingangsspannung $\frac{1}{\sqrt{2}} \doteq -3 \text{ dB}$, man nennt die Grenzfrequenz daher auch den 3 dB-Punkt.

Bei der **Grenzfrequenz** (3 dB-Punkt) ist das Verhältnis von zu Ausgangsspannung Eingangsspannung $1/\sqrt{2} \doteq -3 \text{ dB}$.

In erster Näherung verhält sich der Frequenzgang des OPV wie ein RC-Glied, mit dem Unterschied, dass durch den OPV eine Verstärkung auftritt. Die Verstärkung bei $f = 0$ nennen wir V_0 , die Leerlaufgleichspannungsverstärkung.

$$V(f) = \frac{V_0}{1 + j(f/f_g)} \quad (7.15)$$

Die Frequenzabhängigkeit der Leerlaufverstärkung eines OPV ist in Abbildung 7.8 dargestellt.

Berücksichtigen wir die Frequenzabhängigkeit der Leerlaufverstärkung für den gegengekoppelten OPV, setzen (7.15) in die Gleichung (7.8) ein, dann erhalten wir

$$A(f) = \frac{U_a}{U_e} = \frac{V(f)}{1 + \alpha V(f)} \quad (7.16)$$

$$A(f) = \frac{\frac{V_0}{1 + j(f/f_{gol})}}{1 + \alpha \frac{V_0}{1 + j(f/f_{gol})}} \quad (7.17)$$

Mit den Abkürzungen

$$A_0 = \frac{V_0}{1 + \alpha V_0} \quad f_{gcl} = f_{gol}(1 + \alpha V_0) \quad (7.18)$$

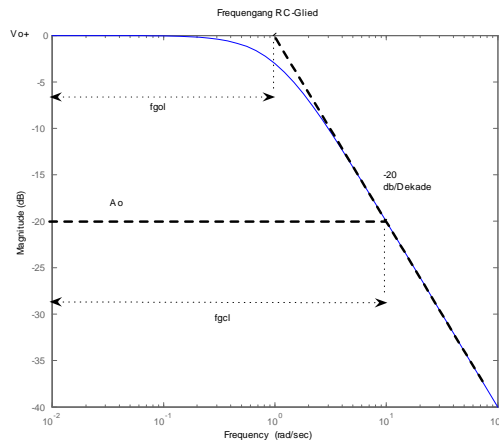


Abbildung 7.8: Frequenzgang ohne und mit Gegenkopplung

wird daraus

$$A(f) = \frac{A_0}{1 + j(f/f_{gcl})} \tag{7.19}$$

f_{gol} ist die Grenzfrequenz der Leerlaufverstärkung, f_{gcl} ist die Grenzfrequenz der gegengekoppelten Schaltung.

Bilden wir das Produkt aus Verstärkung und Bandbreite so erhalten wir

$$A_0 f_{gcl} = \frac{V_0}{1 + \alpha V_0} \cdot f_{gol} (1 + \alpha V_0) = V_0 f_{gol} = 1 \cdot f_T \tag{7.20}$$

Das Produkt aus Verstärkung und Bandbreite ist unabhängig von α und wird Verstärkungsbandbreitenprodukt f_T genannt. Bei der Frequenz f_T ist die Verstärkung 1 oder 0 dB (unity-gain bandwidth). Das Verstärkungs-Band-breitenprodukt von OPVs liegt im Bereich von 0.1MHz bis 1GHz.

Anmerkung 57 Die Bandbreite des gegengekoppelten Verstärkers wird auf Kosten einer geringeren Verstärkung erhöht, das Produkt aus Verstärkung und Bandbreite ist konstant.

In Abbildung 7.8 ist dieser Zusammenhang grafisch dargestellt.

7.1.2 Der reale Operationsverstärker

Die Differenzverstärkung des realen OPV liegt im Bereich zwischen 10^4 bis 10^7 . Die Gleichtaktunterdrückung liegt ebenfalls in diesem Bereich. Die Übertragungskennlinie der realen OPV $U_a = U_a(U_d)$ geht nicht durch den Ursprung, sondern ist um die Offsetspannung verschoben. Werte der Offsetspannung liegen typisch im Bereich von $10\mu V$ bis $1mV$.

Die Eingangswiderstände der Operationsverstärker sind nicht ∞ , es fließen daher Eingangsströme im Bereich von pA bis $10nA$. Durch die Beschaltung (Gegenkopplung) der OPVs werden die hohen Eingangswiderstände noch weiter um die Schleifenverstärkung $g = \frac{V}{A}$ vergrößert. Der Ausgangswiderstand eines OPV ist sehr gering, durch die Gegenkopplung wird dieser Widerstand um die Schleifenverstärkung reduziert.

Das Abweichen von realen OPVs muss durch Ersatzschaltungen berücksichtigt werden, deren Werte aus den Datenblättern der OPVs kommen oder als Modelle vom Hersteller zur Verfügung gestellt werden.

Wie bei allen analogen Schaltungen tritt auch bei OPV Rauschen auf. Das Rauschen lässt sich in Form einer auf den Eingang bezogenen Rauschspannungsdichte beschreiben. Die Rauschspannung ist frequenzabhängig, ist bei niedrigen Frequenzen am höchsten und

Das **Verstärkungs-Bandbreitenprodukt** gibt die Frequenz an, bei der die Verstärkung so weit abnimmt, dass die Amplitude des Ausgangssignals gleich der Amplitude des Eingangssignals ist.

Die (idealen) Annahmen sind bei realen Verstärkern verletzt und müssen durch Ersatzschaltungen modelliert werden.

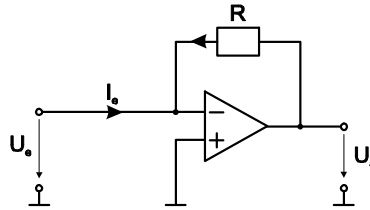


Abbildung 7.9: Stromgesteuerte Spannungsquelle

nimmt bei höheren Frequenzen ab. Darüber hinaus hängt die Rauschspannung noch von Technologie der Schaltung ab. Die Eingangsrauschspannungsdichte liegt im Bereich von $1 - 25nV/\sqrt{Hz}$.

7.2 Gesteuerte Quellen

Die einfachste Aufgabe bei der analogen Signalverarbeitung ist die Verstärkung von Signalen. Die von Sensoren stammenden Signale können als Spannungs- oder Stromquellen auftreten, deren niedrige Signalpegel für die Weiterverarbeitung verstärkt werden müssen.

7.2.1 Spannungsgesteuerte Spannungsquellen

Spannungsgesteuerte Spannungsquellen können durch invertierende oder nichtinvertierende Operationsverstärker realisiert werden.

Spannungsgesteuerte Spannungsquellen mit niedrigem Ausgangswiderstand lassen sich mit Hilfe des invertierenden (Abbildung 7.4) oder nichtinvertierenden (Abbildung 7.5) OPVs realisieren. Die Verstärkung ist durch die externe Beschaltung einstellbar. Der Eingangswiderstand des invertierenden OPV ist R_1 , wobei die Werte von R_1 relativ niedrig sind. Diese Schaltung ist daher nicht für die Verstärkung von Signalquellen mit hohem Innenwiderstand geeignet. Bei Quellen mit hohem Innenwiderstand, die daher nicht belastbar sind, kann der nichtinvertierende OPV eingesetzt werden, dessen Eingangswiderstand sehr hoch ist. Der Nachteil des nichtinvertierenden OPV ist der durch die Gleichtaktaussteuerung entstehende Fehler. Dieser Nachteil tritt beim invertierenden OPV nicht auf.

7.2.2 Stromgesteuerte Spannungsquellen

Eine Stromgesteuerte Spannungsquelle kann aus einem invertierenden OPV durch Nullsetzen des eingangswiderstands erzeugt werden.

Abbildung 7.9 zeigt eine Realisierung mit Hilfe eines OPV.

Wegen der hohen Spannungsverstärkung liegt der invertierende Eingang virtuell auf Masse, wodurch sich ein sehr niedriger Eingangswiderstand ergibt, der ja Voraussetzung für die Stromsteuerung ist. Die Ausgangsspannung dieser Schaltung ist $U_a = (-R)I_e$, da der Strom in den Verstärkereingang vernachlässigt werden kann.

$$U_a = -RI_e \quad (7.21)$$

$$R_e = \frac{R}{V} \quad (7.22)$$

Der Eingangswiderstand der Schaltung ist $R_e = \frac{U_e}{I_e} = \frac{R}{V} \frac{U_a}{U_a}$, wobei V die (sehr große) Leerlaufverstärkung des OPV ist.

7.2.3 Spannungsgesteuerte Stromquellen

Eine Spannungsgesteuerte Stromquelle erhält man, indem man bei einem invertierenden oder nichtinvertierenden OPV den Gegenkopplungswiderstand durch den Verbraucher ersetzt.

Wie in Abbildung 7.10 gezeigt, fließt beim invertierenden und nichtinvertierenden OPV durch den Gegenkopplungswiderstand R_2 der Strom $I_a = U_e/R_1$.

Die Übertragungsfunktion ist also

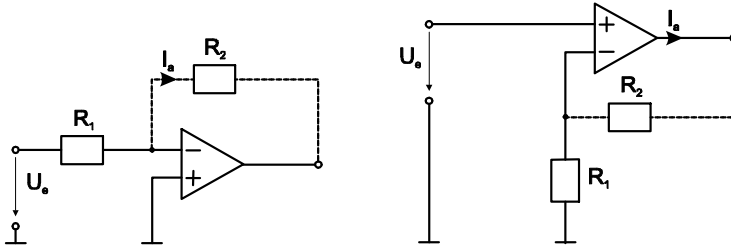


Abbildung 7.10: Spannungsgesteuerte Stromquellen

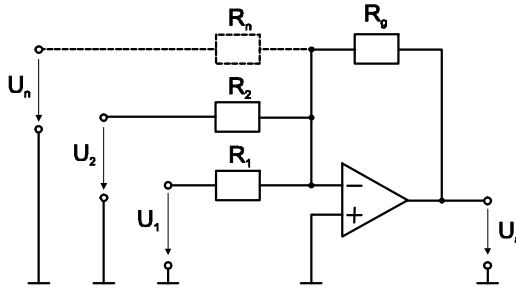


Abbildung 7.11: Summationsverstärker

$$I_a = \frac{1}{R_1} U_e \tag{7.23}$$

und wir haben spannungsgesteuerte Stromquellen, wenn man den Verbraucher als Gegenkopplungswiderstand einsetzt¹. Die Eingangsimpedanz der beiden Schaltungen ist wie beim (nicht)invertierenden OPV.

Der Innenwiderstand der Stromquelle, also der Ausgangswiderstand unserer beiden Schaltungen, sollte theoretisch ∞ sein, was praktisch nicht erreichbar ist. Dazu kommt, dass bei hochohmigen Ausgangswiderständen die praktisch auftretenden Kapazitäten der Schaltung nicht vernachlässigbar sind und der Ausgangswiderstand daher stark frequenzabhängig ist.

7.2.4 Stromgesteuerte Stromquellen

Die invertierende OPV-Schaltung in Abbildung 7.10 hat den Eingangswiderstand $R_e = R_1$. Macht man $R_1 = 0$, dann wird $I_a = I_e$ und man hat eine stromgesteuerte Stromquelle mit sehr guten Eigenschaften.

Eine stromgesteuerte Stromquelle kann aus einer spannungsgesteuerten Stromquelle durch Nullsetzen des Eingangswiderstands erzeugt werden.

7.3 Addierer und Subtrahierer

Der invertierende OPV mit mehreren Eingängen nach Abbildung 7.11 summiert (und invertiert) mehrere Eingangsspannungen.

Durch Anwendung der Kirchhoff'schen Knotenregel erhalten wir

$$\frac{U_1}{R_1} + \frac{U_2}{R_2} + \dots + \frac{U_n}{R_n} = -\frac{U_a}{R_g} \tag{7.24}$$

$$-U_a = \frac{R_g}{R_1} U_1 + \frac{R_g}{R_2} U_2 + \dots + \frac{R_g}{R_n} U_n \tag{7.25}$$

Die Subtraktion kann auf eine Addition zurückgeführt werden, indem man das zu subtrahierende Signal vorher mit einer invertierenden OPV-Schaltung verstärkt.

Ein Addierer oder Subtrahierer ist ein invertierender OPV mit mehreren Eingängen.

¹Für spannungsgesteuerte Stromquellen mit geerdetem Verbraucher müssen andere Schaltungen eingesetzt werden.

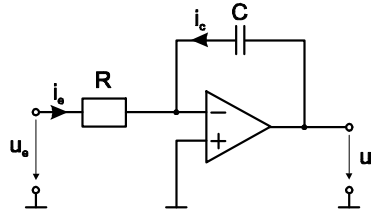


Abbildung 7.12: Integrator

7.4 Integratoren

Ersetzt man bei einem invertierenden OPV den Gegenkopplungswiderstand durch eine Kapazität, bekommt man einen Integrator.

Ersetzt man beim invertierenden OPV den Gegenkopplungswiderstand durch eine Kapazität, wie in Abbildung 7.12 gezeigt, dann liefern die Netzwerkgleichungen die Beziehung

$$i_e + i_c = 0 \quad (7.26)$$

$$i_c = C \frac{du_a}{dt} \quad (7.27)$$

$$\frac{u_e}{R} + C \frac{du_a}{dt} = 0 \quad (7.28)$$

$$u_a = -\frac{1}{RC} \int_0^\tau u_e dt + U_a|_{t=0} \quad (7.29)$$

Die Ausgangsspannung ist das zeitliche Integral der Eingangsspannung, $U_a|_{t=0}$ ist die Anfangsbedingung bzw. die Spannung an C zum Zeitpunkt $t = 0$.

Liegt eine konstante Eingangsspannung am Integrator an, so entsteht am Ausgang eine linear ansteigende Spannung, liegt am Eingang eine Rechteckwechselfspannung an, so entsteht am Ausgang eine Dreiecksspannung.

Legt man die Eingangsspannung $u_e = U_e \cos \omega t$ an, so wird die Ausgangsspannung

$$u_a = -\frac{1}{RC} \int_0^\tau u_e dt = -\frac{U_e}{RC\omega} \sin \omega t \quad (7.30)$$

Das Verhältnis von Ausgangs- zu Eingangsspannung (der Frequenzgang) ist

$$H = \frac{\vec{U}_a}{\vec{U}_e} = -\frac{Z_C}{R} = -\frac{1}{RCs} \Big|_{s=j\omega} \quad (7.31)$$

$$|H| = \frac{1}{RC\omega} \quad (7.32)$$

Der Betrag des Frequenzgangs ist in 7.13 dargestellt.

In der logarithmischen Darstellung der Abbildung 7.13 erkennt man, dass die Steigung des Betrags des Frequenzgangs -20 dB/Dekade (6 dB/Oktave) beträgt. Diese Eigenschaft ist ein einfaches Kennzeichen dafür, ob sich eine Schaltung als Integrator verhält.

Beim realen Integrator können sich Eingangsruhestrom und Offsetspannung sehr ungünstig auswirken, da sich deren Wirkung zeitlich summiert.

7.5 Differentiatoren

Vertauscht man bei einem Integrator R und C , erhält man einen Differentiator.

Vertauscht man bei einem Integrator R und C , wie in Abbildung 7.14 gezeigt, und wenden wir die Knotenregel auf den Summationspunkt an, dann erhalten wir

$$C \frac{du_e}{dt} + \frac{u_a}{R} = 0 \quad (7.33)$$

$$u_a = -RC \frac{du_e}{dt} \quad (7.34)$$

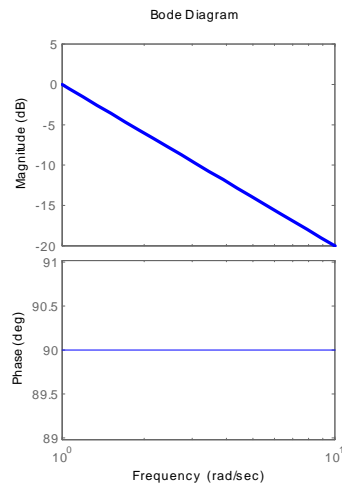


Abbildung 7.13: Bodediagramm des idealen Integrators

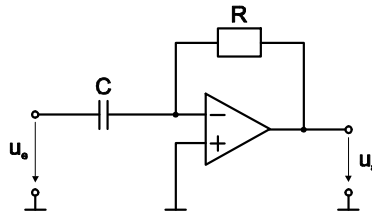


Abbildung 7.14: Differentiator

Die Schaltung verhält sich wie ein Differentiator.

Für sinusförmige Eingangsspannungen $U_e \sin \omega t$ erhalten wir die Ausgangsspannung $u_a = -RC\omega U_e \cos \omega t$. Der Betrag des Frequenzgangs ist $|H| = RC\omega$. Bei doppelt logarithmischer Darstellung eine Gerade mit der Steigung 20 dB/Dekade (beim Integrator -20 dB/Dekade).

Die praktische Realisierung von Differentiatoren bereitet Schwierigkeiten, da große Schwingneigung besteht. (Die Erklärung dafür liegt in der Phasennacheilung des Gegenkopplungsnetzwerkes und der Phasennacheilung des realen OPV, die hier nicht besprochen wurden.)

7.6 Lösung von Differentialgleichungen

Viele technisch physikalische Probleme lassen sich in Form von Differentialgleichungen (DGL) darstellen. Die Lösung dieser Differentialgleichungen lässt sich dadurch finden, dass man sie durch die beschriebenen Analogrechsaltungen nachbildet und die sich einstellende Ausgangsspannung (die "Lösung" der DGL) misst. Um die beim Differentiator angedeuteten Stabilitätsprobleme zu umgehen, formt man die DGL so um, dass ausschließlich Integratoren verwendet werden.

Die Umsetzung einer DGL in eine Schaltung zeigen wir anhand eines Beispiels einer zeitabhängigen DGL 2. Ordnung

$$\tau^2 \frac{d^2 y}{dt^2} + k_1 \tau \frac{dy}{dt} + k_0 y = f\left(\frac{t}{\tau}\right) \quad (7.35)$$

Wir bringen die nicht abgeleiteten Größen auf eine Seite und erhalten

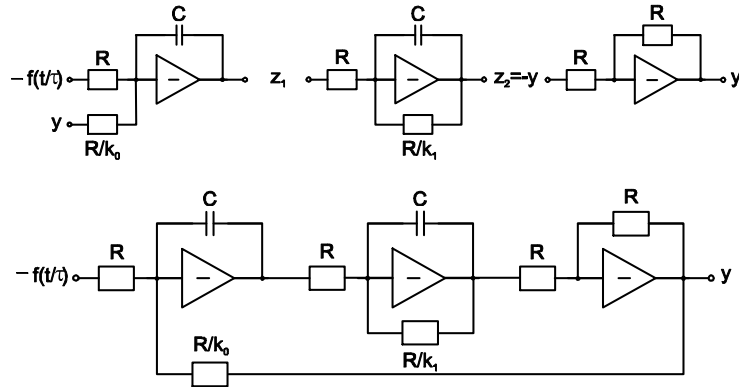


Abbildung 7.15: OPV-Schaltung zur Lösung von DGL 2. Ordnung

$$k_0 y - f\left(\frac{t}{\tau}\right) = -\tau^2 \frac{d^2 y}{dt^2} - k_1 \tau \frac{dy}{dt} \quad (7.36)$$

Jetzt integrieren wir diese Gleichung und dividieren durch $-\tau$

$$\underbrace{-\frac{1}{\tau} \int \left[k_0 y - f\left(\frac{t}{\tau}\right) \right] dt}_{z_1} = \tau \frac{dy}{dt} + k_1 y \quad (7.37)$$

Die linke Seite z der Gleichung lässt sich mit der Integrationsschaltung der Abbildung 7.12 berechnen und wir können sie als einen Schaltungsblock auffassen: Die Eingangsgrößen sind die Erregungsgröße $f\left(\frac{t}{\tau}\right)$ und $k_0 y$, wobei y die zunächst unbekannte und erst zu berechnende Ausgangsgröße ist. Wir erhalten dadurch die DGL

$$z_1 = \tau \frac{dy}{dt} + k_1 y \quad (7.38)$$

$$\frac{1}{\tau} (z_1 - k_1 y) = \frac{dy}{dt} \quad (7.39)$$

die wir wie vorher integrieren. Wir erhalten

$$\underbrace{-\frac{1}{\tau} \int (z_1 - k_1 y) dt}_{z_2} = -y \quad (7.40)$$

Werden Funktionen als Zeitsignale aufgefasst, können Differentialgleichungen gelöst werden, indem sie durch analoge Schaltungen nachgebildet werden und das Zeitsignal, das die Lösung darstellt, gemessen wird.

Diesen Ausdruck können wir mit einem zweiten Summenintegrator integrieren, dessen Eingangsgrößen z_1 und $k_1 y$ sind, wobei auch hier y zunächst unbekannt ist. Die gesuchte Ausgangsgröße y erhalten wir durch Inversion von z_2 .

Abbildung 7.15 zeigt die Teilschaltungen und die Gesamtschaltung zur Lösung der DGL.

Analogrechner lösen Differentialgleichungen auf sehr elegante Art und mit hoher Geschwindigkeit. Parameteränderungen der DGL sind durch Änderung der Bauelementewerte der OPV-Schaltungen sehr leicht durchzuführen. Wegen des geringen Dynamikbereichs analoger Schaltung und der daher nur geringen "Rechengenauigkeit" und wegen des nichtidealen Verhaltens der OPVs, Widerstände und Kapazitäten haben Analogrechner heute ihre Bedeutung verloren und wurden durch Digitalrechner ersetzt, mit denen eine wesentlich höhere Genauigkeit erreichbar ist. Der Hybridrechner EAI SIMSTAR (eine Kombination aus Analog- und Digitalrechner) der TU Wien wurde Ende 1991 außer Betrieb genommen.

7.7 Zusammenfassung

Verstärker sind elektronische Schaltungen die kleine Eingangssignale verstärken, damit sie in nachfolgenden Einheiten verarbeitet werden können. Je nach Aufgabenstellung müssen die Verstärkerschaltungen optimiert werden, was zu einer Vielzahl von Lösungen führt. Heute werden Verstärker nur in seltenen Fällen aus diskreten Bauelementen aufgebaut sondern stehen normalerweise als integrierte Schaltungen zur Verfügung.

Eine Sonderstellung haben die Operationsverstärker, deren Verhalten nicht von der Innenschaltung, sondern von der externen Beschaltung abhängt. Mit Operationsverstärkern können Schaltungen aufgebaut werden, die mathematische Operationen wie Addition, Subtraktion, Multiplikation, Differentiation und Integration mit Signalen durchführen können. Man kann sogar Schaltungen zum Lösen von Differentialgleichungen aufbauen. In der Praxis haben solche Schaltungen durch die Digitaltechnik, die wesentlich leistungsfähiger und vor allem genauer ist, ihre Bedeutung verloren.

Kapitel 8

Analoge Signalverarbeitung

Inhalt

8.1	Analoge Filter	116
8.1.1	Approximation eines idealen Tiefpassfilters	117
8.1.2	Realisierung von analogen Filtern	125
8.1.3	Normierung	126
8.1.4	Frequenztransformation	126
8.2	Nichtlineare analoge Signalverarbeitung	127
8.2.1	Signalgleichrichtung	127
8.2.2	Komparator	127
8.3	Zusammenfassung	129

Mit Mikroprozessoren und (digitalen) Signalprozessoren kann man heute mathematische Operationen in fast beliebiger Genauigkeit ausführen. Die zu verarbeitenden Signale liegen allerdings in der Regel in Form von kontinuierlichen Größen, häufig in Form von elektrischen Spannungen oder Strömen vor. Für die digitale Verarbeitung müssen diese analogen Signale zuerst mit Hilfe von Analog-Digital-Wandlern umgesetzt werden und nach der Signalverarbeitung meistens wieder durch Digital-Analog-Wandler in analoge Form zurückgebracht werden.

In vielen Anwendungen sind die Genauigkeitsanforderungen nicht sehr hoch und man kann kostengünstige Lösungen mit analogen Schaltungen realisieren, wobei die Genauigkeit, die man mit analogen Schaltungen erreichen kann, bei 0.1 % liegt.

Darüber hinaus gibt es Anwendungen bei denen die auftretenden Signalfrequenzen so hoch sind, dass man sie mit heutiger Technologie nicht in Echtzeit A/D-wandeln kann.

Schließlich ist in fast jedem digitalen Signalverarbeitungssystem eine (analoge) Signalverarbeitung auf der A/D-Seite und auf der D/A-Seite erforderlich, um das Frequenzband der Eingangssignale zu begrenzen und um die digitalen Ausgangswerte zu interpolieren.

Die analoge Signalverarbeitung wird daher nie ihre Berechtigung verlieren, wenngleich ihre wirtschaftliche Bedeutung zu Gunsten der digitalen Verarbeitung abnimmt.

Die Entwicklung analoger Schaltungen erfordert sehr viel schaltungstechnische Erfahrung. Trotz guter Simulationsmöglichkeiten entscheidet letztlich die physikalische Realisierung, ob eine Schaltung die gewünschten Spezifikationen erfüllt. Bei komplizierten Schaltungen ist in der Regel immer ein Versuchsaufbau und eine nachfolgende Schaltungsoptimierung nötig. Es sei hier daran erinnert, dass die Modellierung durch R, L und C nicht genau der physikalischen Wirklichkeit entspricht und dass es keine idealen Bauelemente gibt.

Dieses Problem tritt bei digitalen Schaltungen nicht auf. Digitale Systeme kann man sehr exakt modellieren und die Umsetzung in eine Schaltung funktioniert in der Regel auf Anhieb, wenn man von Überlegungsfehlern beim Entwurf absieht.

Im weiteren Verlauf werden wir uns, wie schon bei der Netzwerkanalyse, auf idealisierte Schaltungen beschränken und lediglich das prinzipielle Verhalten der Schaltungen

Analoge Schaltungen können nicht so hohe Genauigkeit erreichen wie digitale Schaltungen. Bei geringen Genauigkeitsanforderungen werden sie oft bevorzugt, da sie kostengünstiger sind.

Analoge Schaltungen haben auch heute noch ihre Bedeutung, obwohl sie in manchen Bereichen zunehmend durch digitale Schaltungen ersetzt werden.

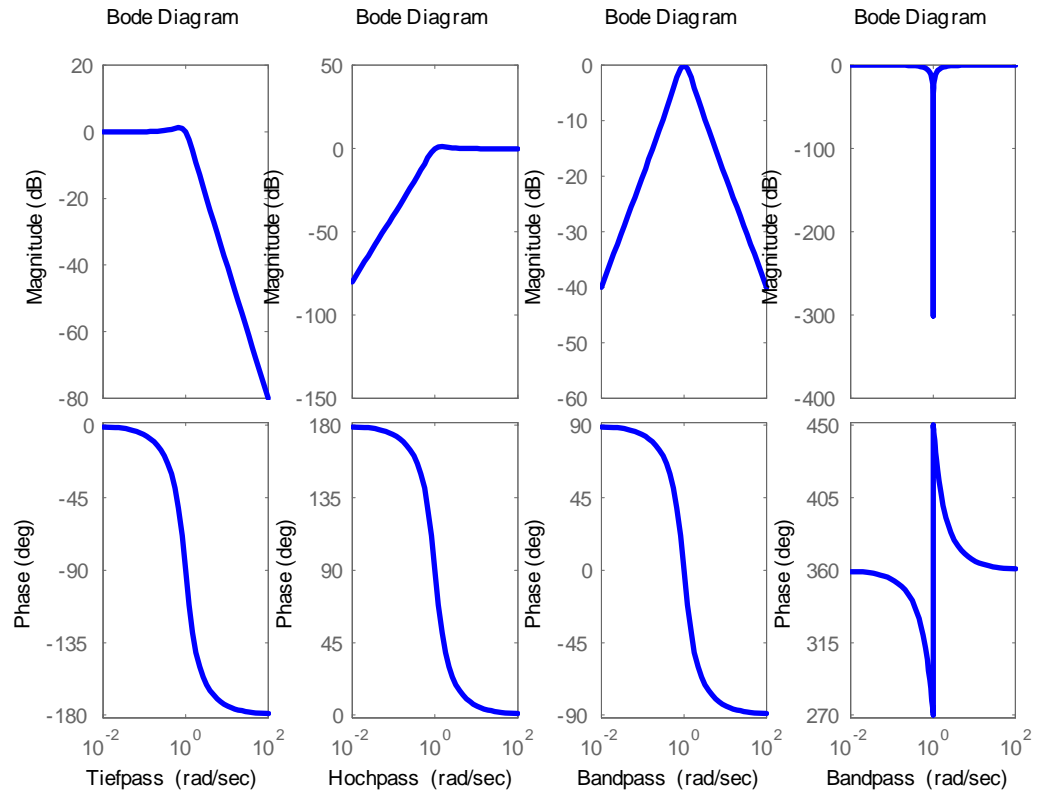


Abbildung 8.1: Frequenzverhalten von Filtern

darstellen.

8.1 Analoge Filter

Analoge **Filter** sind elektrische Netzwerke, die bestimmte **Frequenzanteile** von Signalen **unterdrücken**.

Elektrische Netzwerke werden häufig zur Unterdrückung von Frequenzanteilen in Signalen eingesetzt. Entsprechend nennt man Netzwerke Tiefpassfilter / Tiefpässe, wenn sie nur tiefe Frequenzen ($\omega \in [0, \omega_{grenz}]$) durchlassen, Hochpassfilter / Hochpässe, wenn sie nur hohe Frequenzen durchlassen ($\omega \in [\omega_{grenz}, \omega_{\infty}]$), Bandpassfilter / Bandpässe, wenn sie nur Frequenzen in einem bestimmten Frequenzbereich durchlassen ($\omega \in [\omega_{unten}, \omega_{oben}]$) und Bandsperrungen, wenn sie einen Frequenzbereich unterdrücken ($\omega \notin [\omega_{unten}, \omega_{oben}]$).

Bei einfachen Netzwerken kann man das Übertragungsverhalten leicht aus der Schaltungsstruktur erkennen: Kapazitäten verhalten sich bei niedrigen Frequenzen wie ein Leerlauf, bei hohen Frequenzen wie ein Kurzschluss, Induktivitäten verhalten sich bei niedrigen Frequenzen wie ein Kurzschluss, bei hohen Frequenzen wie ein Leerlauf.

Abbildung 8.1 zeigt das Verhalten der verschiedenen Filtertypen.

Analoge Filter werden entweder mit passiven elektrischen Komponenten (R, L, C) oder mit Operationsverstärkerschaltungen realisiert. Das Übertragungsverhalten der Filter wird durch die Systemfunktion des Filters, d.h. das Verhältnis von Ausgangs- zu Eingangsspannung in Abhängigkeit vom komplexen Frequenzparameter s beschrieben.

$$H(s) = \frac{U_a(s)}{U_e(s)} \quad (8.1)$$

Die Systemfunktion ist eine gebrochen rationale Funktion¹ der Form

¹Wir beschränken uns auf LTI-Systeme.

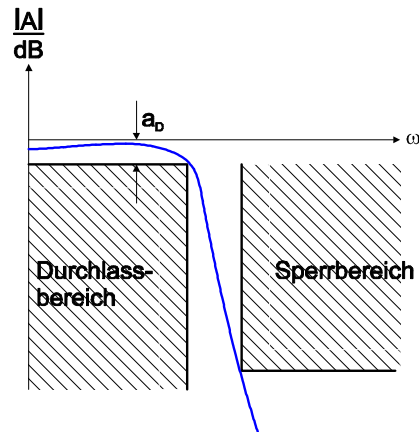


Abbildung 8.2: Toleranzschema eines Filters

$$H(s) = \frac{b_m s^m + b_{m-1} s^{m-1} + b_1 s + b_0}{a_n s^n + a_{n-1} s^{n-1} + a_1 s + a_0} \quad (8.2)$$

Ein Beispiel für die Systemfunktion eines Tiefpassfilters wäre

$$H(s) = \frac{1}{2s^3 + 2s^2 + 2.5s + 1} \quad (8.3)$$

Die Koeffizienten $b_m, \dots, b_0; a_n, \dots, a_0$ der Systemfunktion hängen von den Werten der elektrischen Bauelemente ab und legen damit die Filtereigenschaften fest.

Idealerweise möchte man die gewünschten Frequenzkomponenten eines Signals vollständig durchlassen und die unerwünschten Anteile vollständig unterdrücken. Das ist aber praktisch nicht möglich. Die Übertragungseigenschaften von Filtern werden daher durch ein Toleranzschema festgelegt, das sich aus Qualitätsforderungen der Anwendung ergibt. Abbildung 8.2 zeigt das Toleranzschema eines Tiefpassfilters.

Die Dämpfung des Signals im Durchlassbereich darf nicht größer als ein maximaler Wert a_D [dB] sein, die Dämpfung im Sperrbereich muss größer als ein minimaler Wert a_{SP} [dB] sein. Der Dämpfungsverlauf innerhalb der Toleranzfelder kann einen beliebigen Verlauf annehmen.

Aufgabe des Filterentwurfs ist es, eine Übertragungsfunktion $H = \frac{\vec{U}_a}{\vec{U}_e}$ zu finden, die in das gewünschte Toleranzfeld »passt«. Da bei linearen elektrischen Netzwerken nur gebrochen rationale Funktionen als Übertragungsfunktionen auftreten können, kann die gewünschte und schaltungstechnisch realisierbare Netzwerkfunktion nur eine gebrochen rationale Funktion sein. Aus den Koeffizienten der Netzwerkfunktion kann man wiederum die Bauelementewerte der elektrischen Schaltung ermitteln.

8.1.1 Approximation eines idealen Tiefpassfilters

Ein ideales Tiefpassfilter überträgt alle Frequenzkomponenten im Durchlassbereich und unterdrückt alle Frequenzkomponenten im Sperrbereich. Die Systemfunktion und die Dämpfung eines idealen Tiefpasses sind in Abbildung 8.3 dargestellt.

Diese ideale Filterkennlinie ist nicht realisierbar, wir müssen daher eine gebrochen rationale Funktion finden, die den Betrag der Übertragungsfunktion so gut wie gewünscht annähert. Die Betragsbildung ist eine nichtlineare Operation, mit der wir nicht gut rechnen können. Durch Verwendung der Beziehung

$$|H(j\omega)|^2 = H(j\omega)H(-j\omega) \quad (8.4)$$

erhalten wir aber eine Darstellung mit der wir gut umgehen können. Es ist vorteilhaft, eine neue Funktion $F(j\omega) = 1/H(j\omega)$ einzuführen und $F(j\omega)$ zu approximieren.

Die **Filtereigenschaften** werden durch die **Systemfunktion** eindeutig bestimmt.

Beim Filterentwurf ist meistens ein **Toleranzschema** vorgegeben, das außer dem **Durchlass- und Sperrbereich** auch die maximale Durchlassdämpfung und die minimale Sperrdämpfung angibt.

Die **Kennlinie** eines idealen Filters kann nicht praktisch realisiert werden. Sie kann nur **durch eine gebrochen rationale Funktion angenähert** werden.

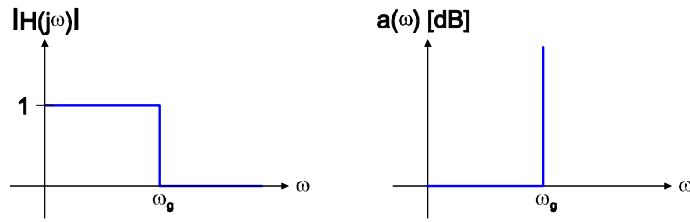


Abbildung 8.3: Frequenzkennlinie eines idealen TP-Filters

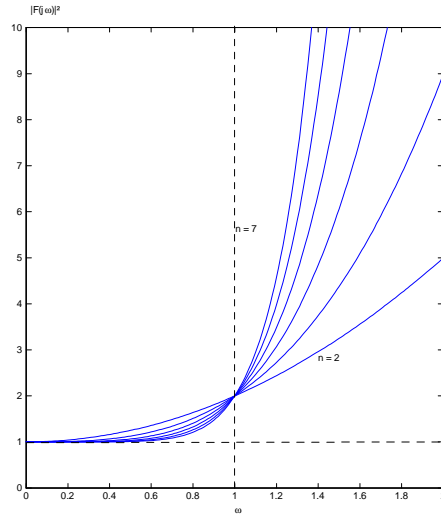


Abbildung 8.4: Approximation durch Potenzansatz

Potenzapproximation

Bei der **Potenzapproximation** wird die ideale Filterkennlinie durch **Parabeln** angenähert.

Wir verwenden für die weitere Rechnung die bekannte Schreibweise in s und wählen den folgenden Approximationsansatz

$$F(s)F(-s) = 1 + s^{2n} \quad \text{für } n = 2, 4, 6, \dots \quad (8.5)$$

$$= 1 - s^{2n} \quad \text{für } n = 1, 3, 5, \dots \quad (8.6)$$

$$= 1 + (-1)^n s^{2n} \quad (8.7)$$

Abbildung 8.4 zeigt den Verlauf der Approximationsfunktion für unterschiedliche n .

Unser Approximationsansatz mit Parabeln n -ter Ordnung zeigt, dass das ideale Verhalten (in Abbildung strichliert gezeichnet) umso besser angenähert wird, je höher die Ordnung ist: Der Verlauf der Approximationsfunktion wird im Durchlassbereich bei höherer Ordnung flacher, der Anstieg im Sperrbereich wird bei höherer Ordnung steiler. Alle Kurven gehen bei $\omega = 1$ – der Grenzfrequenz – durch den Punkt $|F(j\omega)|^2 = 2$, die Verstärkung (Dämpfung) in diesem Punkt beträgt daher -3 dB.

Bei der Approximation haben wir das Produkt $F(s)F(-s)$ angenähert, wir brauchen aber nicht das Produkt, sondern die Funktion $F(s)$. Die Zerlegung in die Teile $F(s)$ und $F(-s)$ führen wir mit Hilfe der Nullstellendarstellung durch.

$$F(s)F(-s) = 1 + (-1)^n s^{2n} = 0 \rightarrow (-1)^n s^{2n} = -1 \quad (8.8)$$

d.h., die Nullstellen liegen auf dem Einheitskreis in der komplexen Ebene. Wir erhalten für

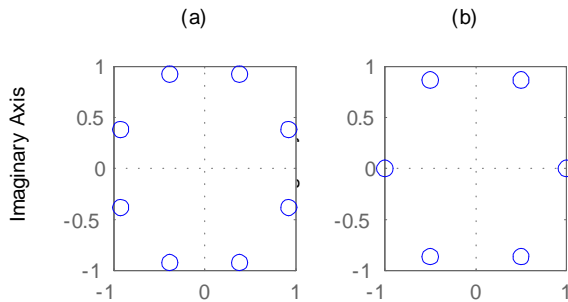


Abbildung 8.5: Nullstellen von (a) $F(s)F(-s) = 1 + s^8$ (b) $F(s)F(-s) = 1 - s^6$

$$\begin{aligned} n = \text{ungerade} & \quad s_{0,k} = e^{jk\pi/n} & \quad k = 1, 2, \dots, 2n \\ n = \text{gerade} & \quad s_{0,k} = e^{j(2k-1)\pi/2n} & \quad k = 1, 2, \dots, 2n \end{aligned} \tag{8.9}$$

Abbildung 8.5 zeigt ein Beispiel für $n = 4$ und $n = 3$.
Das Produkt hat also die Form

$$F(s)F(-s) = (s - s_{0,1})(s - s_{0,2})(s - s_{0,3}) \dots (s - s_{0,2n}) \tag{8.10}$$

Die einzelnen Nullstellen (Klammerausdrücke in 8.10) können (aus mathematischer Sicht) den Funktionen $F(s)$ und $F(-s)$ beliebig zugeordnet werden, wobei aber zwei auf dem Einheitskreis gegenüberliegende Nullstellen nicht der selben Funktion zugeordnet werden können². Die Nullstellen der Hilfsfunktion $F(s)$ sind aber die Polstellen von $H(s)$. Wir wissen, dass bei stabilen Netzwerken alle Pole in der linken offenen komplexen Halbebene liegen müssen. Daher spalten wir das Produkt $F(s)F(-s)$ so auf, dass wir alle Nullstellen in der linken s -Ebene $F(s)$ zuordnen und alle Nullstellen in der rechten s -Ebene $F(-s)$ zuordnen.

Berechnet man daraus die Übertragungsfunktionen $H(s)$ für die zweite bis vierte Ordnung, so erhält man

$$H(s) = \frac{1}{s^2 + 1.4142s + 1} \tag{8.11}$$

$$H(s) = \frac{1}{s^3 + 2s^2 + 2s + 1} \tag{8.12}$$

$$H(s) = \frac{1}{s^4 + 2.613s^3 + 3.4142s + 2.6131s + 1} \tag{8.13}$$

Filter die auf diesem Approximationsansatz beruhen nennt man Potenzfilter (Butterworth Filter).

Beispiel 58 Die Lage der Polstellen der Übertragungsfunktion $H(s)$ der Potenzfilter berechnet die Matlab-Funktion `[z,p,k]=buttap(n)`. Die Darstellung `[z,p,k]` (Zeros, Poles, Gain) kann mit Hilfe der Funktion `[num,den]=zp2tf(z,p,k)` in die Übertragungsfunktion (transfer function) umgerechnet werden, wobei `[num]` die Koeffizienten des Zählerpolynoms und `[den]` die Koeffizienten des Nennerpolynoms enthält. (Die Umrechnung funktioniert auch in die andere Richtung. `[Z,P,K] = TF2ZP(NUM,DEN)` liefert Zeros, Poles, Gain aus Zähler- und Nennerpolynom der Übertragungsfunktion.

Tschebyscheff-Approximation im Durchlassbereich

Zum Unterschied von der Potenzapproximation, die im Durchlassbereich einen flachen Verlauf aufweist, zeigt die Approximation durch Tschebyscheff'sche Polynome im Durchlassbereich ein welliges Verhalten auf, wie Abbildung 8.6 zeigt.

²Eine Nullstelle von $F(s)$ impliziert eine auf dem Einheitskreis gegenüberliegende Nullstelle von $F(-s)$.

Der **Tschebyscheff-Ansatz** weist einen schnelleren Übergang vom Durchlass- in den Sperrbereich als der Potenzansatz auf, der aber mit einer (parametrisierbaren) **Welligkeit im Durchlassbereich** zu bezahlen ist.

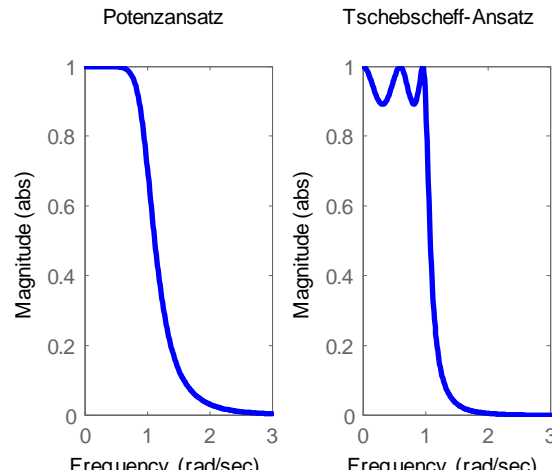


Abbildung 8.6: Frequenzgang Potenz- und Tschebyscheff-Filter 5. Ordnung

Die Welligkeit ist ein Entwurfsparameter und wird so gewählt, dass das gewünschte Filterverhalten im Durchlassbereich erreicht wird. Bei der Welligkeit Null geht das Tschebyscheff-Filter in das Potenzfilter über.

Filter sollen Durchlass- und Sperrbereich möglichst gut trennen, diese Aufgabe wird umso besser erfüllt, je steiler der Übergang zwischen Durchlass- und Sperrbereich ist. Bei gleicher Ordnung ist die Flankensteilheit von Tschebyscheff-Filtern größer als die von Potenzfiltern, wie man in Abbildung 8.6 erkennen kann. Die Flankensteilheit ist auch größer, wenn man eine höhere Welligkeit im Durchlassbereich tolerieren kann.

Tschebyscheff-Filter sind also frequenzselektiver als Potenzfilter.

Beispiel 59 Die Lage der Polstellen der Übertragungsfunktion $H(s)$ der Tschebyscheff-Filter³ berechnet die Matlab-Funktion $[z, p, k] = \text{cheb1ap}(n, r)$. r ist die Welligkeit in dB im Durchlassbereich.

Cauer-Filter

Die **Cauer-Filter** sind die **frequenzselektivsten**, weisen aber eine (parametrisierbare) **Welligkeit im Durchlass- und im Sperrbereich** auf.

Wird eine Tschebyscheff-Approximation im Durchlass- und im Sperrbereich gewählt, dann erhalten wir Cauer- oder elliptische Filter. Bei diesem Approximationsansatz kann die Welligkeit im Durchlass- und im Sperrbereich voneinander unabhängig gewählt werden. Die Cauer-Filter sind die Filter mit der größten Flankensteilheit zwischen Durchlass- und Sperrbereich. Abbildung 8.7 zeigt das Frequenzverhalten eines Cauerfilters mit 1 dB Welligkeit im Durchlassbereich und mindestens 60 dB Dämpfung im Sperrbereich in linearer und doppelt logarithmischer Darstellung. Cauer-Filter haben Übertragungsnullstellen im Sperrbereich, die man in der linearen Darstellung nicht gut erkennen kann. In der logarithmischen Darstellung sind die Nullstellen aber gut erkennbar ($\log 0 = -\infty$).

Beispiel 60 Die Lage der Pol- und Nullstellen der Übertragungsfunktion $H(s)$ der Cauer-Filter⁴ berechnet die Matlab-Funktion $[z, p, k] = \text{ellipap}(n, rp, rs)$. rp, rs sind die Welligkeiten in dB im Durchlass- und Sperrbereich.

Die Bedeutung der Phase

Phasenverschiebungen von Frequenzkomponenten eines Signals verursachen im Allgemeinen eine **Verzerrung seiner Kurvenform**.

Abbildung 8.8 zeigt ein Signal bestehend aus drei sinusförmigen Komponenten. Die Komponenten in (a) und (b) haben gleiche Frequenz, aber unterschiedliche Phase. Abbildung (c) zeigt die Summe der in Abbildung (a) dargestellten Komponenten, Abbildung (d)

³In der englischsprachigen Literatur wird Tschebyscheff mit Chebyshev transkribiert.

⁴Für das Approximationsproblem hat Wilhelm Cauer eine geschlossene Lösung angegeben, bei der vollständige elliptische Integrale erster Gattung auftreten. Daher heißen diese Filter in der englischen Literatur elliptische Filter.

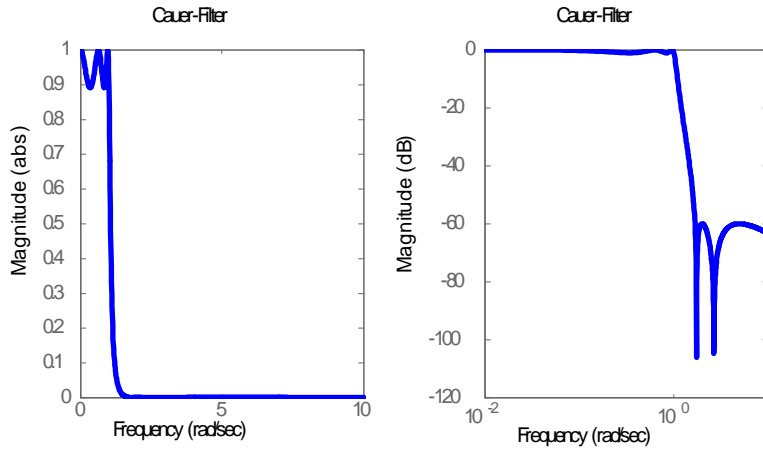


Abbildung 8.7: Cauerfilter 5. Ordnung (linear und doppelt logarithmisch)

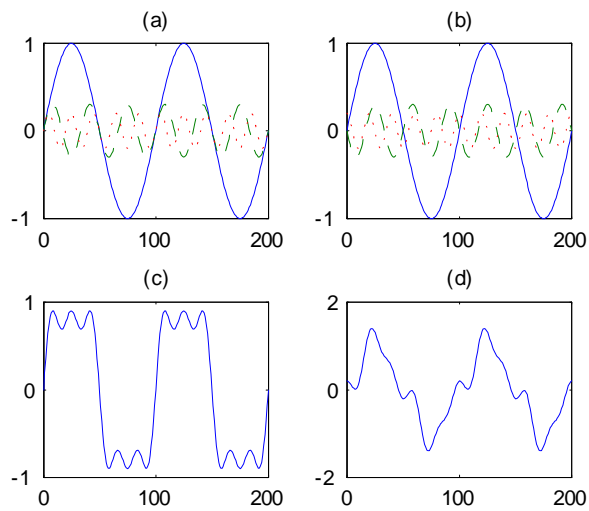


Abbildung 8.8: Sinuskomponenten mit unterschiedlicher Phase

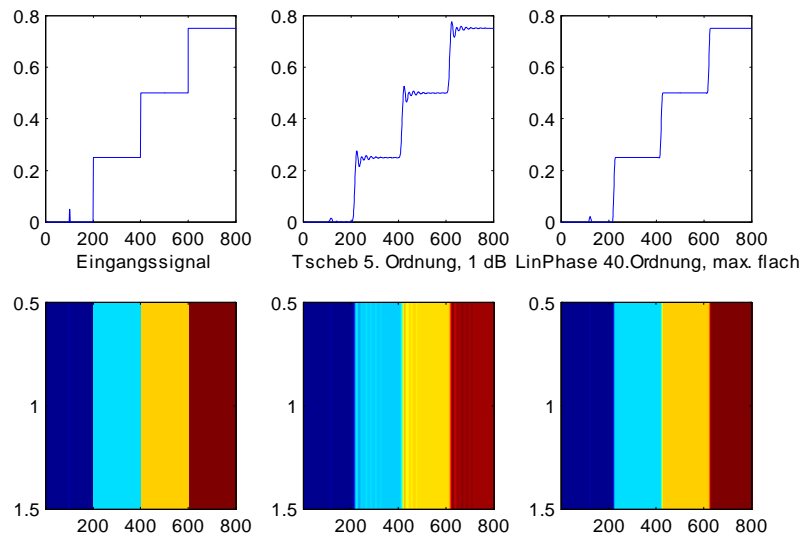


Abbildung 8.9: SZeilensignal in Falschfarbendarstellung

zeigt die Summe der in (b) gezeigten Komponenten. Wie man ohne Zweifel feststellen kann, unterscheiden sich die Signalformen (c) und (d) deutlich voneinander. Dennoch kann das menschliche Ohr keinen Unterschied zwischen (c) und (d) feststellen! Das Ohr kann keine Phasenunterschiede erkennen, die in den Signalen (c) und (d) »kodierte« Klanginformation steckt ausschließlich in der Frequenz. Beim Filtern von Signalen die nur Frequenzinformation enthalten kann das frequenzselektivste Filter – das Cauer-Filter – eingesetzt werden, der Phasengang des Filters ist ohne Bedeutung.

Bei manchen Signalen (z.B. Fernsehsignale) ist ihre **Kurvenform** wichtig, für andere Signale (z.B. Audiosignale) spielt lediglich ihre **Frequenzzusammensetzung** eine Rolle.

Eine völlig unterschiedliche Situation liegt vor, wenn die Information im Zeitbereich, also in der Kurvenform des Signals kodiert ist. Ein gutes Beispiel dafür ist das Zeilensignal eines Fernsehbildes.

Abbildung 8.9 zeigt ein digitales Signal mit einem unerwünschten Störimpuls (a). (Dem digitalen Signal sind Farbwerte zugeordnet, die unterhalb der Signale dargestellt sind.) Durch Filterung soll der Störimpuls unterdrückt werden, die Streifen sollen aber möglichst scharf abgebildet werden.

Bei dieser Aufgabenstellung sind wir im Dilemma: Einerseits möchte man den hochfrequenten Störimpuls unterdrücken, man muss also ein Filter entwerfen, das die (hochfrequenten) Signalkomponenten des Störimpulses unterdrückt. Andererseits möchte man aber die scharf begrenzten Streifen, die ebenfalls hochfrequente Signalkomponenten enthalten, erhalten und müsste daher diese hochfrequenten Komponenten durchlassen. Diese Aufgabenstellung ist also nur mit einem Kompromiss zwischen Störimpulsunterdrückung und Unschärfe der Streifen lösbar.

Ist der **Phasengang** eines Filters **linear**, bleibt die **Kurvenform** des durchgelassenen Signals **erhalten**.

In unserem Beispiel setzen wir für diese Filteraufgabe ein Tschebyscheff-Filter 5. Ordnung ein (b). Der hochfrequente Störimpuls wird abgeschwächt und ist im Bild fast nicht mehr zu sehen, es tritt allerdings ein starkes Überschwingen des Filters auf, das zu sichtbaren Streifen führt. Besser geeignet ist das Linear-Phasenfilter, bei dem zwar die Farbübergänge etwas unscharf werden, aber kein Überschwingen und damit keine Streifen auftreten. Das Linear-Phasenfilter in unserem Beispiel ist 20. Ordnung und viel aufwändiger als das Tschebyscheff-Filter 5. Ordnung. Die lineare Phase ist dafür verantwortlich, dass die Kurvenform der Signale (die Treppenform in unserem Beispiel) erhalten bleibt.

Bemerkung 61 Immer dann, wenn die Kodierung der Information in der Kurvenform des Signals steckt (wie bei Biosignalen oder in Fernsehbildern), ist die Eigenschaft der linearen Phase wichtig. Wenn die Information in den Frequenzkomponenten steckt (wie

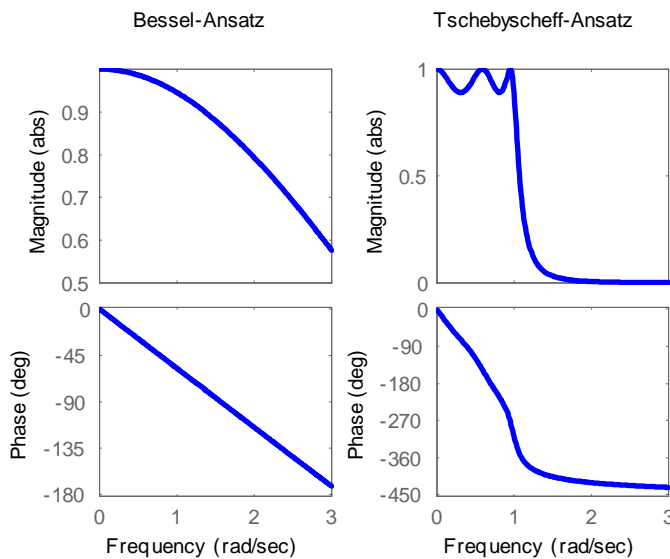


Abbildung 8.10: Bessel- und Tschebyscheff-Tiefpass 5. Ordnung

bei Audiosignalen) ist der Phasenverlauf des Filters von untergeordneter Bedeutung, da das menschliche Ohr nur wenig von der Phase abhängig ist. Hohe Frequenzselektivität (bei gleicher Ordnung des Filters) erzielt man nur mit Filtern, die keine lineare Phase (linearen Phasengang) haben.

Bessel-Filter

Wenn bei der Filterung der Erhalt der Kurvenform eine wichtige Rolle spielt, muss für die Übertragungsfunktion eine linear ansteigende Phase approximiert werden. Mit Hilfe von Bessel-Polynomen kann eine lineare Phase approximiert werden⁵.

Die Bessel-Polynome sind wie folgt definiert

$$B_1 = x + 1 \tag{8.14}$$

$$B_2 = x^2 + 3x + 3 \tag{8.15}$$

$$B_3 = x^3 + 6x^2 + 15x + 15 \tag{8.16}$$

$$B_4 = x^4 + 10x^3 + 45x^2 + 105x + 105 \tag{8.17}$$

$$B_n = (2n - 1)B_{n-1} + x^2B_{n-2} \quad B_0 = 1, B_{-1} = 1/x \tag{8.18}$$

Die Übertragungsfunktion von Bessel-Filtern ist

$$H(s) = \frac{b_0}{B_n(s)} = \frac{b_0}{s^n + b_{n-1}s^{n-1} + \dots + b_1s + b_0} \tag{8.19}$$

Beispiel 62 Die Lage der Polstellen der Übertragungsfunktion $H(s)$ der Bessel-Filter berechnet die Matlab-Funktion `[z,p,k]=besselap(n)`.

Abbildung 8.10 zeigt die Bodediagramme eines Tschebyscheff-Filters 5. Ordnung mit 1 dB Welligkeit und eines Bessel-Filters 5. Ordnung.

Wie man in 8.10 erkennen kann, ist das Phasenverhalten des Bessel-Tiefpasses im dargestellten Bereich linear, allerdings ist die Flankensteilheit des Besselfilters wesentlich geringer als die des Tschebyscheff-Filters. Um eine ähnlich hohe Flankensteilheit wie beim Tschebyscheff-Filter zu erreichen, müsste ein Besselfilter wesentlich höherer Ordnung – mit höherem Hardwareaufwand – eingesetzt werden.

⁵Im Abschnitt über FIR-Filter werden wir digitale Filter kennen lernen, deren Phasenverhalten exakt linear ist.

Bessel-Filter haben einen annähernd **linearen Phasengang**, ihre **Frequenzselektivität** ist aber eher **gering**.

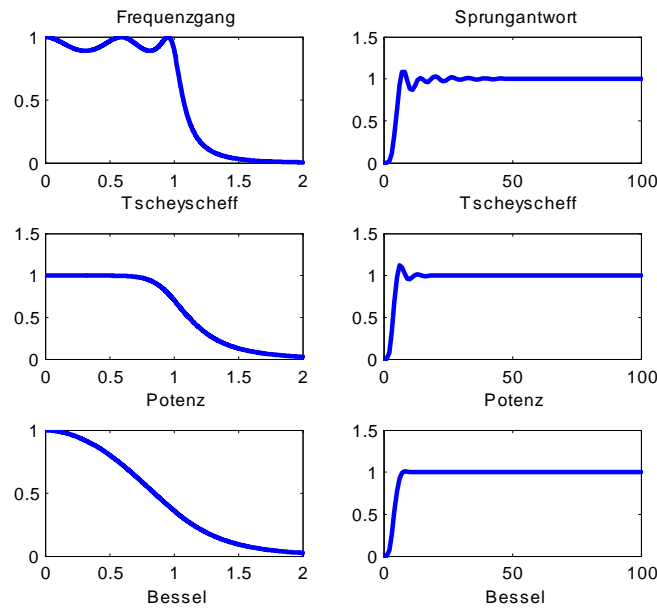


Abbildung 8.11: Frequenzgang und Sprungantworten Filter 5. Ordnung

Vergleich der Filterapproximationen

Filter werden mit zunehmender Ordnung frequenzselektiver.

Filter nach Frequenzselektivität:

1. Cauer
2. Tschebyscheff
3. Potenz
4. Bessel

Das wichtigste Selektionsmerkmal eines Filters ist die Steilheit des Übergangs vom Durchlass- in den Sperrbereich, die Auskunft über die Selektivität des Filters gibt. Je steilflankiger das Filter ist, desto besser wird das gewünschte ideale Filter angenähert.

Die Flankensteilheit eines Filters hängt von der Ordnung (von der Zahl der verwendeten Bauelemente) und von der Approximationsfunktion ab. Je höher die Ordnung, desto größer die Flankensteilheit. Kein realisierbares Filter ist in der Lage, das Ausgangssignal im Sperrbereich vollständig zu unterdrücken.

Die Frequenzselektivität des *Cauer-Filters* ist am größten, die Steilheit des Übergangs vom Durchlass- in den Sperrbereich ist auch von der Welligkeit im Durchlass- und Sperrbereich abhängig. Je größer die Welligkeit im Durchlassbereich und je kleiner die Sperrdämpfung, desto größer die Steilheit. Welche Welligkeit im Durchlassbereich akzeptabel ist und wie hoch die minimale Sperrdämpfung sein muss, hängt von der Anwendung ab. Das *Tschebyscheff-Filter* hat ein ähnliches Verhalten wie das Cauer-Filter, die Dämpfung im Sperrbereich steigt aber monoton an.

Das *Potenzfilter* ist der Sonderfall des Tschebyscheff-Filters für die Welligkeit Null. Die geringste Selektivität (bei gleicher Filterordnung) weist das *Bessel-Filter* auf.

Beim Phasenverhalten liegen die Verhältnisse genau umgekehrt: höchste Frequenzselektivität bedeutet nichtlinearen Phasengang, was zur Folge hat, dass die Kurvenform der Signale stark verzerrt wird. Zur Erhaltung der Kurvenform sind Filter mit angenähert linearem Phasengang – wie beim Besselfilter – erforderlich. Ein gutes Testsignal zur Bewertung des Phasenverhaltens liefert der Eingangssprung. Filter mit linearem Phasengang zeigen praktisch kein Überschwingen (die Kurvenform wird erhalten).

Abbildung 8.11 zeigt Frequenzgang und Sprungantwort von Tschebyscheff-, Potenz- und Bessel-Filter.

Während das Besselfilter kein Überschwingen des Ausgangssignals zeigt, schwingt das Tschebyscheff-Filter stark und langanhaltend über. Die Reaktion eines Filters auf den Sprung – die Verzögerungszeit – hängt von der Ordnung des Filters ab, je höher die Ordnung, desto größer die Verzögerungszeit.

Welcher Filtertyp ausgewählt wird, hängt von der Aufgabenstellung ab. Wir wissen, dass Signale dafür verwendet werden Information zu »kodieren«. Das Kodieren von

Signale, die in ihrer **Kurvenform** Information tragen, müssen mit **linearphasigen Filtern** gefiltert werden.

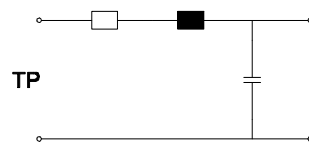
Information kann entweder im Zeitbereich oder im Frequenzbereich erfolgen. Wenn die Information in der Frequenz steckt, können selektive Filter eingesetzt werden, da die Kurvenform keine Rolle spielt. Steckt die Information aber in der Kurvenform, dann müssen Filter mit linearer Phase eingesetzt werden, um die Verzerrungen klein zu halten.

8.1.2 Realisierung von analogen Filtern

LC-Filter

Die Approximation des gewünschten Filterverhaltens liefert die Systemfunktion des Filters. Die Netzwerkanalyse liefert eine gebrochene rationale Funktion. Durch Koeffizientenvergleich können die Bauelementewerte bestimmt werden.

Beispiel 63 Die Potenzapproximation für ein Filter 2. Ordnung liefert die Systemfunktion $H(s) = \frac{1}{s^2 + 1.4142s + 1}$. Die Analyse des folgenden RLC-Netzwerks ergibt die System-



funktion $H(s) = \frac{1}{LCs^2 + RCs + 1}$.

Durch Koeffizientenvergleich erhalten wir daraus

$LC = 1$; $RC = 1.4142$. Daraus können wir die Bauelementewerte⁶ R, L und C ermitteln.

Für den praktischen Entwurf von Filtern gibt es Filterkataloge und CAE-Programme, mit denen die Bauelementewerte ermittelt werden können. Widerstände und Kondensatoren sind in genormten Werten (E-Reihen) erhältlich. Je feiner die Abstufung zwischen den Bauelementen und je präziser die Werte der Bauelemente sind, desto teurer sind die Komponenten. Aber auch bei der feinsten erhältlichen Abstufung ist es nur selten möglich Katalog-Bauelemente zu finden, die genau den für eine Filterapproximation nötigen Wert haben. In diesen Fällen muss durch Zusammenschaltung mehrerer Katalog-Bauelemente der gewünschte Wert »eingestellt« werden. Alternativ können auch abstimmbare Bauelemente (Potentiometer, Drehkondensatoren) verwendet werden, deren Kosten aber deutlich höher liegen als die Kosten der Festwert-Bauelemente.

Induktivitäten sind in der Regel nicht mit den erforderlichen Werten erhältlich und müssen individuell hergestellt werden. LC-Filter werden heute nur mehr bei höheren Frequenzen eingesetzt.

Aktive Filter

LC-Filter sind bei niedrigen Frequenzen wegen der großen Bauelementewerte schlecht herstellbar. Für die Herstellung von Induktivitäten in niedrigen Frequenzbereichen sind sehr viele Windungen erforderlich, was den ohmschen Widerstand der Induktivität erhöht und damit ein Abweichen vom idealen Bauelementverhalten bewirkt. Auf Grund der technologischen Entwicklung der Halbleitertechnik sind heute aktive Bauelemente (Operationsverstärker, OPV) oft kostengünstiger als passive Bauelemente wie z.B. Induktivitäten. Wenn es der Frequenzgang der OPVs zulässt, wird man daher bevorzugt aktive Schaltungen einsetzen.

Abbildung 8.12 zeigt ein aktives Tiefpassfilter 2. Ordnung.

Die Netzwerkanalyse liefert für die Systemfunktion

$$H(s) = \frac{U_a}{U_e} = - \frac{R_2/R_1}{C_1 C_2 R_2 R_3 s^2 + C_1 \left(R_2 + R_3 + \frac{R_2 R_3}{R_1} \right) s + 1} \quad (8.20)$$

⁶Wir haben zwei Gleichungen für drei Unbekannte, können daher den Wert eines Bauelements wählen. Die Wahl erfolgt nach Gesichtspunkten der praktischen Realisierung.

Bei der Implementierung von analogen Filtern sind Bauelemente mit den errechneten Werten meistens nicht erhältlich und müssen durch Ersatzschaltungen nachgebildet werden.

Bei **aktiven Filtern** werden **Operationsverstärker** eingesetzt.

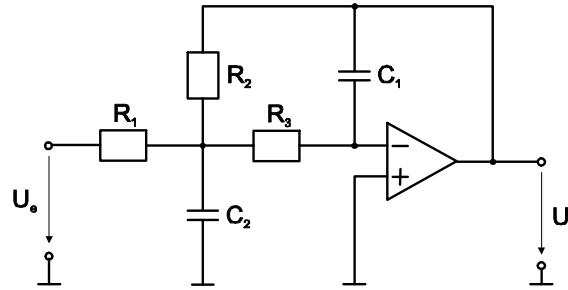


Abbildung 8.12: Aktiver Tiefpass

Die Bauelementewerte können durch Koeffizientenvergleich mit der gewünschten Approximationsfunktion bestimmt werden. Filter höherer Ordnung werden durch Kaskadierung von Filtern 2. Ordnung realisiert. Dazu müssen die Approximationsfunktionen in Terme zweiter Ordnung aufgespalten werden.

Für die praktische Berechnung von aktiven Filtern gibt es Filterkataloge, außerdem stellen die Hersteller von OPVs Filterentwurfsprogramme zu Verfügung.

8.1.3 Normierung

In der **normierten Darstellung** von Netzwerken treten meistens Werte in der **Größenordnung von 1** auf.

Wegen der einfacheren Darstellbarkeit werden Netzwerke häufig normiert, da man es dann mit Zahlen in der Größenordnung von Eins zu tun hat.

Bei der *Widerstandsnormierung* werden alle Größen des Netzwerks durch einen Bezugswiderstand R_0 dividiert, man spricht von der Widerstandsnormierung auf R_0 . Bei der Frequenznormierung wird die komplexe Frequenz s durch eine reelle Bezugsfrequenz ω_0 dividiert. Den Quotienten $s_{\text{nor}\omega} = s/\omega_0$ bezeichnet man als normierte Frequenz. In der Regel werden Widerstands- und Frequenznormierung gleichzeitig angewendet.

unnormiert	R -normiert	ω -normiert	R & ω -normiert
R_{un}	$R_{\text{nor}R} = \frac{R_{\text{un}}}{R_0}$	$R_{\text{nor}\omega} = R_{\text{un}}$	$R_{\text{nor}} = \frac{R_{\text{un}}}{R_0}$
C_{un}	$C_{\text{nor}R} = C_{\text{un}} \cdot R_0$	$C_{\text{nor}\omega} = C_{\text{un}} \cdot \omega_0$	$C_{\text{nor}} = C_{\text{un}} \cdot R_0 \cdot \omega_0$
L_{un}	$L_{\text{nor}R} = \frac{L_{\text{un}}}{R_0}$	$L_{\text{nor}\omega} = L_{\text{un}} \cdot \omega_0$	$L_{\text{nor}} = L_{\text{un}} \cdot \frac{R_0}{\omega_0}$

8.1.4 Frequenztransformation

Durch eine geeignete **Transformation** der der **Systemfunktion** können verschiedene Filtertypen ineinander umgewandelt werden.

Wenn die Systemfunktion eines Tiefpasses bekannt ist, dann kann man mit Hilfe der Frequenztransformation einen Hochpass, einen Bandpass und eine Bandsperre erzeugen. Die Schaltungen findet man, indem man Induktivitäten und Kapazitäten des Tiefpassfilters durch entsprechende Zweipole ersetzt.

Aus dem Tiefpass entsteht ein Hochpass, indem man $s \rightarrow 1/s$ setzt:

$$H_{TP}(s) = \frac{1}{s^3 + 2s^2 + 2s + 1} \rightarrow H_{HP}(s) = \frac{s^3}{1 + 2s + 2s^2 + s^3}$$

Entsprechend wird aus dem Tiefpass ein Bandpass, indem man

$$s \rightarrow \frac{1}{\omega_0 - \omega_u} \left(s + \frac{1}{s} \right) \quad (8.21)$$

setzt und eine Bandsperre, indem man

$$s \rightarrow \frac{\omega_0 - \omega_u}{s + \frac{1}{s}} \quad (8.22)$$

setzt. Die Transformationsbeziehungen auf der Bauelementebene sind in Abbildung 8.13 gezeigt. $B = \omega_0 - \omega_u$

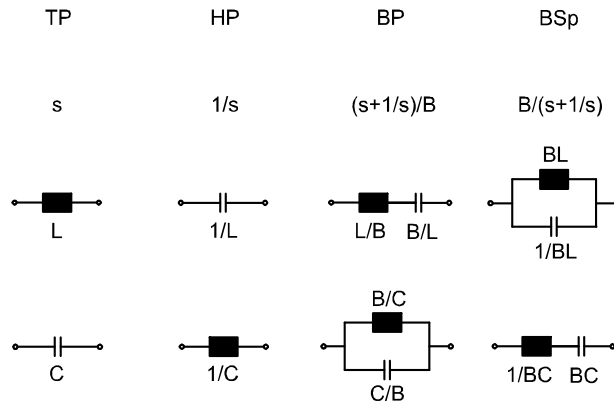


Abbildung 8.13: Frequenztransformation

8.2 Nichtlineare analoge Signalverarbeitung

Neben der einfachen linearen Signalverarbeitung können durch aktive Schaltungen auch Funktionsnetzwerke für Logarithmus, Exponentialfunktion, Sinus- und Kosinusfunktion realisiert werden. Durch geeignete Schaltungen kann man auch Analog- Multiplizieren und die Umwandlung von kartesischen in Polarkoordinaten und umgekehrt durchführen. Diese Schaltungen haben heute nur mehr in Sonderfällen Bedeutung.

8.2.1 Signalgleichrichtung

Die Betragsbildung von Signalen lässt sich auf analoge Weise sehr einfach durch Signalgleichrichtung erreichen. Das dafür verwendete Bauelement ist die (Halbleiter)Diode. Der Zusammenhang zwischen Strom und Spannung an einer Diode ist gegeben durch die Beziehung

$$I_D = I_S \left(e^{\frac{U_D}{nU_T}} - 1 \right) \tag{8.23}$$

$n \approx 1$ bis 2 ist ein Technologieparameter. $U_T = kT/q \approx 26$ mV, ist die sogenannte Temperaturspannung, $I_S \approx 10^{-12}$ bis 10^{-6} nennt man den Sättigungssperrstrom. Wegen des exponentiellen Verhaltens des Diodenstroms und wegen des praktisch vernachlässigbaren Stromes bei negativer Diodenspannung kann man für die Diode ein einfaches Ersatzschaltbild verwenden: In Durchlassrichtung $U_D > 0$ leitet die Diode und hat einen sehr kleinen Widerstand, in Sperrrichtung $U_D < 0$ fließt praktisch kein Strom und die Diode sperrt. Wenn das einfache Schaltermodell der Diode nicht ausreichend ist, kann die Diode durch die Ersatzschaltung nach Abbildung 8.14 dargestellt werden.

Im Durchlassbereich fällt an der Diode die Spannung U_F ab. Für Siliziumdioden ist $U_F \approx 0.7$ V, bei Germanium- und Schottky Dioden ist $U_F \approx 0.3$ V. Bei Betrieb in Durchlassrichtung ist der Schalter geschlossen, in Sperrrichtung ist der Schalter offen. Für genauere Modellierung kann noch der Bahnwiderstand R_B berücksichtigt werden. Den Wert von R_B kann man aus Datenblättern entnehmen oder durch Messungen bestimmen.

Zur Gleichrichtung (Betragsbildung) verwendet man den in Abbildung 8.15 dargestellten Brückengleichrichter.

8.2.2 Komparator

Eine häufig vorkommende Aufgabe der Schaltungstechnik ist der Vergleich von zwei Spannungen. Dafür eignet sich der OPV ohne Gegenkopplung sehr gut. Schaltung und Verhalten sind in Abbildung 8.16 dargestellt.

Solange $U_1 > U_2$ ist die Differenzspannung positiv und auf Grund der hohen Verstärkung des OPV stellt sich der Ausgang auf die maximale Ausgangsspannung (die positive

Eine Eigenschaft der **Diode** ist, dass sie Strom nur in **einer Richtung** »durchlässt«.

Für den **Vergleich zweier Spannungen** eignet sich ein **OPV ohne Gegenkopplung**.

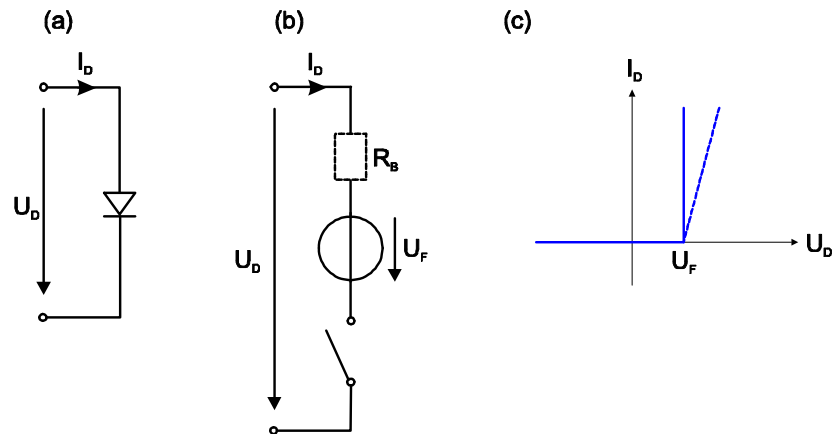


Abbildung 8.14: Diode und Ersatzschaltung (Strichlierte Linien stehen für ein genaueres Modell)

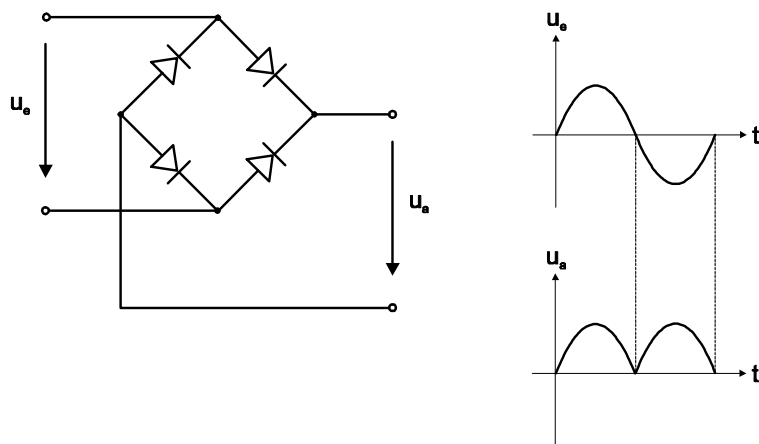


Abbildung 8.15: Gleichrichtung

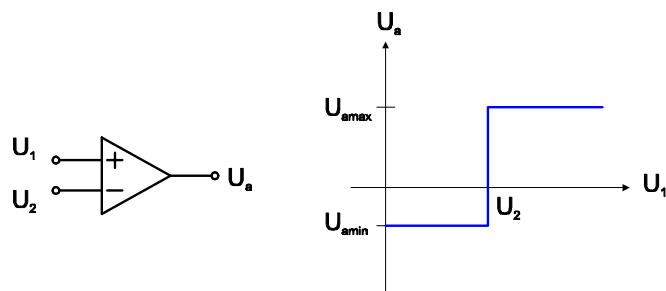


Abbildung 8.16: Komparator

Versorgungsspannung) ein. Bei $U_1 < U_2$ wird die Differenzspannung negativ und es stellt sich die minimale Ausgangsspannung (die negative Versorgungsspannung) ein. Mit Hilfe dieser Schaltung kann man auch eine variable Spannung gegen eine feste Referenzspannung vergleichen.

8.3 Zusammenfassung

Mit Hilfe analoger Schaltungen lassen sich Aufgabenstellungen der Signalverarbeitung im Genauigkeitsbereich von ca. 0.1 % gut und oft kostengünstig lösen.

Beim Aufbau von elektrischen Schaltungen werden analoge Bauelemente verwendet, die einer Fertigungsstreuung unterliegen, analoge Schaltungen müssen daher oft abgeglichen werden. Die Bauelemente zeigen ein Alterungsverhalten, wodurch sich Schaltungseigenschaften im Laufe der Zeit ändern können. Alle diese Probleme können bei der digitalen Signalverarbeitung vermieden werden. Wie aber in den folgenden Abschnitten noch gezeigt wird, brauchen digitale Realisierungen beim Übergang zwischen analoger und digitaler sowie beim Übergang von digitaler in analoge Signaldarstellung immer analoge Filter. Analoge Signalverarbeitung wird daher schon aus diesem Grund immer ihren Platz haben.

Digitale Signale werden durch Abtastung gewonnen. Mit heutiger Technologie kann der Audiobereich und Teile des Videobereichs digital verarbeitet werden, bei höheren Frequenzen kann aber nach wie vor nur analoge Schaltungstechnik eingesetzt werden. Wo immer wirtschaftlich und technisch möglich, wird man wegen der überlegenen Eigenschaften die digitale der analogen Signalverarbeitung vorziehen.

Teil IV

Signalabtastung

Kapitel 9

Signalabtastung

Inhalt

9.1	Vorteile digitaler Signale	134
9.2	Abtastung im Zeitbereich	134
9.3	Mathematische Darstellung des Abtastvorgangs	135
9.4	Aliasing und Folding	138
9.5	Signalrekonstruktion	139
9.6	Filterung	142
9.6.1	Interpolationsfilter	142
9.6.2	Antialiasing Filter	142
9.7	Digitalisierung	143
9.8	Die zeitdiskrete Fourier-Transformation (DTFT)	144
9.9	Abtastung im Frequenzbereich	145
9.10	Die diskrete Fourier-Transformation (DFT)	146
9.11	Systeme mit unterschiedlichen Abtastraten (Multirate Filter)	149
9.12	Zusammenfassung Fouriertransformationen	150
9.13	Zusammenfassung	151

Mit Hilfe der analytischen Darstellung gewinnt man wichtige Einblicke in Signale und Systeme. Für numerische Untersuchungen in einem Rechner ist diese Darstellung aber nicht geeignet, dazu müssen wir die Signale »diskretisieren«, d.h. wir müssen den kontinuierlichen Signalen Proben entnehmen und diese Proben digitalisieren. Durch dieses Vorgehen erhalten wir aus einem kontinuierlichen Signal eine Folge von Zahlen und es ist einleuchtend, dass der Vergleich der beiden Darstellungen nicht einfach ist.

In diesem Kapitel befassen wir uns mit diskreten Signalen. Diskrete Signale entstehen durch »Entnahme« von Signalproben aus kontinuierlichen Signalen. Bei der Entnahme von Signalproben - Abtastung genannt - müssen wir entscheiden, (1) in welchen zeitlichen Abständen die Signalproben entnommen werden und (2) mit welcher Genauigkeit die Signalproben entnommen werden, d. h. wir müssen eine Quantisierung in Zeit und Amplitude vornehmen.

In praktischen Anwendungen wird das kontinuierliche Signal zu den gewünschten Zeitpunkten durch eine Abtast- und Halteschaltung (sample & hold) abgetastet und konstant gehalten, um der Analog-/Digitalwandler-Schaltung genügend Zeit für die Umwandlung zu geben. Man kann leicht sehen, dass sich das vom A/D-Wandler erzeugte digitale Signal, wegen der Quantisierung im Zeit- (Abtastung) und im Amplitudenbereich (A/D-Wandlung), deutlich vom analogen Signal unterscheidet. Dennoch können wir bei geeigneten Quantisierungen eine höhere Qualität der digitalen Signalverarbeitung erreichen, als das mit analogen Methoden möglich ist.

Die Quantisierung führen wir zur Gewinnung eines digitalen Signals durch, um dann Methoden der digitalen Signalverarbeitung einsetzen zu können. Am Ende der (digitalen)

Bei der **Abtastung** werden Signalen zu bestimmten Zeitpunkten **Proben** entnommen. Diese werden im nächsten Schritt durch AD-Wandler in Zahlen umgesetzt.

Signalverarbeitungskette ist in der Regel eine Rückumwandlung des digitalen Signals in ein kontinuierliches Signal notwendig, z. B. um ein Sprachsignal über den Lautsprecher eines CD-Spielers wiederzugeben.

In diesem Zusammenhang stellt sich die Frage, ob es möglich ist, ein kontinuierliches Signal in ein diskretes und wieder zurück in ein kontinuierliches zu wandeln, ohne dass hörbare Qualitätsverluste bei den Signalen auftreten. Die tägliche Erfahrung am Beispiel von Audio-CDs, MP3-Musik, Mobiltelefonen lehrt uns, dass die Umwandlung möglich ist und wir wollen in den nächsten Abschnitten untersuchen, welche Bedingungen dabei einzuhalten sind.

9.1 Vorteile digitaler Signale

Bandbegrenzte analoge Signale (alle praktischen Signale sind bandbegrenzt) können in digitale Signale umgewandelt werden. Die Voraussetzung dafür ist die Einhaltung des Abtasttheorems (bei der Zeitquantisierung) und eine den Qualitätsforderungen angemessene Auflösung (bei der Amplitudenquantisierung). Bei der Umwandlung analoger Signale in digitale Signale entstehen erhebliche Datenmengen, weshalb häufig Datenkompression eingesetzt wird. Bei einer Stereo-Audio-CD wird eine Abtastfrequenz von 44.1 kHz verwendet, die Auflösung beträgt 16 bit, womit sich eine Datenrate von $44.1 \text{ [kHz]} \times 16 \text{ [bit]} \times 2 \text{ [Kanäle]} = 1.4 \text{ Mb/s}$ ergibt, bei der Super-Audio-CD (Direct Stream Digital) beträgt die Datenrate 2.8224 Mb/s, bei der Audio-DVD bis zu $192 \times 24 \times 2 = 9.216 \text{ Mb/s}$.

Mit den heute zur Verfügung stehenden elektronischen Schaltkreisen kann digitale Signalverarbeitung in Echtzeit im Audiobereich, aber nur teilweise im Videobereich bewerkstelligt werden.

Die digitale Signalverarbeitung bietet eine Reihe von Vorteilen, deshalb wird sie, wenn technisch möglich, der analogen Verarbeitung vorgezogen.

Digitale Signale lassen sich im Allgemeinen viel einfacher verarbeiten, übertragen, speichern und vervielfältigen als analoge Signale.

1. Digitale Signale sind weniger stör anfällig als analoge, da nur die Signalzustände Null und Eins existieren und sich die Signale eindeutig rekonstruieren lassen, solange die Störungen unterhalb der Erkennungspegel für Null und Eins liegen.
2. Digitale Signale können bei Störungen auf Übertragungstrecken ohne Qualitätsverlust regeneriert werden und über lange Distanzen übertragen werden.
3. Digitale Signale lassen sich einfacher kodieren (Kompression oder Sicherheit) und verarbeiten als analoge Signale.
4. Die Vervielfältigung digitaler Daten ist sehr einfach. Analoge Daten (Filme, Tonbänder) verlieren bei der Vervielfältigung Qualität durch unvermeidbare Störsignale.
5. Das Speichern von digitalen Signalen ist einfacher und kostengünstiger als das von analogen Signalen.
6. Digitale Hardware lässt sich einfacher und kostengünstiger bauen als analoge Komponenten. Die Kosten digitaler Hardware nehmen laufend durch die Weiterentwicklung der Chiptechnologie ab.

9.2 Abtastung im Zeitbereich

Wir gehen von einem kontinuierlichen Signal aus und wählen der Einfachheit halber die Funktion

$$s(t) = A \cos(\omega t + \varphi) \quad (9.1)$$

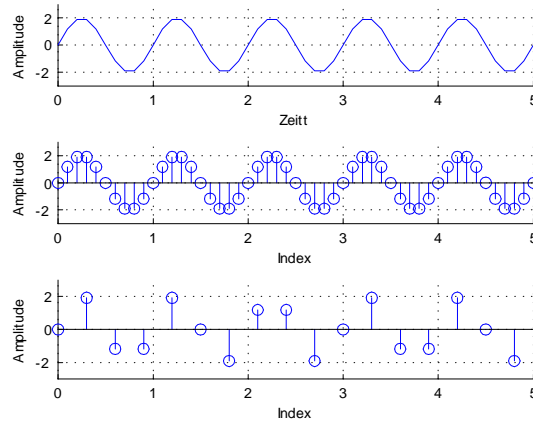


Abbildung 9.1: Signalabtastung

Das (zeit)diskrete¹ Signal erhalten wird durch Entnahme von Proben in äquidistanten Punkten, deren Zeitabstand T_s sei.

$$s[n] = s(nT_s) \quad -\infty < n < \infty \tag{9.2}$$

Als Ergebnis erhalten wir eine Folge von Werten des kontinuierlichen Signals. Die Zahlenwerte der Folge können beliebige Werte annehmen, sind also noch nicht digitalisiert (quantisiert). Aus unserem Ausgangssignal wird dann

$$s[n] = s(nT_s) = A \cos(\omega nT_s + \varphi) = A \cos(\hat{\omega}n + \varphi) \tag{9.3}$$

Die Größe $\hat{\omega}$ bezeichnet man als normalisierte Kreisfrequenz $\hat{\omega} = \omega T_s$. Sie gibt den Winkelunterschied (in rad) zwischen zwei Signalproben an.

Bemerkung 64 Durch die Abtastung geht die Zeitinformation des kontinuierlichen Signals verloren. Schreibt man die Abtastwerte auf eine Liste, kann man nicht erkennen, welche Zeit zwischen den Abtastwerten vergeht. Erst bei Kenntnis des Abtastschrittes T_s kann eine zeitrichtige Rekonstruktion vorgenommen werden. Diese Tatsache verwendet man z.B. bei Synthesizern: Man speichert die Werte der Sinusfunktion von $0 - 90^\circ$ (die restlichen Werte auf 360° lassen sich daraus leicht errechnen) und »schickt« diese Werte mit unterschiedlicher Geschwindigkeit an einen Digital/Analogwandler, um Töne unterschiedlicher Frequenzen zu erzeugen.

Die Dimensionslosigkeit der Folge zeigt sich natürlich auch im mathematischen Ausdruck: Die Kreisfrequenz ω hat die Dimension $[\text{Zeit}^{-1}]$, die Zeit T_s hat die Dimension $[\text{Zeit}]$, ωT_s ist daher dimensionslos.

Abbildung 9.1 zeigt den Zusammenhang zwischen kontinuierlichen und diskreten Signalen. Offensichtlich stellt das Signal mit dichter Abtastfolge das kontinuierliche Signal optisch besser dar als das Signal mit weniger dichter Abtastung. Aber je mehr Abtastwerte entnommen werden, desto mehr Daten müssen verarbeitet und gespeichert werden. Man ist daher daran interessiert möglichst wenige Abtastungen vornehmen zu müssen, will aber das kontinuierliche Signal wieder eindeutig rekonstruieren können.

9.3 Mathematische Darstellung des Abtastvorgangs

Die Abtastung eines Signals $f(t)$ kann dargestellt werden, indem man $f(t)$ mit einer Folge von Einheitsimpulsen $\delta_{0T}(t)$ im Abstand von T_s Sekunden (dem Abtastabstand)

¹Zur Unterscheidung verwenden wir für kontinuierliche Signale runde Klammern, für diskrete Signale eckige Klammern.

Ein **abgetastetes Signal** ist nichts mehr als eine **Folge von Zahlen**. Die **Zeitinformation** geht bei der Abtastung **verloren**.

Mathematisch kann die Abtastung eines Signale durch **Multiplikation** mit einer **Folge von Einheitsimpulsen** modelliert werden.

multipliziert. (Die Abtastung mit δ_0 -Impulsen ist ein theoretisches Konzept, das in elektronischen Schaltungen nur näherungsweise realisiert werden kann.)

$$\begin{aligned}\delta_{0T}(t) &= \sum_k \delta_0(t - kT_s) \\ f_S(t) &= f(t)\delta_{0T}(t) = \sum_k f(t)\delta_0(t - kT_s)\end{aligned}\quad (9.4)$$

Die Impulsfolge $\delta_{0T}(t)$ ist eine periodische Funktion und kann daher in eine Fourierreihe zerlegt werden

$$\delta_{0T}(t) = \sum_{k=-\infty}^{\infty} D_k e^{jk\omega_s t} \quad \omega_s = \frac{2\pi}{T_s} \quad (9.5)$$

$$D_k = \frac{1}{T_s} \int_{-T_s/2}^{T_s/2} \delta_{0T}(t) e^{-jk\omega_s t} dt \quad (9.6)$$

Die Berechnung dieses Integrals liefert $D_k = \frac{1}{T_s}$ und wir erhalten daher

$$\delta_{0T}(t) = \frac{1}{T_s} \sum_{k=-\infty}^{\infty} e^{jk\omega_s t} \quad (9.7)$$

Die Gleichung (9.7) stellt aber nichts anderes dar, als Spektrallinien im Abstand von ω_0 mit den Amplituden $1/T_s$.

In reeller Schreibweise der Gleichung (9.7) erhalten wir

$$\begin{aligned}\delta_{0T}(t) &= \frac{1}{T_s} [1 + 2 \cos \omega_s t + 2 \cos 2\omega_s t + 2 \cos 3\omega_s t + \dots \\ &\quad + 2 \cos k\omega_s t] \quad (k \rightarrow \infty)\end{aligned}\quad (9.8)$$

(Die Sinusanteile heben sich gegenseitig aufgrund der ungeraden Symmetrie des Sinus auf. $\sin(k \cdot j\omega_s t) = -\sin(-k \cdot j\omega_s t)$) Aus Gleichung (9.4) wird

$$f_S(t) = f(t)\delta_{0T}(t) = \frac{1}{T_s} [f(t) + 2f(t) \cos \omega_s t + 2f(t) \cos 2\omega_s t + 2f(t) \cos 3\omega_s t + \dots] \quad (9.9)$$

Das Spektrum $F_s(\omega)$ des abgetasteten Signals $f_S(t)$ ermitteln wir durch die Berechnung der Fouriertransformierten von $f_S(t)$.

$$f(t) \Leftrightarrow F(\omega) \quad (9.10)$$

$$f(t) \cos \omega_s t \Leftrightarrow \frac{1}{2} [F(\omega - \omega_s) + F(\omega + \omega_s)] \quad (9.11)$$

Transformieren wir nun die Terme von Gleichung (9.9) vom Zeitbereich in den Frequenzbereich, so erhalten wir: $F(\omega); F(\omega - 2\omega_s) + F(\omega + 2\omega_s); F(\omega - 3\omega_s) + F(\omega + 3\omega_s); \dots$

$$F_s(\omega) = \frac{1}{T_s} \sum_{n=-\infty}^{\infty} F(\omega - n\omega_s) \quad (9.12)$$

Das Spektrum eines abgetasteten Signals setzt sich aus dem periodisch fortgesetzten Spektrum des Originalsignals zusammen, wobei die Periode der Abtastfrequenz entspricht.

Das Spektrum des abgetasteten Signals setzt sich periodisch im Abstand ω_s fort. Abbildung 9.2 stellt diesen Zusammenhang grafisch dar. (Die Bandbreite erstreckt sich von $-B$ bis $+B$ beträgt also $2B$. Da wir $F(\omega)$ darstellen und $\omega = 2\pi f$ wird daraus $2\pi \cdot 2B$.)

Das Spektrum $F(\omega)$ des Originalsignals $f(t)$ ist im Spektrum des abgetasteten Signals $f_S(t)$ enthalten und kann aus $F_s(\omega)$ durch »Herausschneiden« mit einem idealen Tiefpassfilter fehlerfrei wiederhergestellt werden.

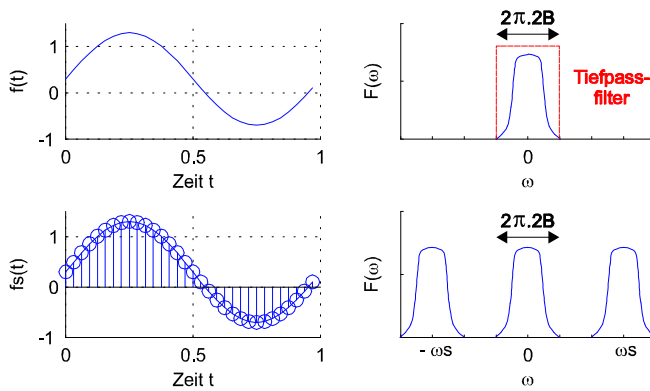


Abbildung 9.2: Periodisches Spektrum

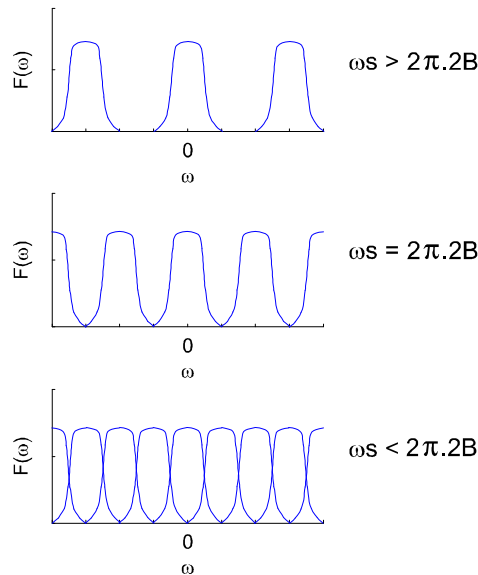


Abbildung 9.3: Überlappen von Spektren

Aus Abbildung 9.2 kann man erkennen, dass die periodische Wiederholung des Spektrums im Abstand von ω_s erfolgt. Je größer ω_s ist, desto weiter rücken die Spektren auseinander. Wird ω_s aber kleiner, dann rücken die Spektren näher aneinander, wenn $\omega_s = 2\pi \cdot 2B$, dann stoßen die Spektren aneinander. Ist $\omega_s < 2\pi \cdot 2B$, dann überlappen die Spektren. Abbildung 9.3 stellt diesen Zusammenhang grafisch dar.

Wie man aus Abbildung 9.3 erkennen kann, lassen sich die Spektren für $\omega_s > 2B$ leicht separieren, für $\omega_s = 2B$ kann eine Trennung nur mit einem rechteckigen (idealen) Filter erfolgen. Derartige Filter lassen sich aber nicht realisieren. Für den Fall dass $\omega_s < 2B$ lassen sich die Spektren nicht mehr trennen.

Daraus folgt das Shannon'sche Abtasttheorem

Satz 65 *Ein kontinuierliches Signal $s(t)$, das keine Frequenzkomponenten größer als f_{\max} enthält, kann exakt aus einer Folge von Proben $s[n] = s(nT_s)$ rekonstruiert werden, wenn die Abtastfrequenz $f_s = 1/T_s$ größer als $2 \times f_{\max}$ ist.*

Das Abtasttheorem gilt nur für *bandbegrenzte* Signale, d.h. für Signale die sinusförmige Komponenten mit einer obersten Frequenz enthalten ($f_{\max} \neq \infty$). Sprachsignale

Damit ein kontinuierliches, bandbegrenzt Signal mit einer Minimalfrequenz von 0 Hz und einer Maximalfrequenz f_{\max} nach der Abtastung fehlerfrei rekonstruiert werden kann, muss die **Abtastfrequenz größer $2 \times f_{\max}$ sein.**

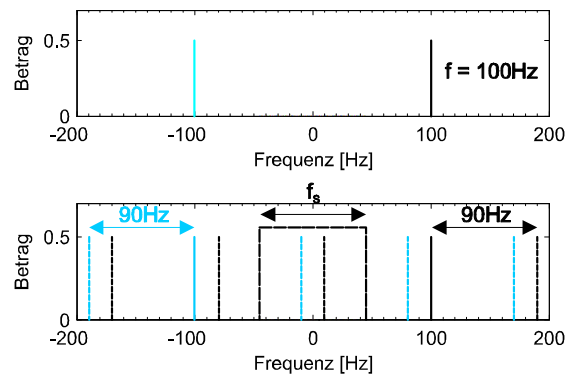


Abbildung 9.4: Aliasing

enthalten Frequenzkomponenten bis 4 kHz, Musiksignale Frequenzkomponenten bis 20 kHz. Die Abtastfrequenz von Audio-CDs beträgt 44.1 kHz, liegt also knapp über den erforderlichen 2×20 kHz.

Die niedrigste Abtastfrequenz, um ein Signal fehlerfrei rekonstruieren zu können, wird Nyquist-Frequenz genannt.

9.4 Aliasing und Folding

Aliasing und Folding tritt auf, wenn das Abtasttheorem verletzt wurde.

Bei **Aliasing** ist die **Frequenz** des **falsch** rekonstruierten Signals **positiv**. Die Abtastfrequenz war kleiner als die größte Frequenzkomponente des Originalsignals.

Beispiel 66 Ein sinusförmiges Signal der Frequenz $f = 100$ Hz werde mit $f_s = 90$ Hz abgetastet. Das 100 Hz-Signal hat im Frequenzbereich Spektrallinien bei ± 100 Hz. Durch die Abtastung entstehen Spektrallinien der periodischen Fortsetzung ($\omega \pm n\omega_s$): 100, 190, 280, 370, ..., 10, -80, -170, ...

Es wird aber auch die negative Frequenz periodisch fortgesetzt: -100, -10, 80, 170, ..., -190, -280, -370.

Wir sehen an diesem Beispiel, dass negative Frequenzkomponenten eine wichtige Rolle spielen. Das abgetastete Signal wird nun durch Filterung rekonstruiert, die Filterbandbreite beträgt $\pm f_s/2 = \pm 45$ Hz. Innerhalb dieses Bandes »findet« das Filter das **Aliassignal** mit den Frequenzen ± 10 Hz, das abgetastete 100 Hz-Signal wird also **falsch** als 10 Hz-Signal rekonstruiert, da das Abtasttheorem verletzt wurde! Abbildung 9.4 stellt die Zusammenhänge grafisch dar.

Bei **Folding** ist die **Frequenz** des **falsch** rekonstruierten Signals **negativ**. Die Abtastfrequenz war zwar größer als die größte Frequenzkomponente des Originalsignals, jedoch nicht groß genug.

Aliasing und Folding kann man bei rotierenden Rädern im Film beobachten.

Beispiel 67 Als zweites Beispiel betrachten wir unser sinusförmiges 100 Hz-Signal, tasten es aber diesmal mit 110 Hz ab. Wir erhalten die Spektralkomponenten: 100, 210, 320, ..., -10, -120, sowie die Spektralkomponenten -100, 10, 120, 230, 340, ..., -210, -320, ...

Die Rekonstruktion des Originalsignals aus dem abgetasteten Signal durch Filterung mit $\pm f_s/2 = \pm 55$ Hz liefert wie im vorigen Beispiel wieder die Signalkomponenten ∓ 10 Hz. Wir sehen aber, dass die Signalkomponenten im Vergleich zum vorherigen Beispiel **gespiegelt** sind! Diese Verletzung des Abtasttheorems nennt man **Folding**.

Beispiel 68 Ein besonders einprägsames Beispiel für die Verletzung des Abtasttheorems tritt bei Filmen auf. Filme sind Folgen von Einzelbildern, die mit einer Rate von 25 Bildern/Sekunde aufgezeichnet werden. Das Auge kann die Einzelbilder nicht mehr auflösen und es entsteht der Eindruck von kontinuierlichen Bewegungen. In der Regel sind diese Bewegungen langsam gegen die Bildfrequenz, das Abtasttheorem wird eingehalten. Am

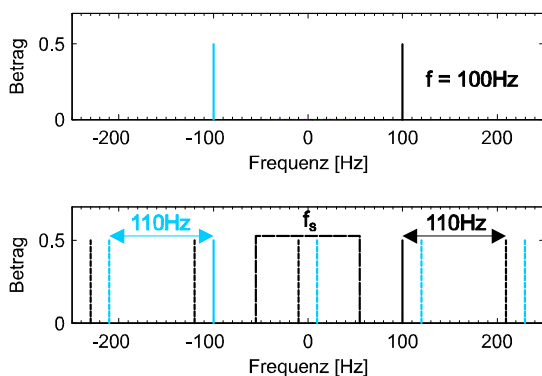


Abbildung 9.5: Folding

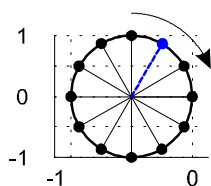


Abbildung 9.6: Drehung eines Speichenrads

Beispiel von Kutschenrädern kann man aber die Verletzung des Abtasttheorems deutlich beobachten: Eine Kutsche bewegt sich, man sieht aber, dass sich die Speichen der Kutsche nur ganz langsam in die Bewegungsrichtung der Kutsche drehen, während sich die Kutsche offensichtlich schneller bewegt. Es kann sogar der Fall eintreten, dass sich die Kutsche vorwärts bewegt, während die Räder scheinbar stehen oder sich nach rückwärts drehen.

Abbildung 9.6 liefert die Erklärung für dieses Phänomen. Wir betrachten die strichliert gezeichnete »Speiche«. Das Rad dreht sich und die Kamera nimmt 25 Bilder/Sekunde auf. Wenn sich die betrachtete Speiche zwischen zwei Bildern um einen geringen Winkel dreht, dann kann man im Film die korrekte Drehung des Rades beobachten (Abtasttheorem eingehalten). Das Rad kann sich aber viel schneller drehen und zwar, zwischen den Abtastungen mehrmals um die eigene Achse (Abtasttheorem verletzt).

Wenn sich die Speiche mehrmals um die eigene Achse dreht und die Bilder aber immer dann genommen werden, wenn sich die Speiche genau n -mal gedreht hat, dann steht das Rad scheinbar.

Wenn sich die Speiche zwar mehrmals um die eigene Achse dreht, die Bilder aber immer dann genommen werden, wenn der Winkel der Speiche $n2\pi + \alpha$ ist, dann dreht sich das Rad scheinbar langsamer. Wir beobachten den Aliasingeffekt wie in unserem vorigem Beispiel, wo aus einem 100 Hz-Signal ein 10 Hz-Signal wurde. Aus einer Raddrehung von vielen Umdrehungen/Sekunde wird eine Raddrehung aus wenigen Umdrehungen/Sekunde. Schließlich tritt noch der Fall auf, dass sich die Speiche mehrmals um die eigene Achse dreht, das Bild aber zu einem Zeitpunkt aufgenommen wird, wo die Speiche etwas vor der vollen Umdrehung steht $n2\pi - \alpha$. In diesem Fall dreht sich das Rad scheinbar rückwärts – also mit negativer Frequenz – und wir beobachten den Effekt des Foldings.

9.5 Signalrekonstruktion

Abgetastete Signale nehmen an den Abtastpunkten Proben aus kontinuierlichen Signalen. Für die Wiederherstellung eines kontinuierlichen Signals müssen daher die »Lücken«

Ein **kontinuierliches Signal** kann durch **Interpolation** zwischen den Abtastwerten aus einem diskreten Signal **erzeugt** werden.

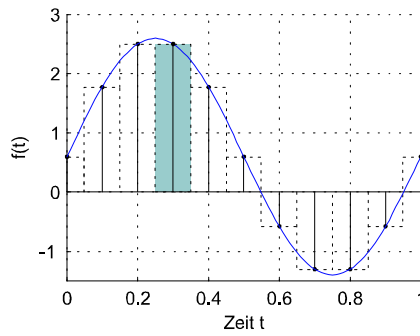


Abbildung 9.7: Rekonstruktion mit Rechteckimpulsen

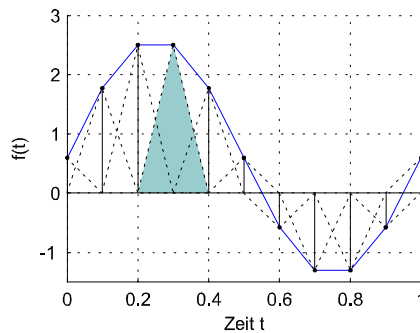


Abbildung 9.8: Rekonstruktion mit Dreieckimpulsen

zwischen den Abtastwerten interpoliert werden.

Die einfachste Form der Interpolation erhält man, indem man den Abtastwert zwischen den Lücken konstant hält. Diese Form der Interpolation kann man auch darstellen, indem man sich das rekonstruierte Signal als eine Folge von gewichteten Rechteckimpulsen vorstellt. Abbildung 9.7 stellt den Zusammenhang grafisch dar.

Die Interpolation mit Rechteckimpulsen führt nur zu einer groben Annäherung an das ursprüngliche Signal. Durch dichteres Abtasten kann man eine bessere Anpassung erreichen, allerdings zum Preis einer höheren Datenmenge.

Eine bessere Form der Interpolation ist die lineare Interpolation zwischen den Abtastpunkten. Die lineare Interpolation kann durch Überlagerung von gewichteten Dreieckimpulsen dargestellt werden, wie in Abbildung 9.8 gezeigt.

Ein **kontinuierliches Signal** kann aus einem diskreten Signal auch **durch (analoge) Filterung erzeugt** werden.

Mit Hilfe von Rechteck- und Dreieckimpulsen stellen wir die Interpolation im *Zeitbereich* dar. Bei der mathematischen Darstellung des Abtastvorgangs in (9.12) und Abbildung 9.2 sehen wir, dass das Spektrum des abgetasteten Signals die periodische Fortsetzung des Spektrums des kontinuierlichen Signals ist. Die Rekonstruktion des kontinuierlichen Signals muss also *fehlerfrei* durch Filterung mit einem idealen Tiefpass-Filter möglich sein. Wir rekonstruieren das Signal also nicht durch *Interpolation im Zeitbereich*, sondern durch *Filterung im Frequenzbereich*, wie in Abbildung 9.9 dargestellt.

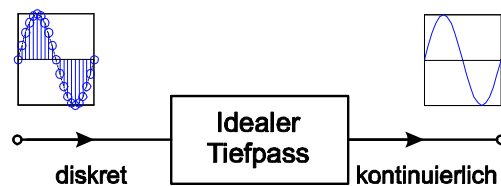


Abbildung 9.9: Interpolation durch ideales Filter

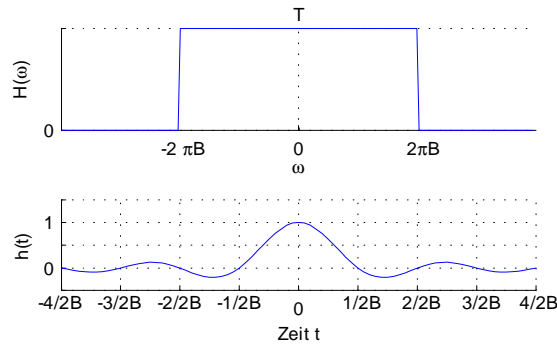


Abbildung 9.10: Impulsantwort des idealen Filters

Wir fragen uns nun, welcher Interpolation im Zeitbereich die ideale Filterung im Frequenzbereich entspricht und schicken dazu das abgestastete Signal durch ein ideales Filter. Ein ideales Tiefpass-Filter lässt im Durchlassbereich alle Frequenzen ungehindert durch, während es im Sperrbereich alle Frequenzen vollständig unterdrückt. Wir berechnen zunächst die Antwort des Filters auf einen Einheitsimpuls. Die Antwort des Filters auf die Impulsfolge $f_s(t)$ berechnen wir dann durch Überlagerung der Antworten auf die gewichteten, zeitlich versetzten Einheitsimpulse.

Die Impulsantwort eines Filters erhält man durch inverse Fouriertransformation des Frequenzgangs $H(\omega) \Rightarrow h(t)$. Die Bandbreite des idealen Filters ergibt sich aus der Nyquist-Frequenz $F_s = 2B$, das Abtastintervall ist dann $T = 1/2B$. Der Frequenzgang des idealen Tiefpassfilters ist

$$H(\omega) = \begin{cases} T & |\omega| < 2\pi B \\ 0 & |\omega| \geq 2\pi B \end{cases} \quad (9.13)$$

$$h(t) = \frac{1}{2\pi} \int_{-2\pi B}^{2\pi B} T e^{j\omega t} d\omega = 2BT \operatorname{sinc}(2\pi Bt) \quad (9.14)$$

$$h(t) = \operatorname{sinc}(2\pi Bt) \quad \text{da bei der Nyquistfrequenz } 2BT = 1 \quad (9.15)$$

Wir treffen wieder auf die Si-Funktion. Während bei der Fouriertransformierten des Rechteckimpulses die Si-Funktion im Frequenzbereich auftritt, tritt bei der Fouriertransformierten des idealen Filters die Si-Funktion im Zeitbereich auf. Abbildung 9.10 zeigt Frequenzgang und Impulsantwort eines idealen Filters².

Aus Abbildung 9.10 können wir sehen, dass $h(t) = 0$ bei allen Vielfachen der Nyquist-Abtastpunkte $t = \pm \frac{n}{2B}$ Null ist, mit Ausnahme von $t = 0$.

Da das ideale Filter das ursprüngliche, kontinuierliche Signal fehlerfrei herstellt, muss auch die Überlagerung der Impulsantworten des idealen Filters das Ausgangssignal fehlerfrei herstellen

$$f(t) = \sum_k f(kT)h(t - kT) = \sum_k f(kT)\operatorname{sinc}(2\pi Bt - k\pi) \quad (9.16)$$

Während die Rekonstruktion des Signal durch Rechteck- und Dreiecksimpulse nur eine ungenaue Wiedergewinnung des Signal ermöglichte, stellt die Rekonstruktion durch überlagerte und gewichtete sinc-Pulse das Signal fehlerfrei her. Abbildung 9.11 zeigt die Signalrekonstruktion (Wegen der besseren Übersichtlichkeit sind die Interpolationspulse nur für jeden zweiten Abtastpunkt gezeichnet.).

²Wie wir sehen, ist die Impulsantwort eines idealen Filters nicht-kausal, d.h. das Filter antwortet bereits vor dem Anlegen des Impulses. Nichtkausale Filter sind nicht realisierbar!

Die **Impulsantwort** des idealen Filters bekommt man durch **inverse Fouriertransformation der Rechteckfunktion** (die Systemfunktion eines idealen Filters ist ein Rechteck).

Die **Filterung** mit einem **idealen Tiefpass** entspricht mathematisch der **Faltung** mit der **sinc-Funktion**.

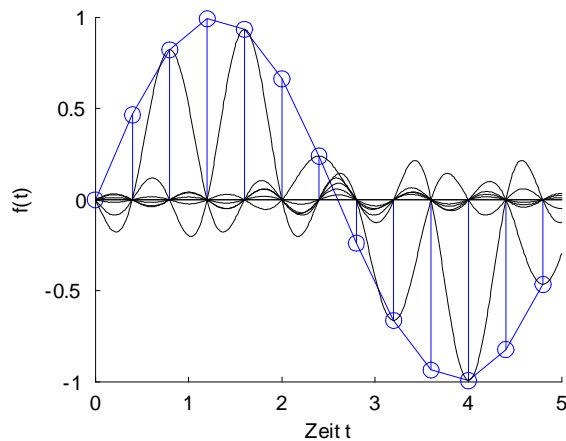


Abbildung 9.11: Rekonstruktion durch sinc-Pulse

9.6 Filterung

9.6.1 Interpolationsfilter

Ein **ideales Filter** ist **nicht-kausal** und daher **nicht realisierbar**.

Ideale (analoge) Tiefpass-Filter sind nicht-kausal und daher nicht realisierbar. Man kann zwar steilflankige analoge Filter mit hoher Dämpfung im Sperrbereich bauen, es ist aber nicht möglich Filter zu realisieren, die die Signale im gesamten Sperrbereich vollständig unterdrücken. Eine praktische Lösung dieses Problems wird dadurch gefunden, dass das Signal mit Abtastfrequenzen größer als der Nyquist-Frequenz abgetastet wird. Damit entstehen Lücken im periodisch fortgesetzten Spektrum und die Anforderungen an die Flankensteilheit des Filters werden geringer. Man erreicht eine praktisch ausreichende Unterdrückung, aber nie die theoretische geforderte vollständige Ausblendung des Sperrbereichs.

9.6.2 Antialiasing Filter

Ein Signal, das **gleichzeitig zeitbegrenzt und bandbegrenzt** ist, **gibt es nicht**.

Eine weiteres Problem besteht darin, dass jedes praktische Signal von endlicher Länge ist. Wie wir von der Fouriertransformation wissen, hat ein Signal endlicher Länge ein unendlich breites Spektrum.

Bemerkung 69 *Ein Signal kann nicht gleichzeitig zeitbegrenzt und bandbegrenzt sein! Ist das Signal zeitbegrenzt, hat es also eine endliche Dauer, dann erstreckt sich das Spektrum von $-\infty$ bis ∞ und ist daher nicht bandbegrenzt. Ist das Signal bandbegrenzt, dann muss sich das Signal über eine unendliche Dauer im Zeitbereich erstrecken, ist also nicht zeitbegrenzt.*

Aus diesem Grund überlappen sich benachbarte Spektren, wie Abbildung 9.12 zeigt.

Antialiasing Filter begrenzen das Band von Signalen vor der Abtastung. Es geht daher immer um **analoge Filter**.

Um das Problem der überlappenden Spektren praktisch in den Griff zu bekommen setzt man (analoge) Antialiasing-Filter ein, die die Bandbreite begrenzen und damit die Überlappung verhindern. Die Antialiasing Filterung muss *vor* der Abtastung erfolgen, da ja erst die Abtastung das Spektrum periodisch fortsetzt. Abbildung 9.13 zeigt das Blockdiagramm eines digitalen Signalverarbeitungssystems.

Das Interpolationsfilter ist das (analoge) Tiefpassfilter nach dem DA/Wandler, das die Werte zwischen den Abtastpunkten im Zeitbereich interpoliert. Im Frequenzbereich entspricht diese Interpolation dem Herausschneiden des gewünschten Spektrums aus dem periodisch fortgesetzten Spektrum.

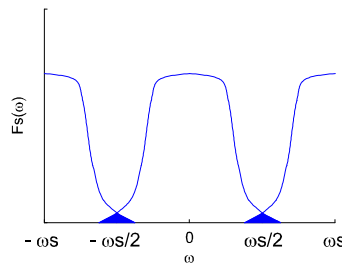


Abbildung 9.12: Überlappung benachbarter Spektren

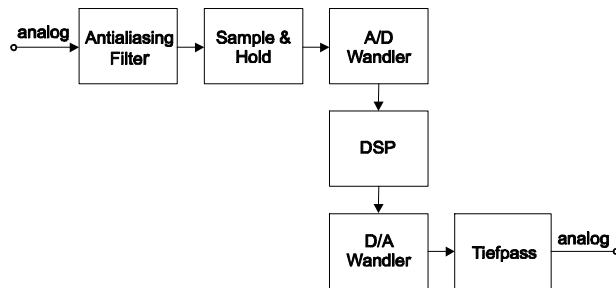


Abbildung 9.13: Blockdiagramm digitale Signalverarbeitung

9.7 Digitalisierung

Nach der Abtastung haben die Abtastwerte noch immer kontinuierliche Werte. Erst durch die Umwandlung der kontinuierlichen Werte in diskrete Werte entstehen digitale (quantisierte) Signale. In praktischen Fällen geschieht das durch Analog-/Digital-Wandlung.

Die Auflösung des Analog-/Digital-Wandlers wird nach Qualitätskriterien ausgewählt. Für Sprachsignale reicht eine Auflösung von 8 bit aus, bei Musiksignalen auf einer Audio-CD beträgt die Auflösung 16 bit. Je geringer die Auflösung des A/D-Wandlers ist, desto stärker weicht das digitale Signal vom analogen Signal ab.

Abbildung 9.14 zeigt das Fehlersignal für niedrigere und höhere Auflösung. Beim Abhören hört man einen »reinen« Ton, der von einem rauschartigen Störsignal überlagert ist, man spricht daher vom Quantisierungsrauschen. Je höher die Auflösung des A/D-Wandlers desto geringer ist der zu hörende Rauschanteil.

Das digitale Signal weicht vom kontinuierlichen, analogen Wert ab, wobei der dabei entstehende Fehler von der Auflösung des Wandlers abhängt. Die Amplitude des Quantisierungsfehler ist maximal Δ , wobei $\Delta = \pm \frac{1}{2} LSB$ der Auflösung des A/D-Wandlers entspricht. Wir können zwar nichts über die Kurvenform des Fehlersignals sagen, sehr wohl aber über die Leistung, die im Fehlersignal steckt und die sich als Rauschen³ bemerkbar macht. Zur Abschätzung des durch die Quantisierung entstehenden Rauschens betrachten wir ein Signal $s(t)$, in dem alle Signalamplituden gleich wahrscheinlich auftreten. Die mittlere Leistung des Fehlersignals ist dann

$$e_{eff}^2 = \frac{1}{\Delta} \int_{-\Delta/2}^{\Delta/2} s^2 ds = \frac{\Delta^2}{12} \tag{9.17}$$

Das Quantisierungsrauschen hat den Effektivwert von

$$e_{eff} = \frac{\Delta}{\sqrt{12}} = \frac{LSB}{\sqrt{12}} \approx 0.29 LSB \tag{9.18}$$

³Rauschsignale sind unregelmäßige Signale die z.B. beim Geräusch eines Wasserfalls, beim Klatschen oder beim elektrischen Strom durch unregelmäßige Teilchenbewegung auftreten.

Bei der **Digitalisierung** wird die **Amplitude** der abgetasteten Signalwerte **quatisiert**.

Die **Differenz** zwischen dem ursprünglichen Signal und dem quantisierten Signal wird **Quantisierungsrauschen** genannt und hat die Amplitude **maximal 1/2 LSB**.

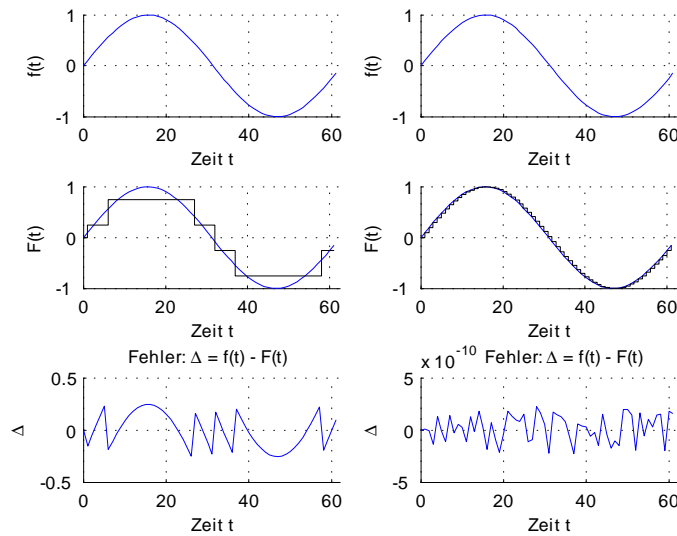


Abbildung 9.14: Amplitudenquantisierung

Bei einem 8-bit Wandler beträgt das effektive Rauschen $0.29/256$ oder $1/900$ des Wertebereichs, bei 12 bit $0.29/4096$ oder $1/14000$ des Wertebereichs, bei 16 bit $0.29/65536$ oder $1/226000$ des Wertebereichs. Jedes analoge Signal enthält aus physikalischen Gründen bereits einen Rauschanteil, die Quantisierung erhöht diesen Rauschanteil in Abhängigkeit von der Auflösung des Wandlers!

Diese Angaben gelten nur für gleichverteilte Quantisierungsfehler und z.B. nicht, wenn sich ein analoges Signal über einen langen Zeitabschnitt nur so gering ändert, dass der digitale Wert konstant bleibt. In derartigen Fällen äußert sich der Quantisierungsfehler nicht als Zufallsrauschen, sondern wie ein (störender) Begrenzungseffekt. Zur subjektiven Verbesserung eines derartigen digitalen Signals wird dem analogen Signal ein kleiner Rauschanteil (mit dem Amplitudenbereich von 2 LSB Spitze-Spitze) hinzugefügt, der bewirkt, dass das digitale Signal nicht konstant bleibt sondern zwischen den benachbarten Werten umschaltet. Dadurch wird der analoge Signalwert im Mittel genauer erreicht und das Signal mit dem hinzugefügten Rauschen wird bei Tonsignalen als natürlicher empfunden als der Begrenzungseffekt. Dieses Verfahren wird Dithering genannt und ist irreversibel.

9.8 Die zeitdiskrete Fourier-Transformation (DTFT)

Die **DTFT** ist die **Fouriertransformation** eines **abgetasteten** (zeitdiskreten) **Signals**. Das **Spektrum** ist dabei **kontinuierlich** und **periodisch**.

Die zeitdiskrete Fourier-Transformation (DTFT) einer Folge $x[n]$ ist definiert

$$X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x[n]e^{-j\omega n} \quad (9.19)$$

Man kann es sich auch als einen Spezialfall der kontinuierlichen Fouriertransformation vorstellen, wo die zu transformierende Funktion aus gewichteten Dirac-Impulsen bei n besteht und sonst überall Null ist.

In der Polardarstellung $X(e^{j\omega}) = |X(e^{j\omega})| e^{j\varphi(\omega)}$ können wir $\varphi(\omega)$ durch $\varphi(\omega) + 2\pi k$, $k \dots$ ganzzahlig ersetzen, ohne das sich $X(e^{j\omega})$ ändert. Da das Argument von $X(e^{j\omega})$ (die komplexe Exponentialfunktion $e^{j\omega}$) selber eine kontinuierliche periodische Funktion von ω ist, muss das Spektrum der DTFT auch kontinuierlich und periodisch fortgesetzt sein.

Beispiel 70 Wir berechnen die DTFT von

$$x[n] = 0.5^n \delta_{-1}[n]$$

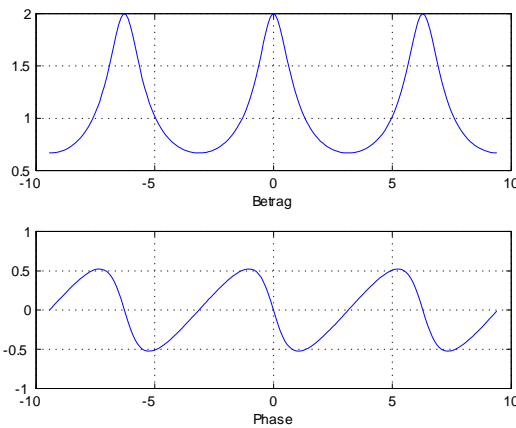


Abbildung 9.15: DTFT von $x[n] = 0.5^n \delta_{-1}[n]$

und erhalten

$$\begin{aligned} X(e^{j\omega}) &= \sum_{n=0}^{\infty} 0.5^n e^{-j\omega n} = \sum_{n=0}^{\infty} (0.5e^{-j\omega})^n \\ &= \frac{1}{1 - 0.5e^{-j\omega}} \end{aligned}$$

Abbildung 9.15 zeigt, dass das Betrags- und Phasenspektrum der abklingenden Sprungfunktion $x[n] = 0.5^n \delta_{-1}[n]$. Wir sehen, dass das Spektrum $X(e^{j\omega})$ einer Folge von Abtastwerten $x[n]$ eine kontinuierliche Funktion ist, die mit Hilfe eines Digitalrechners nicht darstellbar ist, da man ja ∞ viele Punkte für die Darstellung bräuchte. Wir können aber nur eine endliche Zahl von Punkten speichern, müssen daher das Spektrum im Frequenzbereich abtasten.

9.9 Abtastung im Frequenzbereich

Bei der Abtastung im Zeitbereich haben wir gezeigt, dass ein bandbegrenzttes Signal der Bandbreite B Hz aus den Abtastwerten wiederhergestellt werden kann, wenn die Abtastwerte mit einer Abtastfrequenz von $F_s > 2B$ entnommen werden.

Wir untersuchen nun den Abtastvorgang im Frequenzbereich. Während wir es beim Abtasttheorem im Zeitbereich mit *bandbegrenzten* Signalen zu tun hatten, haben wir es beim Abtasttheorem im Frequenzbereich mit *zeitbegrenzten* Signalen zu tun. Wir zeigen, dass das Spektrum $F(\omega)$ eines zeitbegrenzten Signals der Länge τ aus den Abtastwerten von $F(\omega)$, die mit einer Rate von $R > \tau$ Proben pro Hertz entnommen werden, rekonstruiert werden kann. Abbildung (9.16) gibt eine Übersicht über die Zusammenhänge⁴.

$$F(\omega) = \int_{-\infty}^{\infty} f(t)e^{-j\omega t} dt = \int_0^{\tau} f(t)e^{-j\omega t} dt \tag{9.20}$$

Wir rekonstruieren nun $f_P(t)$, ein periodisches Signal, das aus $f(t)$ durch periodische Fortsetzung im Abstand T_0 gebildet wird. Periodische Signale werden durch eine Fourier-Reihe dargestellt:

$$f_P(t) = \sum_{k=-\infty}^{\infty} D_k e^{jk\omega_0 t} \quad \omega_0 = \frac{2\pi}{T_0} \quad \tau < T_0 \tag{9.21}$$

⁴ $F(\omega)$ ist in der Regel eine komplexwertige Funktion, die hier der Einfachheit halber reell dargestellt wird.

Um ein **Spektrum** im **Rechner darstellen** zu können, muss es ebenfalls **abgetastet** werden.

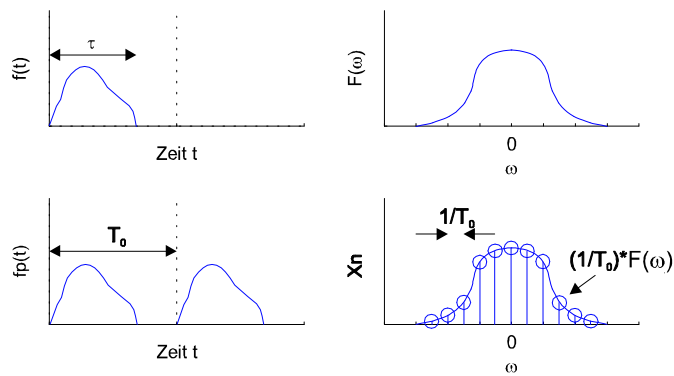


Abbildung 9.16: Abtastung im Frequenzbereich

$$D_k = \frac{1}{T_0} \int_0^{T_0} f(t) e^{-jk\omega_0 t} dt = \frac{1}{T_0} \int_0^\tau f(t) e^{-jk\omega_0 t} dt \quad (9.22)$$

Setzen wir (9.20) in (9.22) ein, erhalten wir

$$D_k = \frac{1}{T_0} F(k\omega_0) \quad (9.23)$$

Durch die inverse Fouriertransformation eines **abgetasteten Spektrums** entsteht ein **periodisches Zeitsignal**.

Das **Spektrum** eines **zeitbegrenzten Signals**, das mit einer genügend hohen Abtastrate **abgetastet** wird, lässt sich **eindeutig** wieder von den Abtastwerten **rekonstruieren**.

Dieses Ergebnis bedeutet, dass die Koeffizienten der Fourier-Reihe für $f_P(t)$ gleich den mit $1/T_0$ multiplizierten Abtastwerten des Spektrums $F(\omega)$ im Abstand von ω_0 sind. Solange $\tau < T_0$ gibt es keine Überlappung und $f(t)$ kann eindeutig aus $f_P(t)$ rekonstruiert werden. Das impliziert, dass auch $F(\omega)$ aus seinen Abtastwerten rekonstruiert werden kann, wenn für die Abtastung die Bedingung $F_0 = (1/T_0), T_0 \geq \tau$ eingehalten wird (F_0 ist dabei der Abstand zwischen zwei Abtastwerten auf der Frequenzachse). Wir können das kontinuierliche Spektrum $F(\omega)$ aus den Abtastwerten von $F(\omega)$ rekonstruieren, wenn die Proben in Abständen nicht größer als $F_0 = 1/\tau$ Hz entnommen werden. Die Abtastrate im Frequenzbereich ist also

$$R = \frac{1}{F_0} \geq \tau \text{ Abtastwerte/Hz.}$$

Ähnlich der Interpolation im Zeitbereich ergibt sich im Frequenzbereich der Zusammenhang

$$F(\omega) = \sum_n F(k\omega_0) \text{sinc} \left(\frac{\omega\tau}{2} - k\pi \right) \quad \omega_0 = \frac{2\pi}{\tau} \quad (9.24)$$

9.10 Die diskrete Fourier-Transformation (DFT)

Die DFT ist ursprünglich aus der Berechnung der Fourier-Koeffizienten und der Fourier-Transformation auf digitalen Rechnern entstanden und ist eines der wichtigsten Werkzeuge der Signalverarbeitung.

Bei der **DFT** wird sowohl im Zeitbereich als auch im Frequenzbereich nur mit **endlich vielen diskreten Werten** gerechnet.

Signale liegen in der Regel nicht in Form von mathematischen Funktionen vor, die Fouriertransformation kann daher nur bei einfachen, idealisierten Signalen analytisch berechnet werden. Für praktische Signale muss die Berechnung des Spektrums numerisch erfolgen, es können daher sowohl im Zeitbereich als auch im Frequenzbereich nur endlich viele diskrete Werte berechnet werden.

- Für die Berechnung der DFT gehen wir von zeitbegrenzten Signalen $f(t)$ der Länge τ aus (Abbildung 9.17(a)). Zeitbegrenzte Signale haben ein Spektrum $F(\omega)$ das nicht unendlich breit ist (Abbildung 9.17(b)).

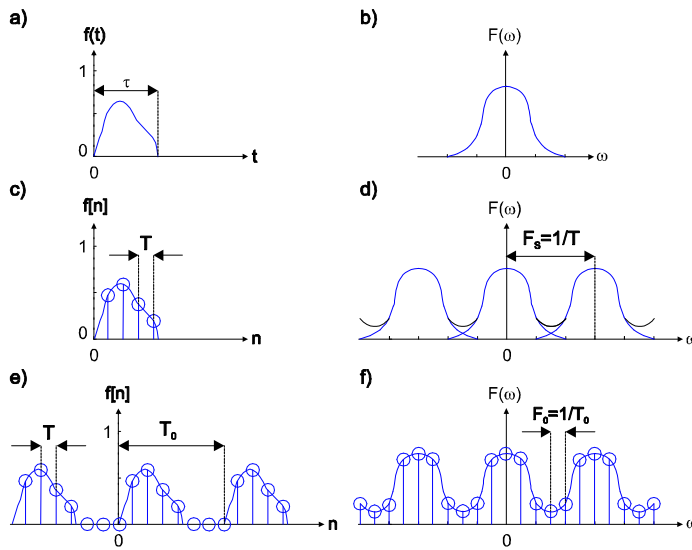


Abbildung 9.17: Abtastung im Zeit- und Frequenzbereich

- Aus dem zeitbegrenzten Signal gewinnen wir das diskrete Signal $f_s(t) = f[n]$ durch Abtastung von $f(t)$ im Abstand $T = 1/F_s$, wie in Abbildung 9.17(c) gezeigt. Durch die Abtastung wird das Spektrum periodisch mit der Periodendauer $F_s = 1/T$ und wir erhalten das Spektrum $F_s(\omega)$, wie man in Abbildung 9.17(d) sieht.
- Die Abtastung des Spektrums im Abstand $F_0 = (1/T_0)$ nach Abbildung 9.17(f) führt zur periodischen Fortsetzung des abgetasteten Zeitsignals, mit der Periode T_0 , wie Abbildung 9.17(e) dargestellt.

Aus Abbildung 9.17 sehen wir, dass ein abgetastetes und periodisch wiederholtes Zeitsignal zu einem abgetasteten periodischen Spektrum führt. Bei der Berechnung der DFT haben wir es also immer mit periodischen Folgen zu tun, wobei wir die Werte immer nur für eine Periode berechnen müssen. (Weil wir wissen, dass sie sich weiter nur mehr wiederholen,).

Bemerkung 71 Beachten Sie, dass es keine mit dem Computer berechenbare Fouriertransformation für endliche digitale Signale gibt. Ist das diskrete Signal zeitbegrenzt, dann entsteht ein kontinuierliches periodisches Spektrum (DTFT), das (vielleicht) analytisch berechnet werden kann, aber nicht in einem Rechner gespeichert werden kann. Das Spektrum kann nur durch Entnahme von Proben (aus einer Periode) gespeichert werden, was bewirkt, dass die Zeitfunktion periodisch fortgesetzt wird. Mit anderen Worten: Die DFT kennt nur periodische Zeitfunktionen und periodische Spektren!

Aus Abbildung 9.17 finden wir noch den wichtigen Zusammenhang, dass die Zahl N_0 der Proben pro Periode im Zeitbereich identisch mit N'_0 , der Zahl der Proben im Frequenzbereich pro Periode ist, da

$$N_0 = \frac{T_0}{T} \quad N'_0 = \frac{F_s}{F_0} \tag{9.25}$$

Da aber

$$F_s = \frac{1}{T} \quad \text{und} \quad F_0 = \frac{1}{T_0} \tag{9.26}$$

erhalten wir

Die **DFT** rechnet nur mit **periodischen Zeitsignalen** und **periodischen Spektren**, obwohl man immer nur eine Periode betrachtet.

$$N_0 = \frac{T_0}{T} = \frac{F_s}{F_0} = N'_0 \quad (9.27)$$

Die Anzahl der Proben des Zeitsignals ist gleich der Anzahl der Proben des Spektrums.

Aus Abbildung 9.17(d,f) sehen wir, dass sich die Spektren durch die periodische Fortsetzung überlappen. Diese Überlappung kann durch Erhöhen der Abtastfrequenz F_s verringert, aber für zeitbegrenzte Signale $f(t)$ nie vollständig eliminiert werden, da deren Spektrum $F(\omega)$ unendlich breit ist. Hätten wir ein bandbegrenztes Signal, träte zwar keine Überlappung im Frequenzbereich auf, aber die periodische Fortsetzung im Zeitbereich durch die Abtastung im Frequenzbereich würde zu einer Überlappung im Zeitbereich führen.

Die DFT weist einen **Fehler** auf, der sich durch Erhöhen der Abtastfrequenz **verringern** lässt, aber **nie verschwinden** kann.

Die bei der numerischen Berechnung der direkten oder inversen Fouriertransformation auftretenden Fehler können wir durch Erhöhen der Abtastfrequenz zwar klein machen, sie lassen sich aber nie vollständig eliminieren!

Für die diskrete Fouriertransformation erhalten wir die Transformationspaare

$$f[n] = \frac{1}{N_0} \sum_{k=0}^{N_0-1} F[k] e^{jk\Omega_0 n} \quad (9.28)$$

$$F[k] = \sum_{n=0}^{N_0-1} f[n] e^{-jk\Omega_0 n} \quad (9.29)$$

$$\Omega_0 = \omega_0 T = \frac{2\pi}{N_0} \quad (9.30)$$

In der Literatur ist häufig eine kompaktere Schreibweise in Gebrauch, bei der man die Ähnlichkeit zu den anderen Fourieroperationen nicht mehr erkennen kann, aber die Symmetrie zwischen DFT und IDFT deutlich wird.

$$f[n] = \frac{1}{N_0} \sum_{k=0}^{N_0-1} F[k] w^{-kn} \quad (9.31)$$

$$F[k] = \sum_{n=0}^{N_0-1} f[n] w^{kn} \quad (9.32)$$

$$w = e^{-j\frac{2\pi}{N_0}} \quad (9.33)$$

Schreibt man die Transformationsbeziehungen in Matrixdarstellung an, so erhält man

$$\underbrace{\begin{bmatrix} F[0] \\ F[1] \\ F[2] \\ \vdots \\ F[N-1] \end{bmatrix}}_{\mathbf{F}} = \underbrace{\begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & w^1 & w^2 & \dots & w^{(N-1)} \\ 1 & w^2 & w^4 & \dots & w^{2(N-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & w^{(N-1)} & w^{2(N-1)} & \dots & w^{(N-1)(N-1)} \end{bmatrix}}_{\mathbf{W}} \cdot \underbrace{\begin{bmatrix} f[1] \\ f[2] \\ f[3] \\ \vdots \\ f[N-1] \end{bmatrix}}_{\mathbf{f}} \quad (9.34)$$

oder

$$\mathbf{F} = \mathbf{W} \cdot \mathbf{f} \quad (9.35)$$

und

$$\begin{bmatrix} f[1] \\ f[2] \\ f[3] \\ \vdots \\ f[N-1] \end{bmatrix} = \frac{1}{N} \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & w^{-1} & w^{-2} & \dots & w^{-(N-1)} \\ 1 & w^{-2} & w^{-4} & \dots & w^{-2(N-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & w^{-(N-1)} & w^{-2(N-1)} & \dots & w^{-(N-1)(N-1)} \end{bmatrix} \cdot \begin{bmatrix} F[0] \\ F[1] \\ F[2] \\ \vdots \\ F[N-1] \end{bmatrix} \quad (9.36)$$

oder

$$\mathbf{f} = \mathbf{W}^{-1} \cdot \mathbf{F} = \frac{1}{N} \mathbf{W}^* \cdot \mathbf{F} \quad (9.37)$$

Aus der Matrixdarstellung kann man die Symmetrie der Operationen erkennen. Bei der Berechnung der (I)DFT durch Matrixmultiplikation werden viele Rechenschritte mehrfach ausgeführt und die Zahl der (komplexen) Rechenoperationen ist $\sim N^2$. Durch Ausnutzen der Symmetrie lässt sich der Rechenaufwand mit Hilfe der schnellen Fouriertransformation (FFT) auf $N \log N$ reduzieren.

Es ist überraschend, dass die Rechenvorschrift für DFT und IDFT, abgesehen vom Vorzeichen und vom Skalierungsfaktor N , vollständig symmetrisch ist, obwohl den Transformationen sehr unterschiedliche Überlegungen zu Grunde liegen.

Signale enthalten nach der Abtastung keine Zeitinformation, sondern sind lediglich Folgen von Zahlen. Auch im Frequenzbereich haben wir es daher lediglich mit Folgen von Zahlen zu tun und man stellt daher das Spektrum über der Zahlenfolge von 0 bis $N/2$ dar⁵. In praktischen Anwendungen ist der Frequenzbereich bekannt, man wählt daher für die Darstellung eine »Frequenzachse«, die dem Problem angepasst ist, wobei folgende Darstellungen gebräuchlich sind

$$0 \leq k \leq N/2 \quad 0 \leq f \leq 0.5 \quad 0 \leq \omega \leq \pi \quad 0 \leq f \leq f_{\text{abtast}}/2 \quad (9.38)$$

Welche Darstellung gewählt wurde, kann in der Regel dem Zusammenhang entnommen werden.

9.11 Systeme mit unterschiedlichen Abtastraten (Multirate Filter)

Durch Verwendung unterschiedlicher Datenraten können die Vorteile analoger und digitaler Systeme kombiniert werden. Wir zeigen die Vorteile am Beispiel der Bearbeitung eines Sprachsignal der Bandbreite 100 Hz bis 3 kHz. Zur Bandbegrenzung setzen wir ein einfaches (und damit kostengünstiges, analoges) RC-Filter ein, tasten das Signal aber mit 64 kS/s ab (Oversampling). Das Spektrum des abgetasteten Signals erstreckt sich bis 32 kHz, das Band von 3 - 32 kHz ist aber bedeutungslos und kann daher mit einem digitalen Tiefpassfilter der Grenzfrequenz 3 kHz unterdrückt werden. Der höhere Aufwand durch das Oversampling und die Realisierung des digitalen Tiefpassfilters kann mit heutiger Technologie kostengünstiger verwirklicht werden als ein selektives analoges Antialiasing-Filter. Schließlich kann die Abtastfrequenz des digitalen Signals von 64 kS/s auf 8 kS/s reduziert werden, indem sieben von acht Abtastwerten entfernt werden. Dieser Vorgang wird *Dezimation* genannt.

Diese Art der Signalverarbeitung (einfaches und billiges analoges RC-Filter, scharfes, aber dennoch einfach und günstig zu realisierendes, digitales Filter) erzielt die selben Ergebnisse wie ein scharfes aber aufwändiges Analogfilter.

Unterschiedliche Abtastraten werden auch bei der *Interpolation* eingesetzt. In ein Datensignal werden Nullen eingefügt. Z.B. wird aus einem analogen 3 kHz Signal das mit 8 kS/s abgetastet wurde durch Einfügen von sieben Nullen nach jedem Abtastwert ein 64 kS/s-Signal. Dieses Signal enthält das 3 kHz Sprachsignal und periodische Fortsetzungen des Spektrums bis 32 kHz. Durch einfach zu realisierende digitale Filterung werden die Frequenzkomponenten über 3 kHz entfernt, durch D/A-Wandlung ein analoges Signal erzeugt, das mit einem einfachen analogen RC-Interpolationsfilter geglättet wird.

Durch Multirate-Filter können (teure) analoge Komponenten durch digitale Signalverarbeitung ersetzt werden. Darüber hinaus kann die Genauigkeit analoger Signalverarbeitung (1 %) bei digitaler Verarbeitung auf den Quantisierungsfehler verbessert werden.

⁵ Aus Symmetriegründen reicht die Darstellung von 0 bis $N/2$, obwohl N Punkte berechnet werden.

Die **FFT** ist ein effizienter **Algorithmus zur Berechnung der DFT**.

Bei diesen Systemen werden die guten Eigenschaften von **analogen** und **digitalen** Filtern **kombiniert**.

Eine höhere Abtastrate verringert die Anforderungen an Antialiasing- und Interpolationsfilter. Die hohe Abtastrate ist oft kostengünstiger zu realisieren als leistungsfähige analoge Filter.

9.12 Zusammenfassung Fouriertransformationen

Wir fassen die Transformationsbeziehungen zwischen Zeit- und Frequenzbereich zusammen. Abbildung 9.18 zeigt die zugehörigen Signalformen im Zeitbereich. Man nennt $F(\omega)$ die (direkte) Fouriertransformation von $f(t)$ und $f(t)$ die inverse Fouriertransformation von $F(\omega)$.

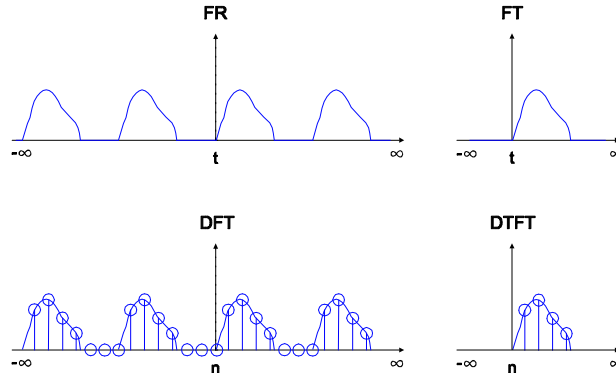


Abbildung 9.18: Fourierreihe (FR), Fouriertransformation (FT), Zeitdiskrete FT (DTFT), Diskrete FT (DFT)

Fourierreihe (FR)

Kontinuierliche, periodische (unendliche) Zeitfunktion $f(t) \Leftrightarrow$ (un)endlich breites Linienspektrum $F[k]$.

$$f(t) = \sum_{k=-\infty}^{+\infty} F[k] e^{jk\omega_0 t} \quad (9.39)$$

$$F[k] = \frac{1}{T} \int_0^T f(t) e^{-jk\omega_0 t} dt \quad (9.40)$$

Fouriertransformation (FT)

Aperiodische, kontinuierliche Zeitfunktion $f(t) \Leftrightarrow$ unendliches kontinuierliches Spektrum $F(\omega)$.

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) e^{j\omega t} d\omega \quad (9.41)$$

$$F(\omega) = \int_{-\infty}^{\infty} f(t) e^{-j\omega t} dt \quad (9.42)$$

Zeitdiskrete Fouriertransformation (DTFT)

Aperiodische, diskrete Zeitfunktion $x[n] \Leftrightarrow$ periodisches kontinuierliches Spektrum $X(e^{j\omega})$.

$$x[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(e^{j\omega}) e^{j\omega n} d\omega \quad (9.43)$$

$$X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x[n] e^{-j\omega n} \quad (9.44)$$

Diskrete Fouriertransformation (DFT)

Periodische, diskrete Zeitfunktion $f[n] \Leftrightarrow$ diskretes, periodisches Spektrum $F[k]$.

$$f[n] = \frac{1}{N} \sum_{k=0}^{N-1} F[k] e^{jk\hat{\omega}_0 n} \quad \hat{\omega}_0 = \omega T = \frac{2\pi}{N} \quad (9.45)$$

$$F[k] = \sum_{n=0}^{N-1} f[n] e^{-jk\hat{\omega}_0 n} \quad (9.46)$$

Die Transformationsbeziehungen stellen den Zusammenhang zwischen Zeit- und Frequenzbereich her und unterscheiden sich lediglich dadurch, dass bei kontinuierlichen Funktionen Integrale, bei diskreten Funktionen Summen auftreten.

Bei der Zerlegung einer periodischen, kontinuierlichen Funktion ist die Annäherung nichtstetiger Funktionen zwar so genau, dass die Energiedifferenz zwischen Original- und Fourierintegral gegen Null geht, es tritt aber das Gibbs'sche Phänomen auf. Bei diskreten Signalen ist die Fourierzerlegung exakt, d.h. es tritt keine Differenz zwischen Original- und Fourierfolge auf.

Signale können periodisch oder aperiodisch, kontinuierlich oder diskret sein. Je nach Signaltyp verwenden wir die Fourierreihe (periodisch/kontinuierlich), die Fouriertransformation (aperiodisch/kontinuierlich), die Zeitdiskrete Fouriertransformation (aperiodisch/diskret) oder die Diskrete Fouriertransformation (periodisch/diskret).

In allen Fällen sind die Aufbaufunktionen Sinusfunktionen bzw. -folgen, die sich von $-\infty$ bis ∞ erstrecken. Kontinuierliche, aperiodische Signale brauchen unendlich viele Frequenzkomponenten zur Darstellung der Zeitfunktion.

Es gibt keine Fourierdarstellung für diskrete Signale endlicher Länge. Man behilft sich dadurch, dass man das zeitbegrenzte Signal mit Nullen fortsetzt oder indem man die Zeitfunktion periodisch fortsetzt. Im Fall des Auffüllens mit Nullen gelangt man zur DTFT, die zur Darstellung der Zeitfunktion eine unendliche Zahl von Sinuskomponenten benötigt, was bedeutet, dass es unmöglich ist, die DTFT mit einem Rechneralgorithmus zu berechnen. Die einzige Möglichkeit die uns zur Verfügung steht, ist die DFT, weil Rechner nur mit diskreten Signalen endlicher Länge, die periodisch fortgesetzt sind, rechnen können.

9.13 Zusammenfassung

In unserer Umwelt treten in der Regel zeit- und amplituden-kontinuierliche Signale auf. Beispiele dafür sind Temperatur-, Druck-, akustische, ... Signale. Die Verarbeitung dieser Signale kann in »analoger« Form durch Messwertaufnehmer, Verstärker, analoge Filter, ... erfolgen.

Analoge Systeme beruhen in ihrer Wirkungsweise auf Massen und elektrischen Ladungen und den zwischen ihnen wirksamen physikalischen Kräften. Die Beschreibung bzw. mathematische Modellierung derartiger Systeme erfolgt mit den bekannten Gesetzen der Physik.

Beispiel 72 *Wir betrachten die Schallplatte als Vertreter eines analogen Systems. Die Schallschwingungen eines Mikrofons werden in spiralförmigen, zum Mittelpunkt der Schallplatte verlaufenden Rillen gespeichert. Bei der Wiedergabe wird die Nadel des Tonabnehmers analog zur Schallamplitude bewegt. Diese mechanische Bewegung wird in ein elektrisches Signal umgewandelt, verstärkt und über Lautsprecher wiedergegeben.*

Im Gegensatz zu analogen Systemen stehen digitale Systeme, die auf abstrakten Größen – Zahlen – und der Manipulation dieser Zahlen beruhen. (Die Manipulation der Zahlen wird letztlich durch physikalische (analoge) Prozesse bewerkstelligt, die digital abstrahiert werden.)

Analoge Systeme können groß im Vergleich zu digitalen Systemen sein (Vinylschallplatte \Leftrightarrow MP3-Spieler) und mehr Energie verbrauchen. Analoge Systeme können schwer modifiziert werden, um sie für andere Aufgaben einsetzbar zu machen. Analoge

Systeme sind auf Grund ihrer physikalischen Arbeitsweise anfällig für Fehler und Störungen. Dennoch ist der weit überwiegende Teil unserer Systeme analog. Wo es technologisch und wirtschaftlich möglich ist, werden digitale Systeme eingesetzt.

Digitale Systeme sind aus dem Bedürfnis entstanden, analoge Systeme mit Digitalrechnern zu untersuchen und erst etwa 50 Jahre alt. Bei der numerischen Berechnung von analogen Systemen müssen auf Grund der Begrenzungen der Digitalrechner (Speichermöglichkeit und Zahlendarstellung) die Signale im Zeit- und im Amplitudenbereich quantisiert werden. Die Quantisierung im Amplitudenbereich wird nach Qualitäts Gesichtspunkten durchgeführt, je feiner die Auflösung desto geringer ist die Abweichung vom Originalsignal bzw. das Quantisierungsrauschen. Die Quantisierung im Zeitbereich muss so gewählt werden, dass das abgetastete Signal am Ende einer Verarbeitungskette wieder in ein analoges Signal umgewandelt werden kann. Die Abtastfrequenz wird durch das Abtasttheorem festgelegt und muss größer als die doppelte maximale im Signal enthaltene Frequenzkomponente sein. Es lassen sich also nur bandbegrenzte Signale nach Abtastung rekonstruieren. Diese Bedingung trifft aber für alle technischen Signale zu.

Die mathematische Darstellung von Signalen führt in eine abstrakte Welt mit Eigenschaften, die nicht immer leicht verständlich sind.

- Nur Signale unendlicher Dauer haben eine eindeutige Frequenz (Spektrallinie)
- Zeitbegrenzte Signale haben ein unendlich breites (kontinuierliches) Spektrum.
- Die Frequenz eines zeitbegrenzten Signals lässt sich nicht exakt bestimmen und ist »unscharf«.
- Durch die Abtastung im Zeitbereich (Darstellung auf Digitalrechnern) wird das Spektrum des abgetasteten Signals periodisch.
- Durch die Abtastung im Frequenzbereich (Darstellung auf Digitalrechnern) wird die Zeitfunktion periodisch.
- Mit anderen Worten: Auf Digitalrechnern lassen sich weder zeitbegrenzte noch bandbegrenzte Signale darstellen, sondern nur periodische Signale. Begrenzung in Zeit und Frequenz entsteht dadurch, dass man nur eine Periode betrachtet.
- Die digitale (numerische) Darstellung von analogen Signalen ist nur mit begrenzter Genauigkeit möglich. Die Genauigkeit kann aber so hoch gewählt werden, dass keine merkbare Abweichung des digitalen vom analogen Signals feststellbar ist.

Kapitel 10

AD- und DA-Umsetzer

Inhalt

10.1	Abtast-Halte-Glieder	153
10.2	AD-Umsetzer	154
10.2.1	Parallelverfahren	154
10.2.2	Wägeverfahren	155
10.2.3	Zählverfahren	155
10.3	Digital-Analog-Umsetzer	156
10.4	Zusammenfassung	158

Analog-Digital-Umsetzer (Analog to Digital Converter, ADC) erzeugen aus einer elektrischen Spannung eine Zahl, die zu dieser Spannung proportional ist. Umgekehrt erzeugen Digital-Analog-Umsetzer (Digital to Analog Converter, DAC) aus einer Zahl eine Ausgangsspannung, die zur Zahl proportional ist.

ADCs und DACs werden heute als Komponenten geliefert oder sind Subsysteme von digitalen Signalprozessoren. Diese Komponenten sind monolithisch integrierte Bauelemente, die für ihre Funktion teilweise externe Bauelemente benötigen. Es gibt eine Vielzahl von ADCs und DACs, die für spezielle Aufgaben optimiert sind. Die folgende Darstellung gibt einen Überblick über die Funktionsweise und die wichtigsten heute erreichbaren technischen Daten.

AD- und DA-Umsetzer verwandeln elektrische Spannungen in Zahlen und Zahlen in elektrische Spannungen.

10.1 Abtast-Halte-Glieder

Wir werden noch sehen, dass die Abtastung von Signalen nach den Regeln des Abtasttheorems erfolgen muss. Zu jedem Abtastzeitpunkt muss eine analoge Probe des Signals genommen werden und in eine Zahl umgesetzt werden. Da sich Signale, bzw. die sie realisierenden elektrischen Spannungen, zeitlich ändern, muss der Abtastwert während der Wandlung gespeichert – gehalten – werden (Sample and Hold, SH). Die Speicherung muss analog erfolgen und die in Abbildung 10.1 gezeigte Prinzipschaltung erläutert die Funktionsweise.

Abtast-Halte-Glieder (Sample and Hold) **speichern** den **Momentenwert** einer zeitlich veränderlichen Spannung.

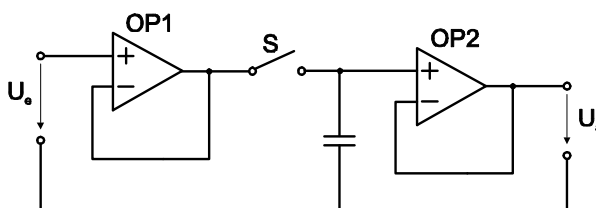


Abbildung 10.1: Abtast- und Halteschaltung

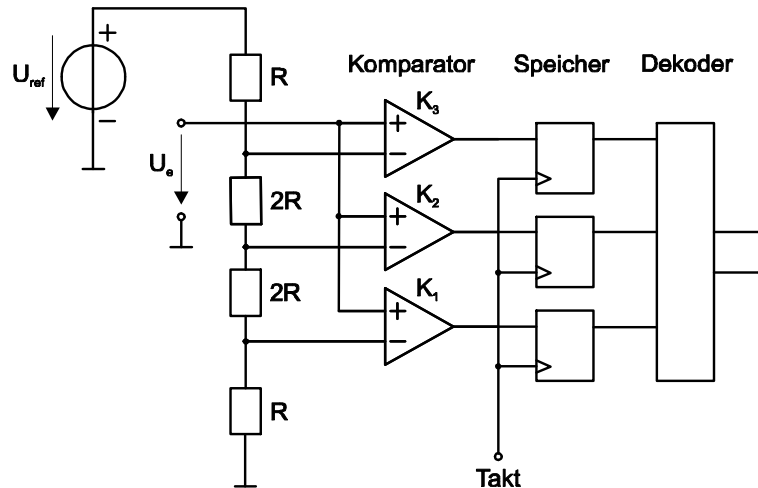


Abbildung 10.2: Blockdiagramm 2bit-Parallelwandler

Der $OP1$ und $OP2$ haben hochohmige Eingänge und belasten daher weder die Eingangsspannungsquelle noch den Speicher(Halte)-Kondensator. Bei geschlossenem Schalter S folgt die Spannung an C der Eingangsspannung. Wird der Schalter geöffnet, dann bleibt die Ladung – und damit die Spannung – am Kondensator konstant, da der nachfolgende $OP2$ den Kondensator wegen seines hochohmigen Eingangs nicht entlädt.¹ Der Schalter S wird als Halbleiterschalter realisiert, um eine hohe Schaltgeschwindigkeit zu erreichen. Die Aufladezeit des Kondensators wird durch das Produkt aus dem Schalterwiderstand und der Kapazität $t_E \sim RC$ bestimmt. Die Spannung am Kondensator steigt exponentiell nach der Beziehung $u_c(t) = U(1 - e^{-\frac{t}{RC}})$ an. Damit u_C auf 0.1 % genau ist, beträgt die Aufladezeit $t_E = 6.9 \times RC$. Für schnelle Einstellzeiten muss dieses Produkt möglichst klein sein. Da sich die Schalterwiderstände von Halbleiterschaltern im geschlossenen Zustand aus technologischen Gründen nicht beliebig klein machen lassen, muss für kleine Einstellzeiten C klein gemacht werden. Damit wird aber auch die gespeicherte Ladung klein und das nichtideale Verhalten von $OP2$ wirkt sich stärker aus. Mit heutiger Technologie sind Einstellzeiten auf 0.1% unter 10 ns zu erreichen.

10.2 AD-Umsetzer

Je nach erforderlicher Wandelzeit und Auflösung werden unterschiedliche Verfahren eingesetzt.

10.2.1 Parallelverfahren

Beim (schnellen aber aufwendigen) **Parallelverfahren** wird die Eingangsspannung **gleichzeitig** mit jeder möglichen Ausgangsspannung verglichen.

Mit Parallelverfahren können Umsetzfrequenzen von 10 – 1000 MHz bei Genauigkeiten von 4 – 10 bit erreicht werden. Bei Parallelverfahren wird die Eingangsspannung gleichzeitig mit mehreren abgestuften Referenzspannungen mit Hilfe von Komparatoren verglichen. Abbildung 10.2 zeigt den prinzipiellen Aufbau.

Die Spannungsteilerkette $R - 2R - 2R - R$ bildet die Referenzspannungen $U_{ref} \times (\frac{1}{6}, \frac{3}{6}, \frac{5}{6})$, die die Quantisierung der Eingangsspannung festlegen. Liegt z.B. eine Eingangsspannung $U_{ref}/3$ an, dann werden die Ausgänge der Komparatoren $K1 = 1, K2 = 0, K3 = 0$ in den Speicherelementen gespeichert und im nachfolgenden Dekoder in die gewünschte Zahlendarstellung umgewandelt. In unserem Beispiel ist die Auflösung des Wandlers 2 bit.

Der Aufwand des Parallelverfahrens ist sehr hoch, da man für eine N -stufige Auflösung ($\lceil \log_2(N) \rceil$ bit) $N - 1$ Komparatoren benötigt.

¹Praktische Operationsverstärker weichen vom idealen Verhalten ab, die vorgestellten Überlegungen erklären daher nur das prinzipielle Verhalten.

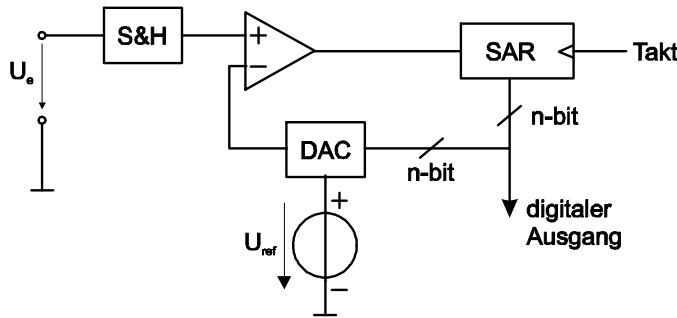


Abbildung 10.3: Blockdiagramm Wandler nach Zählverfahren

10.2.2 Wägeverfahren

Das Wägeverfahren kommt mit nur einem Komparator aus, benötigt allerdings eine längere Wandelzeit, da Binärstellen nacheinander ermittelt werden. Abbildung 10.3 zeigt den prinzipiellen Aufbau der Schaltung.

Die Funktionsweise der Schaltung ist wie folgt: Zum Messbeginn wird der Inhalt des Successive Approximation Registers auf Null gesetzt. Dann wird das höchste Bit im SAR auf Eins gesetzt. Über den DAC wird dieser Wert, der der halben Referenzspannung entspricht, an den Komparator gelegt und geprüft, ob die Eingangsspannung größer als $U_{ref}/2$ ist. Wenn ja, bleibt das höchste Bit gesetzt, wenn nein, wird es auf Null gesetzt. Im nächsten Schritt wird der Vorgang mit dem zweithöchsten Bit wiederholt. Der Vergleichsvorgang wird solange wiederholt, bis alle Bits durchlaufen sind.² Der Umwandlungsvorgang wird also schrittweise, für jedes Bit ein Umwandlungsschritt, gemacht und dauert daher im Vergleich zum Parallelverfahren länger, benötigt aber nur einen Komparator. SAR-ADCs erreichen Umwandlungsfrequenzen bis zu 300 MHz bei Wortbreiten bis zu 8 bit. Bei höherer Auflösung ändern sich die Wandlungszeiten etwa proportional mit der Auflösung.

Beim **Wägeverfahren** wird eine **binäre Suche** nach dem Wert der Eingangsspannung durchgeführt.

10.2.3 Zählverfahren

ADCs auf Basis des Zählverfahren verlangen den geringsten Schaltungsaufwand von allen drei Verfahren, benötigen aber Umwandlungszeiten die zwischen 0.001 – 1 sek liegen und sind daher nur für langsam veränderliche Signale geeignet, erreichen aber eine hohe Auflösung. Abbildung 10.4 zeigt den prinzipiellen Schaltungsaufbau.

Vor Beginn der Messung ist der Schalter $S1$ offen, der Schalter $S2$ geschlossen. Die Ausgangsspannung des OP, der als Integrator beschaltet ist, ist daher Null. Der Zähler ist auf Null gesetzt. Zu Beginn der Messung wird $S1$ an den Messeingang gelegt und die (positiv angenommene) Eingangsspannung wird integriert. Der Ausgang des OP wird negativ, der Komparatorausgang wird Eins und gibt über das &-Glied den Takteingang an den Zähler. Der Zähler läuft t_1 bis zum Überlauf. Wenn der Zähler überläuft wird $S1$ an die (negative) Referenzspannung gelegt, die integriert wird, wodurch die Ausgangsspannung U_i ansteigt. Die Integration der Referenzspannung ist nach t_2 beendet, wenn U_i bis auf Null angestiegen ist, der Komparator auf Null geht und damit über das &-Glied den Zähler stoppt. Der Zählerstand ist gleich der Zahl der Taktimpulse während t_2 und damit proportional der Eingangsspannung, da

$$-\frac{1}{RC}U_e t_1 - \frac{1}{RC}U_{ref} t_2 = 0 \quad (10.1)$$

Ist die Dauer eines Zählimpulses gleich T , dann beträgt die Zeitdauer $t_1 = (Z_{max} + 1)T$, wobei Z_{max} die höchste Zahl des Zählers ist. Die Zeitdauer $t_2 = ZT$, wobei Z die

Beim **Zählverfahren** wird gemessen, wie groß das **Integral der Eingangsspannung** über ein fixes Zeitintervall ist.

²Dieser Vorgang entspricht dem Wägen mit einer Gleichgewichtswaage und Meßgewichten, daher der Name Wägeverfahren. Algorithmisch gesehen entspricht dieses Verfahren der binären Suche.

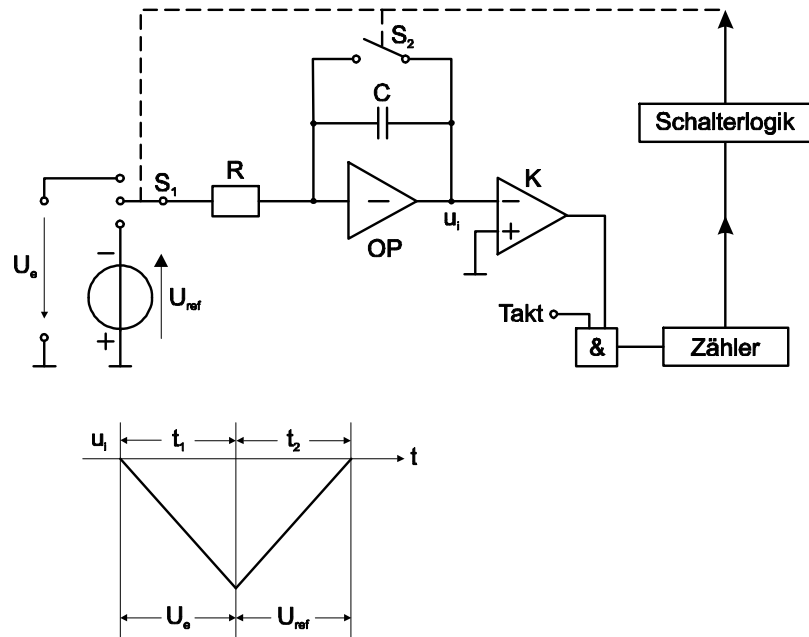


Abbildung 10.4: Blockdiagramm Dual-Slope-Wandler

Zahl der Zählimpulse während der Integrationsphase der Referenzspannung ist. Setzen wir ein dann wird

$$-\frac{1}{RC}U_e(Z_{\max} + 1)T - \frac{1}{RC}U_{ref}ZT = 0 \quad (10.2)$$

Wir sehen, dass sich RC und T aus dieser Gleichung kürzt und dass wir den folgenden Zusammenhang erhalten

$$Z = -\frac{U_e}{U_{ref}}(Z_{\max} + 1) \quad (10.3)$$

Wir sehen, dass in die Messung weder die Zeitkonstante RC , noch die Taktfrequenz $1/T$ eingehen. Es muss lediglich sichergestellt werden, dass die Taktfrequenz während der Messung konstant bleibt. Das lässt sich mit vergleichsweise einfachen Taktgeneratoren erreichen. Die Genauigkeit der Messung hängt allerdings von der Genauigkeit der Referenzspannung U_{ref} ab.

Mit diesem Verfahren lässt sich eine Auflösung von 18 bit bei einer Genauigkeit von $\pm 0.006\%$ und einer Wandlerfrequenz von einigen zehn Hz erreichen.

10.3 Digital-Analog-Umsetzer

Wie beim ADC gibt es auch beim DAC die drei Umsetzungsverfahren Parallel-, Wäge- und Zähl-, wobei alle drei Verfahren darauf beruhen, die Ausgangsspannung durch Spannungsteilung aus einer Referenzspannung zu bilden.

Abbildung 10.5 zeigt den prinzipiellen Aufbau des **Parallelverfahrens**.

Bei diesem Verfahren werden alle möglichen Ausgangsspannungen mit Hilfe eines Spannungsteilers erzeugt. Durch einen Decoder³ wird derjenige Schalter geschlossen, der zur gewünschten Ausgangsspannung führt. Der Aufbau von Spannungsteilern nach dem Parallelverfahren ist außerordentlich kritisch, da sehr viele Widerstände erforderlich sind, die exakt gleich groß sind. Darüber hinaus ist die Anzahl der Schalter gleich der Zahl der gewünschten Ausgangsspannungen Z_{\max} . Dieses Verfahren wird daher nur in Sonderfällen eingesetzt.

³Der Decoder kann einfach durch einen Demultiplexer realisiert werden.

Der **Parallel-DA-Wandler** hat **alle** möglichen **Ausgangsspannungen** »vorbereitet« und wählt immer die gewünschte aus.

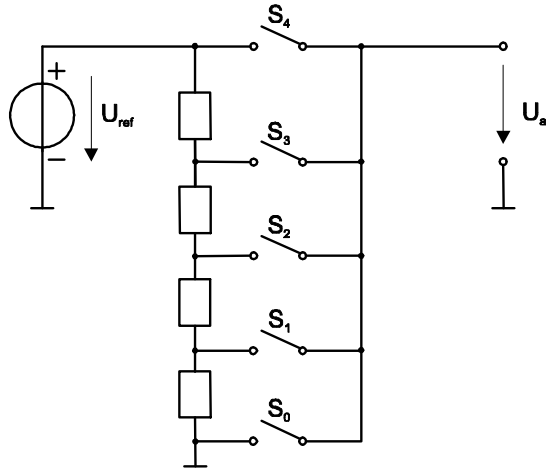


Abbildung 10.5: DA-Parallelwandler

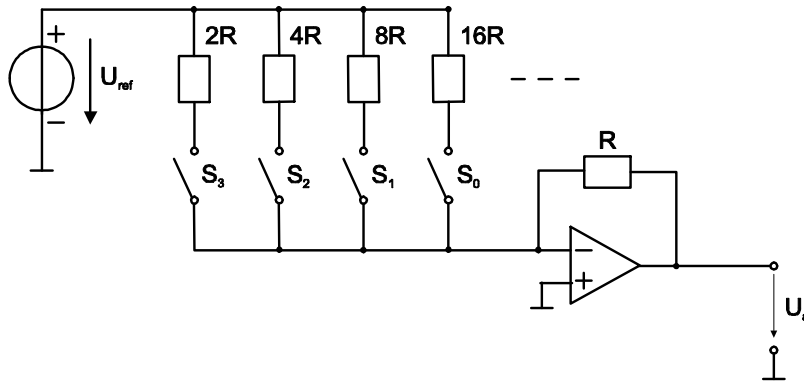


Abbildung 10.6: DA-Wandler nach dem Wägeverfahren

Das **Wägeverfahren** ist in Abbildung 10.6 dargestellt und kommt mit $ld(Z_{max})$ Schaltern aus.

Die Ausgangsspannung wird über binär gewichtete Widerstände aufsummiert, die Schalter S_i werden entsprechend der Binärzahl geschlossen. Für 0001 ist der Schalter S_0 geschlossen und wir erhalten durch den invertierenden Summierverstärker die Ausgangsspannung $U_a = -\frac{1}{16}U_{ref}$. Im allgemeinen Fall erhalten wir

$$U_a = - \left(\frac{1}{2} \begin{Bmatrix} 1 \\ 0 \end{Bmatrix} + \frac{1}{4} \begin{Bmatrix} 1 \\ 0 \end{Bmatrix} + \frac{1}{8} \begin{Bmatrix} 1 \\ 0 \end{Bmatrix} + \frac{1}{16} \begin{Bmatrix} 1 \\ 0 \end{Bmatrix} \right) U_{ref} \quad (10.4)$$

Das einfachste und langsamste Umsetzungsverfahren ist das **Zählverfahren**, wie in Abbildung 10.7 dargestellt.

Der Schalter S wird periodisch geöffnet und geschlossen. Das RC -Glied integriert

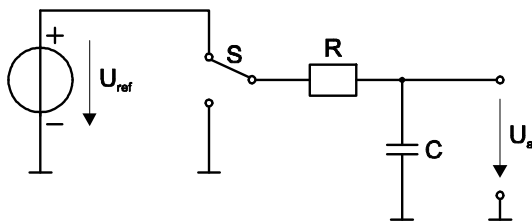


Abbildung 10.7: DA-Wandler nach dem Zählverfahren

Beim **Wägeverfahren** entsteht die Ausgangsspannung als **Summe** von **Teilspannungen**, deren Größe den einzelnen Binärstellen des Eingangsdatenworts entspricht.

Beim **Zählverfahren** wird ein **PWM-Signal** erzeugt und durch einen Kondensator geglättet.

(bildet den Mittelwert) der Eingangsspannung. Durch die Wahl des Tastverhältnisses des Schalters kann die gewünschte Ausgangsspannung eingestellt werden. Diese Schaltung kann nur für einfache und langsame Anwendungen eingesetzt werden.

In allen drei Umsetzungsvarianten ist die Stabilität der Referenzspannung bestimmend für die Genauigkeit der Ausgangsspannung. Durch nichtideales Verhalten der Schalter und OPVs können Abweichungen vom gewünschten monotonen Verhalten der DACs auftreten. Die dargestellten Widerstandsnetzwerke stellen Idealisierungen dar. In praktischen Schaltungen sind mit den Widerständen und Schaltern immer (unerwünschte) Kapazitäten verbunden, die das dynamische Verhalten von schnellen Umsetzern beeinflussen, sodass es beim Umschalten von einer Zahl zur nächsten zu unerwünschten Störimpulsen kommen kann.

10.4 Zusammenfassung

AD- und DA-Umsetzer bilden die Schnittstelle zwischen analoger und digitaler Welt. Die Umsetzverfahren werden nach der erforderlichen Umwandlungsgeschwindigkeit und Genauigkeit ausgewählt. Die schnellste Umsetzung ist mit aufwendigen und daher teuren Parallelverfahren möglich, mit denen man Umsetzungen bis zu einem GHz und Genauigkeiten bis zu 10 bit erreichen kann. Die parallelen Verfahren haben auch den höchsten Leistungsbedarf von allen Umwandlungsverfahren.

Wenn die Umwandlungsgeschwindigkeit geringer sein kann, werden Wägeverfahren eingesetzt, die für Wandlungsfrequenzen von 10 kHz bis mehreren 10 MHz, bei Genauigkeiten von 8 – 16 bit verwendet werden.

Die langsamsten aber genauesten Verfahren sind die Zählverfahren, mit denen Genauigkeiten bis zu 20 bit erzielt werden können und bei denen die Wandlerfrequenz zwischen 1 Hz – 1 kHz liegt.

ADCs und DACs werden überwiegend als monolithisch integrierte Bausteine angeboten, die teilweise mit externen Komponenten beschaltet werden müssen. Diskrete Schaltungen werden nur in Ausnahmefällen aufgebaut.

Teil V

Digitale Signalverarbeitung

Kapitel 11

FIR-Filter

Inhalt

11.1 Allgemeine FIR-Filter	162
11.2 Die Impulsantwort	163
11.2.1 Die Faltungssumme	164
11.3 Frequenzgang von FIR-Filtern	166
11.3.1 Frequenzgang und Überlagerungssatz	170
11.3.2 Einschaltvorgänge	170
11.4 Blockdiagramm von FIR-Filtern	171
11.4.1 Realisierung von Filtern	173
11.5 Kaskadieren von Filtern	174
11.5.1 Vom idealen zum realen Filter	175
11.6 Zusammenfassung	179

Ein Filter ist ein System, das die Eigenschaften eines Signals verändert. Häufig wird der Begriff Filter und System synonym verwendet.

Wir beginnen die Beschreibung von diskreten (digitalen) Systemen mit den Finite Impulse Response Filtern, in Deutsch Nicht-rekursive Filter genannt.

Unter einem digitalen System verstehen wir eine Anordnung, die mit einem Transformationsprozess eine Eingangs-Folge in eine Ausgangs-Folge umwandelt. Abbildung 11.1 zeigt ein Blockdiagramm eines zeitdiskreten Systems.



Abbildung 11.1: Zeitdiskretes System

Als Beispiel für den Transformationsprozess eines Filters nehmen wir die Bildung des gleitenden Mittelwerts über drei Punkte einer Folge: In jedem Zeitpunkt wird der Ausgangswert aus drei zeitlich zurückliegenden¹ Punkten der Eingangsfolge gebildet.

n	-2	-1	0	1	2	3	4	5	6	7	8
$x[n]$ =			0	1	3	2	-1	2	0	0	0
$y[n]$			0	$\frac{1}{3}$	$\frac{4}{3}$	2	$\frac{4}{3}$	1	$\frac{1}{3}$	$\frac{2}{3}$	0

Der Ausgangswert zum Zeitpunkt $n = 0$ wird durch Mittelung der Werte $x[0, -1, -2] = [0, 0, 0]$ gebildet. Die Eingangsfolge hat zu den Zeitpunkten $x[-1, -2]$ keine definierten Werte, wir nehmen an, dass die Folge $x[n]$ zum Zeitpunkt $n[0]$ beginnt und setzen die

Digitale Filter setzen **Eingangszahlenfolgen** in **Ausgangszahlenfolgen** um.

Ein **Beispiel** eines FIR-Filterns ist das **Mittelwertfilter**.

¹Wir können bei kausalen Systemen nur Werte, die in der Vergangenheit liegen $[n - 1], [n - 2], \dots$ verwenden. Gehen auch Werte die in der Zukunft liegen $[n + 1], [n + 2], \dots$ in die Berechnung des Ausgangswerts ein, dann spricht man von nicht-kausalen Systemen. Echtzeitsysteme sind immer kausal.

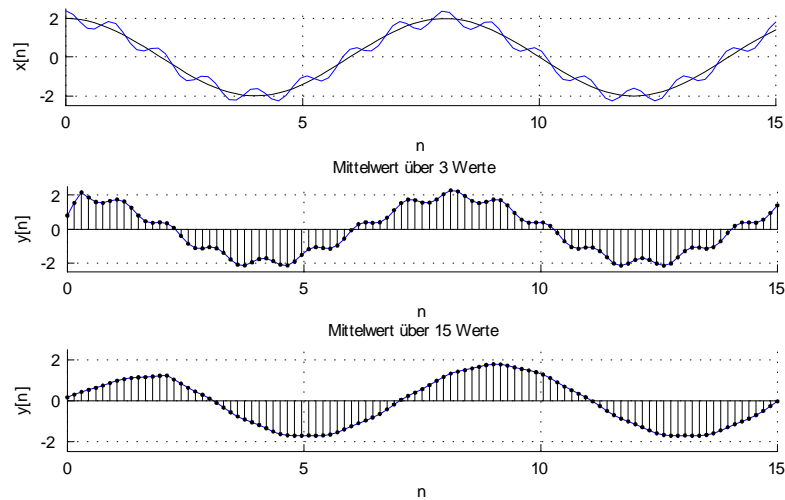


Abbildung 11.2: Mittelwertfilter

Werte für $n < 0$ auf Null. Der Ausgangswert zum Zeitpunkt $n = 3$ wird aus den Werten $[2, 3, 1] \mapsto \frac{6}{3} = 2$ gebildet, u.s.w.

Wir können den Transformationsprozess von Eingangsfolge in Ausgangsfolge durch den folgenden Zusammenhang beschreiben

$$y[n] = \frac{1}{3} (x[n] + x[n-1] + x[n-2]) \quad (11.1)$$

Um die Filterwirkung eines Mittelwertfilters besser verstehen zu können, erzeugen wir ein Signal, das aus einer Komponente niedriger und aus einer Komponente hoher Frequenz besteht.

$$x[n] = 2 \cos(\hat{\omega}n) + \frac{1}{5} \cos(8\hat{\omega}n) \quad (11.2)$$

Die Eingangsfolge $x[n]$ filtern wir mit zwei Filtern, einmal wird der Mittelwert über 3 Werte gebildet, das zweite Mal über 15 Werte..

Abbildung 11.2 stellt die Eingangsfolge und die beiden Ausgangsfolgen graphisch dar.

Aus Abbildung 11.2 können wir erkennen, dass am Anfang und am Ende des Ausgangssignals das Filter »gefüllt« bzw. »geleert« wird, diese Abschnitte nennt man Einschwing- und Abklingphase des Filters.

Wir sehen, dass die höhere Frequenzkomponente geglättet (unterdrückt) wird. Wir können aus den Abbildungen schließen, dass das Ausmaß der Unterdrückung höherer Frequenzen davon abhängt, über wieviele Punkte der Mittelwert gebildet wird.

Ein Filterverhalten, das niedrige (tiefe) Frequenzen ungedämpft überträgt, während es hohe Frequenzen dämpft nennt man *Tiefpass-Filter*.

Das **Mittelwertfilter** hat ein **Tiefpassverhalten**.

11.1 Allgemeine FIR-Filter

In unserem Beispiel haben wir den Transformationsprozess zwischen Eingangs- und Ausgangssignal über die Mittelwertbildung festgelegt. Für Mittelwertfilter können wir allgemein schreiben:

$$y[n] = \frac{1}{N} (x[n] + x[n-1] + \dots + x[n - (N-1)]) = \frac{1}{N} \sum_{k=0}^{N-1} x[n-k] \quad (11.3)$$

Beim Mittelwertfilter haben wir für die Berechnung des Ausgangswertes alle Werte der Eingangfolge mit dem gleichen Wert $1/N$ multipliziert. Lassen wir bei der Berechnung der Ausgangfolge für jeden Wert der Eingangfolge einen eigenen Koeffizienten zu, dann erhalten wir die Filtergleichung für allgemeine FIR-Filter:

$$y[n] = \sum_{k=0}^M b_k x[n-k] \quad (11.4)$$

Die Größe M nennt man die Ordnung des Filters. Die Zahl der Werte, die in die Berechnung der Ausgangfolge eingehen ist also immer $M+1$. Das Mittelwertfilter ist ein Tiefpass-Filter, andere Filter-Charakteristika (Durchlassen von tiefen oder hohen Frequenzen, selektives Unterdrücken von Frequenzen oder selektives Durchlassen von Frequenzen) können durch die Wahl geeigneter Koeffizienten b_k erreicht werden.

Aus Gleichung (11.4) sieht man, das FIR-Filter lineare, zeitinvariante Systeme sind, weil bei der Transformation des Eingangssignals nur lineare Operationen (Multiplikation mit einer Konstante b_k und Addition) verwendet werden. Die Linearität kann auch rechnerisch leicht nachgewiesen werden:

$$\sum_{k=0}^M b_k (u \cdot x_1[n-k] + v \cdot x_2[n-k]) \quad (11.5)$$

$$= \sum_{k=0}^M b_k (u \cdot x_1[n-k]) + \sum_{k=0}^M b_k (v \cdot x_2[n-k]) \quad (11.6)$$

$$= u \cdot \sum_{k=0}^M b_k x_1[n-k] + v \cdot \sum_{k=0}^M b_k x_2[n-k] \quad (11.7)$$

Da bei jedem Berechnungsschritt jeder der Eingangswerte immer um die selbe Zeitkonstante² verschoben wird und die Koeffizienten b_k sich mit der Zeit nicht ändern, sind diese Systeme auch zeitinvariant.

11.2 Die Impulsantwort

Systeme können eine unendliche Anzahl von Signalen transformieren, wir können aber das Verhalten eines Systems nicht für jedes Eingangssignal berechnen und suchen daher einen Weg, wie wir die Eigenschaften von Systemen einfacher darstellen können.

- Wir wissen bereits, dass komplizierte Signale aus sinusförmigen Komponenten oder aus Einheitsimpulsen zusammengesetzt werden können. Wir können uns also jedes beliebige Eingangssignal aus sinusförmigen Komponenten oder aus Einheitsimpulsen »aufgebaut« vorstellen.
- Außerdem wissen wir, dass bei linearen zeitinvarianten Systemen der Überlagerungssatz gilt. Die Systemantwort auf ein beliebiges Eingangssignal lässt sich durch Überlagerung der Systemantworten auf die Aufbausignale ermitteln.
- Für Untersuchung der Eigenschaften von Systemen ist es daher ausreichend, wenn wir die Antwort des Systems auf die aufbauenden Funktionen kennen.

Wir befassen uns zunächst mit der Impulsantwort und gehen von Abbildung 11.3 aus.

Die Einheitsimpuls-Folge, kurz Einheitsimpuls genannt, ist für diskrete Systeme definiert

$$\delta_0[n] = \begin{cases} 1 & n = 0 \\ 0 & n \neq 0 \end{cases} \quad (11.8)$$

²Diskrete Systeme kennen streng genommen keinen Zeitbegriff, weil sie lediglich Folgen von Zahlen verarbeiten. Trotzdem wurde hier wegen der engen Beziehung zwischen Zahlenfolgen und abgetasteten Zeitsignalen der (dimensionslose!) Folgeparameter n als Zeit bezeichnet.

Beim FIR-Filter M -ter Ordnung entspricht jeder **Ausgangswert** einer **Linearkombination** der letzten $M+1$ **Eingangswerte**.

FIR-Filter sind **lineare zeitinvariante** Systeme.

Da es sich bei **FIR-Filtern** um **LTI-Systeme** handelt, gilt auch hier der **Überlagerungssatz**.

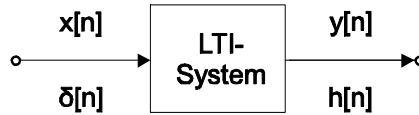


Abbildung 11.3: LTI-System

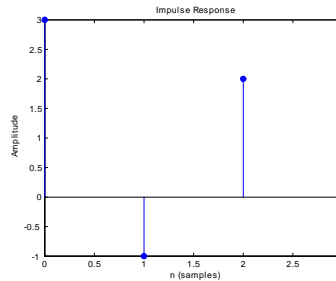


Abbildung 11.4: Impulsantwort FIR-Filter

Zur Berechnung der Impulsantwort setzen wir als Eingangssignal $x[n] = \delta_0[n]$ und erhalten daraus das Ausgangssignal. Da die Impulsantwort einen besonderen Stellenwert bei der Untersuchung von LTI-Systemen hat, gibt man ihr die eigene Abkürzung $h[n]$.

$$y[n] = h[n] = \sum_{k=0}^M b_k \delta_0[n - k] = \begin{cases} b_n & n = 0, 1, \dots, M \\ 0 & \text{sonst} \end{cases} \quad (11.9)$$

Die **Impulsantwort** eines **FIR-Filters** setzt sich aus dessen **Filterkoeffizienten** zusammen.

Das Anlegen eines Einheitspulses an ein FIR-Filter bewirkt, dass am Ausgang die Folge der Filterkoeffizienten b_k erscheint. Da die Zahl der Koeffizienten endlich ist, ist auch die Dauer der Impulsantwort endlich, daher der Name **Finite Impulse-Response** Filter.

Unter Verwendung der Impulsantwort kann man Gleichung (11.4) auch in anderer Form anschreiben und erhält

$$y[n] = \sum_{k=0}^M h[k]x[n - k] \quad (11.10)$$

Gleichung (11.10) nennt man **Faltungssumme**, der Ausgang eines LTI-Systems wird berechnet, indem man die Eingangsfolge mit der Impulsantwort »faltet«.

Wie jedes LTI-System, wird auch ein FIR-Filter **durch** die **Impulsantwort vollständig beschrieben**.

Bemerkung 73 Kennt man die Impulsantwort eines Systems, dann kann man die Systemantwort für beliebige Eingangssignale berechnen. Mit der Kenntnis der Impulsantwort wissen wir alles über das Übertragungsverhalten eines Systems.

Beispiel 74 Matlab stellt die Funktion `impz(B,A)` zur Berechnung der Impulsantwort eines Filters zu Verfügung. Für ein FIR-Filter mit den Koeffizienten $B = [3, -1, 2, 1]$ erhalten wir die Impulsantwort in Abbildung 11.4.

$H = \text{impz}(B, 1)$, und H ist der Spaltenvektor der Impulsantwort. Wird kein Ergebnis zugewiesen, dann wird die Impulsantwort auf den Bildschirm ausgegeben. Für FIR-Filter ist $A = [1]$ und die Impulsantwort gleich den Koeffizienten b_k der Filtergleichung. Die Funktion `impz` ist daher besser für IIR-Filter geeignet, die wir später kennen lernen werden.

11.2.1 Die Faltungssumme

Das **Ausgangssignal** entsteht durch **Faltung** des **Eingangssignals** mit der **Impulsantwort**.

Die Faltung ist eine Rechenvorschrift, die aus der Impulsantwort und dem Eingangssignal das Ausgangssignal erzeugt. Die Faltung ist die wichtigste Operation bei der Untersuchung

des Systemverhaltens im Zeitbereich, wir werden die Faltung daher genau untersuchen. Anstelle der Schreibweise

$$y[n] = \sum_{k=0}^M h[k]x[n - k]$$

ist auch die Kurzschreibweise

$$y[n] = x[n] * h[n] \tag{11.11}$$

gebräuchlich.

Wir können die Operation der Faltung aus zwei Blickwinkeln betrachten:

1. Wir untersuchen, wie jeder Wert des *Eingangssignals* zu allen Punkten des Ausgangssignals beiträgt.
2. Wir untersuchen, wie jeder Wert des *Ausgangssignals* aus allen Punkten des Eingangssignals gebildet wird.

Betrachtungsweise 1 hilft die Faltung aus Konzeptsicht zu verstehen, Betrachtungsweise 2 beschreibt lediglich den Rechenvorgang bei der Faltung.

Wirkung des Eingangssignals

Das Eingangssignal denken wir uns durch die Summe von gewichteten, zeitversetzten Impulsfunktionen gebildet.

Jeder Abtastwert des Eingangssignals erzeugt seine eigene – mit der Amplitude des Eingangssignals gewichtete – Impulsantwort. Die Antwort des Systems ist die Summe aller gewichteten und zeitversetzten Impulsantworten. Wir zeigen den Zusammenhang an folgendem Beispiel:

$$x[n] = [1, 2, 3, -2, 2, 1] \quad h[n] = [3, -1, 2, 1] \tag{11.12}$$

$x[0] = 1$	$y[0] = 1 \cdot h[n - 0]$	3	-1	2	1	0	0	0	0	0	0
$x[1] = 2$	$y[1] = 2 \cdot h[n - 1]$	0	6	-2	4	2	0	0	0	0	0
$x[2] = 3$	$y[2] = 3 \cdot h[n - 2]$	0	0	9	-3	6	3	0	0	0	0
$x[3] = -2$	$y[3] = -2 \cdot h[n - 3]$	0	0	0	-6	2	-4	-2	0	0	0
$x[4] = 2$	$y[4] = 2 \cdot h[n - 4]$	0	0	0	0	6	-2	4	2	0	0
$x[5] = 1$	$y[5] = 1 \cdot h[n - 5]$	0	0	0	0	0	3	-1	2	1	0
	$y[n] = \sum x[n]h[n - k]$	3	5	9	-4	16	0	1	4	1	0

Zusammensetzung des Ausgangssignals

Aus Gleichung (11.10) sehen wir, dass sich jeder Wert des Ausgangssignals aus gewichteten und zeitversetzten Eingangssignalen zusammensetzt. Das Eingangssignal wird mit dem entsprechenden Wert der Impulsantwort gewichtet.

$h[0] = 3$	$y[0] = 3 \cdot x[n - 0]$	3	6	9	-6	6	3	0	0	0	0
$h[1] = -1$	$y[1] = -1 \cdot x[n - 1]$	0	-1	-2	-3	2	-2	-1	0	0	0
$h[2] = 2$	$y[3] = 2 \cdot x[n - 2]$	0	0	2	4	6	-4	4	2	0	0
$h[3] = 1$	$y[4] = 1 \cdot x[n - 3]$	0	0	0	1	2	3	-2	2	1	0
	$y[n] = \sum h[k]x[n - k]$	3	5	9	-4	16	0	1	4	1	0

Wie man sieht liefern beide Betrachtensweisen identische Ergebnisse, da die Faltung kommutativ ist.

Die Faltung ist kommutativ.

$$x[n] * h[n] = h[n] * x[n] \tag{11.13}$$

Aus mathematischer Sicht macht es keinen Unterschied welches Signal das Eingangssignal und welches die Impulsantwort ist. Aus physikalischer Sicht ist nur die Überlagerung der Systemantworten auf die Impulse von denen sich das Eingangssignal zusammensetzt

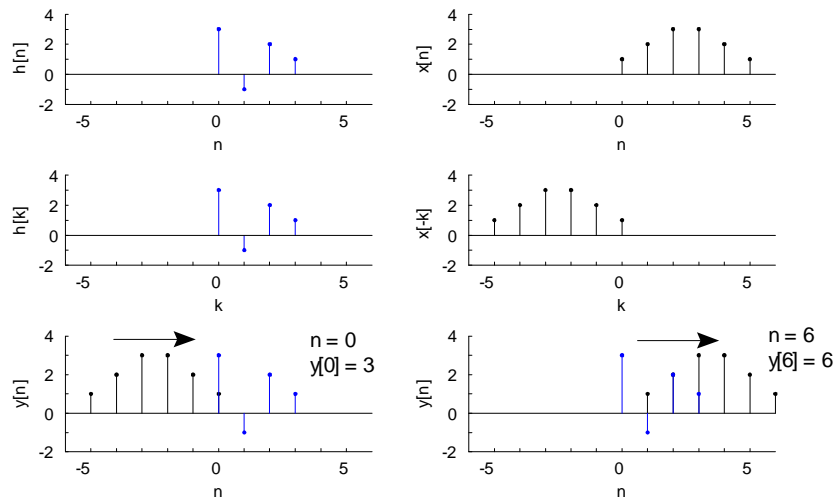


Abbildung 11.5: Faltung

sinnvoll, die Zusammensetzung des Ausgangssignals aus gewichteten, zeitversetzten Eingangssignalen liefert lediglich eine Vorschrift zur Berechnung der Faltungssumme.

Abbildung 11.5 stellt den Faltungsvorgang grafisch dar. Insbesondere wird deutlich, woher der Begriff Faltung stammt: $x[-k]$ (d) entsteht durch Umklappen (Falten) von $x[k]$. Das Ergebnis der Faltung ist das Ausgangssignal $y[n]$, wobei n als Parameter in die Zeitsignale $x[]$ und $h[]$ eingeht. Um eine eindeutige Schreibweise zu ermöglichen, muss daher die Variable k eingeführt werden und Eingangsfolge und Impulsantwort werden zu $x[k]$ und $h[k]$ (Bild Mitte links und Mitte rechts).

$h[k]$ wird mit $x[-k]$ multipliziert, die Produkte werden summiert und ergeben $y[0]$ (links unten). $x[-k]$ wird um einen Schritt nach rechts verschoben ($x[1-k]$), mit $h[k]$ multipliziert und die Summe der Produkte liefert $y[1]$. $x[-k]$ wird um zwei Schritte verschoben, die Rechenvorschrift liefert $y[2]$, u.s.w. (Teilbild rechts unten) zeigt die Operation nach sechsmaliger Verschiebung.

Beispiel 75 Die Faltungssumme kann mit Matlab durch die Funktion `conv` berechnet werden.

```
 $x = [1, 2, 3, -2, 2, 1]$       $h = [3, -1, 2, 1]$ 
```

```
conv(x,h) => 3 5 9 -4 16 0 1 4 1
```

Alternativ könnte die Funktion `filter` verwendet werden. $y = \text{filter}(B,A,x)$ berechnet für ein Filter mit den Koeffizienten B und der Eingangsfolge $x[n]$, die Ausgangsfolge $y[n]$. Für FIR-Filter ist $A = [1]$ und wir erhalten

```
 $B = [3 -1 2 1]$ ; filter(B,1,[x,0 0 0 0]) => 3 5 9 -4 16 0 1 4 1 0
```

11.3 Frequenzgang von FIR-Filtern

Für die Darstellung eines Filters im Frequenzbereich wird die komplexe Exponentialfunktion verwendet.

Einheitsimpuls und komplexe Exponentialfunktion sind die wichtigsten Testsignale für die Untersuchung von Systemeigenschaften. Die Verwendung des Einheitsimpuls führt zur Faltungssumme und zur Darstellung von Systemeigenschaften im Zeitbereich. Die komplexe Exponentialfunktion führt zum Frequenzgang und zur Darstellung im Frequenzbereich. Sowohl die Darstellung im Zeitbereich, als auch die Darstellung im Frequenzbereich beschreiben ein System vollständig, sind aber unterschiedliche Betrachtensweisen der Systemeigenschaften. Je nach Aufgabenstellung ist die eine oder die andere Darstellung besser geeignet.

Zur Berechnung des Frequenzgangs gehen wir von der allgemeinen Gleichung der FIR-Filter aus

$$y[n] = \sum_{k=0}^M b_k x[n-k] = \sum_{k=0}^M h[k] x[n-k] \quad (11.14)$$

Das Eingangssignal ist die komplexe Exponentialfolge

$$x[n] = A e^{j(\hat{\omega}n + \varphi)} = A e^{j\varphi} e^{j\hat{\omega}n} \quad -\infty < n < \infty \quad (11.15)$$

Die Größe $\hat{\omega}$ nennt man die normierte Kreisfrequenz³.

$$\hat{\omega} = \omega T_s \quad T_s \dots \text{Abtastperiode} \quad (11.16)$$

Wir setzen (11.15) in die Gleichung des FIR-Filter ein und erhalten

$$y[n] = \sum_{k=0}^M b_k A e^{j\varphi} e^{j\hat{\omega}(n-k)} \quad (11.17)$$

$$= \left(\sum_{k=0}^M b_k e^{-j\hat{\omega}k} \right) A e^{j\varphi} e^{j\hat{\omega}n} \quad (11.18)$$

Wir sehen, dass $y[n]$ aus dem Ausdruck in Klammern, multipliziert mit der Eingangsfolge $x[n]$ besteht. Den Ausdruck in Klammern, die Funktion

$$H(\hat{\omega}) = \sum_{k=0}^M b_k e^{-j\hat{\omega}k} = \sum_{k=0}^M h[k] e^{-j\hat{\omega}k} \quad (11.19)$$

nennt man den Frequenzgang des Systems. $H(\hat{\omega})$ ⁴ kann für jede Eingangsfrequenz $\hat{\omega}$ berechnet werden und ist im Allgemeinen komplexwertig. Gleichung (11.19) ist nichts anderes als die DTFT von $h[k]$. Wir sehen, dass (analog zu kontinuierlichen Systemen) der Frequenzgang eines FIR-Filters die (zeitdiskrete) Fouriertransformierte seiner Impulsantwort ist.

Als Beispiel berechnen wir den Frequenzgang eines Filters mit den Koeffizienten $b_k = [1, 2, 1]$

$$H(\hat{\omega}) = 1e^{-j\hat{\omega}0} + 2e^{-j\hat{\omega}1} + 1e^{-j\hat{\omega}2} = 1 + 2e^{-j\hat{\omega}} + 1e^{-j2\hat{\omega}}$$

$H(\hat{\omega})$ könnte man durch einsetzen von Werten für $\hat{\omega}$ ausrechnen, wobei $H(\hat{\omega})$ für alle $\hat{\omega}$, mit Ausnahme von $\hat{\omega} = 0$, komplexe Werte annimmt. Real-/Imaginärteil oder Betrag/Phase von $H(\hat{\omega})$ kann man über $\hat{\omega}$ auftragen und erhält damit den Frequenzgang des Filters. Wir finden für unser Beispiel auch eine analytische Lösung durch Herausheben von $e^{-j\hat{\omega}}$

$$H(\hat{\omega}) = e^{-j\hat{\omega}} (e^{-j\hat{\omega}} + e^{j\hat{\omega}} + 2)$$

Verwenden wir die Euler'sche Beziehung

$$\cos \alpha = \frac{e^{j\alpha} + e^{-j\alpha}}{2}$$

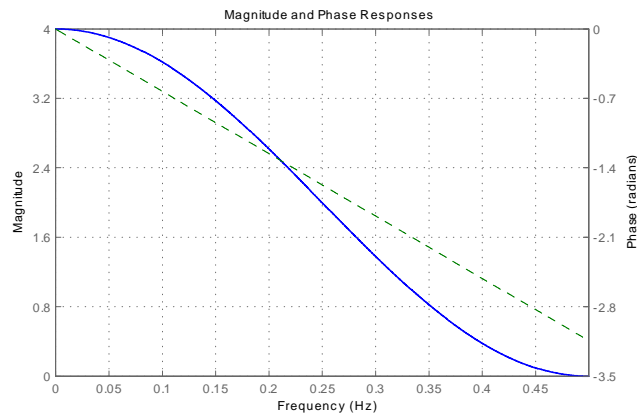
so wird daraus

$$H(\hat{\omega}) = (2 + 2 \cos \hat{\omega}) e^{-j\hat{\omega}} \quad (11.20)$$

³Durch die Abtastung geht, wie wir gesehen haben, die Zeitinformation verloren. Aus der Kreisfrequenz ω [1/sek] wird dadurch die dimensionslose Größe $\hat{\omega}$. In diskreten Systemen haben wir es immer mit der Größe $\hat{\omega}$ zu tun. Wegen der einfacheren Schreibweise wird aber in der Literatur häufig für $\hat{\omega}$ nur ω geschrieben.

⁴Der Frequenzgang ist eine kontinuierliche Funktion, die für beliebige $\hat{\omega}$ definiert ist, daher die Schreibweise in runden Klammern.

Der **Frequenzgang** eines FIR-Filters ist die (zeitdiskrete) **Fouriertransformierte seiner Impulsantwort**.

Abbildung 11.6: Frequenz- und Phasengang $b_k = [1, 2, 1]$

Die komplexwertige Funktion (11.20) lässt sich leicht in Betrag und Phase trennen und wir erhalten

$$|H(\hat{\omega})| = (2 + 2 \cos \hat{\omega}) \quad \angle H(\hat{\omega}) = -\hat{\omega} \quad (11.21)$$

Abbildung 11.6 zeigt den Frequenzgang in Betrag und Phase (strichliert).

Wir können erkennen, dass die Amplituden niedriger Frequenzen weniger, während Amplituden höherer Frequenzen stärker abgeschwächt werden. Das FIR-Filter mit den Koeffizienten $b_k = [1, 2, 1]$ verhält sich wie ein Tiefpass-Filter. Die Phase ist eine lineare Funktion der Frequenz.

Wie wir sehen, hängt der Frequenzgang des Filters

$$H(\hat{\omega}) = \sum_{k=0}^M b_k e^{-j\hat{\omega}k} \quad (11.22)$$

nur von den Koeffizienten b_k ab und kann für beliebige Eingangsfrequenzen aus der Beziehung (11.22) berechnet werden.

Aus der Beziehung

$$y[n] = \left(\sum_{k=0}^M b_k e^{-j\hat{\omega}k} \right) A e^{j\varphi} e^{j\hat{\omega}n} \quad (11.23)$$

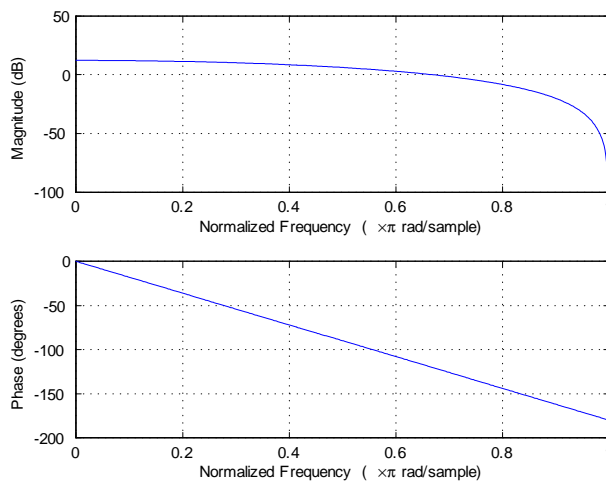
$$= H(\hat{\omega}) \cdot \underbrace{A e^{j\varphi}}_{\text{komplexe Eingangsamplitude}} \cdot e^{j\hat{\omega}n} \quad (11.24)$$

$$= \underbrace{\{H(\hat{\omega}) A e^{j\varphi}\}}_{\text{komplexe Ausgangsamplitude}} \cdot e^{j\hat{\omega}n} \quad (11.25)$$

können wir erkennen, dass die komplexe Exponentialfunktion beim Durchgang durch das Filter erhalten bleibt. Für das Eingangssignal $A e^{j\varphi} e^{j\hat{\omega}n}$ ändert sich lediglich die (komplexe) Amplitude von $A e^{j\varphi} \rightarrow H(\hat{\omega}) A e^{j\varphi}$. Aus der komplexen Zahl $A e^{j\varphi}$ wird durch Multiplikation mit dem komplexen Wert $H(\hat{\omega})$ die Amplitude der komplexen Ausgangsfunktion.

Bemerkung 76 Die komplexe Exponentialfunktion ist die einzige Signalform, die das lineare System ohne Änderung der Kurvenform durchläuft. Es ändern sich lediglich Amplitude und Phase, die Eigenschaft, sinusförmige Funktion zu sein, bleibt erhalten! Der Ausdruck $y[n] = H(\hat{\omega})x[n]$ gilt **nur** für die Eingangsfunktion $x[n] = A e^{j\varphi} e^{j\hat{\omega}n}$ und hat für keine andere Eingangsfunktion eine Bedeutung!

Die **komplexe Exponentialfunktion verändert** beim Durchlaufen eines FIR-Filters **nur** ihre **Amplitude**. Die Ausgangsamplitude erhält man **durch Multiplikation** der Eingangsamplitude **mit dem Frequenzgang**.

Abbildung 11.7: Amplituden- und Phasengang von $B = [1, 2, 1]$

Der komplexwertige Frequenzgang wird grafisch entweder in Real- und Imaginärteil

$$H(\hat{\omega}) = \operatorname{Re}\{H(\hat{\omega})\} + j \operatorname{Im}\{H(\hat{\omega})\} \quad (11.26)$$

oder in Betrag und Phase

$$H(\hat{\omega}) = |H(\hat{\omega})| e^{j\angle H(\hat{\omega})} \quad (11.27)$$

dargestellt. Für die Berechnung des Übertragungsverhaltens von Filtern ist die Darstellung in Betrag und Winkel die geeignetere, da man sich in der Regel dafür interessiert, welchen Einfluss das Filter auf die Amplitude (Verstärkung oder Abschwächung) der Frequenzkomponenten hat. Die Wirkung auf die Phase ist in der Regel weniger wichtig.

Bemerkung 77 Nicht jedes System das auf eine sinusförmige Eingangsschwingung eine sinusförmige Ausgangsschwingung erzeugt muss linear sein. Ein gutes Beispiel dafür ist der Phasenregelkreis. Bei dieser Schaltung wird ein freilaufender Oszillator auf eine externe Sinusschwingung synchronisiert. Der Phasenregelkreis ist aber nichtlinear und der Überlagerungssatz gilt nicht, wie man sich leicht durch Anlegen von zwei Sinusschwingungen an den Phasenregelkreis überlegen kann: Der interne Oszillator kann nur auf eine der beiden Sinuskomponenten synchronisiert werden.

Beispiel 78 Die Berechnung von Amplituden- und Phasengang kann einfach mit der Matlabfunktion `freqz` durchgeführt werden. `freqz(B,A)` berechnet den (komplexen) Wert der Systemfunktion $H(z) = \frac{B(z)}{A(z)}$. Für FIR-Filter ist $A(z) = 1$ und wir erhalten für das Beispiel

`B = [1 2 1]; H = freqz(B,1);`

H ist ein Spaltenvektor des (komplexen) Frequenzgangs. Wird `freqz(B,1)` keinem Ergebnis zugewiesen, dann wird der Amplitudengang (in logarithmischer Darstellung in dB) und der Phasengang auf dem Bildschirm angezeigt, wie Abbildung 11.7 zeigt.

Sinusschwingungen sind ein gutes Werkzeug zur praktischen Überprüfung der Linearität eines Systems, indem man an den Eingang ein Sinussignal anlegt und den Ausgang mit einem Oszilloskop beobachtet.

- Als erstes wird beobachtet, ob das Ausgangssignal sinusförmig ist.
- Dann wird die Amplitude des Eingangssignal verändert und beobachtet, ob das Signal sinusförmig bleibt.

In den meisten Anwendungen ist der **Amplitudengang** eines Filters von **größerer Bedeutung** als der **Phasengang**.

Ein Übertragungsverhalten, das die Form von Sinusschwingungen behält, ist ein Indiz, aber kein Beweis für die Linearität eines Systems.

- Dann wird die Frequenz des Eingangssignals verändert und das Ausgangssignal beobachtet. Amplitude und Phase des Ausgangssignals werden sich ändern, bei bestimmten Frequenzen möglicherweise sogar Null werden, die Kurvenform bleibt aber bei linearen Systemen sinusförmig.

Diese messtechnische Untersuchung ist selbstverständlich kein strenger Beweis dafür, dass das System linear ist. Für praktische Zwecke kann man aber gut abschätzen, ob das System (im untersuchten Bereich) linear ist.

11.3.1 Frequenzgang und Überlagerungssatz

Signale lassen sich aus sinusförmigen Schwingungen zusammensetzen. Wir können gedanklich jede einzelne Frequenzkomponente durch ein LTI-System schicken, die Amplituden der Frequenzkomponenten des Ausgangssignals $\{H(\hat{\omega})Ae^{j\varphi}\}$ berechnen und das Ausgangssignal aus diesen Frequenzkomponenten zusammensetzen.

Ein einfaches Beispiel erläutert die Rechenschritte :

An unserem FIR-Filter mit den Koeffizienten $b_k = [1, 2, 1]$ liegt das Eingangssignal

$$x[n] = 2 + 3 \cos\left(\frac{\pi}{4}n - \frac{\pi}{2}\right) + 2 \cos\left(\frac{9}{10}\pi n\right).$$

Als ersten Schritt berechnen wir den Frequenzgang für die normierten Kreisfrequenzen 0^5 , $\pi/4$ und 0.9π und erhalten die Werte:

$$H(0) = 4 \quad H\left(\frac{\pi}{4}\right) = 3.414\angle(-0.7854) \quad H(0.9\pi) = 0.097\angle(-2.8274)$$

Mit diesen Werten multiplizieren wir die (komplexen) Amplituden des Eingangssignals und erhalten als Ausgangssignal:

$$y[n] = 4 \cdot 2 + 3.414 \cdot 3 \cos\left(\frac{\pi}{4}n - \frac{\pi}{2} - 0.7854\right) + 0.097 \cdot 2 \cos\left(0.9n - \frac{\pi}{2} - 2.8274\right)$$

Die **Berechnung** des Ausgangssignals kann auch im **Frequenzbereich** erfolgen, indem man mit Hilfe des Frequenzgangs die **Ausgangsamplitude jeder Frequenzkomponente** berechnet.

Das Ausgangssignal ist die Zeitfunktion $y[n]$, die Berechnung des Ausgangssignals erfolgt aber im Frequenzbereich. Die Berechnung des Ausgangssignals im Zeitbereich wäre mathematisch aufwändiger und würde die Berechnung der Faltungssumme erforderlich machen. Im Frequenzbereich ist die Berechnung sehr einfach, es sind lediglich die (komplexen) Amplituden des Eingangssignals mit dem Frequenzgang zu multiplizieren.

Beispiel 79 Die Funktion $\text{freqz}(B, A)$ berechnet den Frequenzgang, kann aber auch für die Berechnung des Übertragungsfaktors diskreter Frequenzen verwendet werden, wenn man die gewünschten (Kreis)Frequenzen bei der Berechnung übergibt $[H, w] = \text{freqz}(B, A, w)$

$$[H, w] = \text{freqz}([1 \ 2 \ 1], 1, [0, (\pi/4), (\pi*9/10)])$$

$$[4, (2.4142 - 2.4142i), (-0.0931 - 0.0302i)] = [4e^{j0}, 3.41420e^{-j0.7854}, 0.0979e^{-j2.8274}]$$

11.3.2 Einschaltvorgänge

Einschaltvorgänge werden mathematisch mit Hilfe der **Sprungfunktion** modelliert.

Der Frequenzgang hat nur für sinusförmige Eingangssignale eine Bedeutung. Sinusförmige Signale erstrecken sich von $-\infty < n < \infty$, jedes praktische Signal hat aber einen Anfang, es wird »eingeschaltet«. Wir wollen nun untersuchen, wie sich das »Einschalten« auf das Systemverhalten auswirkt. Zur mathematischen Darstellung einer eingeschalteten komplexen Exponentialfunktion multiplizieren wir die Funktion mit der Sprungfunktion Die Sprungfunktion ist definiert

$$\delta_{-1}[n] = \begin{cases} 1 & n \geq 0 \\ 0 & n < 0 \end{cases} \quad (11.28)$$

Abbildung 11.8 zeigt die Sprungfunktion.

⁵Ein konstantes Signal kann man auch als ein Kosinussignal mit der Frequenz 0 darstellen.

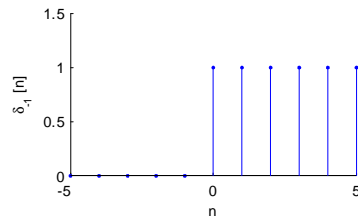


Abbildung 11.8: Sprungfunktion

Die eingeschaltete komplexe Exponentialfunktion lässt sich darstellen

$$x[n] = X e^{j\hat{\omega}n} \delta_{-1}[n] = \begin{cases} X e^{j\hat{\omega}n} & 0 \leq n \\ 0 & n < 0 \end{cases} \quad (11.29)$$

Für den Ausgang des Systems müssen wir daher schreiben

$$y[n] = \sum_{k=0}^M b_k X e^{j\hat{\omega}(n-k)} \delta_{-1}[n-k] \quad (11.30)$$

Wie man sich leicht überlegen kann, hat das Ausgangssignal drei Zeitabschnitte.

- Für $n < 0$ liegt kein Eingangssignal an, der Ausgang ist daher Null.
- Für die Zeit $0 \leq n < M$ »schwingt« das Filter ein. Erst wenn für die Berechnung des Ausgangssignals $M + 1$ Werte des von Null verschiedenen Eingangssignals ($n - M \geq 0$) zur Verfügung stehen, ist der Einschwingvorgang abgeschlossen.
- Ab dem Zeitpunkt $n \geq M$ verhält sich das Filter so, als ob ständig eine komplexe Exponentialfunktion angelegen wäre. Das Filter hat keine längere »Erinnerung« als der Ordnung des Filters entspricht.

FIR-Filter haben immer eine **endliche Einschwing- und Abklingphase**.

Für die diese drei Abschnitte erhalten wir folgende Beziehungen:

$$y[n] = \begin{cases} 0 & n < 0 \\ \left(\sum_{k=0}^n b_k e^{-j\hat{\omega}k} \right) X e^{j\hat{\omega}n} & 0 \leq n < M \\ \left(\sum_{k=0}^M b_k e^{-j\hat{\omega}k} \right) X e^{j\hat{\omega}n} & n \geq M \end{cases} \quad (11.31)$$

Bemerkung 80 Immer dann, wenn eine Änderung des Eingangssignals auftritt (Amplitude oder Phase oder Frequenz) findet ein neuer Einschwingvorgang des Filters statt, der die Dauer M hat. Die Berechnung des Ausgangssignals während des Einschwingvorgangs muss im Zeitbereich durch Lösen der Filtergleichung erfolgen. Mehr dazu erfahren wir im Kapitel über IIR-Filter.

Ein einfaches Beispiel zeigt den Einschwingvorgang des FIR-Tiefpasses mit den Koeffizienten $b_k = \{1, 2, 4, 2, 1\}$ an den wir ein eingeschaltetes Kosinussignal anlegen. Abbildung 11.9 zeigt das Eingangs- und Ausgangssignal.

11.4 Blockdiagramm von FIR-Filtern

Betrachtet man die Filtergleichung des FIR-Filters

$$y[n] = \sum_{k=0}^M b_k x[n-k] \quad (11.32)$$

Bei der Realisierung von FIR-Filtern müssen die Operationen **Verzögerung** (Speicherung), **Multiplikation** und **Addition** durchgeführt werden.

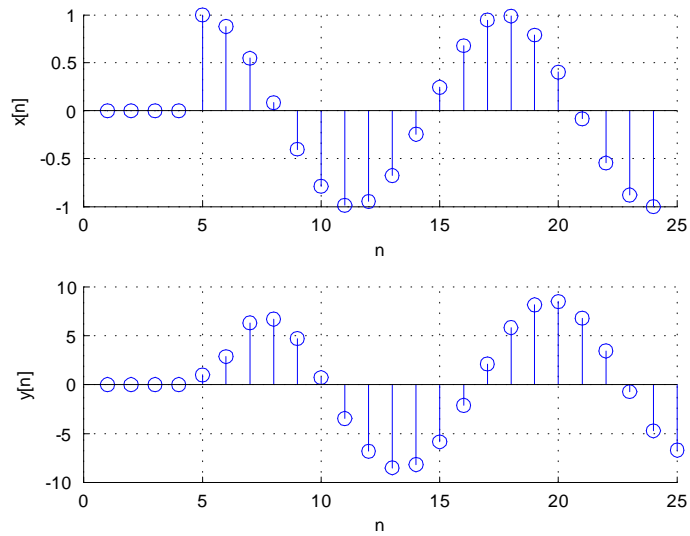


Abbildung 11.9: Eingeschaltetes Kosinussignal

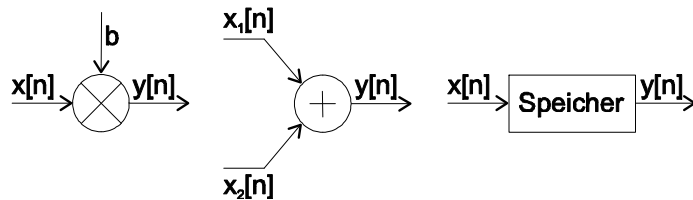


Abbildung 11.10: Recheneinheiten digitaler Filter

so sieht man, dass für die Berechnung des Ausgangssignals drei Rechenoperationen erforderlich sind: Speicherung (Verzögerung) der Werte des Eingangssignals $x[n-1], x[n-2], \dots, x[n-M]$, Multiplikation mit den Filterkoeffizienten b_k und Summenbildung $\sum_{k=0}^M$. Zur Durchführung dieser Operationen brauchen wir die entsprechenden Recheneinheiten, die in Abbildung 11.10 dargestellt sind.

Aus der Filtergleichung lässt sich damit das Blockdiagramm eines FIR-Filters zeichnen, wie Abbildung 11.11 zeigt.

$$y[n] = b_0x[n] + b_1x[n-1] + b_2x[n-2] + b_3x[n-3] \tag{11.33}$$

Das Blockdiagramm ist ebenso eine eindeutige Darstellung eines FIR-Filters wie Impulsantwort und Frequenzgang!

Die Filterstruktur nach Abbildung 11.11 nennt man *Direktform*, da die Filterkoeffizi-

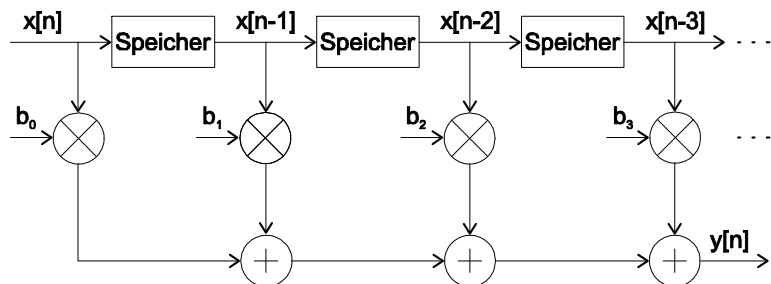


Abbildung 11.11: Blockdiagramm Direktform FIR-Filter

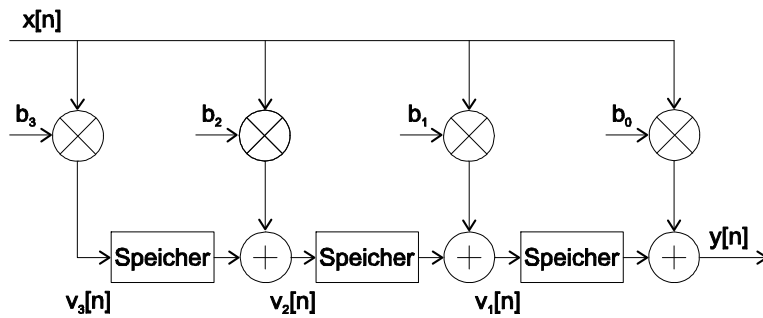


Abbildung 11.12: Blockdiagramm FIR Transponierte Form

enten direkt aus der Filtergleichung genommen werden können. Eine weitere Realisierung ist in Abbildung 11.12 gezeigt.

Die Filtergleichung ergibt sich aus:

$$\begin{aligned} y[n] &= b_0 x[n] + v_1[n-1] \\ v_1[n] &= b_1 x[n] + v_2[n-1] \\ v_2[n] &= b_2 x[n] + v_3[n-1] \\ v_3[n] &= b_3 x[n] \end{aligned}$$

Diese Filtergleichungen können durch Einsetzen gelöst werden:

$$\begin{aligned} y[n] &= b_0 x[n] + b_1 x[n-1] + v_2[n-2] \\ v_2[n] &= b_2 x[n] + b_3 x[n-1] \\ &\downarrow \\ y[n] &= b_0 x[n] + b_1 x[n-1] + b_2 x[n-2] + b_3 x[n-3] \end{aligned}$$

Die Realisierung nach Abbildung (11.12) nennt man *transponierte Form*.

An den Blockdiagrammen sieht man, wie das Eingangssignal mit der Impulsantwort des Filters gefaltet wird. Die direkte und die transponierte Form repräsentieren die im Abschnitt 11.2.1 vorgestellten zwei verschiedenen Betrachtungsweisen der Faltung.

FIR-Filter können in **direkter** oder **transponierter Form** realisiert werden.

11.4.1 Realisierung von Filtern

Zur Realisierung von (FIR)-Filtern stehen drei Möglichkeiten zur Verfügung. Welche Form der Implementierung gewählt wird hängt vom Einsatzgebiet und von wirtschaftlichen Überlegungen ab.

1. Realisierung des FIR-Algorithmus durch einen Allzweckrechner
Die Realisierung in einem Allzweckrechner kann sehr schnell durchgeführt werden, eignet sich daher gut für die Untersuchung der Eigenschaften der Rechenvorschrift des Filters. Nachteil ist, dass Allzweckrechner groß, für ausschließliche Berechnung der Filteraufgabe zu teuer und zu schwer sind und einen hohen Stromverbrauch haben, was sie für praktische Signalverarbeitungsaufgaben in der Regel unbrauchbar macht.
2. Realisierung durch einen Signalprozessor
Wirtschaftlich sinnvolle Lösungen bieten sich in vielen Fällen durch Signalprozessoren an. Bei der Analyse der Rechenaufgaben, die ein digitales Filter durchführen muss, stellt man fest, dass die Operationen Multiplikation und Addition sehr häufig vorkommen. Standardprozessoren sind nicht für Multiplikationen optimiert

Welche Art der technischen Umsetzung von digitalen Filtern gewählt wird, hängt von den Anforderungen der konkreten Anwendung ab.

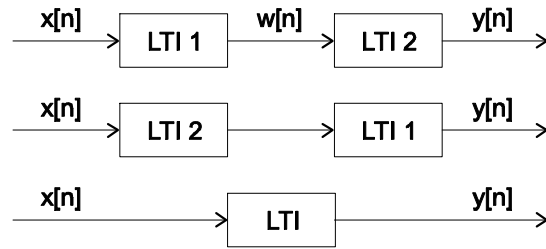


Abbildung 11.13: Kaskadieren von FIR-Filtern

und daher langsam. Signalprozessoren optimieren vor allem die Operation Multiplizieren/Add und erreichen für diese Aufgabenstellung eine sehr hohe Rechenleistung. Darüber hinaus sind in den meisten Signalprozessoren noch die Funktionen der Signal-Abtastung und -Rekonstruktion integriert. Es gibt eine Vielzahl von Signalprozessoren, die in der Regel für spezielle Aufgabenstellungen maßgeschneidert sind.

3. Realisierung in Hardware

Die Realisierung in dedizierter Hardware führt zu sehr kleinen Lösungen in Form von Chips, die schnell sind und wenig Leistung brauchen. HW-Realisierungen sind erst ab Stückzahlen in der Größenordnung von mehreren Hunderttausend wirtschaftlich herzustellen und nur eingeschränkt veränderbar, d.h. Anzahl und Wert der Filterkoeffizienten können nicht oder nur eingeschränkt geändert werden.

11.5 Kaskadieren von Filtern

Abbildung 11.13 zeigt zwei kaskadierte LTI-Systeme.

Bei der Zusammenschaltung im *Zeitbereich* erhalten wir:

$$y[n] = (x[n] * h_1[n]) * h_2[n] \quad (11.34)$$

$$y[n] = (x[n] * h_2[n]) * h_1[n] \quad (11.35)$$

$$y[n] = x[n] * (h_1[n] * h_2[n]) \quad (11.36)$$

Da die **Faltung** (im Zeitbereich) und die **Multiplikation** (im Frequenzbereich) **kommutativ** und **transitiv** sind, können Filter in **beliebiger Reihenfolge** hintereinander **geschaltet** werden und durch einfach berechenbare Ersatzsysteme **ersetzt** werden.

Da die Faltung kommutativ ist, spielt die Reihenfolge der Kaskadierung $LTI_1 \rightarrow LTI_2$ oder $LTI_2 \rightarrow LTI_1$ keine Rolle. Die Impulsantwort eines einzelnen Ersatzsystems ist $h[n] = h_1[n] * h_2[n]$.

Bei der Zusammenschaltung im *Frequenzbereich* erhalten wir:

$$y_1[n] = H_2(\omega) (H_1(\omega) e^{j\hat{\omega}n}) = H_2(\omega) H_1(\omega) e^{j\hat{\omega}n} \quad (11.37)$$

$$y_2[n] = H_1(\omega) (H_2(\omega) e^{j\hat{\omega}n}) = H_2(\omega) H_1(\omega) e^{j\hat{\omega}n} \quad (11.38)$$

$$y[n] = (H_2(\omega) H_1(\omega)) e^{j\hat{\omega}n} \quad (11.39)$$

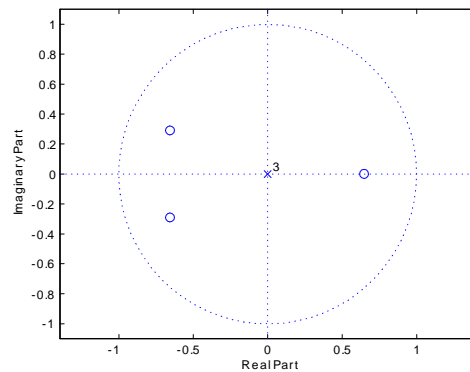
Der Frequenzgang eines einzelnen Ersatzsystems ist $H_2(\omega) H_1(\omega)$.

Anstelle der Faltung im Zeitbereich tritt die Multiplikation im Frequenzbereich.

$$h_1[n] * h_2[n] \Leftrightarrow H_1(\omega) H_2(\omega) \quad (11.40)$$

Beispiel 81 Den⁶ Zusammenhang $H(\omega) = H_1(\omega) H_2(\omega)$ können wir dazu verwenden, ein System höherer Ordnung in Teilsysteme niedriger Ordnung aufzuspalten, was Vorteile bei der Implementierung bringt. Dazu müssen wir $H(\omega)$ in die Nullstellendarstellung bringen und die Nullstellen dann den Teilfunktionen $H_1(\omega)$ und $H_2(\omega)$ zuordnen.

⁶Dieses Beispiel ist etwas vorgegriffen und sollte erst nach dem Kapitel über Pol- und Nullstellen gelesen werden. Inhaltlich passt es aber an diese Stelle.



Mit Hilfe der Matlabfunktion `tf2zpk`, wandeln wir die Polynomdarstellung in die Nullstellendarstellung um

$B = [3 \ 2 \ -1 \ -1]$; $A = [1]$;

$[z, p, k] = \text{tf2zpk}(B, 1)$ liefert einen Spaltenvektor für die Nullstellen. Die Polstellen von FIR-Filtern liegen im Nullpunkt. Da das Polynom $B(j\omega) = 3 + 2(j\omega)^{-1} - (j\omega)^{-2} - (j\omega)^{-3}$ dritter Ordnung ist erhalten wir drei Nullstellen

$z' \Rightarrow 0.6465, -0.6566 - 0.2907i, -0.6566 + 0.2907i$

aus denen wir ein Subsystem 1. Ordnung mit der reellen und ein Subsystem 2. Ordnung mit den beiden konjugiert komplexen Nullstellen bilden und zwar

$z1 = z([1;])$; $z2 = z([2 \ 3;])$; $p1 = [0]$; $p2 = [0; 0]$;

Das Subsystem 1, bestehend aus den Nullstellen $z1$ und das Subsystem 2, bestehend aus den Nullstellen $z2$, wandeln wir mit Hilfe der Funktion `zp2tf` aus der Nullstellendarstellung in die Polynomdarstellung um und erhalten

$H1 = \text{zp2tf}(z1, p1, 1)$; $H2 = \text{zp2tf}(z2, p2, k)$;

$H1(j\omega) = 1.0000 - 0.6465(j\omega)^{-1}$; $H2(j\omega) = 3.0000 + 3.9395(j\omega)^{-1} + 1.5468(j\omega)^{-2}$

$H(j\omega) = H1(\omega)H2(\omega)$ erhalten wir durch Multiplikation der Systemfunktionen der Teilsysteme

$\text{conv}(H1, H2) \Rightarrow 3 \ 2 \ -1 \ -1$

Zur Veranschaulichung der Trennung zeigen wir die Nullstellen (o) und Polstellen (x) von $H(z)$

11.5.1 Vom idealen zum realen Filter

Der Frequenzgang des idealen Tiefpassfilters ist

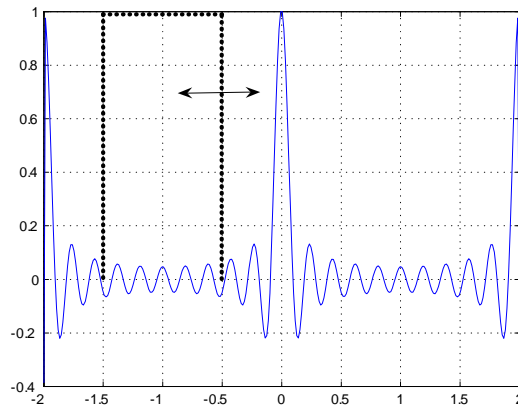
$$H_{iTP}(\hat{\omega}) = \begin{cases} 1 & |\hat{\omega}| \leq \omega_c \\ 0 & \omega_c < |\hat{\omega}| \leq \pi \end{cases} \quad (11.41)$$

Die Impulsantwort eines Filters erhält man durch inverse Fouriertransformation des Frequenzgangs $H(\hat{\omega}) \Rightarrow h[n]$.

$$h_{iTP} = \frac{1}{2\pi} \int_{-\pi}^{\pi} H_{iTP}(\hat{\omega}) e^{j\hat{\omega}n} d\hat{\omega} = \frac{1}{2\pi} \left(\frac{e^{j\omega_c n}}{jn} - \frac{e^{-j\omega_c n}}{jn} \right) \dots = \frac{\sin \omega_c n}{\pi n}, \quad -\infty < n < \infty \quad (11.42)$$

Die Impulsantwort eines idealen Filters ist nicht kausal, d.h. das Filter antwortet bereits vor dem Anlegen des Impulses. Nichtkausale Filter sind in Echtzeit nicht realisierbar! Um zu einem realisierbaren Filter zu gelangen, schneiden wir die Impulsantwort bei $-M \leq n \leq M$ ab und verschieben um M nach rechts. Die Koeffizienten des kausalen Tiefpasses sind dann

Das **ideale Tiefpassfilter** kann **nicht realisiert** werden, weil es **nicht kausal** ist.

Abbildung 11.14: $W_R(\hat{\omega})$ und $H_{iTP}(\hat{\omega})$ in normierter Darstellung

$$h_{TP} = \begin{cases} \frac{\sin[\omega_c(n-M)]}{\pi(n-M)} & 0 \leq n \leq N-1, \quad N = 2M+1 \\ 0 & \text{außerhalb} \end{cases} \quad (11.43)$$

Die Fouriertransformation der **durch Abschneiden »kausalisierten« Impulsantwort** des idealen Filters ergibt den Frequenzgang des **realen Filters**.

Die durch das Abschneiden der Impulsantwort h_{iTP} entstehende Impulsantwort h_{TP} ist nicht mehr die Impulsantwort des idealen Filters. Wir müssen daher den Frequenzgang des zur Impulsantwort h_{TP} gehörigen realen Filters berechnen:

Die Operation des Abschneiden bei $-M \leq n \leq M$ kann auch durch Multiplikation der Impulsantwort des idealen Filters mit einem Rechteckfenster (im Zeitbereich) dargestellt werden.

$$h_{TP} = h_{iTP} \cdot w_R \quad (11.44)$$

Wir berechnen den Frequenzgang des Filters H_{TP} durch Fouriertransformation der Impulsantwort. Die Fouriertransformierte von h_{iTP} kennen wir bereits, sie ist der Frequenzgang des idealen Filters. Die Fouriertransformierte der Rechteckfunktion haben wir bereits im Abschnitt über die Signalrekonstruktion kennengelernt und dort gesehen, dass die *Si*-Funktion auftritt. Aus der Multiplikation im Zeitbereich wird die Faltung im Frequenzbereich und wir erhalten zusammengefasst

$$H_{TP}(\hat{\omega}) = H_{iTP}(\hat{\omega}) * W_R(\hat{\omega}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} H_{iTP}(\hat{\omega}') W_R(\hat{\omega} - \hat{\omega}') d\hat{\omega}' \quad (11.45)$$

$$w_R[n] = \begin{cases} 1 & 0 \leq |n| \leq M \\ 0 & \text{außerhalb} \end{cases} \quad (11.46)$$

$$W_R(\hat{\omega}) = \sum_{n=-M}^M e^{j\hat{\omega}n} = \frac{\sin\left[\frac{(2M+1)\hat{\omega}}{2}\right]}{\sin\frac{\hat{\omega}}{2}} \quad (11.47)$$

Gleichung (11.47) ist die Fouriertransformierte des Rechteckfensters, die periodische sinc-Funktion (Diriclet'sche Funktion), wie in Abbildung 11.14 dargestellt.

Der Frequenzgang von $H_{TP}(\hat{\omega})$ ergibt sich durch Faltung der in Abbildung 11.14 dargestellten Funktionen. Das Ergebnis ist in Abbildung 11.15 dargestellt.

Der Frequenzgang des realisierbaren Filters $H_{TP}(\hat{\omega})$ weicht vom (idealen) rechteckförmigen Frequenzgang ab. Je größer M gewählt wird, desto genauer kann das Rechteck angenähert werden, es bleibt aber immer eine Abweichung (Gibbs'sches Phänomen im Frequenzbereich), ähnlich der Abweichung bei der Fourierzerlegung von Rechteckschwingungen.

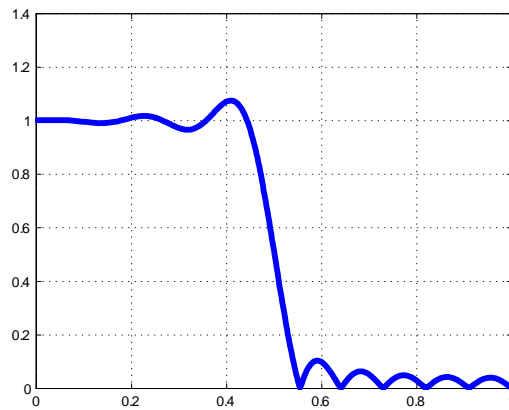


Abbildung 11.15: Ergebnis der Faltung $|H_{TP}(\hat{\omega})|$

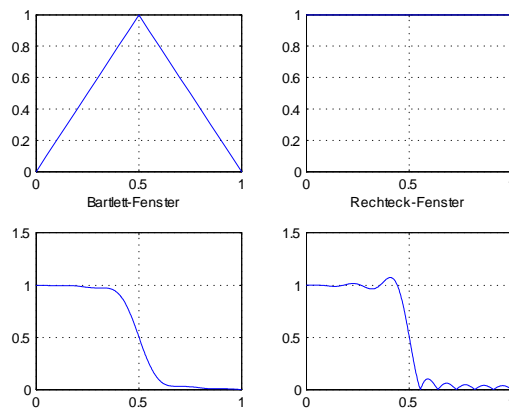


Abbildung 11.16: Vergleich Tiefpass mit Dreieck- und Rechteckfenster

Die Welligkeit im Durchlassbereich des Filters $H_{TP}(\hat{\omega})$ entsteht durch den abrupten Übergang des Rechteckfensters im Zeitbereich. Wählt man einen »sanfteren«, linearen Übergang mit einem Dreiecksfenster, verschwindet die Welligkeit im Durchlassbereich, wie Abbildung 11.16 zeigt.

Es gibt eine Reihe weiterer Fenster (Hamming, Hann, Blackmann, ...), die den Übergang mittels einer modifizierten Kosinusfunktion realisieren. Abbildung 11.17 vergleicht das Rechteck- mit dem Hamming-Fenster

$$w_{\text{Hamming}}[n] = 0.54 + 0.46 \cos\left(\frac{2\pi n}{2M+1}\right) \quad -M \leq n \leq M \quad (11.48)$$

Wie wir sehen, ist beim Hamming-Filter die Welligkeit im Durchlassbereich verschwunden, allerdings ist die Flankensteilheit des Übergangs zwischen Durchlass- und Sperrbereich beim Hamming-Filter geringer als beim Rechteck-Filter.

Die bisher betrachteten Fensterfunktionen haben keinen wählbaren Parameter, der es erlauben würde, die Filtereigenschaften zu beeinflussen. Das Kaiser-Fenster verwendet die Besselfunktion erster Art nullter Ordnung

$$w_{\text{Kaiser}} = \frac{I_0\left[\beta\sqrt{1-(n/M)^2}\right]}{I_0(\beta)}, \quad -M \leq n \leq M \quad (11.49)$$

Es gibt verschiedene Arten die Impulsantwort des idealen Filters zu »beschneiden«. Welche benutzt wird, hängt von der Anwendung ab.

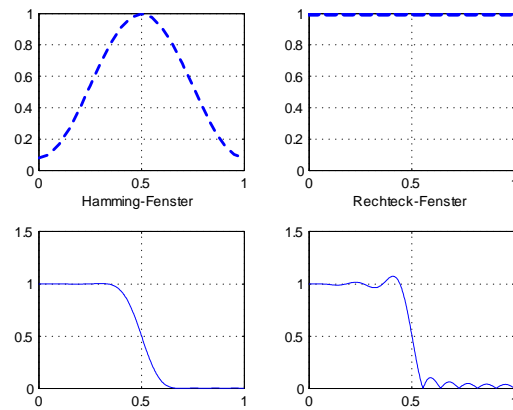


Abbildung 11.17: Tiefpass mit Hamming- bzw. Rechteckfenster

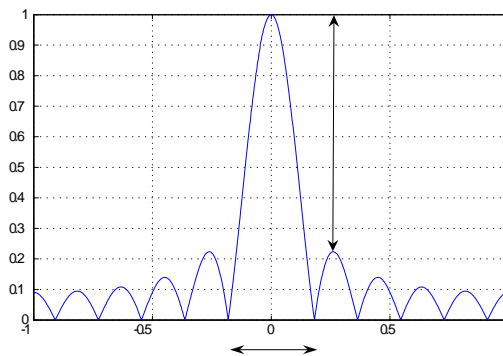


Abbildung 11.18: Fenstereigenschaften

mit dem Parameter β . Erhöhen von β macht die Hauptkeule breiter und verringert die Amplituden der Nebenkeulen (erhöht die Dämpfung im Sperrbereich). Typische Werte von β liegen zwischen $4 < \beta < 9$. Für $\beta = 0$ geht das Kaiserfenster in das Rechteckfenster über.

Um die unterschiedlichen Fensterfunktion vergleichen zu können, geht man von der Kurvenform ihrer Fouriertransformierten, wie in Abbildung 11.18 gezeigt, aus.

Dabei interessiert vor allem die Breite der Hauptkeule (der erste Nulldurchgang) und die Amplitude der größten Nebenkeule.

Fenster	B_{Haupt}	A_{Neben}
Rechteck	$4\pi/(2M + 1)$	-13 dB
Hann	$8\pi/(2M + 1)$	-32 dB
Hamming	$8\pi/(2M + 1)$	-43 dB
Blackmann	$12\pi/(2M + 1)$	-58 dB

Wie man obiger Tabelle entnehmen kann, ist die Amplitude der größten Nebenkeule umso kleiner, je breiter die Hauptkeule ist. Für ein Tiefpassfilter gilt daher: Je höher die Sperrdämpfung, desto weniger steilflankig ist der Übergang vom Durchlass- in den Sperrbereich.

In ähnlicher Weise können die Filterfunktionen von Hochpass, Bandpass und Bandsperrung ermittelt werden, indem man die Impulsantworten der idealen Filter ermittelt, abschneidet und kausal macht.

Neben der Berechnung der Filterkennlinie mit Hilfe von Fensterfunktionen kann die Filterfunktion auch durch Approximation im Durchlass- und Sperrbereich gefunden werden, wie z.B. beim Dolph-Tschebyscheff oder Remez-Entwurf.

Da der Rechenaufwand zur Berechnung von Filterfunktionen hoch ist, bedient man sich entsprechender Programme wie z.B. der Matlab Filterdesign Toolbox.

11.6 Zusammenfassung

FIR-Filter realisieren Rechenvorschriften, die ausschließlich Werte der Eingangsfolge $x[n]$ verwenden. Durch Wahl der Filterkoeffizienten b_k kann das gewünschte Verhalten des Filters (Tiefpass, Hochpass, Bandpass, Bandsperre) verwirklicht werden.

Zur Untersuchung der Eigenschaften von (FIR-)Filtern verwendet man die Impulsantwort und den Frequenzgang. Beliebige Eingangssignale lassen sich aus Impulsfolgen bzw. komplexen Exponentialfolgen zusammensetzen und bei LTI-Systemen mit dem Überlagerungssatz berechnen.

Die Impulsantwort – und daher auch die Einschwing- und Abklingzeit – von FIR-Filtern hat endliche Dauer. FIR-Filter können lineare Phase haben, eine Eigenschaft, die für die Erhaltung der Kurvenform von Signalen wichtig ist. FIR-Filter sind immer stabil und lassen sich gut in Hardware und Software realisieren. FIR-Filter sind bei gleicher Ordnung weniger selektiv als die später vorgestellten IIR-Filter oder umgekehrt zur Realisierung einer gewünschten Filterselektivität ist bei FIR-Filtern eine höhere Ordnung als bei IIR-Filtern erforderlich, was zu höheren Implementierungskosten führt.

Kapitel 12

Die z-Transformation

Inhalt

12.1 Der z -Bereich	182
12.2 Eigenschaften der z -Transformation	183
12.3 Blockdiagramme und Systemfunktion	184
12.4 Inverse Filterung	185
12.5 Zusammenhang z -Bereich und Frequenzbereich	185
12.5.1 Pole und Nullstellen von $H(z)$	186
12.6 Zusammenfassung	189

In den Abschnitten über die Impulsantwort und den Frequenzgang haben wir die Antwort eines Systems auf den Einheitspuls und die komplexe Exponentialfunktion ermittelt. Das Eingangssignal $x[n]$ wird durch die Systemantwort $T\{x[n]\}$ in das Ausgangssignal transformiert. Es ist zu beachten, dass die ermittelten Beziehungen $T\{x[n]\}$, also $h[n]$ und $H(\omega)$ nur für den Einheitsimpuls bzw. die komplexe Exponentialfunktion eine Bedeutung haben!

Wir untersuchen nun das Systemverhalten für eine weitere Eingangsgröße und zwar für

$$x[n] = z^n \quad \text{für alle } n \quad (12.1)$$

wobei z eine beliebige komplexe Zahl ist.

Während der Einheitsimpuls und ein sinusförmiges Eingangssignal (in der mathematischen Darstellung der komplexen Exponentialfunktion) einen Bezug zur physikalischen Welt haben, ist das beim Eingangssignal z^n nicht der Fall. In Anlehnung an die Bezeichnungen Zeitbereich und Frequenzbereich, nennen wir die Darstellung mit z den z -Bereich.

Der z -Bereich ist eine abstrakte mathematische Welt, die eine Reihe von Eigenschaften hat, die bei der Untersuchung von Systemeigenschaften sehr hilfreich sind:

- Durch die z -Transformation werden aus Differenzgleichungen gebrochen rationale Funktionen, es werden damit Polynome für die Systembeschreibung eingeführt. Polynome sind »gutartige« Funktionen, mit den sich leicht rechnen lässt. Algebraische Operationen wie Division, Multiplikation und Faktorisierung von Polynomen entsprechen dem Zerlegen bzw. Zusammensetzen von LTI Systemen.
- Die rechnerisch aufwendige Operation der Faltung geht über in die Multiplikation.
- Gebrochen rationale Funktionen bestehen aus einem Zähler- und einem Nennerpolynom. Aus der Lage der Wurzeln (Nullstellen) dieser Polynome lassen sich Eigenschaften digitaler Systeme ableiten, insbesondere sind dadurch Aussagen über die Stabilität möglich.
- Die z -Transformation hat für diskrete Systeme dieselbe Bedeutung wie die Laplace-Transformation für kontinuierliche Systeme.

Eine weitere sehr wichtige »**Testfunktion**« für digitale Systeme ist die Funktion z^n

Der z -**Bereich** kann nicht so einfach physikalisch interpretiert werden wie der Frequenzbereich, hat aber viele **Vorteile aus mathematischer Sicht**.

Zur Beschreibung von Systemen werden je nach Aufgabenstellung unterschiedliche Darstellungen gewählt:

Der *Zeitbereich* ist der Bereich der »wirklichen« Signale, er wird durch Differenzgleichungen beschrieben und von der Impulsantwort $h[n]$ und Faltung $y[n] = h[n] * x[n]$ regiert.

Der *Frequenzbereich* ist der Bereich der »Töne« – der Spektren – wird durch Polynome beschrieben und vom Frequenzgang $H(\omega)$ und der Multiplikation $Y[\omega] = H(\omega)X[\omega]$ regiert.

Der **Frequenzbereich** ist im z -Bereich enthalten.

Der z -Bereich ist eine abstrakte Welt, die durch Pole und Nullstellen beschrieben und von der Systemfunktion $H(z) = \frac{Y(z)}{X(z)}$ regiert wird. Im z -Bereich wird $x[n] = z^n$ als Eingangsgröße gewählt, wobei z eine beliebige komplexe Zahl und daher auch $z = e^{j\omega}$ sein kann. Wie wir sehen werden »liegt« der Frequenzbereich daher im z -Bereich und ist ein Sonderfall des z -Bereichs. Dennoch wird zwischen den beiden Bereichen unterschieden, da sie unterschiedlichen Charakter haben.

12.1 Der z -Bereich

Die Systemeigenschaften im z -Bereich ermitteln wir wieder über die Beziehung zwischen Eingangs- und Ausgangssignal

$$y[n] = \sum_{k=0}^M b_k x[n-k] = \sum_{k=0}^M b_k z^{[n-k]} = \left(\sum_{k=0}^M b_k z^{-k} \right) z^n \quad (12.2)$$

Die Funktion

$$H(z) = \sum_{k=0}^M b_k z^{-k} = \sum_{k=0}^M h[k] z^{-k} \quad (12.3)$$

nennt man Systemfunktion. Für die Impulsfunktion erhalten wir die wichtige Beziehung

$$h[n] = \sum_{k=0}^M b_k \delta_0[n-k] \Leftrightarrow H(z) = \sum_{k=0}^M b_k z^{-k} \quad (12.4)$$

Die **Systemfunktion** eines digitalen Systems ist die z -Transformierte seiner **Impulsantwort**.

Bemerkung 82 Die Systemfunktion $H(z)$ ist die z -Transformierte der Impulsantwort $h[n]$.

Als Beispiel berechnen wir die Systemfunktion des FIR-Tiefpasses mit den Koeffizienten $b_k = [1, 2, 1]$

$$H(z) = b_0 z^{-0} + b_1 z^{-1} + b_2 z^{-2} = 1 + 2z^{-1} + z^{-2} = \quad (12.5)$$

$$= 1 + \frac{2}{z} + \frac{1}{z^2} = \frac{z^2 + 2z + 1}{z^2} = \frac{Y(z)}{X(z)} \quad (12.6)$$

Die Systemfunktion ist eine gebrochen rationale Funktion mit dem Zählerpolynom $Y(z) = z^2 + 2z + 1$ und dem Nennerpolynom $X(z) = z^2$.

Durch die z -Transformation wird aus der Differenzgleichung

$$\begin{aligned} h[n] &= b_0 \delta_0[n] + b_1 \delta_0[n-1] + b_2 \delta_0[n-2] \\ h[n] &\Rightarrow H(z) \\ \delta_0[n] &\Rightarrow 1 \\ \delta_0[n-1] &\Rightarrow z^{-1}, \quad \delta_0[n-2] \Rightarrow z^{-2} \end{aligned}$$

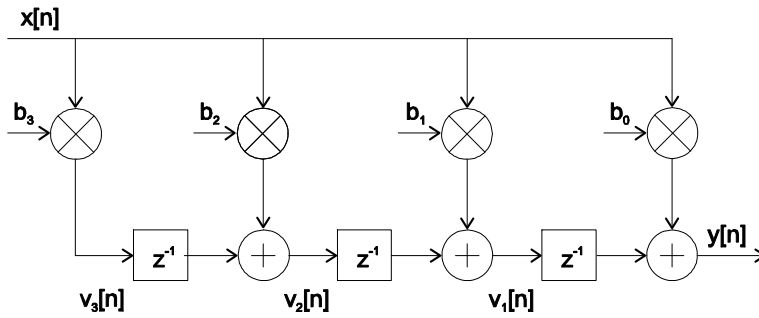


Abbildung 12.1: Darstellung mit z^{-1}

die gebrochene rationale Funktion

$$H(z) = \frac{z^2 + 2z + 1}{z^2}$$

Für die z -Transformierte der Eingangsfolge $x[n]$ erhalten wir

$$x[n] = \sum_{k=0}^N x[k]\delta_0[n - k] \Leftrightarrow X(z) = \sum_{k=0}^M x[k]z^{-k} \tag{12.7}$$

Die Eingangsfolge $x[n] = [1, 2, 3, -1, 2, -3]$ hat die z -Transformierte

$$X(z) = 1 + 2z^{-1} + 3z^{-2} - 1z^{-3} + 2z^{-4} - 3z^{-5}.$$

Die z -Transformierte des Einheitsimpulses ist

$$\delta_0[n] \Leftrightarrow 1 \tag{12.8}$$

12.2 Eigenschaften der z -Transformation

Die z -Transformation ist eine lineare Transformation und es gilt

$$ax_1[n] + bx_2[n] \Leftrightarrow aX_1(z) + bX_2(z) \tag{12.9}$$

Die Multiplikation einer Folge mit z^{-1} verschiebt die Folge um einen Schritt, wie das folgende Beispiel zeigt

$$X(z) = z^{-1} (1 + 2z^{-1} + 3z^{-2} - 1z^{-3} + 2z^{-4} - 3z^{-5}) = \tag{12.10}$$

$$= 0z^0 + 1z^{-1} + 2z^{-2} + 3z^{-3} - 1z^{-4} + 2z^{-5} - 3z^{-6} \tag{12.11}$$

Diesen Zusammenhang können wir anschreiben (Verschiebung im Zeitbereich)

$$x[n - 1] \Leftrightarrow z^{-1}X(z) \tag{12.12}$$

Für eine Verzögerung um n_0 Punkte erhalten wir

$$x[n - n_0] \Leftrightarrow z^{-n_0}X(z) \tag{12.13}$$

Wie wir wissen, sind für die Realisierung von digitalen Filtern Verzögerungsglieder (Speicher) erforderlich. Die Abbildung 12.1 zeigt eine Realisierung. Die Eigenschaft (12.12) der z -Transformation wird in Blockdiagrammen häufig zur Kennzeichnung der Verzögerungsglieder verwendet, wie Abbildung 12.1 zeigt.

Streng genommen ist diese Darstellung nicht korrekt, da im Blockdiagramm die Zeitbereichnotation $x[n]$ und die z -Bereichnotation z^{-1} in einer gemeinsamen Darstellung verwendet wird. Dennoch ist diese Darstellung gebräuchlich.

Die z -Transformation ist **linear**.

Die **Verschiebung** im **Originalbereich** entspricht einer **Multiplikation mit z^{-1}** im **z -Bereich**

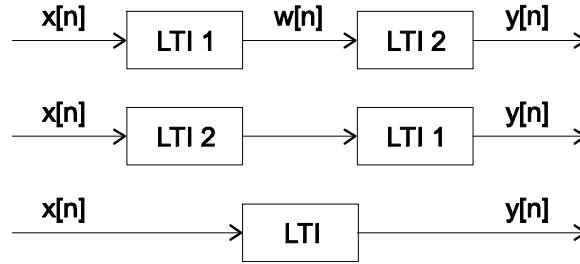


Abbildung 12.2: Reihenfolge der Kaskadierung

Die Impulsantwort eines Verzögerungselements ist nichts anderes als die Verschiebung des Impulses um einen Schritt in der Folge.

$$h[n] = \delta_0[n - 1] \quad (12.14)$$

Die Verzögerung um einen Schritt berechnet man im Zeitbereich durch Faltung des Eingangssignals mit der Impulsantwort

$$y[n] = x[n] * h[n] = x[n] * \delta_0[n - 1] = x[n - 1] \quad (12.15)$$

Im z -Bereich erhalten wir für das Verzögerungselement:

$$Y(z) = z^{-1}X(z) \quad (12.16)$$

Der Verschiebung im Zeitbereich um n_0 entspricht die Multiplikation mit z^{-n_0} im z -Bereich.

Im Allgemeinen gilt

$$y[n] = x[n] * h[n] = \sum_{k=0}^M h[k]x[n - k]. \quad (12.17)$$

Eine **Faltung** im **Originalbereich** wird durch die z -Transformation zu einer **Multiplikation** im **z -Bereich**.

Die z -Transformation ist eine lineare Transformation, es gilt daher der Überlagerungssatz und wir können schreiben

$$Y(z) = \sum_{k=0}^M h[k] (z^{-k} X(z)) = \left(\sum_{k=0}^M h[k] z^{-k} \right) X(z) = H(z)X(z) \quad (12.18)$$

Der Faltung im Zeitbereich entspricht die Multiplikation im Frequenzbereich.

$$y[n] = h[n] * x[n] \Leftrightarrow Y(z) = H(z)X(z) \quad (12.19)$$

12.3 Blockdiagramme und Systemfunktion

Aus (12.17) und der Kommutativität der Faltung folgt, dass die Systemfunktion zweier kaskadierter System $H(z) = H_1(z)H_2(z) = H_2(z)H_1(z)$ ist. Die Reihenfolge der Kaskadierung ist ohne Belang, wie Abbildung 12.2 zeigt.

Polynome können durch Aufspalten in ihre Wurzeln (Nullstellen) in Polynome niedrigerer Ordnung zerlegt werden. Wir betrachten die Systemfunktion

$$H(z) = 1 - 2z^{-1} + 2z^{-2} - z^{-3} \quad (12.20)$$

(12.20) hat eine Wurzel bei $z = 1$ und kann daher zerlegt werden in:

$$H(z) = (1 - z^{-1})(1 - z^{-1} + z^{-2}) \quad (12.21)$$

Das System 3. Ordnung lässt sich durch Faktorisierung in ein System 1. Ordnung und ein System 2. Ordnung aufspalten, wie Abbildung 12.3 zeigt.

Mit Hilfe der **Nullstellendarstellung** der Systemfunktion $H(z)$ können **große Systeme** aus **kleineren Teilsystemen** zusammengesetzt werden.

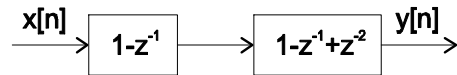


Abbildung 12.3: Faktorisierung von Systemen

12.4 Inverse Filterung

Signale werden beim Durchgang durch Übertragungskanäle beeinflusst und in vielen Fällen möchte man diese Signaländerung aufheben, um das Originalsignal wieder herzustellen. Dazu muss man aus dem Ausgangssignal und der Impulsantwort das Eingangssignal ermitteln. Diese Aufgabenstellung nennt man inverse Filterung oder Deconvolution. Die Gleichung $y[n] = h[n] * x[n]$ muss nach $x[n]$ aufgelöst werden.

Die inverse Filterung kann in der Zeitbereichsdarstellung in der Regel nicht gelöst werden. Im z -Bereich kann man schreiben:

$$Y(z) = H_1(z)H_2(z)X(z) = H(z)X(z) = X(z) \tag{12.22}$$

$H_1(z)$ ist der Übertragungskanal, $H_2(z)$ ist das korrigierende System. Um das ursprüngliche Signal wieder herzustellen muss gelten:

$$H_1(z)H_2(z) = 1 \tag{12.23}$$

Beispiel 83 Als praktisches Beispiel der inversen Filterung soll die Nachbearbeitung (digitalisierter) alter Schallplattenaufnahmen betrachtet werden. Alte Aufnahmen wurden mit primitiven Aufnahmesystemen durchgeführt, die die Signale verzerrten. Die Entzerrung kann gemacht werden, wenn die Impulsantwort des ursprünglichen Systems bekannt ist. Leider ist das in der Regel nicht der Fall und man muss die Impulsantwort des Aufnahmesystems annehmen und durch Experimente ein akzeptables Signal wiederherstellen.

Beispiel 84 Das folgende Beispiel erläutert die mathematischen Zusammenhänge:

$$H_1(z) = (1 - z^{-1} + 0.5z^{-2}) \tag{12.24}$$

$$H_2(z) = \frac{1}{H_1(z)} = \frac{1}{(1 - z^{-1} + 0.5z^{-2})} = \tag{12.25}$$

$$= \frac{z^2}{z^2 - z + 0.5} \tag{12.26}$$

H_2 hebt die Wirkung von H_1 auf, die Systemfunktion des inversen Filters H_2 ist eine gebrochene rationale Funktion. Die Systemfunktion von FIR-Filtern hat immer die Form

$$H_{FIR}(z) = \frac{b_n z^n + b_{n-1} z^{n-1} + \dots + b_1 z + b_0}{z^n} \tag{12.27}$$

Das FIR-Filter in unserem Beispiel kann also nicht durch ein FIR-Filter invertiert werden. Wir werden im Kapitel über IIR-Filter eine Lösung für dieses Problem finden.

12.5 Zusammenhang z -Bereich und Frequenzbereich

Für die Darstellung der Eigenschaften eines digitalen Systems haben wir folgende Beziehungen gefunden:

$$z - \text{Bereich} \quad \Leftrightarrow \quad \hat{\omega} - \text{Bereich}$$

$$H(z) = \sum_{k=0}^M b_k z^{-k} \quad \Leftrightarrow \quad H(\hat{\omega}) = \sum_{k=0}^M b_k e^{-j\hat{\omega}k}$$

Bei der **inversen Filterung** werden **Signalverzerrungen** »rückgängig gemacht«.

Die Wirkung eines FIR-Filters kann durch ein nachgeschaltetes FIR-Filter nicht mehr behoben werden.

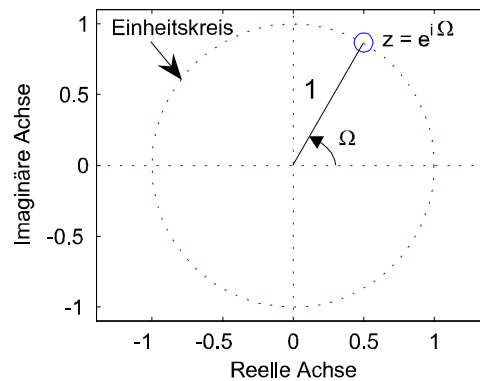


Abbildung 12.4: Einheitskreis in der komplexen Ebene

Im z -Bereich kann z beliebige Werte annehmen, also auch die Werte $z = e^{j\hat{\omega}}$. Setzen wir in $H(z)$ für $z = e^{j\hat{\omega}}$ ein dann wird aus $H(z)|_{z=e^{j\hat{\omega}}} = H(\hat{\omega})$.

$z = e^{j\hat{\omega}}$ hat den Betrag $|z| = 1$ und die Punkte von z liegen daher in der komplexen Ebene auf dem Kreis mit dem Radius $|z| = 1$ und dem Winkel $\angle z = \hat{\omega}$. Abbildung (12.4) stellt diesen Zusammenhang grafisch dar.

Man gelangt vom z -Bereich zum Frequenzbereich, indem man für $z = e^{j\hat{\omega}}$ wählt.

Über jedem Punkt der komplexen Ebene ist der zugehörige Wert von $H(z)$ – bzw. von $H(\hat{\omega})$ wenn z auf dem Einheitskreis liegt – aufgetragen. Die Funktionswerte sind im Allgemeinen komplexe Werte. Trägt man über der komplexen Ebene z den Betrag von $|H(z)|$ auf, dann erhält man ein »Gebirge« in axonometrischer Darstellung. Denkt man sich über dem Einheitskreis einen Zylinder, so gibt der Schnitt des Zylinders mit dem »Gebirge« den Verlauf des Betrags des Frequenzgangs. Schneidet man den Zylinder der Länge nach auf und rollt ihn ab, so erhält man die bekannte Darstellung des Frequenzgangs.

Der **Frequenzgang** befindet sich in der z -Ebene **über dem Einheitskreis**.

Beispiel 85 Das folgende Beispiel in Abbildung 12.5 stellt den Betrag der Funktion $H(z) = \frac{1}{1-0.9z^{-1}}$ dar. Diese Funktion wird an der Stelle $z = 0.9$ Null, $H(z)$ wird daher an dieser Stelle unendlich. (In der Zeichnung ist die Unendlichkeitsstelle abgeschnitten.) Der Schnitt des Zylinders über dem Einheitskreis mit $H(z)$ ist in Abbildung 12.5 dick gezeichnet, in Abbildung 12.6 ist nur der Zylinder gezeigt.

Beispiel 86 Schneidet man den Zylinder in Abbildung 12.6 auf und rollt ihn ab, so erhält man die bekannte Darstellung des Frequenzgangs. In Abbildung 12.7 ist nur die Hälfte des »Zylinders« (0 bis π) gezeichnet.

12.5.1 Pole und Nullstellen von $H(z)$

Nullstellen des **Zählerpolynoms** von $H(z)$ entsprechen den **Nullstellen** von $H(z)$. Nullstellen des **Nennerpolynoms** heißen **Polstellen** von $H(z)$.

Die Systemfunktion eines diskreten LTI-Systems ist im Allgemeinen eine gebrochene rationale Funktion der Form:

$$H(z) = \frac{b_0z + b_1z + b_2z^2 + \dots + b_mz^m}{a_0z + a_1z + a_2z^2 + \dots + a_nz^n} = \frac{Y(z)}{X(z)} \quad (12.28)$$

Die Nullstellen des Zählerpolynoms $Y(z)$ nennt man Nullstellen von $H(z)$. An den Nullstellen des Nennerpolynoms $X(z)$ wird $H(z) \rightarrow \infty$, diese Punkte nennt man Polstellen von $H(z)$.

Durch die Lage der Pole und Nullstellen ist die Systemfunktion bis auf einen konstanten Faktor eindeutig dargestellt. Die grafische Darstellung von Polen und Nullstellen nennt man Pol-/Nullstellendarstellung. Nullstellen werden als Kreise, Polstellen als Kreuze dargestellt.

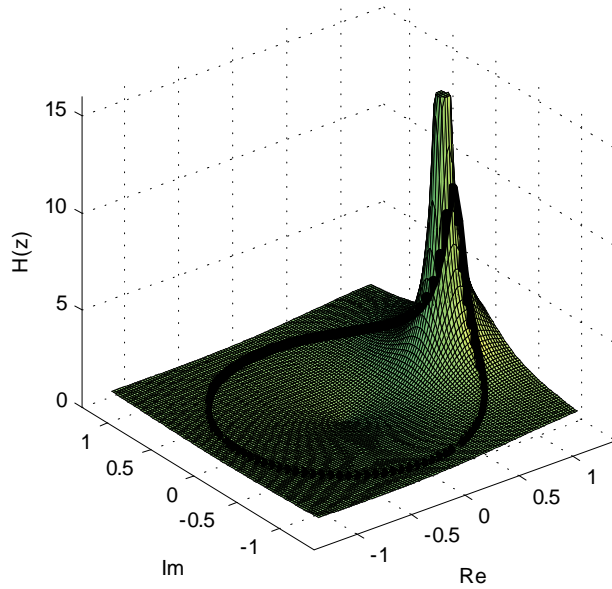


Abbildung 12.5: Betrag von $H(z) = \frac{1}{1-0.9z^{-1}}$

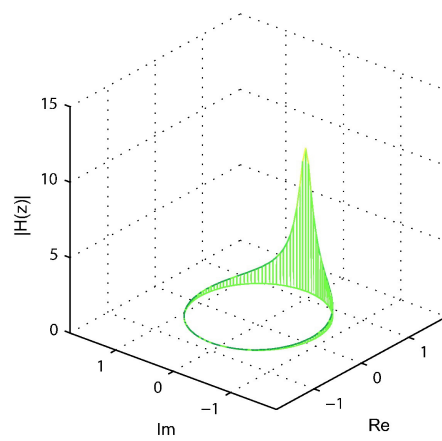


Abbildung 12.6: Zylinder über Einheitskreis

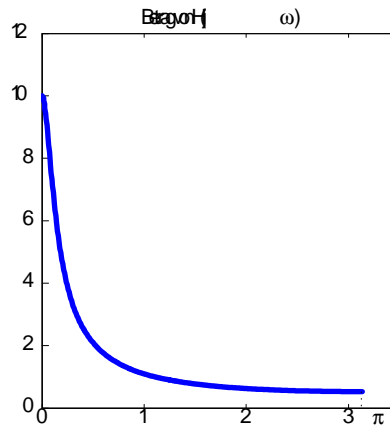
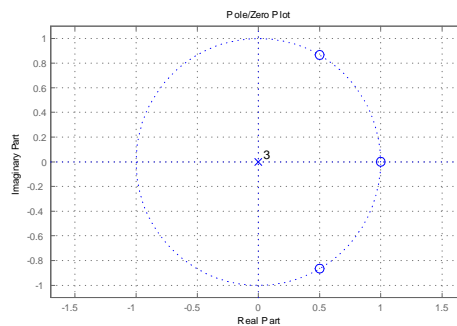
Abbildung 12.7: Betrag des Frequenzgangs von $H(z) = \frac{1}{1-0.9z^{-1}}|_{z=j\omega}$ 

Abbildung 12.8: Pol-/Nullstellen-Darstellung

Beispiel 87 Als Beispiel ist das PN-Diagramm der Systemfunktion

$$H(z) = 1 - 2z^{-1} + 2z^{-2} - z^{-3} = \frac{z^3 - 2z^2 + 2z - 1}{z^3}$$

in Abbildung 12.8 dargestellt. Es wurde mit den folgenden Matlab-Befehlen erzeugt:
 $B = [1, -2, 2, -1]$; $A = [1]$; $zplane(B, A)$

Die Nullstellen des Zählers liegen bei

$$\begin{aligned} z^3 - 2z^2 + 2z - 1 &= (z - 1)(z - e^{j\pi/3})(z - e^{-j\pi/3}) = 0 \\ z_0 &= 1, e^{\pm j\pi/3} \end{aligned}$$

Die Polstellen (Nullstellen des Nenners) $z^3 = 0$ liegen (dreifach) im Nullpunkt. Bei FIR-Filtern liegen die Polstellen immer im Nullpunkt. Bei LTI-Systemen nach (12.28) treten die Pole in allgemeiner Lage auf, wir werden den Einfluss der Pole auf das Systemverhalten im Abschnitt über IIR-Filter kennenlernen.

Durch die **Pol- und Nullstellen** ist die **Systemfunktion** bis auf einen konstanten Faktor **eindeutig bestimmt**.

Nullstellen auf dem Einheitskreis bedeuten, dass die zugehörige Frequenzkomponente am Ausgang des Systems nicht auftritt, also vollständig unterdrückt wird. Erinnerung: $H(z) = 0$ für ein z auf dem Einheitskreis $\iff H(\hat{\omega}) = 0$ für die entsprechende normierte Kreisfrequenz $\hat{\omega}$.

12.6 Zusammenfassung

Durch die z -Transformation werden Funktionen vom Zeitbereich (n -Bereich) in den Bildbereich (z -Bereich) abgebildet. Durch diese Abbildung werden aus Differenzgleichungen im n -Bereich Polynome im z -Bereich. Wie wir im Kapitel über die IIR-Filter noch sehen werden, können Differenzgleichungen in den z -Bereich transformiert werden, dort als algebraische Gleichungen gelöst werden und stellen nach Rücktransformation in den n -Bereich die Lösung der Differenzgleichungen im Originalbereich dar. Dadurch gelangt man zu einer systematischen Lösung der Differenzgleichungen. Die häufig vorkommenden Transformationsbeziehungen zwischen n - und z -Bereich gibt es in Tabellenform, was die Rechnung erheblich erleichtert.

Die Polynome im z -Bereich bzw. die Lage der Pole und Nullstellen gibt Aufschluss über das Verhalten der Systemfunktion¹.

Eine Bewegung auf dem Einheitskreis von 0 bis $\pm\pi$ entspricht dem Durchlaufen der (komplexen) Frequenz von $0 \leq \hat{\omega} \leq \omega_G$. Da es sich um abgetastete Systeme handelt, ist die höchste Frequenz $\omega_G = 0.5\omega_{\text{abtast}}$.² Für $z = e^{j\hat{\omega}}$ wird aus der Darstellung im z -Bereich die Darstellung im Frequenzbereich. Nullstellen auf dem Einheitskreis entsprechen also Nullstellen der Übertragungsfunktion.³

¹Für stabile zeit-diskrete Systeme müssen die Polstellen innerhalb des Einheitskreises liegen, wie wir bei den IIR-Filtern sehen werden. Bei zeit-kontinuierlichen Systemen müssen die Pole in der linken offenen Halbebene des s -Bereichs liegen.

²Der Bewegung auf dem Einheitskreis entspricht bei zeit-kontinuierlichen Systemen die Bewegung auf der imaginären Achse von 0 bis $\pm j\infty$ in der s -Ebene.

³Entsprechend liegen bei zeit-kontinuierlichen Systemen die Übertragungsnulstellen auf der imaginären Achse.

Kapitel 13

IIR-Filter

Inhalt

13.1 Vom FIR- zum IIR-Filter	191
13.2 Zeitverhalten von IIR-Filtern	192
13.3 Eigenschaften von IIR-Filtern	193
13.3.1 Linearität und Zeitinvarianz	193
13.3.2 Impulsantwort	194
13.3.3 Sprungantwort	195
13.3.4 Systemfunktion von IIR-Filtern	197
13.3.5 Unendliche Folgen im z -Bereich	197
13.4 Realisierung von IIR-Filtern	198
13.4.1 Erste Direktform	198
13.4.2 Zweite Direktform	199
13.4.3 Transponierte Form	200
13.5 Pole und Nullstellen	200
13.5.1 Stabilität und Lage der Pole	201
13.6 Frequenzgang von IIR-Filtern	201
13.7 3D-Darstellung der Systemfunktion	202
13.8 Die inverse z-Transformation	203
13.8.1 Partialbruchzerlegung	205
13.9 Einschwingvorgänge und Stabilität	208
13.10 Konjugiert komplexe Pole	210
13.10.1 z -Transformierte komplexer Pole	211
13.11 Blockdiagramm der Partialbruchzerlegung	212
13.12 Darstellung von Systemen	212
13.12.1 Zeitbereich	212
13.12.2 z -Bereich	213
13.12.3 Frequenzbereich	213
13.12.4 Blockdiagramm	215
13.13 Zusammenfassung	215

In diesem Kapitel werden wir eine weitere Klasse von diskreten Filtern kennen lernen, die Infinite Impulse Response Filter.

13.1 Vom FIR- zum IIR-Filter

FIR Filter verwenden zur Berechnung des Ausgangswertes $y[n]$ nur Werte des Eingangssignals $x[n]$.

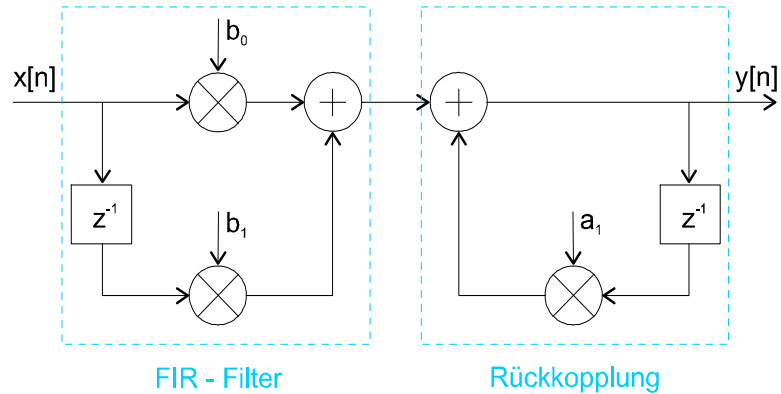


Abbildung 13.1: Blockdiagramm IIR-Filter

$$y[n] = \sum_{k=0}^M b_k x[n-k]$$

IIR-Filter verwenden (im Unterschied zu FIR-Filtern) auch Werte der **Ausgangsgröße**, um den nächsten Ausgangswert zu berechnen.

Beim Rekursivfilter werden aber auch Werte des Ausgangssignals zur Berechnung herangezogen, wie das einfache Beispiel in Abbildung (13.1) zeigt.

Wie man sieht, besteht das Filter in Abbildung 13.1 aus einem FIR-Filter erster Ordnung und aus einem nachgeschalteten Rückkopplungsteil. Die beschreibende Gleichung des Filters der Abbildung (13.1) lautet

$$y[n] = b_0 x[n] + b_1 x[n-1] + a_1 y[n-1] \quad (13.1)$$

In Gleichung (13.1) tritt neben $y[n]$ auch die (zeitverzögerte) Ausgangsgröße $y[n-1]$ auf, eine derartige Gleichung nennt man Differenzgleichung. Im allgemeinen Fall lautet die Differenzgleichung für ein IIR-Filter

$$y[n] = \underbrace{\sum_{l=1}^N a_l y[n-l]}_{\text{Ausgangswerte}} + \underbrace{\sum_{k=0}^M b_k x[n-k]}_{\text{Eingangswerte}} \quad (13.2)$$

Die **Ordnung** von IIR-Filtern ist gleich der **Anzahl der Rückführungsterme**.

Bei IIR-Filtern bezeichnet man N – die Zahl der Rückführungsterme – als Ordnung des Systems.

13.2 Zeitverhalten von IIR-Filtern

Das Zeitverhalten von IIR-Filtern untersuchen wir zunächst an einem einfachen Beispiel eines Systems erster Ordnung und mit dem Rückführungskoeffizienten $a_1 = 0.8$

$$y[n] = 0.8y[n-1] + x[n] \quad (13.3)$$

An unser Beispielsystem legen wir das Eingangssignal

$$x[n] = 2\delta[n] - 3\delta[n-1] + 2\delta[n-3] \quad (13.4)$$

und berechnen das Ausgangssignal durch Einsetzen in die Differenzgleichung. Für $n = 0$ erhalten wir

$$y[0] = 0.8y[0-1] + x[0] \quad (13.5)$$

Es wird angenommen, dass alle Werte des Eingangs- und Ausgangssignals zu den Zeitpunkten $n < 0$ gleich Null sind.

Der Wert $y[-1]$ ist zunächst nicht definiert. Wir nehmen an, dass das System vor An-

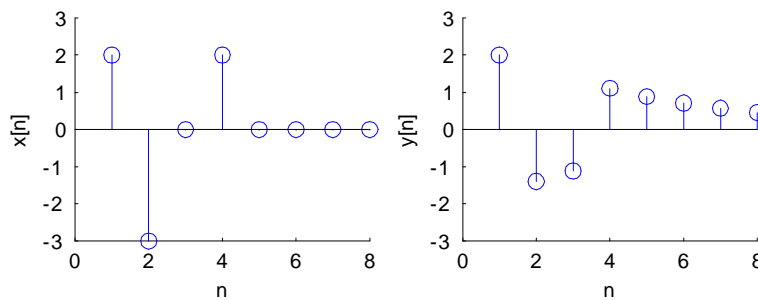


Abbildung 13.2: Eingangs- & Ausgangssignal IIR-Filter

legen des Eingangssignals in Ruhelage war und dass alle Verzögerungselemente (Speicher) mit dem Wert Null geladen sind. Dann erhalten wir $y[-1] = 0$.

Bemerkung 88 *Im weiteren Verlauf werden wir generell annehmen, dass das Eingangssignal zum Zeitpunkt $n_0 = 0$ »eingeschaltet« wird und dass die Systeme vor dem Anfangszeitpunkt $n_0 = 0$ in Ruhelage sind, dass also folgende Anfangsbedingungen gelten:*

$$\begin{aligned} x[n] &= 0 \quad \text{für } n < 0 \\ y[n] &= 0 \quad \text{für } n < 0 \end{aligned} \quad (13.6)$$

Unter diesen Annahmen erhalten wir für unser Beispiel

$$\begin{aligned} y[0] &= 0.8y[-1] + x[0] = 0.8 \cdot 0 + 2 = 2 \\ y[1] &= 0.8y[0] + x[1] = 0.8 \cdot 2 - 3 = -1.4 \\ y[2] &= 0.8y[1] + x[2] = 0.8 \cdot (-1.4) + 0 = -1.12 \\ y[3] &= 0.8y[2] + x[3] = 0.8 \cdot (-1.12) + 2 = 1.104 \end{aligned} \quad (13.7)$$

Ab diesem Zeitpunkt ist das Eingangssignal Null.

$$\begin{aligned} y[4] &= 0.8y[3] + x[4] = 0.8 \cdot 1.104 + 0 = 0.883 \\ y[5] &= 0.8y[4] + x[5] = 0.8 \cdot 0.883 + 0 = 0.707 \\ &\vdots = \vdots \end{aligned}$$

Abbildung 13.2 stellt den Zusammenhang zwischen Eingangs- und Ausgangssignal dar¹.

Aus Gleichung (13.7) können wir erkennen, dass das Ausgangssignal für $n > 3$ nach der Differenzgleichung $y[n] = 0.8y[n-1]$ abklingt, für die wir folgende Lösung finden

$$y[n] = y[3] \cdot (0.8)^{n-3} \quad \text{für } n \geq 3 \quad (13.8)$$

Aus Gleichung (13.8) sehen wir, dass $y[n]$ erst für $n \rightarrow \infty$ gegen Null geht, dass die Antwort dieses Filters also unendlich lang dauert, daher auch der Name *infinite impulse response*.

13.3 Eigenschaften von IIR-Filtern

13.3.1 Linearität und Zeitinvarianz

Man kann zeigen, dass für ein IIR-System nach Gleichung (13.2), wenn die Anfangsbedingungen nach Gleichung (13.6) erfüllt sind, der Überlagerungssatz gilt und derartige Systeme daher linear und zeitinvariant sind.

IIR-Filter sind **linear** und **zeitinvariant**.

¹In der Abbildung sind beide Signale wegen der besseren Lesbarkeit um 1 nach rechts verschoben.

13.3.2 Impulsantwort

Bei den FIR-Filtern haben wir gesehen, dass die Impulsantwort ein Filter eindeutig beschreibt. Beliebige Eingangssignale können durch Überlagerung von gewichteten, zeitversetzten Impulsen dargestellt werden, das Ausgangssignal ist aus gewichteten und zeitversetzten Impulsantworten zusammengesetzt.

Das **Ausgangssignal** kann auch bei IIR-Filtern durch die **Faltung** des **Eingangssignals** mit der **Impulsantwort** berechnet werden.

Auch IIR-Filter sind lineare, zeitinvariante Systeme, es gilt daher der Überlagerungssatz und die Berechnung der Antwort eines Systems auf eine beliebige Eingangsfolge erfolgt mit Hilfe der Faltung

$$y[n] = \sum_{k=-\infty}^{\infty} x[k]h[n-k] \quad (13.9)$$

Da sowohl das Eingangssignal als auch die Impulsantwort zeitlich nicht begrenzt sind, muss die Faltungssumme im Allgemeinen über den gesamten Zeitraum von $-\infty$ bis ∞ berechnet werden.

Für ein System erster Ordnung

$$y[n] = a_1y[n-1] + b_0x[n] \quad (13.10)$$

ist die Differenzgleichung der Impulsantwort

$$h[n] = a_1h[n-1] + b_0\delta_0[n] \quad (13.11)$$

Wir müssen eine Lösung für Gleichung (13.11) finden, wobei das Problem auftritt, dass die Lösung – also die Folge $h[n]$ – in Gleichung (13.11) in der Form $h[n]$ und in der Form $h[n-1]$ steckt. Zum Unterschied von den FIR-Filtern kann man die Lösung bei IIR-Filtern also nicht unmittelbar anschreiben.²

Für die **direkte Berechnung der Impulsantwort** eines IIR-Filters muss eine lineare **Differenzgleichung** mit konstanten Koeffizienten gelöst werden.

Wie bei Differentialgleichungen müssen wir auch bei Differenzgleichungen den Lösungsansatz »erraten«. Die Betrachtung von (13.8) legt folgenden Lösungsansatz nahe

$$h[n] = \begin{cases} b_0(a_1)^n & \text{für } n \geq 0 \\ 0 & \text{für } n < 0 \end{cases} \quad (13.12)$$

Den Vorgang des Einschaltens stellen wir durch die Sprungfunktion dar und können daher schreiben

$$\delta_{-1}[n] = \begin{cases} 1 & \text{für } n \geq 0 \\ 0 & \text{für } n < 0 \end{cases} \quad (13.13)$$

$$h[n] = b_0(a_1)^n \delta_{-1}[n] \quad (13.14)$$

Zum Beweis der Korrektheit der Lösung setzen wir in die Differenzgleichung (13.11) ein. Wir zeigen die Gültigkeit zunächst bei $n = 0$ und dann für $n > 0$ für den Lösungsansatz und die DGL

$$\text{(Lösung) } n = 0 : h[0] = b_0(a_1)^0 = b_0 \quad \checkmark$$

$$\text{(DGL) } n = 0 : h[n] = a_1h[n-1] + b_0\delta_0[n] = a_1(0) + b_0\delta_0[0] = b_0 \quad \checkmark$$

$$\text{(Lösung) } n > 0 : h[n] = b_0(a_1)^n \quad \checkmark$$

$$\text{(DGL) } n > 0 : h[n] = a_1h[n-1] + (0)$$

$$n > 0 : b_0(a_1)^n = a_1b_0(a_1)^{n-1} = b_0(a_1)^n \quad \checkmark$$

Die Lösung (13.12) erfüllt die Differenzgleichung (13.11)

²Bei zeit-kontinuierlichen Systemen erster Ordnung tritt anstelle der Differenzgleichung eine Differentialgleichung auf. Die gesuchte Zeitfunktion $h(t)$ tritt in einer Differentialgleichung erster Ordnung in der Form von $h(t)$ und $\dot{h}(t) = d[h(t)]/dt$ auf.

Das Lösen von Differenzgleichungen durch »Erraten« ist mühselig und verlangt viel Erfahrung. Wir werden in einem späteren Abschnitt ein direktes Verfahren mit Hilfe der z -Transformation und Partialbruchzerlegung kennenlernen.

13.3.3 Sprungantwort

Die Berechnung der Systemantwort auf ein beliebiges Eingangssignal kann wie beim FIR-Filter durch Lösen der Differenzgleichung oder mit Hilfe der Faltung erfolgen. Die Lösung über die Faltung ist bei IIR-Filtern wegen der unendlich langen Impulsantwort in der Regel schwierig oder unmöglich, außer in Sonderfällen, wo es eine geschlossene Lösung gibt. Ein derartiges Beispiel liefert die Sprungantwort eines IIR-Filters erster Ordnung.

Ein IIR-Filter ist auch durch die **Sprungantwort** vollständig beschrieben.

$$y[n] = a_1 y[n-1] + b_0 x[n] \tag{13.15}$$

$$\delta_{-1}[n] = \begin{cases} 1 & \text{für } n \geq 0 \\ 0 & \text{für } n < 0 \end{cases} \tag{13.16}$$

Zunächst berechnen wir die Sprungantwort durch Einsetzen in die Differenzgleichung und punktweises Berechnen des Ausgangssignals

n	$x[n]$	$y[n]$
0	1	b_0
1	1	$b_0 + b_0(a_1)$
2	1	$b_0 + b_0(a_1) + b_0(a_1)^2$
3	1	$b_0(1 + a_1 + a_1^2 + a_1^3)$
4	1	$b_0(1 + a_1 + a_1^2 + a_1^3 + a_1^4)$
\vdots	1	\vdots

Das Entwicklungsgesetz für $y[n]$ lässt sich leicht erkennen

$$y[n] = b_0 (1 + a_1 + a_1^2 + a_1^3 + a_1^4 + \dots + a_1^n) = b_0 \sum_{k=0}^n a_1^k \tag{13.17}$$

Für die Summe der geometrischen Reihe $\sum_{k=0}^n a_1^k$ finden wir in einer Formelsammlung

$$\sum_{k=0}^n q^k = \begin{cases} n + 1 & \text{für } q = 1 \\ \frac{q^{n+1} - 1}{q - 1} & \text{sonst} \end{cases} \tag{13.18}$$

und erhalten dann

$$y[n] = b_0 \frac{1 - a_1^{n+1}}{1 - a_1} \quad \text{für } n \geq 0, \quad \text{wenn } a_1 \neq 1 \tag{13.19}$$

Bei der Untersuchung von $y[n]$ müssen wir drei Fälle unterscheiden:

1. Wenn $|a_1| > 1$, dann dominiert der Zählerausdruck a_1^{n+1} und $y[n]$ steigt über alle Grenzen. In diesem Fall spricht man von einem unstablen System.
2. Wenn $|a_1| < 1$, dann geht der Ausdruck a_1^{n+1} gegen Null für $n \rightarrow \infty$, man spricht von einem stabilen System. Wenn $n \rightarrow \infty$ dann wird

$$\lim_{n \rightarrow \infty} y[n] = \lim_{n \rightarrow \infty} b_0 \frac{1 - a_1^{n+1}}{1 - a_1} = \frac{b_0}{1 - a_1} \tag{13.20}$$

3. Wenn $|a_1| = 1$ ist eine weitere Untersuchung erforderlich:
 Für $a_1 = 1$ erhalten wir $y[n] = (n + 1)b_0$ und $y[n]$ wächst mit $n \rightarrow \infty$ über alle Grenzen.
 Für $a_1 = -1$ wechselt $y[n]$ zwischen b_0 und Null und zwar $y[n] = b_0$ für gerades n , $y[n] = 0$ für ungerades n .

Die Sprungantwort eines IIR-Filters erster Ordnung kann noch durch »Erraten und Überprüfen« berechnet werden.

Die **Rückführungskoeffizienten** a_k sind für die **Stabilität** von IIR-Filtern ausschlaggebend.

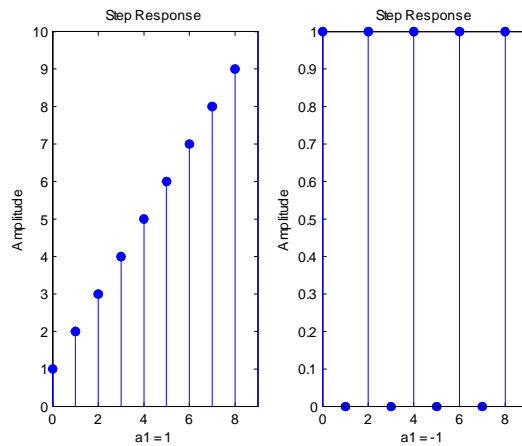


Abbildung 13.3: Sprungantwort für $a_1 = 1$ und $a_1 = -1$

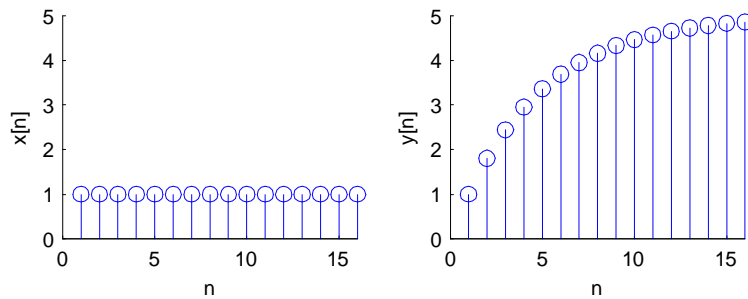


Abbildung 13.4: Sprungantwort IIR-Filter

Abbildung 13.3 zeigt die Sprungantworten für diese beiden Fälle.

Mehr über Stabilität werden wir bei der Untersuchung von IIR-Filtern im z -Bereich erfahren.

Für das IIR-Filter $y[n] = 0.8y[n - 1] + x[n]$ ist die Sprungantwort in Abbildung 13.4 dargestellt.

Nach der Berechnung der Sprungantwort durch Einsetzen in die Differenzgleichung wollen wir die Berechnung mit Hilfe der Faltung ausführen.

$$y[n] = x[n] * h[n] = \sum_{k=-\infty}^{\infty} x[k]h[n - k] \tag{13.21}$$

Da bei IIR-Filtern sowohl Eingangssignal als auch Impulsantwort unendliche Dauer haben, ist die Berechnung der Faltung in der Regel nicht einfach. In unserem Beispiel kann aber eine geschlossene Lösung gefunden werden.

Wir falten das Eingangssignal $\delta_{-1}[n]$ mit der Impulsantwort $h[n] = b_0(a_1)^n\delta_{-1}[n]$ und erhalten

$$y[n] = \sum_{k=-\infty}^{\infty} \delta_{-1}[k]b_0(a_1)^{n-k}\delta_{-1}[n - k] \tag{13.22}$$

Da $\delta_{-1}[k] = 0$ für $k < 0$ und $\delta_{-1}[n - k] = 0$ für $n - k < 0$ (bzw. $n < k$) muss die Faltung nicht von $-\infty$ bis ∞ durchgeführt werden sondern nur von Null bis n und wir erhalten

Bei einem IIR-Filter (erster Ordnung) kann die Sprungantwort auch durch Faltung berechnet werden.

$$y[n] = \begin{cases} 0 & \text{für } n < 0 \\ \sum_{k=0}^n b_0(a_1)^{n-k} & \text{für } n \geq 0 \end{cases} \quad (13.23)$$

Wir formen um

$$y[n] = \sum_{k=0}^n b_0(a_1)^{n-k} = b_0(a_1)^n \sum_{k=0}^n (a_1)^{-k} \quad (13.24)$$

Unter Verwendung der Reihensumme aus Gleichung (13.18) wird daraus

$$y[n] = b_0(a_1)^n \frac{1 - \left(\frac{1}{a_1}\right)^{n+1}}{1 - \left(\frac{1}{a_1}\right)} = b_0 \frac{1 - a_1^{n+1}}{1 - a_1} \quad (13.25)$$

und damit dasselbe Ergebnis wie in Gleichung (13.19).

13.3.4 Systemfunktion von IIR-Filtern

Mit Hilfe der z -Transformation haben wir die Systemfunktion für FIR-Filter eingeführt und gesehen, dass der Faltung im Zeitbereich die Multiplikation im z -Bereich entspricht.

$$y[n] = h[n] * x[n] \quad \Leftrightarrow \quad Y(z) = H(z)X(z) \quad (13.26)$$

Dieser Zusammenhang gilt auch für die IIR-Filter. Zum Unterschied von den FIR-Filtern, bei denen die Systemfunktion immer ein Polynom in z^{-1} ist (z.B. $H(z) = 1 + 2z^{-1} + 2z^{-2} + z^{-3}$), ist die Systemfunktion von IIR-Filtern immer eine gebrochen rationale Funktion, also eine Funktion, die aus Zähler- und Nennerpolynom besteht. Als Beispiel nehmen wir wieder ein IIR-Filter 1.Ordnung

$$y[n] = a_1 y[n-1] + b_0 x[n] + b_1 x[n-1] \quad (13.27)$$

Durch z -Transformation wird aus der Differenzgleichung (13.27)

$$Y(z) = a_1 z^{-1} Y(z) + b_0 X(z) + b_1 z^{-1} X(z) \quad (13.28)$$

Nach Umformung erhalten wir

$$H(z) = \frac{Y(z)}{X(z)} = \frac{b_0 + b_1 z^{-1}}{1 - a_1 z^{-1}} = \frac{B(z)}{A(z)} \quad (13.29)$$

Für ein allgemeines IIR-Filter erhalten wir

$$H(z) = \frac{Y(z)}{X(z)} = \frac{b_0 + b_1 z^{-1} + b_2 z^{-2} + \dots + b_m z^{-m}}{1 - a_1 z^{-1} - a_2 z^{-2} - \dots - a_n z^{-n}} = \frac{B(z)}{A(z)} \quad (13.30)$$

Im Zähler stehen die Vorwärtskopplungsterme (FIR-Eingangsteil), im Nenner Eins plus die negativen Rückkopplungsterme.

13.3.5 Unendliche Folgen im z -Bereich

Wie wir gesehen haben, ist die Impulsantwort von IIR-Filtern »infinite«, es treten also unendliche Folgen auf, die auch im z -Bereich dargestellt werden müssen. Wir wollen zunächst die z -Transformation der sehr wichtigen Folge $h[n] = a^n \delta_{-1}[n]$ untersuchen.

$$H(z) = \sum_{n=0}^{\infty} a^n z^{-n} = \sum_{n=0}^{\infty} (az^{-1})^n \quad (13.31)$$

Gleichung (13.31) ist eine geometrische Reihe. In einer Formelsammlung finden wir

Die **Systemfunktion** von IIR-Filtern ist immer eine **gebrochen rationale Funktion**, die aus Zähler- und Nennerpolynom besteht.

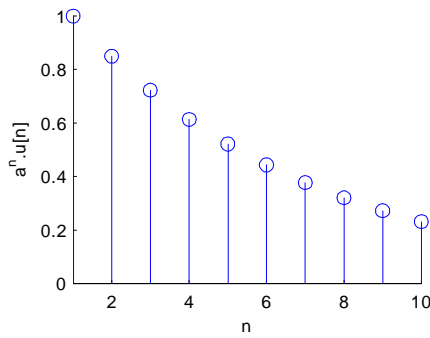


Abbildung 13.5: $y[n] = 0.7^n \delta_{-1}[n]$

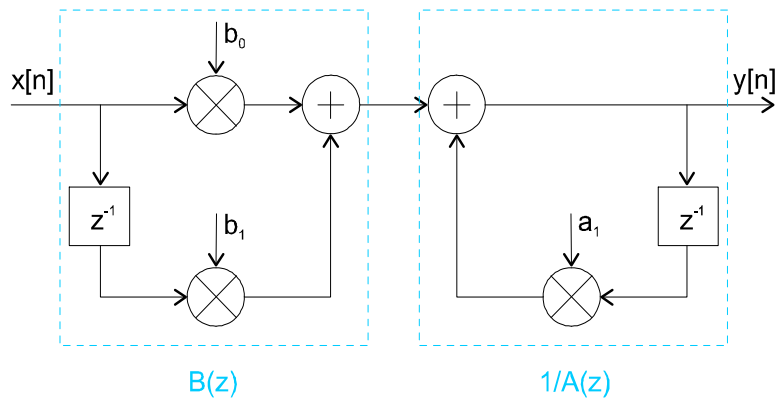


Abbildung 13.6: Blockdiagramm IIR-Filter

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1-x} \quad \text{für } |x| < 1 \tag{13.32}$$

Die z -Transformierte von $a^n \delta_{-1}[n]$ ist $\frac{1}{1-az^{-1}}$.

und daraus

$$H(z) = \sum_{n=0}^{\infty} a^n z^{-n} = \frac{1}{1-az^{-1}} \quad |a| < 1 \tag{13.33}$$

Wir haben ein wichtiges Transformationspaar gefunden

$$\text{Zeitbereich} \quad \Leftrightarrow \quad z - \text{Bereich}$$

$$a^n \delta_{-1}[n] \quad \Leftrightarrow \quad \frac{1}{1-az^{-1}} \tag{13.34}$$

Abbildung 13.5 zeigt den Verlauf von $a^n \delta_{-1}[n]$ für $a = 0.7$.

13.4 Realisierung von IIR-Filtern

13.4.1 Erste Direktform

Die erste **Direktform** eines IIR-Filters ist die **intuitive Realisierung** des Filters nach der **Differenzgleichung**.

Die erste Direktform lässt sich unmittelbar aus der Differenzgleichung ableiten. Abbildung 13.6 zeigt ein Blockdiagramm für ein IIR-Filter erster Ordnung.

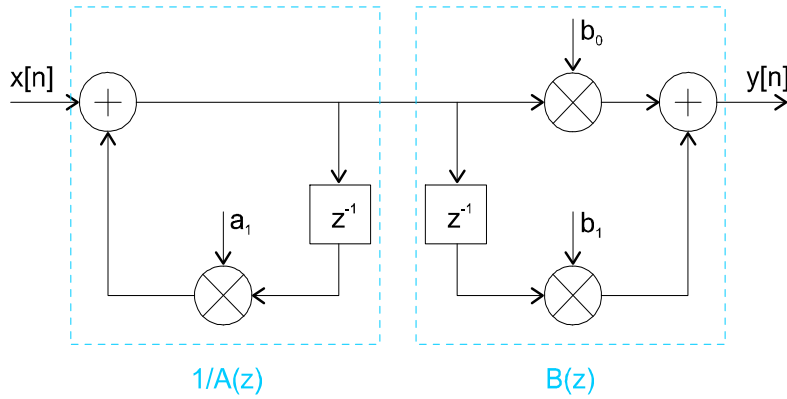


Abbildung 13.7: IIR-Filter 2. Direktform

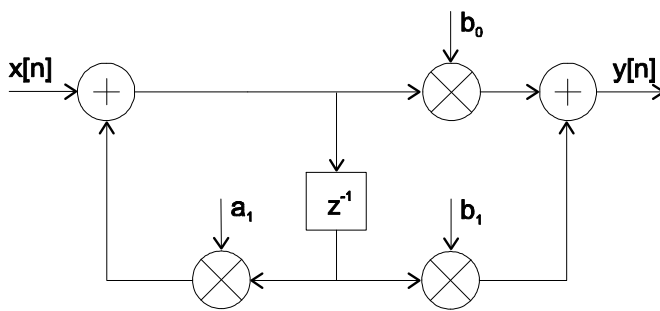


Abbildung 13.8: IIR-Filter 2. Direktform

13.4.2 Zweite Direktform

Wir wissen, dass bei kaskadierten LTI-Systemen die Reihenfolge der Kaskadierung ohne Einfluss auf die Systemfunktion ist.

$$H(z) = B(z) \cdot \frac{1}{A(z)} = \frac{1}{A(z)} \cdot B(z) \tag{13.35}$$

Damit ergibt sich das Blockdiagramm für ein IIR-Filter erster Ordnung in zweiter Direktform nach Abbildung 13.7.

Wie man aus Abbildung 13.7 leicht erkennen kann, führen die beiden Verzögerungsglieder identische Operationen aus, ein Verzögerungsglied kann also eingespart werden, wie in Abbildung 13.8 gezeigt ist.

Zur Beschreibung von Systemstrukturen und Algorithmen verwendet man wegen ihrer kompakten Form häufig Signalflussgraphen. Abbildung 13.9 zeigt den Signalflussgraphen eines allgemeinen IIR-Filter der zweiten Direktform.

Bei der **zweiten Direktform** können **Verzögerungsglieder eingespart** werden.

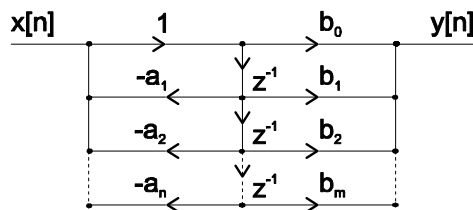


Abbildung 13.9: Signalflussgraph IIR-Filter 2. Direktform

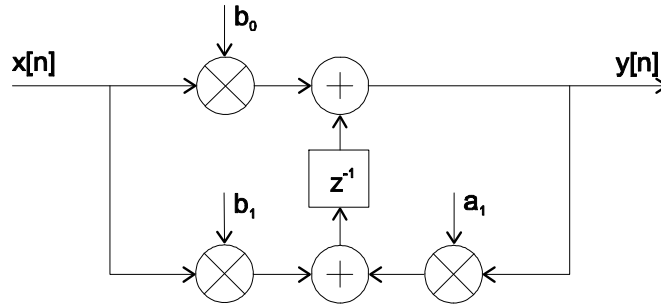


Abbildung 13.10: Transponierte Form

13.4.3 Transponierte Form

Blockdiagramm 13.10 zeigt die transponierte Form eines IIR-Filter erster Ordnung.

Die transponierte Form erhält man aus der Direktform durch folgende topologische Umwandlung:

1. Die Multiplizierer bleiben unverändert.
2. Alle Signalpfeile kehren ihre Richtung um.
3. Aus den Verzweigungspunkten werden Addierer, aus den Addierern Verzweigungspunkte.
4. Eingang wird Ausgang, Ausgang wird Eingang.

Eine weitere Realisierungsform werden wir im Kapitel über die Partialbruchzerlegung kennenlernen.

Die **verschiedenen Realisierungsformen** produzieren das **gleiche Ergebnis** nur unter der Annahme, dass **numerische Fehler vernachlässigt** werden können.

Die gezeigten Realisierungen brauchen die gleiche Anzahl von Additionen und Multiplikationen und produzieren dasselbe Ausgangssignal. Das gilt allerdings nur dann, wenn wir eine präzise Arithmetik voraussetzen. Bei der Implementierung in Hardware oder Software treten auf Grund der endlichen Genauigkeit bei der Zahlendarstellung Rundungsfehler auf, was dazu führt, dass jede Implementierung sich geringfügig anders verhält. Die Implementierung von digitalen Filtern bedarf der genauen Untersuchung der Rundungsfehler und Zahlenüberläufe.

13.5 Pole und Nullstellen

Die Systemfunktion eines allgemeinen IIR-Filters lautet

$$H(z) = \frac{Y(z)}{X(z)} = \frac{b_0 + b_1 z^{-1} + b_2 z^{-2} + \dots + b_m z^{-m}}{1 - a_1 z^{-1} - a_2 z^{-2} - \dots - a_n z^{-n}} = \frac{B(z)}{A(z)} \quad (13.36)$$

Ein IIR-Filter wird durch die Lage der **Pole und Nullstellen** bis auf einen konstanten Faktor **eindeutig beschrieben**.

Die Lage der Wurzeln des Zählerpolynoms $B(z)$ (Nullstellen von $H(z)$) und des Nennerpolynoms $A(z)$ (Polstellen von $H(z)$) erlaubt Aussagen über die Eigenschaften des Filters und hat großen Einfluß auf das Filterverhalten.

Für ein IIR-Filter erster Ordnung finden wir folgende Pole und Nullstellen

$$H(z) = \frac{b_0 + b_1 z^{-1}}{1 - a_1 z^{-1}} = \frac{b_0 z + b_1}{z - a_1} \quad (13.37)$$

$$b_0 z + b_1 = 0 \quad \Rightarrow \quad z_0 = -\frac{b_1}{b_0} \quad (13.38)$$

$$z - a_1 = 0 \quad \Rightarrow \quad z_\infty = a_1 \quad (13.39)$$

$$H(z)|_{z=-b_1/b_0} = 0 \text{ (Nullstelle)} \quad H(z)|_{z=a_1} \rightarrow \infty \text{ (Polstelle)} \quad (13.40)$$

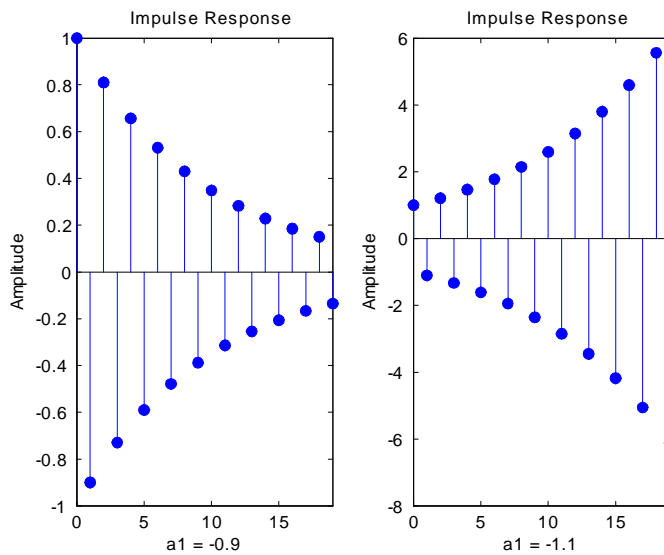


Abbildung 13.11: Stabiles $a_1 = -0.9$ und unstabiles $a_1 = -1.1$ Verhalten

13.5.1 Stabilität und Lage der Pole

Wir betrachten ein IIR-Filter erster Ordnung mit folgender Systemfunktion

$$H(z) = \frac{b_0 + b_1 z^{-1}}{1 - a_1 z^{-1}} \tag{13.41}$$

Die Impulsantwort dieses Systems ist

$$h[n] = b_0(a_1)^n \delta_{-1}[n] + b_1(a_1)^{n-1} \delta_{-1}[n-1] = \begin{cases} 0 & \text{für } n < 0 \\ b_0 & \text{für } n = 0 \\ (b_0 + b_1/a_1)a_1^n & \text{für } n \geq 1 \end{cases} \tag{13.42}$$

Aus Gleichung (13.42) sehen wir, dass $h[n]$ abklingt, wenn $|a_1| < 1$. Der Pol der Systemfunktion $H(z)$ liegt bei $z = a_1$. Die Bedingung $|a_1| < 1$ bedeutet im z -Bereich, dass der Pol innerhalb des Einheitskreises in der komplexen Ebene liegt. Abbildung 13.11 zeigt die Impulsantwort für ein stabiles und ein instabiles System.

Ein **IIR-Filter** ist genau dann **stabil**, wenn sich alle **Pole** in der komplexen Ebene **innerhalb des Einheitskreises** befinden.

Bemerkung 89 Man kann die obige Betrachtung verallgemeinern und sagen, dass ein LTI-System, das die Anfangsbedingungen aus Gleichung (13.6) erfüllt und dessen Pole innerhalb des Einheitskreises der z -Ebene liegen, immer stabil ist.

13.6 Frequenzgang von IIR-Filtern

Der Frequenzgang beschreibt die Antwort eines Systems auf die Eingangsfolge $x[n] = e^{j\hat{\omega}n}$. Wenn die Systemfunktion eines Filters gegeben ist, dann kann der Frequenzgang daraus ermittelt werden, indem man $z = e^{j\hat{\omega}}$ setzt.

Der Frequenzgang wird aus der Systemfunktion ermittelt, indem $z = e^{j\hat{\omega}}$ gesetzt wird.

$$H(\hat{\omega}) = H(z)|_{z=e^{j\hat{\omega}}} \tag{13.43}$$

Für ein IIR-Filter erster Ordnung erhalten wir

$$H(z) = \left(\frac{b_0 + b_1 z^{-1}}{1 - a_1 z^{-1}} \right) \Big|_{z=e^{j\hat{\omega}}} \Rightarrow H(e^{j\hat{\omega}}) = \frac{b_0 + b_1 e^{-j\hat{\omega}}}{1 - a_1 e^{-j\hat{\omega}}} \tag{13.44}$$

Der Frequenzgang (13.44) ist eine komplexwertige Funktion, die in der Regel in Betrag und Phase dargestellt wird. Den Betrag eines IIR-Filter erster Ordnung erhalten wir aus

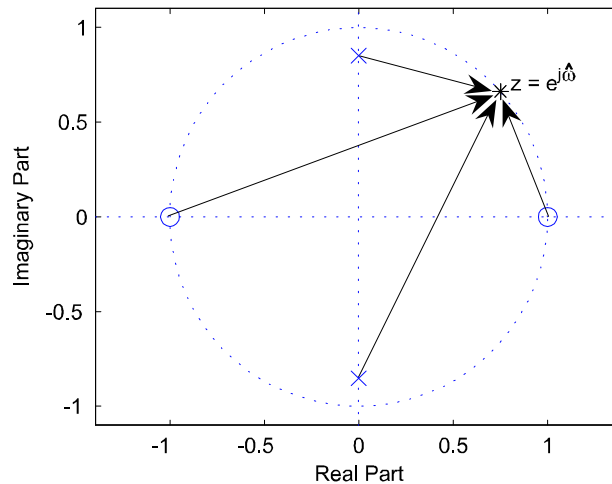


Abbildung 13.12: Grafische Ermittlung des Frequenzgangs

$$|H(e^{j\hat{\omega}})|^2 = H(e^{j\hat{\omega}})H^*(e^{j\hat{\omega}}) = \quad (13.45)$$

$$= \frac{|b_0|^2 + |b_1|^2 + 2b_0b_1 \cos \hat{\omega}}{1 + |a_1|^2 - 2a_1 \cos \hat{\omega}} \quad (13.46)$$

Für die Phase erhalten wir

$$\varphi(\hat{\omega}) = \arctan \frac{-b_1 \sin \hat{\omega}}{b_0 + b_1 \cos \hat{\omega}} - \arctan \frac{a_1 \sin \hat{\omega}}{1 - a_1 \cos \hat{\omega}} \quad (13.47)$$

Weder der Ausdruck für den Betrag, noch für die Phase sind besonders aussagekräftig und können bestenfalls als Ausgangspunkt für die Berechnung der graphischen Darstellung genommen werden.

Eine andere Möglichkeit besteht darin, mit Hilfe der Pole und Nullstellen den ungefähren Verlauf des Frequenzgangs zu skizzieren. Zur Erläuterung der Vorgehensweise gehen wir von einem einfachen IIR-Filter mit zwei Polen und zwei Nullstellen aus.

$$H(z) = \frac{(z - z_1)(z - z_2)}{(z - p_1)(z - p_2)} \quad (13.48)$$

$$|H(e^{j\hat{\omega}})| = \frac{|e^{j\hat{\omega}} - z_1| |e^{j\hat{\omega}} - z_2|}{|e^{j\hat{\omega}} - p_1| |e^{j\hat{\omega}} - p_2|} \quad (13.49)$$

Die Terme $|e^{j\hat{\omega}} - z_i|$ sind Längen der Zeiger von den Nullstellen z_i zum Punkt $e^{j\hat{\omega}}$, der auf dem Einheitskreis liegt, die Terme $|e^{j\hat{\omega}} - p_i|$ sind Längen der Zeiger von den Polstellen p_i zum Punkt $e^{j\hat{\omega}}$. Abbildung 13.12 zeigt den Zusammenhang.

Der Betrag des Frequenzgangs $|H(e^{j\hat{\omega}})|$ ist das Produkt der Längen der Zeiger von den Nullstellen zum Punkt $e^{j\hat{\omega}}$, dividiert durch das Produkt der Längen der Zeiger von den Polstellen zum Punkt $e^{j\hat{\omega}}$. Unter Verwendung dieser Beziehung kann man den Betrag des Frequenzgangs punktweise ermitteln.

13.7 3D-Darstellung der Systemfunktion

Der **Betrag der Systemfunktion** kann als **Fläche über der komplexen Zahlenebene** (der z -Ebene) dargestellt werden.

Eine für Verständniszwecke anschauliche Darstellung liefert die 3D-Grafik des Betrags der Systemfunktion.

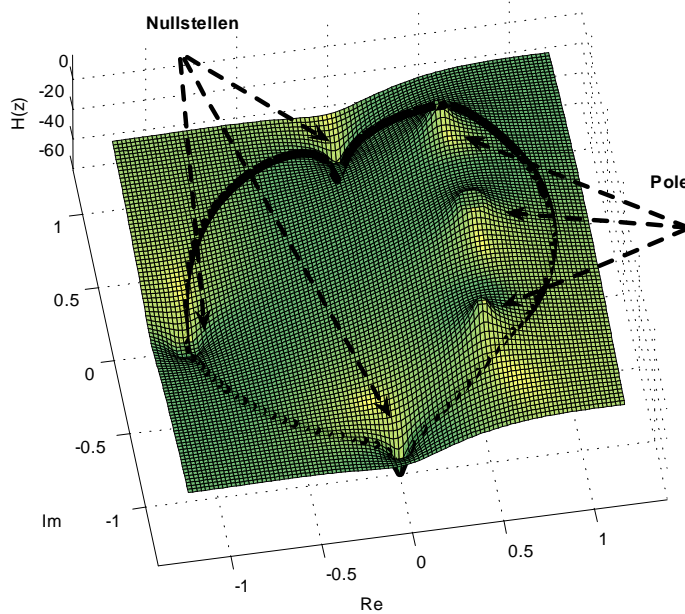


Abbildung 13.13: 3D-Plot des Betrags eines IIR-Filters 3.Ordnung

Abbildung 13.13 zeigt den 3D-Plot des Betrags der Systemfunktion eines IIR-Filters mit drei Pol- und drei Nullstellen. Der Betrag ist logarithmisch dargestellt, um die Nullstellen besser erkennen zu können.

Das zu Abbildung 13.13 gehörende Pol-/Nullstellendiagramm ist in Abbildung 13.14 gezeigt.

Der Frequenzgang des Filters] (in linearem Maßstab) der Abbildung 13.13 ist in Abbildung 13.15 gezeigt. (Da Nullstellen auf dem Einheitskreis liegen, muss die Übertragungsfunktion für die zugehörigen Frequenzen Null sein, wie man auch im Frequenzgang bei 0.5 und 1 sehen kann. In der logarithmischen Darstellung in Abbildung 13.13 wird der Wert bei diesen Frequenzen $-\infty$, da $\log(0) = -\infty$.)

Der **Frequenzgang** befindet sich **über dem Einheitskreis** in der z -Ebene.

Die Sprungantwort dieses Filters zeigt Abbildung 13.16.

13.8 Die inverse z –Transformation

Wir fassen unser Wissen über die z –Transformation zusammen

$$\text{Linearität } ax_1[n] + bx_2[n] \Leftrightarrow aX_1(z) + bX_2(z) \quad (13.50)$$

$$\text{Verschiebung im Zeitbereich } x[n - n_0] \Leftrightarrow z^{-n_0} X(z) \quad (13.51)$$

$$\text{Einheitsimpuls } \delta_0[n] \Leftrightarrow 1 \quad (13.52)$$

$$\text{Verschobener Einheitsimpuls } \delta_0[n - n_0] \Leftrightarrow z^{-n_0} \quad (13.53)$$

$$\text{Faltung } x[n] * h[n] \Leftrightarrow X(z)H(z) \quad (13.54)$$

$$a^n \delta_{-1}[n] \Leftrightarrow \frac{1}{1 - az^{-1}} \quad (13.55)$$

$$\text{Sprungfunktion } \delta_{-1}[n] \Leftrightarrow \frac{1}{1 - z^{-1}} \quad (13.56)$$

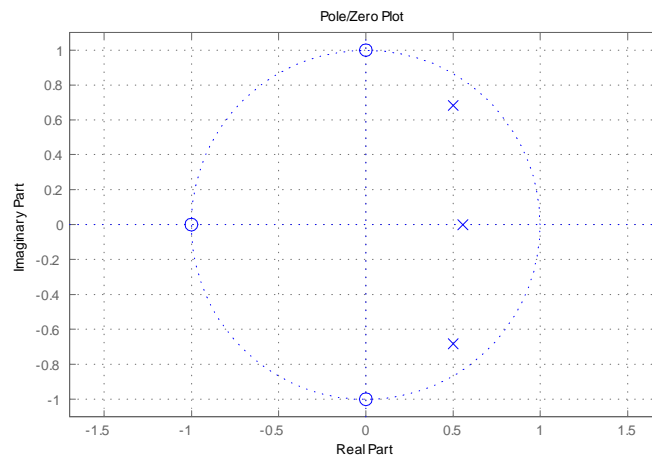


Abbildung 13.14: Pol-/Nullstellendiagramm des IIR-Filters 3. Ordnung

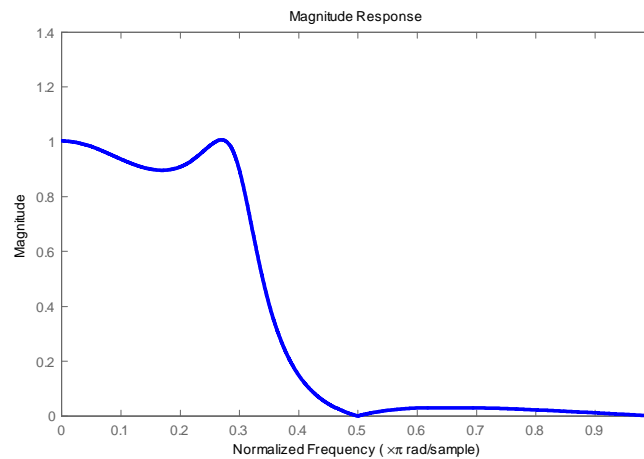


Abbildung 13.15: Frequenzgang des Filters 3. Ordnung

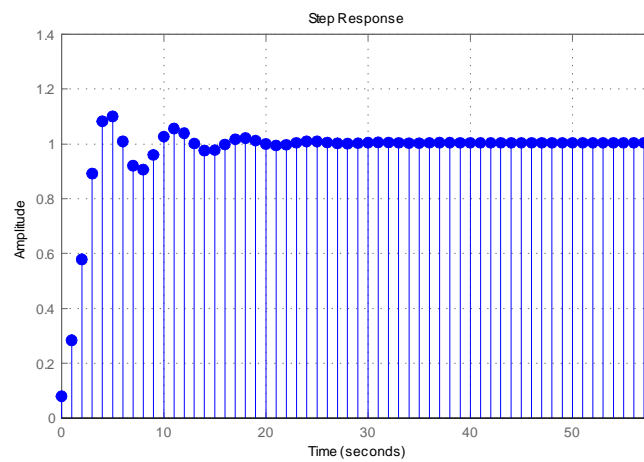


Abbildung 13.16: Sprungantwort IIR-Filter 3. Ordnung

13.8.1 Partialbruchzerlegung

Mit Hilfe der Partialbruchzerlegung wird eine Systemfunktion höherer Ordnung in eine Summe von gebrochen rationalen Funktionen 1. Ordnung³ zerlegt.

$$H(z) = \frac{B(z)}{A(z)} = \frac{B(z)}{(1-p_1z^{-1})(1-p_2z^{-1})\dots(1-p_nz^{-1})} = \quad (13.57)$$

$$= \frac{R_1}{(1-p_1z^{-1})} + \frac{R_2}{(1-p_2z^{-1})} + \dots + \frac{R_n}{(1-p_nz^{-1})} + K \quad (13.58)$$

Wegen der Linearität der z -Transformation und unter Verwendung der Beziehungen

$$\frac{1}{1-az^{-1}} \Leftrightarrow a^n \delta_{-1}[n] \quad \text{und} \quad R \cdot H(z) \Leftrightarrow R \cdot h[n] \quad (13.59)$$

kann man $H(z)$ in $h[n]$ transformieren und damit die Impulsantwort von $H(z)$ ohne »Erraten« der Lösung der Differenzgleichung auf algebraischem Weg finden.

Ein Beispiel erläutert die Vorgangsweise. Zuerst zerlegen wir das Nennerpolynom in seine Wurzelarstellung⁴

$$H(z) = \frac{b_0 + b_1z^{-1}}{(1-p_1z^{-1})(1-p_2z^{-1})} \quad (13.60)$$

Daraus finden wir die Partialbruchdarstellung

$$H(z) = \frac{R_1}{(1-p_1z^{-1})} + \frac{R_2}{(1-p_2z^{-1})} \quad (13.61)$$

Die Größen R_1 und R_2 nennt man Residuen von $H(z)$. Die Bestimmung der Residuen R_1 und R_2 kann durch Ausmultiplizieren und Koeffizientenvergleich erfolgen. Aus (13.61) erhalten wir zwei Gleichungen mit den Unbekannten R_1 und R_2 .

Es gibt aber einen eleganteren und schnelleren Weg. Wir formen um, indem wir beide Seiten der Gleichung (13.60) und (13.61) mit $(1-p_1z^{-1})$ multiplizieren

$$\begin{aligned} H(z)(1-p_1z^{-1}) &= \frac{b_0 + b_1z^{-1}}{(1-p_2z^{-1})} = \\ &= R_1 + R_2 \frac{(1-p_1z^{-1})}{(1-p_2z^{-1})} \end{aligned} \quad (13.62)$$

Gleichung (13.62) muss für alle z gelten, insbesondere auch an der Stelle des Pols $z = p_1$ und wir berechnen den Wert an dieser Stelle

$$\begin{aligned} H(z)(1-p_1z^{-1})|_{z=p_1} &= \frac{b_0 + b_1z^{-1}}{(1-p_2z^{-1})}|_{z=p_1} = \\ &= R_1 + R_2 \frac{(1-p_1z^{-1})}{(1-p_2z^{-1})}|_{z=p_1} = R_1 \end{aligned} \quad (13.63)$$

Der Ausdruck $(1-p_1z^{-1})|_{z=p_1}$ wird Null, damit verschwindet der letzte Bruch und wir erhalten

$$R_1 = \frac{b_0 + b_1/p_1}{(1-p_2/p_1)}$$

³Fasst man konjugiert komplexe Pole zusammen, um reelle Koeffizienten für die Partialbruchzerlegung zu erhalten, dann entstehen Teilfunktionen 2. Ordnung.

⁴Die Berechnung der Nullstellen eines allgemeinen Polynomes kann nur bis zur 3. Ordnung, in Sonderfällen auch bis zu 4. Ordnung geschlossen erfolgen. Bei höheren Ordnungen können die Nullstellen nur mehr numerisch berechnet werden. Jedes Polynom der Ordnung n hat n Nullstellen, die entweder reell oder konjugiert komplex sind.

Durch die **Partialbruchzerlegung** wird ein System höherer Ordnung in Systeme 1. oder 2. Ordnung zerlegt.

Es gibt mehrere Methoden, um die Partialbruchzerlegung durchzuführen.

In ähnlicher Weise können wir R_2 berechnen und erhalten

$$R_2 = \frac{b_0 + b_1/p_2}{1 - p_1/p_2}$$

Beispiel 90 Wir führen die Rechnung mit einem Zahlenbeispiel durch

$$H(z) = \frac{-1 + 1.7z^{-1}}{1 - 0.4z^{-1} - 0.21z^{-2}} \quad (13.64)$$

$$H(z) = \frac{-1 + 1.7z^{-1}}{(1 + 0.3z^{-1})(1 - 0.7z^{-1})} \quad (13.65)$$

Daraus finden wir die Partialbruchdarstellung

$$H(z) = \frac{R_1}{(1 + 0.3z^{-1})} + \frac{R_2}{(1 - 0.7z^{-1})} \quad (13.66)$$

Wir formen um, indem wir beide Seiten der Gleichungen mit $(1 + 0.3z^{-1})$ multiplizieren

$$H(z)(1 + 0.3z^{-1}) = \frac{-1 + 1.7z^{-1}}{(1 - 0.7z^{-1})} = \quad (13.67)$$

$$= R_1 + R_2 \frac{(1 + 0.3z^{-1})}{(1 - 0.7z^{-1})} \quad (13.68)$$

Wir berechnen nun den Wert des obigen Ausdrucks an der Stelle des Pols $z = -0.3$

$$H(z)(1 + 0.3z^{-1})|_{z=-0.3} = \frac{-1 + 1.7z^{-1}}{(1 - 0.7z^{-1})}|_{z=-0.3} = \quad (13.69)$$

$$= R_1 + R_2 \frac{(1 + 0.3z^{-1})}{(1 - 0.7z^{-1})}|_{z=-0.3} = R_1 \quad (13.70)$$

$$R_1 = -2 \quad (13.71)$$

$$H(z)(1 - 0.7z^{-1})|_{z=0.7} = \frac{-1 + 1.7z^{-1}}{(1 + 0.3z^{-1})}|_{z=0.7} = \quad (13.72)$$

$$= R_1 \frac{(1 - 0.7z^{-1})}{(1 + 0.3z^{-1})}|_{z=0.7} + R_2 = R_2 \quad (13.73)$$

$$R_2 = 1 \quad (13.74)$$

$$H(z) = \frac{-2}{(1 + 0.3z^{-1})} + \frac{1}{(1 - 0.7z^{-1})} \quad (13.75)$$

Durch Rücktransformation unter Verwendung von

$$a^n \delta_{-1}[n] \Leftrightarrow \frac{1}{1 - az^{-1}} \quad (13.76)$$

erhalten wir die Impulsantwort

$$h[n] = (-2)(-0.3)^n \delta_{-1}[n] + 0.7^n \delta_{-1}[n] \quad (13.77)$$

Zur Darstellung berechnen wir $h[n]$ numerisch mit der Matlab-Funktion `impz`
 $B = [-1 \ 1.7]$; $A = [1 \ -0.4 \ -0.21]$; `impz(B,A)` und erhalten Abbildung 13.17

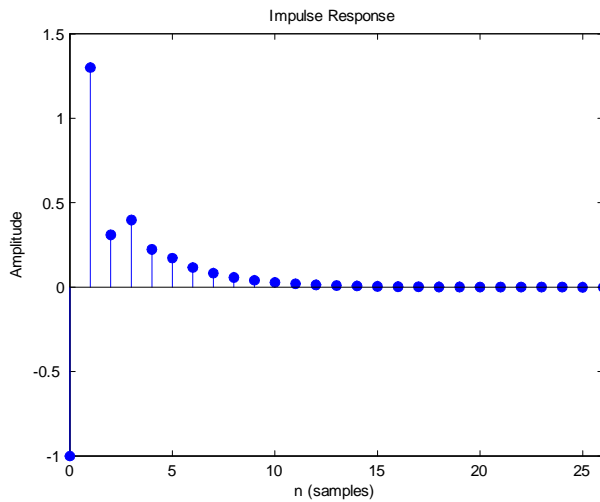


Abbildung 13.17: Impulsantwort $H(z) = \frac{-1+1.7z^{-1}}{1-0.4z^{-1}-0.21z^{-2}}$

Beispiel 91 Im vorigen Beispiel war der Zählergrad kleiner als der Nennergrad. Wenn das nicht der Fall ist, muss zuerst Zähler- durch Nennerpolynom dividiert werden, um einen kleineren Zählergrad zu erreichen. Ein Beispiel zeigt den Rechenvorgang

$$H(z) = \frac{1 + 0.9z^{-1} - 0.42z^{-2}}{1 - 0.4z^{-1} - 0.21z^{-2}} \tag{13.78}$$

Polynomdivision

$$(-0.42z^{-2} + 0.9z^{-1} + 1) : (-0.21z^{-2} - 0.4z^{-1} + 1) = 2 \tag{13.79}$$

$$-(-0.42z^{-2} - 0.8z^{-1} + 2) : \tag{13.80}$$

$$(1.7z^{-1} - 1) : \tag{13.81}$$

$$H(z) = 2 + \frac{-1 + 1.7z^{-1}}{1 - 0.4z^{-1} - 0.21z^{-2}} = 2 + \frac{-1 + 1.7z^{-1}}{(1 + 0.3z^{-1})(1 - 0.7z^{-1})} \tag{13.82}$$

Der weitere Verlauf der Rechnung ist wie im vorigen Beispiel. Die Impulsantwort ist

$$h[n] = 2\delta[n] + (-2)(-3)^n\delta_{-1}[n] + 0.7^n\delta_{-1}[n] \tag{13.83}$$

Zusammenfassung Partialbruchzerlegung

Wir fassen die Schritte bei der inversen z-Transformation zusammen:

1. Zählergrad M muss kleiner Nennergrad N sein. (Wenn das nicht der Fall ist, muss vorher durch Polynomdivision die Bedingung $M < N$ hergestellt werden.)
2. Bestimmen der Nullstellen und Darstellung des Nennerpolynoms in der Form $(1 - p_k z^{-1})$, $k = 1, 2, \dots, N$
3. Partialbruchzerlegung von $H(z)$

$$H(z) = \sum_{k=1}^N \frac{R_k}{(1 - p_k z^{-1})} \quad R_k = H(z)(1 - p_k z^{-1})|_{z=p_k} \tag{13.84}$$

Die Rücktransformation ergibt die Impulsantwort

$$h[n] = \sum_{k=1}^N R_k (p_k)^n \delta_{-1}[n] \tag{13.85}$$

Die **Partialbruchzerlegung** ist nur dann möglich, wenn der **Zählergrad kleiner als der Nennergrad** ist. Is es nicht der Fall, ist eine Polynomdivision nötig.

13.9 Einschwingvorgänge und Stabilität

Bei **IIR-Filtern** ist die **Einschwingzeit unendlich lang**.

Beim Anlegen einer eingeschalteten Sinusschwingung an FIR-Filter haben wir gesehen, dass sich das Filter nach der Einschwingzeit so verhält, also ob die Sinusschwingung ständig angelegt gewesen wäre. Die Dauer der Einschwingzeit ist gleich der Anzahl der Verzögerungsglieder des Filters. Die Berechnung des Einschwingvorgangs haben wir so durchgeführt, dass wir während der Einschwingphase die Differenzgleichung des Filters (mit der eingeschalteten Sinusschwingung als Eingangssignal) gelöst haben, nach der Einschwingzeit konnten wir den weiteren Verlauf mit Hilfe der Frequenzantwort des Filters bestimmen.

Für die Berechnung des Einschwingverhaltens von IIR-Filtern bedienen wir uns der z -Transformation. Wir erläutern die Vorgangsweise an einem System erster Ordnung

$$H(z) = \frac{b_0}{1 - a_1 z^{-1}} \quad (13.86)$$

Die mathematische Darstellung der eingeschalteten komplexen Exponentialfunktion der normierten Kreisfrequenz $\hat{\omega}_0$ lautet

$$x[n] = e^{j\hat{\omega}_0 n} \delta_{-1}[n] \quad (13.87)$$

Unter Verwendung des Transformationspaars

$$a^n \delta_{-1}[n] \Leftrightarrow \frac{1}{1 - az^{-1}} \quad (13.88)$$

wird daraus

$$X(z) = \frac{1}{1 - e^{j\hat{\omega}_0} z^{-1}} \quad (13.89)$$

Das Ausgangssignal ist

$$Y(z) = H(z)X(z) = \frac{b_0}{1 - a_1 z^{-1}} \cdot \frac{1}{1 - e^{j\hat{\omega}_0} z^{-1}} = \frac{b_0}{(1 - a_1 z^{-1})(1 - e^{j\hat{\omega}_0} z^{-1})} \quad (13.90)$$

Wir führen die Partialbruchzerlegung durch und transformieren dann in den Zeitbereich

$$Y(z) = \frac{\left(\frac{b_0 a_1}{a_1 - e^{j\hat{\omega}_0}}\right)}{(1 - a_1 z^{-1})} + \frac{\left(\frac{b_0}{1 - a_1 e^{-j\hat{\omega}_0}}\right)}{(1 - e^{j\hat{\omega}_0} z^{-1})} \quad (13.91)$$

$$y[n] = \underbrace{\left(\frac{b_0 a_1}{a_1 - e^{j\hat{\omega}_0}}\right) (a_1)^n \delta_{-1}[n]}_{\text{Transienter Anteil}} + \underbrace{\left(\frac{b_0}{1 - a_1 e^{-j\hat{\omega}_0}}\right) e^{j\hat{\omega}_0 n} \delta_{-1}[n]}_{\text{Eingeschwungener Zustand}} \quad (13.92)$$

Die **Systemantwort** eines IIR-Filters besteht aus einem **transienten Anteil** und einem **Anteil des eingeschwungenen Systems**.

Gleichung (13.92) besteht aus zwei Termen: Der erste Term ist die Folge $(\dots)(a_1)^n$, der zweite Term ist das Eingangssignal – die komplexe Exponentialfunktion – multipliziert mit dem Frequenzgang des Filters.

Der erste Term klingt für $n \rightarrow \infty$ ab, wenn $|a_1| < 1$ und wird daher transienter Anteil genannt. Wir sprechen von einem stabilen System, wenn der Einschwinganteil abklingt. Das immer dann der Fall, wenn der Pol (die Pole) der Systemfunktion innerhalb des Einheitskreises in der z -Ebene liegen. Liegt der Pol außerhalb des Einheitskreises, dann steigt $(a_1)^n$ an und das Filter ist nicht stabil.

Beispiel 92 Wir zeigen die Zusammenhänge an einem Beispiel. Für das Tiefpassfilter erster Ordnung mit den Koeffizienten $b_0 = 1$, $a_1 = -0.9$ und der komplexen Exponentialfunktion $\hat{\omega}_0 = 2\pi/10$ erhalten wir durch Einsetzen in (13.92)

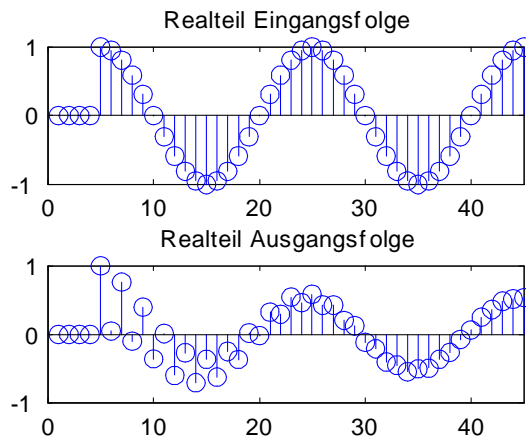


Abbildung 13.18: Antwort des Systems $H(z) = \frac{1}{1+0.9z^{-1}}$ auf eingeschaltetes Kosinussignal

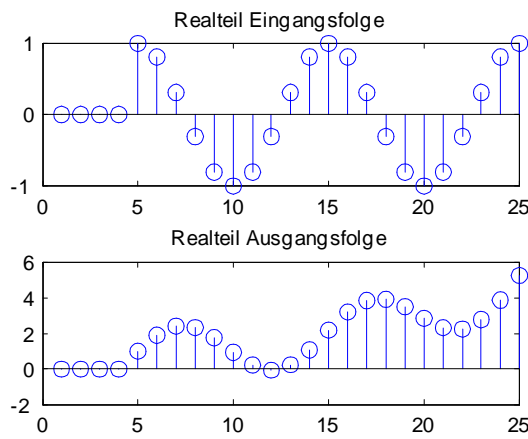


Abbildung 13.19: Instabiles Filter Polstelle bei $z_P = 1.1$

$$\begin{aligned}
 y[n]_1 &= \left(\frac{-0.9}{-0.9 - e^{j0.2\pi}} \right) (-0.9)^n \delta_{-1}[n] = 0.4980 e^{-j0.3313} (-0.9)^n \delta_{-1}[n] \\
 y[n]_2 &= \left(\frac{1}{1 + 0.9e^{-j0.2\pi}} \right) e^{j0.2\pi n} \delta_{-1}[n] = 0.5533 e^{j0.2971} e^{j0.2\pi n} \delta_{-1}[n] = \\
 &= 0.5533 \cos(0.2\pi n + 0.2971) \delta_{-1}[n] + j0.5533 \sin(0.2\pi n + 0.2971) \delta_{-1}[n] \\
 y[n] &= y[n]_1 + y[n]_2
 \end{aligned}$$

Abbildung 13.18 zeigt den Realteil des Eingang- und Ausgangssignals.

Abbildung 13.19 zeigt ein instabiles Filter mit einem Pol bei $a_1 = 1.1$.

Im allgemeinen Fall erhalten wir für die eingeschaltete komplexe Exponentialfolge $x[n] = e^{j\tilde{\omega}_0 n} \delta_{-1}[n]$ die Ausgangsfolge eines Systems N -ter Ordnung mit Polstellen bei p_k

Bei einem stabilen Filter klingt der transiente Anteil mit der Zeit ab.

$$y[n] = \sum_{k=1}^N R_k(p_k)^n \delta_{-1}[n] + H(e^{j\tilde{\omega}_0 n}) e^{j\tilde{\omega}_0 n} \delta_{-1}[n] \quad (13.93)$$

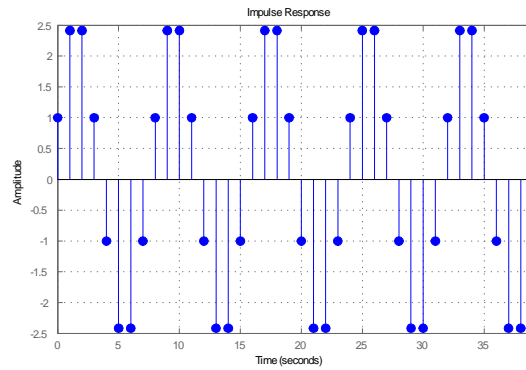


Abbildung 13.20: Impulsantwort IIR-Filter, Pole auf dem Einheitskreis

13.10 Konjugiert komplexe Pole

Bisher haben wir Pole auf der reellen Achse betrachtet, Pole können aber nicht nur reell sondern auch konjugiert komplex auftreten.

Beispiel 93 Als Beispiel untersuchen wir ein System mit konjugiert komplexen Polen auf dem Einheitskreis bei $\hat{\omega} = \pi/4$

$$H(z) = \frac{1 + z^{-1}}{(1 - e^{j\pi/4}z^{-1})(1 - e^{-j\pi/4}z^{-1})} = \frac{1 + z^{-1}}{1 - 1.4142z^{-2} + z^{-2}} \quad (13.94)$$

Die Partialbruchzerlegung ergibt

$$H(z) = \frac{1.3066e^{-j1.1781}}{(1 - e^{j\pi/4}z^{-1})} + \frac{1.3066e^{j1.1781}}{(1 - e^{-j\pi/4}z^{-1})} \quad (13.95)$$

Die Impulsantwort erhalten wir durch inverse z -Transformation

$$\begin{aligned} h[n] &= 1.3066e^{-j1.1781}e^{(j\pi/4)n}\delta_{-1}[n] + \\ &\quad + 1.3066e^{j1.1781}e^{-j(\pi/4)n}\delta_{-1}[n] \\ h[n] &= 2 * 1.3066 \cos\left(\frac{\pi}{4}n - 1.1781\right)\delta_{-1}[n] \end{aligned} \quad (13.96)$$

Pole können entweder **reell** sein oder in **konjugiert komplexen** Paaren auftreten.

Wie Gleichung (13.96) zeigt, erzeugt das mit einem Puls angeregte System eine Sinusschwingung! Die Impulsantwort ist die Sinusschwingung der Gleichung (13.96), dennoch ist dieses System nicht stabil, da es für das Eingangssignal der Frequenz $\pi/4$ ein unbegrenztes Ausgangssignal erzeugt (Resonanz).

Abbildung 13.20 zeigt die Impulsantwort unseres Beispiels.

Beispiel 94 Als weiteres Beispiel untersuchen wir ein System mit Polen innerhalb des Einheitskreises.

$$H(z) = \frac{1 + z^{-1}}{1 - z^{-1} + 0.5z^{-2}} = \frac{1 + z^{-1}}{\left(1 - \frac{1}{\sqrt{2}}e^{j\pi/4}z^{-1}\right)\left(1 - \frac{1}{\sqrt{2}}e^{-j\pi/4}z^{-1}\right)} \quad (13.97)$$

$$= \frac{1.5811e^{-j1.2490}}{\left(1 - \frac{1}{\sqrt{2}}e^{j\pi/4}z^{-1}\right)} + \frac{1.5811e^{j1.2490}}{\left(1 - \frac{1}{\sqrt{2}}e^{-j\pi/4}z^{-1}\right)} \quad (13.98)$$

$$h[n] = 2 \cdot 1.5811 \left(\frac{1}{\sqrt{2}}\right)^n \cos\left(\frac{\pi}{4}n - 1.249\right) \quad (13.99)$$

Die Impulsantwort von $\frac{1+z^{-1}}{1-z^{-1}+0.5z^{-2}}$ zeigt Abbildung 13.21.

Wie man in Abbildung 13.21 erkennen kann, klingt die Impulsantwort ab. Wir haben ein stabiles System, da die Pole innerhalb des Einheitskreises liegen.

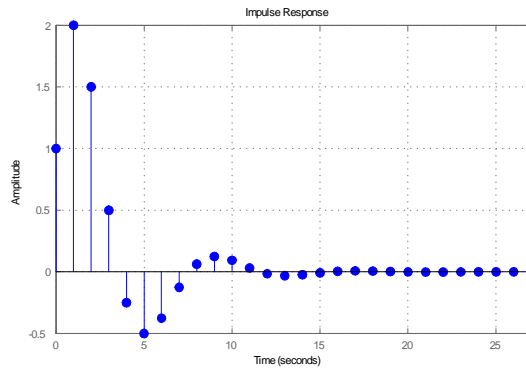


Abbildung 13.21: Impulsantwort

13.10.1 z-Transformierte komplexer Pole

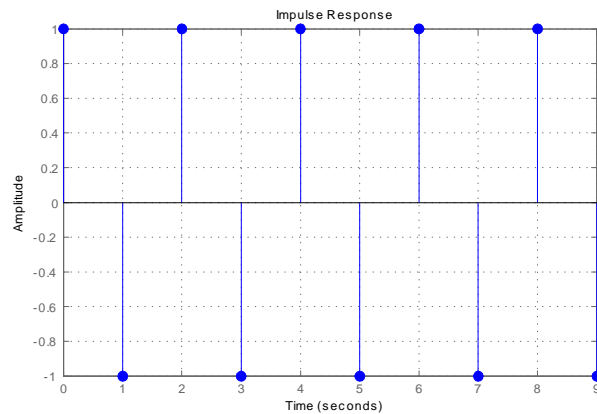
Die Transformationsbeziehungen für komplexe Pole lauten allgemein

$$\begin{aligned}
 (\sin \omega_0 n) \delta_{-1}[n] &\Leftrightarrow \frac{(\sin \omega_0) z^{-1}}{1 - 2(\cos \omega_0) z^{-1} + z^{-2}} \\
 (\cos \omega_0 n) \delta_{-1}[n] &\Leftrightarrow \frac{1 - (\cos \omega_0) z^{-1}}{1 - 2(\cos \omega_0) z^{-1} + z^{-2}} \\
 (r^n \cos \omega_0 n) \delta_{-1}[n] &\Leftrightarrow \frac{1 - (r \cos \omega_0) z^{-1}}{1 - (2r \cos \omega_0) z^{-1} + r^2 z^{-2}} \\
 (r^n \sin \omega_0 n) \delta_{-1}[n] &\Leftrightarrow \frac{(r \sin \omega_0) z^{-1}}{1 - (2r \cos \omega_0) z^{-1} + r^2 z^{-2}} \\
 K \cdot r^n \cos(\omega_0 k + \varphi) \delta_{-1}[n] &\Leftrightarrow \frac{0.5K e^{j\varphi}}{1 - r e^{j\omega_0} z^{-1}} + \frac{0.5K e^{-j\varphi}}{1 - r e^{-j\omega_0} z^{-1}} \\
 K \cdot r^n \cos(\omega_0 k + \varphi) \delta_{-1}[n] &\Leftrightarrow \frac{b_0 + b_1 z^{-1}}{1 + 2a z^{-1} + r^2 z^{-2}} \\
 K &= \sqrt{\frac{b_1^2 r^2 + b_0^2 - 2a b_0 b_1}{r^2 - a^2}} \\
 \omega_0 &= \arccos\left(\frac{-a}{r}\right) \\
 \varphi &= \arctan \frac{ab_0 - b_1}{r^2 - a^2}
 \end{aligned}$$

Beispiel 95 Als Beispiel berechnen wir die Impulsantwort von $H(z) = \frac{1}{1+z^{-2}}$. Die Tabelle der Transformationsbeziehungen legt den Lösungsansatz $h[n] = \cos \omega_0 n$ nahe.

$$\begin{aligned}
 \cos(\omega_0 n) &\Leftrightarrow \frac{1 - (\cos \omega_0) z^{-1}}{1 - 2(\cos \omega_0) z^{-1} + z^{-2}} \\
 \cos \omega_0 &= 0 \quad \omega_0 = \arccos 0 = \pi/2 \\
 h(n) &= \cos\left(\frac{\pi}{2} n\right)
 \end{aligned}$$

Abbildung 13.22 zeigt die Impulsantwort unseres Beispiels.

Abbildung 13.22: Impulsantwort $H(z) = \frac{1}{1+z^{-1}}$

13.11 Blockdiagramm der Partialbruchzerlegung

Ein IIR-Filter kann mit Hilfe der **Partialbruchzerlegung** als eine **Parallelschaltung** von Filtern **1. und 2. Ordnung** realisiert werden.

Die Systemfunktion eines IIR-Filters in Partialbruchzerlegung führt auf die Form

$$H(z) = b_{00} + \frac{b_{01}}{1 - a_{11}z^{-1}} + \dots + \frac{b_{02} + b_{12}z^{-1}}{1 - a_{12}z^{-1} - a_{22}z^{-2}} + \dots \quad (13.100)$$

mit dem konstanten Faktor b_{00} , einem Block für einen reellen Pol und einem Block für die konjugiert komplexen Pole. Abbildung 13.23 zeigt das zugehörige Blockdiagramm.

Für mehrere einfache bzw. konjugiert komplexe Pole sind die entsprechenden Blöcke parallel zu schalten.

13.12 Darstellung von Systemen

Ein System kann auf mehrere Arten eindeutig beschrieben werden.

Wir haben mehrere Formen der Darstellung von Systemen kennengelernt. Jede dieser Formen beschreibt das System eindeutig, welche Darstellung gewählt wird, hängt von der Aufgabenstellung ab.

Die einzelnen Darstellungsformen werden am Beispiel eines IIR-Filters 2. Ordnung zusammenfassend dargestellt.

13.12.1 Zeitbereich

$$y[n] = a_1y[n-1] + a_2y[n-2] + b_0x[n] + b_1x[n-1] + b_2x[n-2] \quad (13.101)$$

Im Zeitbereich wird das Eingangs-/Ausgangsverhalten des Filters dargestellt. Der Zeitbereich dient der Darstellung des »wirklichen« (physikalischen) Signals. Das Ausgangssignal wird entweder durch Lösung der Differenzgleichung (13.101) gefunden oder durch Faltung des Eingangssignals mit der Impulsantwort des Systems ermittelt.

$$y(n) = x[n] * h[n] \quad (13.102)$$

Bei FIR-Filtern ist die Impulsantwort die Folge der Koeffizienten

$$h[n] = \sum_{k=0}^M b_k \delta_0[n-k] \quad (13.103)$$

Bei IIR-Filtern ist die Impulsantwort die Lösung der Differenzgleichung

$$h[n] = a_1h[n-1] + a_2h[n-2] + b_0\delta_0[n] + b_1\delta_0[n-1] + b_2\delta_0[n-2] \quad (13.104)$$

Diese DGL lässt sich im Zeitbereich in der Regel nur schwer lösen, kann aber mit Hilfe der z -Transformation und der Partialbruchzerlegung leicht gelöst werden.

Zeitbereich:

Impulsantwort,
Sprungantwort,
Differenzgleichung

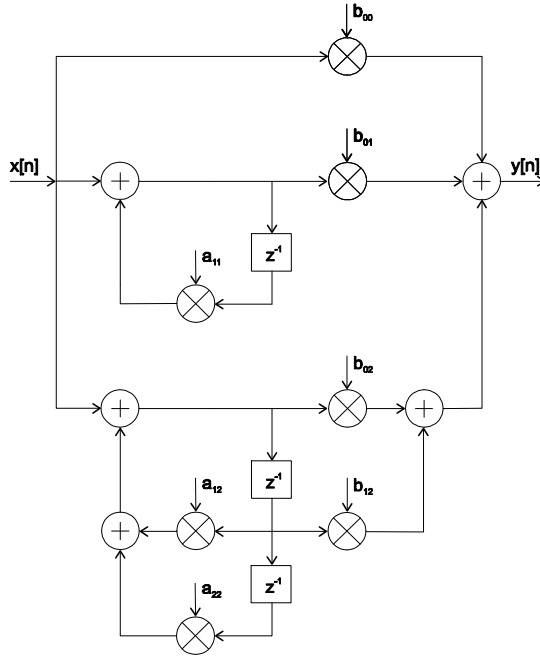


Abbildung 13.23: Blockdiagramm Partialbruchrealisierung

13.12.2 z-Bereich

$$H(z) = \frac{b_0 + b_1 z^{-1} + b_2 z^{-2}}{1 - a_1 z^{-1} - a_2 z^{-2}} \tag{13.105}$$

Die Systemfunktion $H(z)$ lässt sich durch z -Transformation der Differenzengleichung (13.104) ermitteln. Diese Transformation ist sehr einfach durchzuführen. Der z -Bereich ist ein abstrakter Bereich, der wenig anschaulich ist, sich aber gut für Aufgabenstellungen im Bereich der Analyse und Synthese von Filtern eignet.

z-Bereich:
Systemfunktion,
PN-Diagramm

Die Systemfunktion wird neben der Polynomdarstellung auch häufig in der Pol-/Nullstellenform dargestellt

$$H(z) = K \frac{(z - z_1)(z - z_2)}{(z - p_1)(z - p_2)} \tag{13.106}$$

In dieser Form sind insbesondere Aussagen über die Stabilität des Filters möglich. Anstelle der Darstellung wie in Gleichung (13.106) wird häufig eine grafische Form gewählt, die man Pol/Nullstellendiagramm nennt. Abbildung 13.24 zeigt ein Beispiel einer PN-Darstellung.

Durch Transformation in den z -Bereich ist die Ermittlung der Impulsantwort des Filters einfach möglich. Die inverse z -Transformierte von $H(z)$ ist die Impulsantwort $h[n]$.

13.12.3 Frequenzbereich

Der Frequenzgang beschreibt die Antwort des Systems auf die komplexe Exponentialfunktion. Der Frequenzbereich ermöglicht eine Aussage über das Übertragungsverhalten von Filtern für »Töne« bzw. Spektren und erlaubt die Beschreibung des Durchlass- und Sperrverhaltens des Filters. Der Frequenzbereich ist ein Sonderfall des z -Bereichs für $z = e^{j\hat{\omega}}$.

Frequenzbereich:
Frequenzgang

$$H(e^{j\hat{\omega}}) = \frac{b_0 + b_1 e^{-j\hat{\omega}} + b_2 e^{-j2\hat{\omega}}}{1 - a_1 e^{-j\hat{\omega}} - a_2 e^{-j2\hat{\omega}}} \tag{13.107}$$

Abbildung 13.25 zeigt den Frequenzgang eines IIR-Filters.

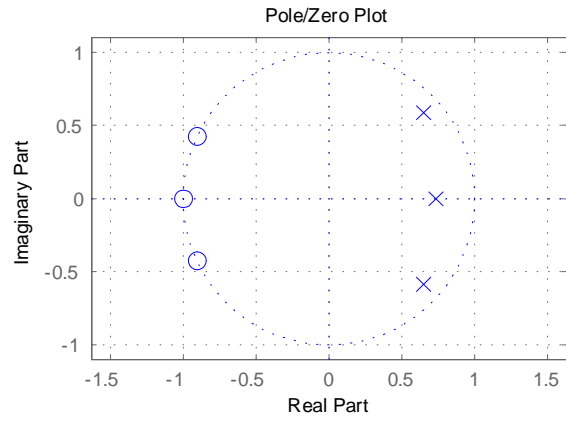


Abbildung 13.24: Pol-/Nullstellendarstellung

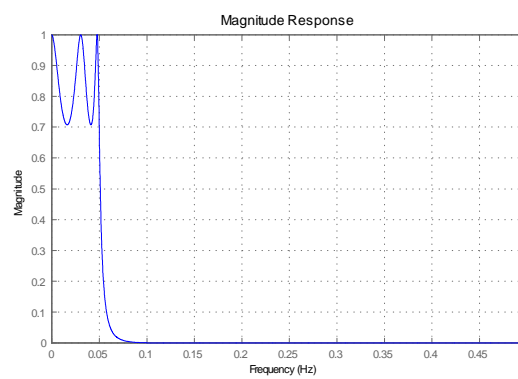


Abbildung 13.25: Elliptischer Tiefpass 5. Ordnung

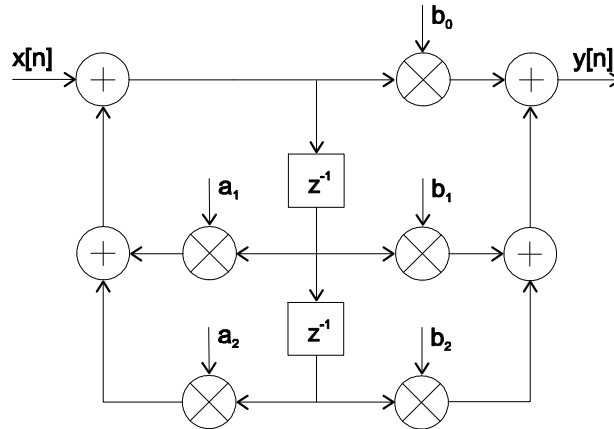


Abbildung 13.26: Blockdiagramm IIR-Filter 2. Ordnung

13.12.4 Blockdiagramm

Die Darstellung als Blockdiagramm eignet sich vor allem zur schematischen Darstellung der Hardwareblöcke eines Filters. Abbildung 13.26 zeigt ein Blockdiagramm eines IIR-Filters 2. Ordnung.

Hardware:
Blockdiagramm

13.13 Zusammenfassung

IIR oder rekursive Filter verwenden zur Berechnung der Ausgangsgröße neben Werten der Eingangsgröße auch Werte der Ausgangsgröße. Das führt in der Zeitbereichsdarstellung zu Differenzgleichungen. Während diese Differenzgleichungen numerisch leicht zu lösen sind, kann eine analytische Lösung im Zeitbereich nur durch Erraten und Überprüfen eines Lösungsansatzes gefunden werden. Ein zielführenderes Verfahren ist die Verwendung der z -Transformation. Durch diese Abbildung entstehen aus den Differenzgleichungen im Originalbereich (n -Bereich) Polynome im Bildbereich (z -Bereich). Im z -Bereich kann die Systemantwort $Y(z) = H(z)X(z)$ leicht ermittelt werden, wenn $H(z)$ und $X(z)$ bekannt sind. $y[n]$ wird durch inverse z -Transformation von $Y(z)$ ermittelt, wobei man sich vorteilhafterweise der Partialbruchzerlegung und bekannter Transformationspaare aus Tabellen bedient. Mit Hilfe dieses Verfahrens ist eine systematische Lösung der Differenzgleichungen – ohne einen Lösungsansatz erraten zu müssen – möglich.

Die inverse z -Transformierte der Systemfunktion $H(z)$ ist die Impulsantwort, da die z -Transformierte des Einheitsimpulses $\delta_0[n] \Leftrightarrow 1$ ist.

Der Frequenzgang eines LTI-Systems lässt sich dadurch berechnen, dass für $z = e^{j\omega}$ gesetzt wird. Der Frequenzbereich liegt innerhalb des z -Bereichs, $z = e^{j\omega}$, also auf dem Einheitskreis.

Die Pole und Nullstellen der Systemfunktion lassen Aufschlüsse über das Verhalten eines Systems zu. Für stabile Systeme müssen die Pole innerhalb des Einheitskreises liegen, Nullstellen auf dem Einheitskreis bedeuten Nullstellen im Frequenzgang, d.h. Eingangsfrequenzen, die den Nullstellen entsprechen werden vom System vollständig unterdrückt.

IIR-Filter haben im Vergleich zu FIR-Filtern bei gleicher Ordnung einen steileren Übergang vom Durchlassbereich in den Sperrbereich und sind daher selektiver. Der Phasengang von IIR-Filtern ist allerdings nichtlinear, sie sind daher weniger geeignet für Signale deren Kurvenform bei der Filterung erhalten bleiben muss.

Teil VI
Anhänge

Anhang A

Komplexe Zahlen

Inhalt

A.1	Darstellung komplexer Zahlen	220
A.2	Die Euler'schen Formeln	222
A.3	Rechenregeln für komplexe Zahlen	224
	A.3.1 Regeln für kartesische Koordinaten	224
	A.3.2 Regeln für Polarkoordinaten	225
A.4	Geometrische Betrachtung	226
A.5	Achtung Phase	229
A.6	Zusammenfassung	230

Komplexe Zahlen treten in der Schule zum ersten Mal bei der Lösung von quadratischen Gleichungen auf. Wir nehmen die Gleichung $x^2 + 6x + 25$ als Beispiel. Diesen Gleichungstyp können wir mit folgender Formel lösen:

$$x^2 + px + q = 0 \quad \Rightarrow \quad x_{1,2} = -\frac{p}{2} \pm \sqrt{\left(\frac{p}{2}\right)^2 - q} \quad (\text{A.1})$$

Für unsere Gleichung erhalten wir $x_{1,2} = -3 \pm \sqrt{9 - 25} = -3 \pm \sqrt{-16}$ und sehen, dass diese Gleichung keine Lösung im Reellen hat, da die Zahl unter der Wurzel negativ ist. Durch die Einführung der imaginären Einheit¹

$$j = \sqrt{-1} \quad (\text{A.2})$$

wird der Zahlenbereich erweitert und wir finden als Lösung die komplexen Zahlen $x_{1,2} = -3 \pm j4$.

In diesem Kapitel wird der Umgang mit komplexen Zahlen zusammenfassend dargestellt, insbesondere sehen wir, dass

- das Rechnen mit komplexen Zahlen den Regeln der Algebra folgt (Die imaginäre Einheit i wird wie eine Konstante behandelt.),
- die Euler'sche Formel eine Brücke zwischen der Welt der Algebra und der Welt der Geometrie bildet,
- die Zeigerdarstellung ein mächtiges Werkzeug liefert, um sinusförmige Signale darzustellen, die wiederum eine wichtige Rolle in der Elektrotechnik und Signalverarbeitung spielen.

Die Bezeichnung „komplexe Zahlen“ verleitet dazu, an komplizierte Zahlen zu denken. Komplexe Zahlen vereinfachen jedoch die Rechnung erheblich und der „imaginäre“ Anteil hilft wesentlich bei der Vorstellung und Darstellung elektrotechnischer Zusammenhänge.

¹In der Mathematik wird die imaginäre Einheit mit i abgekürzt. In der Elektrotechnik wird die imaginäre Einheit mit j abgekürzt, um Verwechslungen zu vermeiden, da i die Abkürzung für den Strom ist.

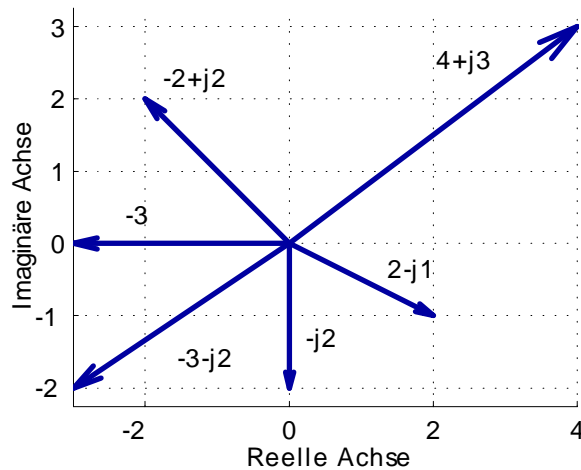


Abbildung A.1: Zeiger in der komplexen Ebene

A.1 Darstellung komplexer Zahlen

Komplexe Zahlen sind aus reellen und imaginären Zahlen zusammengesetzt. Eine komplexe Zahl wird dargestellt als

$$z(x, y) = x + jy = \operatorname{Re}\{z\} + j \operatorname{Im}\{z\} \quad (\text{A.3})$$

wobei x und y reelle Zahlen sind und Realteil und Imaginärteil von z heißen.

Anschaulich werden die komplexen Zahlen als Punkte der Ebene mit den Koordinaten x und y dargestellt. Diese Ebene wird komplexe oder Gauß'sche Zahlenebene genannt. Die x -Achse stellt dabei die reellen Zahlen, die y -Achse die imaginären Zahlen dar. Eine komplexe Zahl wird in der Gauß'schen Zahlenebene als Pfeil vom Koordinatenursprung zum Punkt mit den Koordinaten (x, y) in der komplexen Ebene dargestellt, wie Abbildung A.1 zeigt.

Für diesen "Pfeil" wählen wir den Begriff *Zeiger*, ein Zeiger stellt eine komplexe Zahl in der Zahlenebene dar. In der Literatur wird bei der Darstellung komplexer Zahlen auch häufig die Bezeichnung Vektor verwendet. Unter Vektoren versteht man in Physik und Elektrotechnik gerichtete Größen wie Kräfte, Beschleunigungen oder Feldstärken. Komplexe Zahlen sind *keine* gerichteten Größen, zur Unterscheidung von Vektoren wird daher der Begriff Zeiger verwendet.

Neben der Darstellung komplexer Zahlen in kartesischen Koordinaten ist auch die Darstellung in Polarkoordinaten möglich, wie die Abbildung A.2 zeigt. In polarer Form ist ein Zeiger durch seine Länge r und seine Richtung (φ) definiert. Die Richtung (der Winkel) kann entweder im Grad- oder im Bogenmaß angegeben werden. Beim Rechnen mit Winkeln, ist das Bogenmaß praktischer, für die Vorstellung ist das Gradmaß anschaulicher, wir müssen daher oft vom Bogenmaß in das Gradmaß umgerechnen. Der Kreisumfang (gemessen in rad) beträgt 2π und entspricht dem Winkel von 360° , der Umrechnungsfaktor vom Bogen- in das Gradmaß beträgt daher

$$\varphi [\text{rad}] \times (360/2\pi) = \varphi [\text{rad}] \times (180/\pi) \rightarrow \varphi [^\circ]. \quad (\text{A.4})$$

Da das Rechnen in Bogenmaß einfacher ist, die Vorstellung aber im Gradmaß einfacher ist, wird häufig der Umrechnungsfaktor beim Anschreiben von Winkeln explizit angeschrieben. Der Winkel von 45° erscheint dann als $45 \cdot \frac{\pi}{180}$ im Bogenmaß.

Für die Polardarstellung kann die Schreibweise $z = r \angle \varphi$ gewählt werden. Die Darstellung von Zeigern in kartesischen Koordinaten (Real- und Imaginärteil) und Polarkoordinaten (Betrag und Winkel) ist selbstverständlich äquivalent: $3 + j4 \Leftrightarrow 5 \angle 37^\circ$

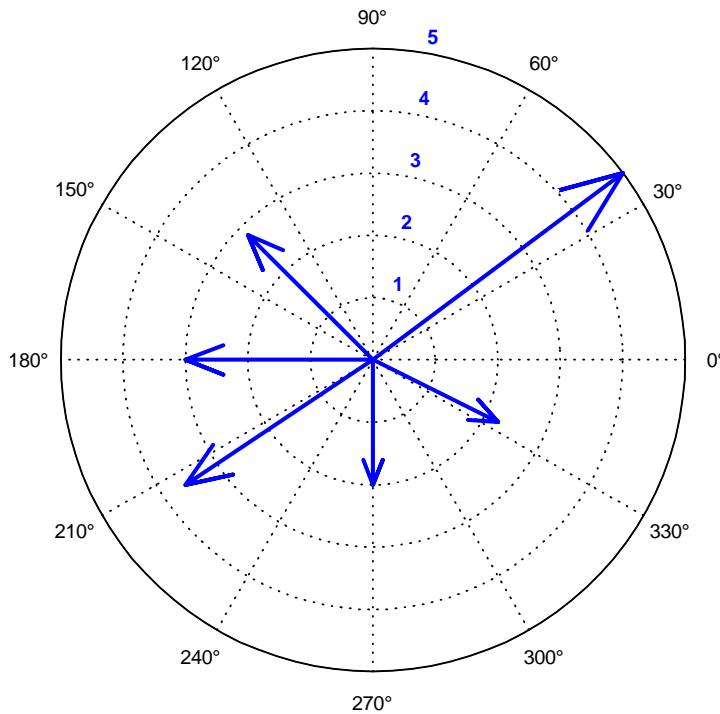


Abbildung A.2: Zeiger in Polardarstellung

Beispiel 96 Zahlen = $[4 + 3j, -2 + 2j, -3, -3 - 2j, -2j, 2 - 1j]$;
 abs(Zahlen) \Rightarrow 5.0 2.8 3.0 3.6 2.0 2.2
 Winkel = $(180/\pi) * \text{angle}(\text{Zahlen}) \Rightarrow$ 37 135 180 -146 -90 -27

Welche Form der Darstellung verwendet wird, hängt von der Rechenaufgabe ab: Addition und Subtraktion lassen sich einfacher mit Real- und Imaginärteil berechnen, Multiplikation und Division einfacher mit Betrag und Winkel. Im Zuge einer Rechnung wechselt man daher immer wieder die Koordinatensysteme.

Abbildung A.3 stellt die Zusammenhänge zwischen kartesischen und Polarkoordinaten graphisch dar

$$\begin{aligned}
 x &= r \cos \varphi; & y &= r \sin \varphi; & z &= r \cos \varphi + j (r \sin \varphi) & (A.5) \\
 r &= \sqrt{x^2 + y^2}; & \tan \varphi &= \frac{y}{x} = \frac{\text{Imaginärteil}}{\text{Realteil}}; & \varphi &= \arctan \frac{y}{x}
 \end{aligned}$$

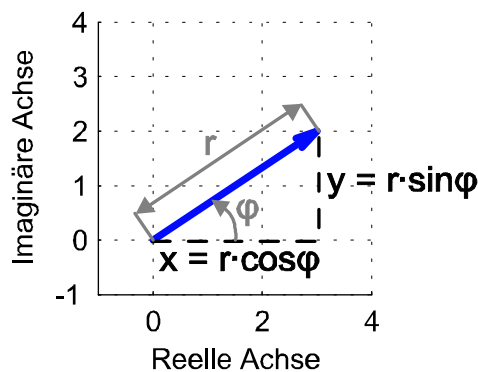


Abbildung A.3: Polar - kartesisch

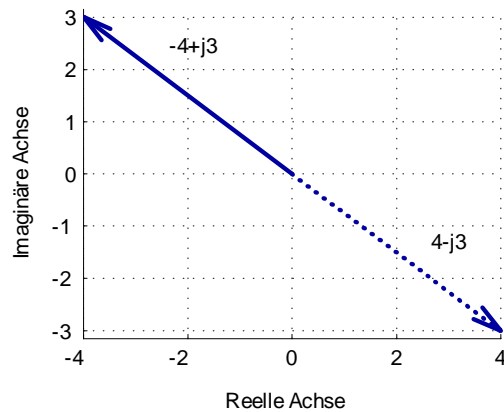


Abbildung A.4: Zeiger im 2. Quadranten

Bemerkung 97 Bei der Umrechnung von kartesischen Koordinaten in Polarkoordinaten ist Vorsicht geboten, da der Arcustangens nur zwischen -90° und $+90^\circ$ definiert ist. Abbildung A.4 zeigt worauf geachtet werden muss. Bei der Berechnung des Winkels φ des Zeigers $-4 + j3$ erhalten wir $\varphi = \arctan(\text{Im}/\text{Re}) = \arctan(-\frac{3}{4}) = -0.6435 \text{ rad} = -36.87^\circ$. Wir sehen aber aus unserer Zeichnung, dass der Winkel zwischen 90° und 180° liegen muss!

Das Argument des Arcustangens ist in unserem Beispiel negativ, da der Realteil negativ und der Imaginärteil positiv ist. Das Argument wäre aber auch dann negativ, wenn der Realteil positiv und der Imaginärteil negativ wäre. In Abbildung A.4 ist dieser Fall durch den punktierten Zeiger dargestellt. Da der arctan nur zwischen -90° und $+90^\circ$ definiert ist, ist -0.6435 auch der Winkel des Zeigers $4 - j3$!

Unser Zeiger liegt aber im zweiten Quadranten, der korrekte Winkel beträgt daher $\pi - 0.6435 \approx (180^\circ - 36.87^\circ)$. Dieses Problem tritt immer auf, wenn Zeiger im zweiten oder dritten Quadranten liegen.

Beim händischen Rechnen mit komplexen Zahlen kann man sich leicht verrechnen, Matlab stellt aber eine Reihe von Hilfsfunktionen zur Verfügung, die das Rechnen erleichtern.

Matlab-Funktion	Eigenschaft	Beispiel
<code>real(z)</code>	Realteil von z	<code>real(3-4i) ⇒ 3</code>
<code>imag(z)</code>	Imaginär von z	<code>imag(3-4i) ⇒ -4</code>
<code>abs(z)</code>	Betrag von z	<code>abs(3-4i) ⇒ 5</code>
<code>angle(z)</code>	Winkel von z	<code>angle(3-4i) ⇒ -0.9273</code>
<code>conj(z)</code>	z konjugiert komplex	<code>conj(3-4i) ⇒ 3+4i</code>

In der Matlab-Funktion `angle()` ist der Winkel von $-\pi$ bis π definiert, das arctan Problem tritt daher nicht auf.

A.2 Die Euler'schen Formeln

Mit der Schreibweise $z = r\angle\varphi$ können Zeiger in Polarform dargestellt werden. Es gibt aber eine wesentlich elegantere und leistungsfähigere Form der Darstellung, die durch die Euler'sche Formel gegeben ist.

$$e^{j\varphi} = \cos \varphi + j \sin \varphi \quad (\text{A.6})$$

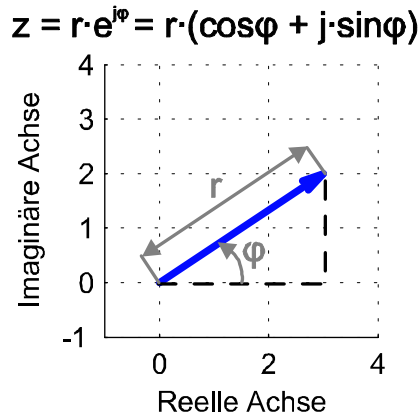


Abbildung A.5: Geometrische Interpretation der Euler'schen Formel

Die Euler'sche Formel lässt sich durch Reihenentwicklung wie folgt beweisen²

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \dots \tag{A.7}$$

Durch Einsetzen des Arguments $j\varphi$ erhalten wir

$$\begin{aligned} e^{j\varphi} &= 1 + j\varphi - \frac{\varphi^2}{2!} - j\frac{\varphi^3}{3!} + \frac{\varphi^4}{4!} + j\frac{\varphi^5}{5!} + \dots \\ &= \underbrace{1 - \frac{\varphi^2}{2!} + \frac{\varphi^4}{4!} - \dots}_{\cos \varphi} + j \underbrace{\left(\varphi - \frac{\varphi^3}{3!} + \frac{\varphi^5}{5!} - \dots\right)}_{+j \sin \varphi} \end{aligned} \tag{A.8}$$

$$\begin{aligned} \sin \varphi &= \varphi - \frac{\varphi^3}{3!} + \frac{\varphi^5}{5!} - \dots \\ \cos \varphi &= 1 - \frac{\varphi^2}{2!} + \frac{\varphi^4}{4!} - \dots \end{aligned}$$

Die obigen Zusammenhänge sind rein *algebraisch* hergeleitet, wir können aber die Euler'sche Formel (A.6) auch *geometrisch* interpretieren. Die Funktion $e^{j\varphi}$ ist äquivalent mit der Polardarstellung $1\angle\varphi$, stellt also einen Zeiger der Länge 1 und mit dem Winkel φ dar. Der Zeiger $r\angle\varphi$ ist also äquivalent mit $re^{j\varphi}$ wie in Abbildung A.5 gezeigt.

Bemerkung 98 *Beachten Sie, dass der Beweis der Richtigkeit der Euler'schen Formel auf Basis der Algebra erbracht wurde! Die Zahlendarstellung in der komplexen Ebene liefert die geometrische Interpretation. Die Euler'sche Formel bildet eine Brücke zwischen der Welt der Algebra und der Welt der Geometrie!*

Formen wir die Euler'sche Formel um, dann gelangen wir nach einigen Zwischenschritten zu einer zweiten Form der Darstellung

$$\begin{aligned} e^{j\varphi} &= \cos \varphi + j \sin \varphi \\ e^{-j\varphi} &= \cos \varphi - j \sin \varphi \\ + \text{---} &= \text{---} \text{---} \text{---} \text{---} \text{---} \\ e^{j\varphi} + e^{-j\varphi} &= 2 \cos \varphi \end{aligned}$$

²Bei $e^x = 1 + x + \frac{x^2}{2!} + \dots$ handelt es sich um eine Potenzreihe, die für alle x aus dem Konvergenzbereich $(-\infty, \infty)$ absolut und gleichmäßig konvergiert. Deshalb darf man die Reihe umordnen und sie konvergiert gegen den gleichen Grenzwert.

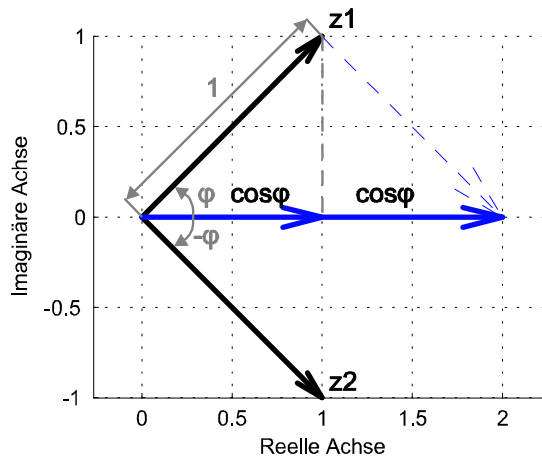


Abbildung A.6: Rechts- und linksdrehende Zeiger

Wie man leicht ausrechnen kan, lässt sich auch der Sinus durch $e^{j\varphi}$ ausgedrückt. Zusammengefasst erhalten wir die inversen Euler'schen Formeln:

$$\cos \varphi = \frac{e^{j\varphi} + e^{-j\varphi}}{2}; \quad \sin \varphi = \frac{e^{j\varphi} - e^{-j\varphi}}{2j} \quad (\text{A.9})$$

Abbildung A.6 zeigt, dass der Kosinus durch die Zusammensetzung von zwei Zeigern (mit positivem und negativem Winkel φ) gebildet wird. Aus den komplexwertigen Funktionen $e^{j\varphi}$ und $e^{-j\varphi}$ entsteht die reellwertige Funktion $\cos(\varphi)$.

Die Kreisfunktionen Sinus und Kosinus lassen sich wie folgt darstellen

$$\begin{aligned} \cos \varphi &= \operatorname{Re}\{e^{j\varphi}\} \quad \text{oder} \quad \cos \varphi = \frac{e^{j\varphi} + e^{-j\varphi}}{2} \\ \sin \varphi &= \operatorname{Im}\{e^{j\varphi}\} \quad \text{oder} \quad \sin \varphi = \frac{e^{j\varphi} - e^{-j\varphi}}{2j} \end{aligned}$$

Bemerkung 99 Die Darstellung der reellwertigen und anschaulichen Kreisfunktionen durch komplexwertige und wenig anschauliche Exponentialfunktionen mag zunächst verwirrend sein, bringt aber beträchtliche Vereinfachungen bei der Rechnung.

Die Rechenregeln für den Umgang mit Zeigern können sowohl über die Algebra als auch über die Geometrie gefunden werden. Wir fassen zunächst die algebraischen Rechenregeln zusammen und geben dann eine geometrische Interpretation zur Erleichterung des Verständnisses.

A.3 Rechenregeln für komplexe Zahlen

Die Rechenregeln für die komplexen Zahlen folgen den Rechenregeln der Algebra, wobei lediglich zu beachten ist, dass die mit j gekennzeichneten Imaginärteile eigene Größen darstellen und dass $j^2 = -1$.

A.3.1 Regeln für kartesische Koordinaten

Addition

$$z_1 + z_2 = (x_1 + jy_1) + (x_2 + jy_2) = (x_1 + x_2) + j(y_1 + y_2) \quad (\text{A.10})$$

$$(-2 + 3i) + (6 - 9i) = 4 - 6i$$

Subtraktion

$$z_1 - z_2 = (x_1 + jy_1) - (x_2 + jy_2) = (x_1 - x_2) + j(y_1 - y_2) \quad (\text{A.11})$$

$$(-2 + 3i) - (6 - 9i) = -8 + 12i$$

Multiplikation

$$z_1 \times z_2 = (x_1 + jy_1) \times (x_2 + jy_2) \quad (\text{A.12})$$

$$= x_1x_2 + jx_1y_2 + jy_1x_2 + j^2y_1y_2 \quad (\text{A.13})$$

$$= (x_1x_2 - y_1y_2) + j(x_1y_2 + y_1x_2)$$

$$(-2 + 3i) \times (6 - 9i) = 15 + 36i$$

Konjugiert komplexe Zahl

$$z_1^* = z_1' = (x_1 + jy_1)^* = (x_1 - jy_1) \quad (\text{A.14})$$

$$(-2 + 3i)^* = -2 - 3i$$

Division

$$z_1 \div z_2 = (x_1 + jy_1) \div (x_2 + jy_2) \quad (\text{A.15})$$

$$\frac{z_1}{z_2} = \frac{z_1 z_2^*}{z_2 z_2^*} = \frac{z_1 z_2^*}{|z_2|^2}$$

$$= \frac{(x_1 + jy_1)(x_2 - jy_2)}{x_2^2 + y_2^2} \quad (\text{A.16})$$

$$= \frac{(x_1x_2 + y_1y_2) + j(x_2y_1 - x_1y_2)}{x_2^2 + y_2^2}$$

$$\frac{(-2+3i)}{(6-9i)} = -\frac{1}{3}$$

A.3.2 Regeln für Polarkoordinaten

Während Addition und Subtraktion in kartesischen Koordinaten einfach durch Addition bzw. Subtraktion der Real- und Imaginärteile durchgeführt werden, sind Multiplikation und Division in kartesischen Koordinaten unübersichtlich und können viel einfacher in Polarkoordinaten berechnet werden.

Multiplikation

$$z_1 \times z_2 = r_1 e^{j\varphi_1} \times r_2 e^{j\varphi_2} = r_1 r_2 e^{j(\varphi_1 + \varphi_2)} \quad (\text{A.17})$$

$$2e^{j30^\circ} \times 3e^{j60^\circ} = 6e^{j90^\circ} = j6$$

Konjugiert komplexe Zahl

$$z_1^* = (r_1 e^{j\varphi_1})^* = r_1 e^{-j\varphi_1} \quad (\text{A.18})$$

$$(2e^{j30^\circ})^* = 2e^{-j30^\circ}$$

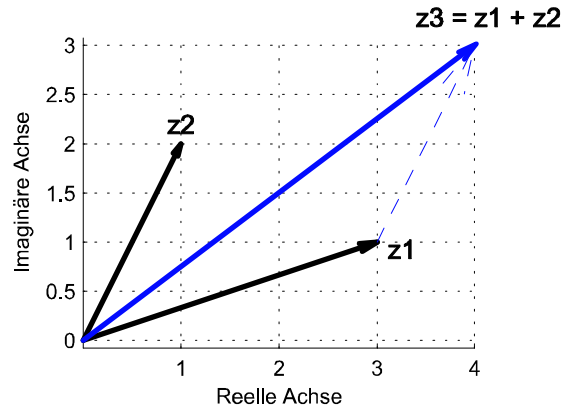


Abbildung A.7: Zeigeraddition

Division

$$z_1 \div z_2 = \frac{r_1 e^{j\varphi_1}}{r_2 e^{j\varphi_2}} = \frac{r_1}{r_2} e^{j(\varphi_1 - \varphi_2)} \quad (\text{A.19})$$

$$2e^{j30^\circ} \div 3e^{j60^\circ} = \frac{2}{3} e^{-j30^\circ}$$

Multiplikation und Division sind in Polarkoordinaten sehr einfach durchführbar, es müssen lediglich die Beträge multipliziert (dividiert) und die Winkel addiert (subtrahiert) werden. Die Addition und Division in Polarkoordinaten kann nur durch Umwandlung in kartesische Koordinaten, Durchführung von Addition bzw. Subtraktion und Rückumwandlung in Polarkoordinaten durchgeführt werden.

Matlab arbeitet mit komplexen Zahlen, die Rechenoperationen Addition (+), Subtraktion (-), Multiplikation (*) und Division (/) gelten auch für komplexes Argument. Die Eingabe ist in kartesischer und Polardatstellung möglich, die Ausgabe erfolgt immer kartesisch und muss bei Bedarf umgerechnet werden.

```

z1 = -2.0000 + 3.0000i      3*exp(i*45*pi/180) = 2.1213 + 2.1213i
z2 = 6.0000 - 9.0000i      [phi,r] = cart2pol(6,-9)
z1+z2 = 4.0000 - 6.0000i   r = 10.8167      phi = -0.9828
z1-z2 = -8.0000 + 12.0000i
z1*z2 = 15.0000 + 36.0000i = 3.6056*exp(2.1588i)*10.8167*exp(-0.9828i)
z1/z2 = -0.3333

```

A.4 Geometrische Betrachtung

Abbildung A.7 zeigt die *Addition* der Zeiger $z_1 + z_2 = z_3$. z_1 und z_2 werden wie im Kräfteparallelogramm zum Zeiger z_3 zusammengesetzt.

Die *Subtraktion* von $z_1 - z_2$ ist nichts anderes als $z_1 + (-z_2)$.

$(-z_2)$ ist der in die gegengesetzte Richtung zeigende Zeiger z_2 . Die Abbildung A.8 stellt die Zeigersubtraktion graphisch dar.

Die *Multiplikation* der Zeiger $z_1 z_2$ stellt man am besten über die Polarform dar: $|z_1| e^{j\varphi_1} |z_2| e^{j\varphi_2} = |z_1| |z_2| e^{j(\varphi_1 + \varphi_2)}$. Abbildung A.9 stellt diesen Zusammenhang graphisch dar. Die Multiplikation mit der Zahl $e^{j\varphi_2}$ bedeutet eine Drehung des Zeiger $|z_1| e^{j\varphi_1}$ um den Winkel φ_2 . Die Multiplikation einer Zahl mit j entspricht einer Drehung um 90° , da $j = e^{j\pi/2}$.

Auch die *Division* kann am besten über die Polarform dargestellt werden.

$$z_3 = \frac{z_1}{z_2} = \frac{r_1}{r_2} e^{j(\varphi_1 - \varphi_2)}$$

Abbildung A.10 stellt die Division graphisch dar. Ein interessanter Fall tritt auf, wenn $z_1 = r_1 e^{j\varphi}$ und $z_2 = r_2 e^{j(\varphi + 90^\circ)}$ also einen Winkel von 90° einschließen. $z_3 = z_1/z_2$ wird dann $-j(r_1/r_2)$.

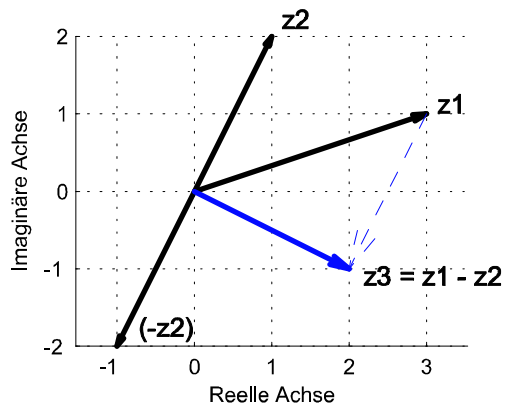


Abbildung A.8: Subtraktion von Zeigern

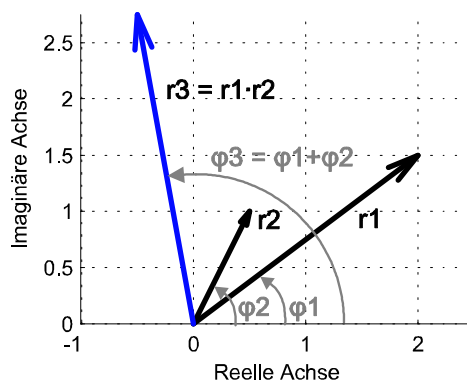


Abbildung A.9: Zeigermultiplikation

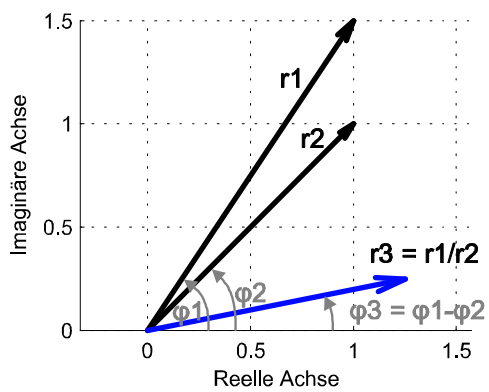


Abbildung A.10: Zeigerdivision

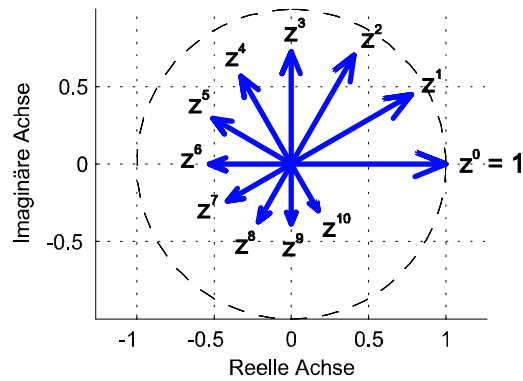


Abbildung A.11: Potenzen der komplexen Zahl

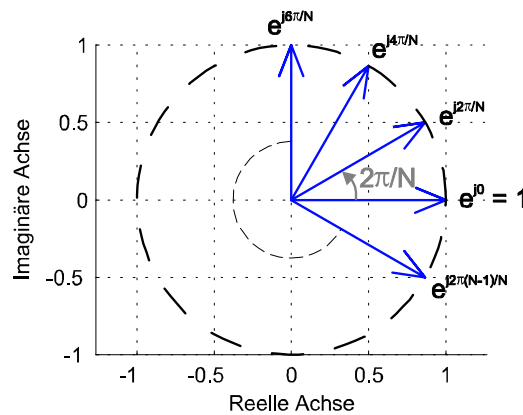


Abbildung A.12: Wurzeln von 1

Ganzzahlige *Potenzen* einer komplexen Zahl kann man einfach über die Polarform berechnen

$$z^N = (re^{j\varphi})^N = r^N e^{jN\varphi}$$

Abbildung A.11 zeigt die Potenzen z^1, z^2, \dots, z^{10} , für die Zahl $z = 0.9e^{j30^\circ}$.

Es fehlt noch die *Wurzel* einer (komplexen) Zahl. Als Beispiel betrachten wir $\sqrt[4]{1}$. Vom Fundamentalsatz der Algebra wissen wir, dass das Polynom $z^4 - 1 = 0$ vier Wurzeln haben muss. Die Lösung 1 finden wir leicht, auch die Lösung -1 ist nicht schwer zu bestimmen, etwas Nachdenken ist aber für die weiteren Lösungen $\pm j$ erforderlich. Allgemein gilt die Beziehung:

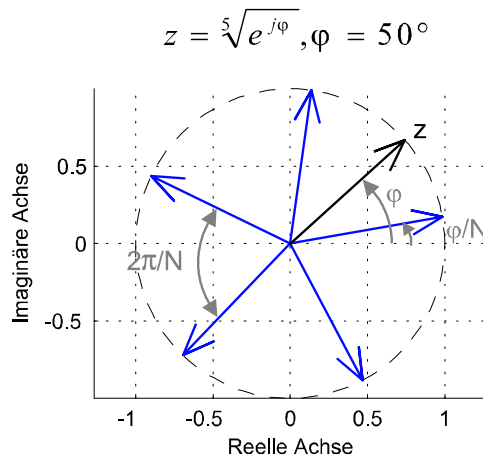
$$\sqrt[N]{1} = e^{j\frac{2\pi n}{N}} \quad n = 0, 1, 2, \dots, N-1$$

Das bedeutet, dass die Wurzeln gleichverteilt auf dem Einheitskreis liegen und dass der Winkel zwischen den Wurzeln $\varphi = 2\pi/N$ beträgt. Abbildung A.12 stellt diesen Zusammenhang dar.

Im allgemeinen Fall ist aber nicht $\sqrt[N]{1}$ zu bilden, sondern die Wurzel einer beliebigen komplexen Zahl. Wir finden hier folgenden Zusammenhang:

$$z = \sqrt[N]{re^{j\varphi}} = \sqrt[N]{r} e^{j(\frac{\varphi}{N} + \frac{2\pi n}{N})} \quad n = 0, 1, 2, \dots, N-1$$

Abb. A.13 zeigt ein Beispiel für eine Wurzel fünfter Ordnung, wobei der Betrag der Zahl, aus der die Wurzel gezogen wird, mit 1 angenommen ist. Die Lösungen liegen daher auf dem Einheitskreis, die erste Wurzel liegt bei $e^{j\varphi/N}$, die weiteren Wurzeln sind um den Winkel $2\pi/N$ gedreht. Für eine Zahl z mit Betrag $\neq 1$, wäre der Radius des Kreises auf dem die Wurzeln liegen $\sqrt[N]{|z|}$.

Abbildung A.13: Wurzel aus e^{j50°

A.5 Achtung Phase

Bei der Darstellung nach Betrag und Winkel treten paar Besonderheiten auf, die zu beachten sind.

- Bei der Berechnung der Winkelfunktionen wird normalerweise das Bogenmaß verwendet. Damit vermeidet man die umständliche Darstellung des Winkels im Gradmaß, die die Zahlenbasis 360 für Grad, 60 für Minuten und Sekunden und 10 für Bruchteile von Sekunden verwendet, z.B. $30^\circ 45' 48.1234''$. Dafür handelt man sich schwer vorstellbare Bogenwinkel für aus dem Gradmaß gut bekannte Winkel ein, z.B. $0.7854 = \pi/4 = 45^\circ$. Daher wird die Umrechnung oft implizit angeschrieben, z.B. $\sin(\alpha^\circ * \pi/180)$.
- Wenn der Realteil Null ist und der Imaginärteil einen endlichen Wert hat, tritt bei der Berechnung der Phase eine Division durch Null auf, da $\varphi = \arctan \frac{\text{Imaginärteil}}{\text{Realteil}}$, was in der Berechnung entsprechend berücksichtigt werden muss. Wenn die Zahl rein imaginär ist, ist die Phase $\pi/2$ bei positivem Imaginärteil und $-\pi/2$ bei negativem Imaginärteil.
- Der Arcustangens ist nur für $-90^\circ \leq \varphi \leq 90^\circ$ definiert. Für die komplexe Zahl $1+1i$ erhalten wir den Winkel $\varphi = \arctan \frac{1}{1} = 0.7854 \text{ [rad]} = 45^\circ$. Für die komplexe Zahl $-1 - 1i$ erhalten wir ebenfalls den Winkel $\varphi = \arctan(\frac{-1}{-1}) = \arctan(1) = 0.7854 \text{ [rad]} = 45^\circ$, obwohl die korrekte Phase (Zahl liegt im dritten Quadranten) $\varphi = -135^\circ$ beträgt. Dieses Problem tritt immer auf, wenn die komplexe Zahl im zweiten oder dritten Quadranten liegt. Sind Real- und Imaginärteil negativ, dann muss vom berechneten Winkel 180° abgezogen werden, wenn der Realteil negativ ist und der Imaginärteil positiv ist, dann müssen 180° addiert werden. Die MATLAB-Funktion `angle(H)` liefert die Phase vorzeichenrichtig im Bogenmaß.
- Die Phase ist wegen der Periodizität der Sinusfunktion lediglich im Bereich von $0 \leq \varphi \leq 2\pi$ definiert. Wenn sich ein Zeiger um $\varphi + n \cdot 2\pi$ dreht, dann ist die Phase dennoch nur φ . Dieser Sachverhalt führt dazu, dass bei Darstellungen des Phasenverlaufs die Phase zwischen $-\pi$ und π liegt, man aber auf Grund des physikalischen Zusammenhangs weiß, dass die Phase $\varphi + n \cdot 2\pi$ beträgt. Man wählt daher häufig eine Darstellung, die die Phase streckt (unwrapping) und dadurch die Sprungstellen vermeidet. Abbildung A.14 zeigt die Darstellung der Phase zwischen $\pm 180^\circ$ und die unwrapped-Darstellung.

Beispiel 100 Als Ergebnis einer Berechnung erhalten wir die komplexen Werte H
`plot((180/pi)*angle(H))`

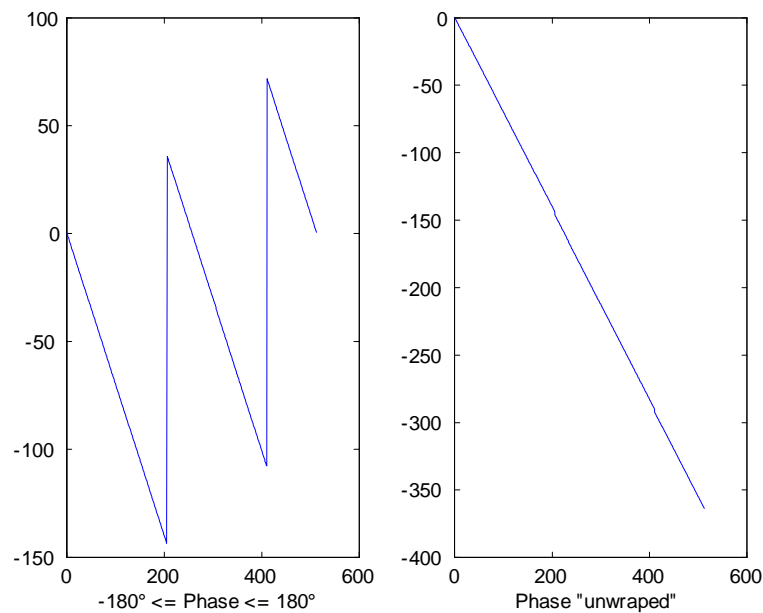


Abbildung A.14: Phasendarstellung

```

xlabel('-180° <= Phase <= 180° ')
subplot(1,2,2)
plot(unwrap((180/pi)*angle(H)))
xlabel('Phase "unwrapped"')

```

- π und $-\pi$ stellen dieselbe Phase dar. Durch Rundungsfehler in der Rechnung können Sprünge in der Phase auftreten, die zu Phasensprüngen zwischen benachbarten Punkten führen können.
- Wenn die Amplitude einer komplexen Zahl Null (oder sehr klein) ist, hat die Phase keine Bedeutung und kann sinnlose Werte annehmen.
- Der Betrag ist definitionsgemäß eine positive Zahl. Dennoch läßt man gelegentlich einen »negativen« Betrag zu, um die Darstellung zu vereinfachen. Abbildung A.15 zeigt den Zusammenhang am Beispiel der Funktion $\sin(x)/x$.

A.6 Zusammenfassung

Mit Hilfe der komplexen Darstellung lässt sich die – für Elektrotechnik und Signalverarbeitung sehr wichtige – Sinus(Kosinus)funktion in eine kompakte Form bringen, mit der man wesentlich besser rechnen kann als mit der reellen Darstellung. Durch geometrische Interpretation der komplexen Zahlen gelangt man zur Zeigerdarstellung, mit der sich die Operationen mit komplexen Zahlen anschaulich erläutern lassen.

Komplexe Zahlen können in kartesischer oder Polarform angeschrieben werden. Je nach Rechnung ist die eine oder die andere Form zweckmäßiger, im Zuge von umfangreicheren Berechnungen ist es daher häufig nötig zwischen den Koordinatensystemen zu wechseln. Matlab kennt komplexe Zahlen und stellt Routinen zur Durchführung der Berechnung und Darstellung komplexer Zahlen zur Verfügung.

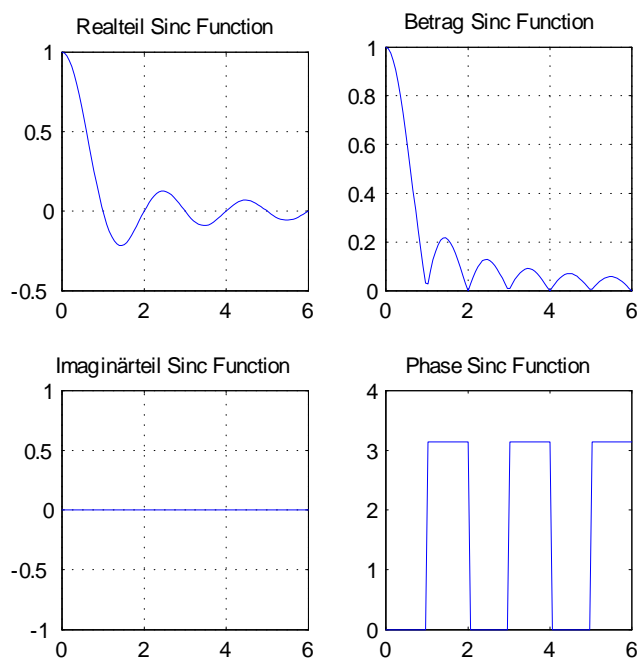


Abbildung A.15: »Negativer« Betrag erzeugt $Phase = 0$

Anhang B

MATLAB

Inhalt

B.1	Der Workspace	234
B.1.1	Hilfe und Orientierung	234
B.1.2	Bildschirmausgabe	234
B.1.3	Namensgebung	234
B.1.4	Zahlen	234
B.1.5	Skalare, Vektoren, Matrizen	234
B.2	Elementare Operationen	236
B.3	Elementare komplexe Operationen	236
B.4	Graphiken in Matlab	236
B.5	Programmierung	237
B.5.1	Kommentare	237
B.5.2	Scripts	237
B.5.3	Funktionen	237
B.5.4	Kontrollstrukturen	237
B.5.5	Ein- und Ausgabe	238
B.6	Kontinuierliche LTI-Systeme	238
B.6.1	Approximation analoger Filter	239
B.7	Diskrete Systeme	239
B.7.1	Approximation digitaler Filter	240
B.8	Praktischer Filterentwurf	240
B.9	Signalverarbeitung	240
B.10	Zusammenfassung	241

Die bei der Untersuchung von kontinuierlichen und diskreten linearen zeitunabhängigen Systemen verwendeten Konzepte und Operationen sind mit Handrechnung schwer beherrschbar. Mit MATLAB steht eine Arbeitsumgebung zur Verfügung, die die Mühen der Berechnung abnimmt und so den Blick für die Eigenschaften und das Verhalten von Systemen freihält und zum vertiefenden Experimentieren mit LTI-Systemen einlädt.

MATLAB steht für MATrix LABoratory und ist eine Arbeitsumgebung für numerische Berechnungen und zur Visualisierung von Daten. MATLAB stellt eine umfangreiche Sammlung von Funktionen für mathematische, statistische und technische Berechnungen zur Verfügung. Toolboxen erweitern die Standard-MATLAB-Funktionen für spezielle Arbeitsgebiete wie Digitale Signalverarbeitung, Bildverarbeitung, Regelungstechnik, ...

Für die Untersuchung von kontinuierlichen LTI-Systemen verwenden wir die *Control Systems Toolbox*, für diskrete LTI-Systeme die *Signal Processing Toolbox*. Für spezielle Aufgaben bei der Untersuchung von digitalen Filtern steht als Erweiterung der Signal Processing Toolbox die *Filter Design Toolbox* zur Verfügung.

Die folgenden Kapitel fassen die wichtigsten Befehle zusammen, ausführliche Dokumentation steht durch die Kommandozeilenfunktion `help` bzw. `help fname`, bzw. über das Benützermenü Help (F1) zur Verfügung.

B.1 Der Workspace

B.1.1 Hilfe und Orientierung

<code>help</code>	zeigt eine Liste von Programmen, die in der MATLAB-Umgebung installiert sind
<code>help topic</code>	zeigt Hilfeinformation zu MATLAB-Funktionen an
<code>demo</code>	startet ein Demo-Programm zu einer Vielzahl von MATLAB-Funktionen
<code>who</code>	listet alle aktuellen Variablen
<code>whos</code>	listet alle aktuellen Variablen einschließlich deren Dimensionen
<code>what</code>	listet alle aktuellen MATLAB Dateien (Funktionen und Scripts)
<code>clear</code>	löscht alle Variablen im Workspace
<code>diary »filename«</code>	protokolliert eine MAT-LAB Sitzung <code>diary on</code> <code>diary off</code>
<code>save variable</code>	speichert alle Variablen im File <code>variable.mat</code>
<code>load variable.mat</code>	läd alle Variablen aus dem File <code>variable.mat</code>
<code>save variable A,B,C</code>	speichert die Variablen A, B, C im File <code>variable.mat</code>
<code>quit</code>	beendet MATLAB

B.1.2 Bildschirmausgabe

Das Ergebnis wird standardmäßig mit fünf signifikanten Stellen ausgegeben und ist immer in der Variable `ans` gespeichert. Nach der Ausgabe wird am Bildschirm eine Leerzeile eingefügt. Wird eine Eingabezeile mit einem `;` beendet, unterbleibt die Ausgabe auf den Bildschirm.

<code>format compact</code>	unterdrückt die Leerzeile
<code>format loose</code>	aufheben von <code>compact</code>
<code>format long e</code>	
<code>format short</code>	
<code>format short e</code>	
<code>format bank</code>	zwei Stellen nach dem Komma

B.1.3 Namensgebung

MATLAB unterscheidet zwischen Groß- und Kleinbuchstaben.

B.1.4 Zahlen

MATLAB rechnet grundsätzlich mit komplexen Zahlen.

```
sqrt(-1) ==> 0 + 1.0000i
(3+4i)+(6-7i) ==> 9.0000 - 3.0000i
exp(0.5*pi*i) ==> 0.0000 + 1.0000i
exp(0.5*pi*1j) ==> 0.0000 + 1.0000i
```

Bei der Eingabe kann sowohl `i` als auch `j` als imaginäre Einheit gewählt werden. Die Ausgabe erfolgt immer mit `i`. Die Eingabe kann in Polarform (exponentieller Form) erfolgen, die Ausgabe ist immer in Real- und Imaginärteil.

Bemerkung 101 *Achtung bei der Verwendung von `i` oder `j` als Variable:*

```
i = 3
4+3*i ==> 13 % i wurde auf 3 gesetzt!
aber 4+3i ==> 4.0000 + 3.0000i
Wiederherstellen von i als imaginäre Einheit: i = sqrt(-1)
Um Fehler zu vermeiden, Schreibweise 4 + 3*1i verwenden.
```

B.1.5 Skalare, Vektoren, Matrizen

MATLAB kennt intern nur Matrizen, ein Skalar ist eine `1x1` Matrix. Vektoren gibt es nur als einzeilige oder einspaltige Matrix, häufig Zeilenvektor (`1 x N`) oder Spaltenvektor (`N x 1`) genannt. Die Dimension kann mit `size(N)` abgefragt werden.

```
A = 1
size(A) ==> 1 1
B=[1 2 3]
size(B) ==> 1 3
```

Eingabe Zeilenvektoren

```
zv = [2 4 -5 2.345 -9.1] oder zv = [2,4,6,1.23,0]
```

Eingabe von Spaltenvektoren

```
b = [2;4;6;1.23;0] % Elemente durch ; getrennt.
Die Umwandlung von Zeilen- in Spaltenvektor erfolgt durch Transponieren.
zv.' ==>
2.0000
4.0000
-5.0000
2.3450
-9.1000
```

Achtung: Es gibt zwei Formen von **transpose**: `.'` ist das nichtkonjugierte transpose, `'` ist das konjugierte transpose.

Eingabe von Matrizen

```
Mat = [1 2 3;4 5 6;7 8 9]
```

Vektor- und Matrix-Generatoren

Vektor aus fortlaufenden Zahlen

```
X = Anfang:Schrittweite:Ende
X = 1:2:9 ==> 1 3 5 7 9
aber X = 1:2:8 ==> 1 3 5 7
```

Wenn keine Schrittweite angegeben ist, wird die Standardschrittweite »1« gesetzt.

```
5:10 ==> 5 6 7 8 9 10
```

```
alpha = exp(i*[0:pi/4:2*pi]) % erzeugt Zeiger von Null bis 2*pi im Anstand von pi/4
```

```
alpha ==> (1.0000),(0.7071 + 0.7071i),(0.0000 + 1.0000i),...
```

```
Linspace(Anfang, Ende, Anzahl_Punkte) % erzeugt Anzahl_Punkte zwischen Anfang und Ende
```

```
linspace(0, 5, 4) ==> 0 1.6667 3.3333 5.0000
```

```
Nullmatrix = zeros(Zeilen,Spalten)
```

```
ones(1,5) % erzeugt Zeilenvektor 1x5 mit den Elementen 1
```

```
Einheitsmatrix = eye(7)
```

```
x=0:N; y=sin(2*pi*x/N); % erzeugt eine Sinusschwingung einer Periode mit N Punkten
```

```
TestGleich = rand(1,8) % erzeugt Zeilenvektor aus gleichverteilten Zufallszahlen
```

```
TestNormal = randn(4,7) % erzeugt Matrix aus normalverteilten Zufallszahlen
```

Zusammenfügen von Arrays

```
array = [array1,array2] % fügt array2 an array1 in Zeilenrichtung (Dimension 2) an
```

```
Alternative cat(2, array1,array2)
```

```
array = [array1;array2] % fügt array2 an array1 in Spaltenrichtung (Dimension 1) an
```

```
Alternative cat(1, array1,array2)
```

Die Array-Dimensionen müssen verträglich sein!

Indizierung von Arrays

Die Indizierung erfolgt über entsprechende Indexstrukturen
`vec = [10 20 30 40 50 60]; ind = [1 3 5];`

```
vec(ind) ==> 10 30 50
mat(1:2,1:2) % indiziert 1. und 2. Zeile und 1. und 2. Spalte
mat(:,1:2) % indiziert alle Zeilen und 1. und 2. Spalte
mat(1:2,:) % indiziert 1. und 2. Zeile und alle Spalten
mat(:) % erzeugt einen Spaltenvektor durch Aneinanderfügen der Spaltenvektoren der Matrix
```

Ermittlung der Dimensionen von Arrays

```
size(array) % liefert Zeilenvektor mit den Dimensionen des Arrays
length(vec) % Länge des Vektors
```

B.2 Elementare Operationen

```
+ - * / ^ % Addition, Subtraktion, Multiplikation, Division, Potenz
sqrt(N) % Berechnet Quadratwurzel von N,
wenn N komplex, dann wird nur ein komplexes Ergebnis berechnet.
Die elementaren Operationen sind immer Matrizenoperationen!
mat*mat % ist eine Matrizenmultiplikation
elementweise Operationen erfolgen über .-Operation:
mat.*mat oder mat./mat
sin(x), cos(x), tan(x), asin(x), acos(x), atan(x) % Winkelfunktionen
```

B.3 Elementare komplexe Operationen

<code>real(z)</code>	Realteil von z
<code>imag(z)</code>	Imaginärteil von z
<code>abs(z)</code>	Betrag von z
<code>angle(z)</code>	Phasenwinkel in rad
<code>conj(z)</code> oder z'	konjugiert komplexe Zahl
<code>complex(a,b)</code>	erzeugt komplexe Zahl aus Realteil a und Imaginärteil b
<code>[x,y] = pol2cart(phi,R)</code>	erzeugt kartesische aus Polarkoordinaten
<code>[phi,R] = cart2pol(x,y)</code>	erzeugt Polar- aus kartesischen Koordinaten

B.4 Graphiken in Matlab

`figure(N)` % dient zur Darstellung von Ergebnissen in mehreren Fenstern
wenn `figure(N)` nicht verwendet wird, dann wird die Ausgabe immer in dasselbe Fenster geleitet.

`subplot(m,n,p)` % dient zur Darstellung mehrere Ergebnisse im Fenster `figure(N)`
 $m \times n$ ist die Fenstermatrix, p ist das Fenster in das die Darstellung gestellt wird. Die Nummerierung der Darstellungsfenster erfolgt zeilenweise.

`hold on` % verhindert, dass ein neuer Plot den alten löscht; damit wird das Überlagern von Darstellungen möglich.

`hold off` % ein neuer Plot löscht die vorherige Darstellung

`clf` % löscht die Darstellung und erzeugt leeres Fenster

`plot(x,y,s)` % zeichnet die Werte von y über x

s ist eine Stringvariable die aus einem oder mehreren Elementen besteht und die Art der Darstellung festlegt.

y	yellow	.	point	-	solid
m	magenta	o	circle	:	dotted
c	cyan	x	x-mark	-.	dashdot
r	red	+	plus	--	dashed
g	green	*	star		
b	blue	s	square		
w	white	d	diamond		
k	black	v	triangle down		
		^	triangle up		
		<	triangle left		
		>	triangle right		
		p	pentagram		
		h	hexagram		

```

plot(x,y,'y-') % zeichnet eine gelbe durchgehende Linie
stem(x,y) % zeichnet Datenreihe y an den von x gegebenen Punkten
compass(x,y) % zeichnet Vektor mit den Komponenten x, y vom Koordinatenursprung in Polarkoordinaten
compass(z) % zeichnet komplexe Zahl z in Polarkoordinaten
title 'Titel'; xlabel 'x-Achse'; ylabel 'y-Achse'

```

B.5 Programmierung

B.5.1 Kommentare

Kommentare beginnen mit % und können auch am Ende einer Programmzeile stehen

B.5.2 Scripts

Scripts haben keine Argumente sondern sind lediglich eine Sammlung von MATLAB-Befehlen. Sie werden über den Namen des Skript-Files aufgerufen.

B.5.3 Funktionen

Funktionen müssen das Schlüsselwort `function` in der ersten Zeile enthalten.

```

function erg = lvect(x,y)
% Berechnet die Länge des Vektors mit den Koordinaten x,y
% mit dem Pythagoräischen Lehrsatz
erg = sqrt(x.^2+y.^2)
Funktionen werden mit dem Funktionsnamen aufgerufen.

```

B.5.4 Kontrollstrukturen

MATLAB stellt die üblichen Kontrollstrukturen zu Verfügung

```

for k=1:N
    Anweisungsblock
end
if Bedingung
    Anweisungsblock 1
elseif Bedingung
    Anweisungsblock 2
else
    Anweisungsblock 3
end
while Bedingung
    Anweisungsblock
end

```

B.5.5 Ein- und Ausgabe

`disp('Eingabe der Werte')`

`x=input('Zahl der Koeffizienten: ')`

Darüber hinaus stellt MATLAB umfangreiche Bibliotheken für die Oberflächenprogrammierung zur Verfügung.

B.6 Kontinuierliche LTI-Systeme

LTI-Systeme werden im s -Bereich in Polynom- oder PN-Form dargestellt.

$$F(s) = \frac{B(s)}{A(s)} = \frac{b_m s^m + b_{m-1} s^{m-1} + \dots + b_1 s + b_0}{a_n s^n + a_{n-1} s^{n-1} + \dots + a_1 s + a_0} \quad (\text{B.1})$$

$$F(s) = k \frac{(s - \beta_1)(s - \beta_2) \dots (s - \beta_m)}{(s - \alpha_1)(s - \alpha_2) \dots (s - \alpha_n)} \quad (\text{B.2})$$

Die Systemfunktion wird in den MATLAB Workspace durch die Befehle

`F = tf([bm, bm-1, ..., b1, b0], [an, an-1, ..., a1, a0])` bzw.

`F = zp([β1, β2, ..., βm], [α1, α2, ..., αn], k)`

einggegeben und steht dann als sogenanntes LTI-Objekt zur Verfügung. Mit der Funktion

`get(F)` können alle Eigenschaften des LTI-Objektes abgefragt, mit `wert=F.eigenschaft{1}`

(z.B. `F.num{1}`) kann die gewünschte Eigenschaft abgefragt werden. Die Eigenschaft lässt

sich umgekehrt mit `set(F, 'eigenschaft', wert)` setzen.

<code>impulse(F)</code>	berechnet die Impulsantwort und zeigt sie auf dem Schirm an
<code>[y,t] = impulse(F)</code>	berechnet die Impulsantwort und weist sie <code>y</code> und <code>t</code> zu
<code>step(F)</code>	berechnet die Sprungantwort und zeigt sie auf dem Schirm an
<code>lsim(F,x,t)</code>	berechnet Zeitanwort auf <code>x(t)</code>
<code>freqs(B,A)</code>	berechnet den Frequenzgang
<code>bode(F)</code>	berechnet Amplituden- und Phasengang von <code>F</code>
<code>[Z,P,k] = tf2zp(Zae,Nen)</code>	berechnet Pole, Nullstellen und <code>k</code> aus Polynomdarstellung ¹
<code>[Zae,Nen] = zp2tf(Z,P,k)</code>	berechnet die Polynomdarstellung aus Polen und, Nullstellen
<code>P = pole(F)</code>	liefert die Pole von <code>F</code>
<code>Z = zero(F)</code>	liefert die Nullstellen von <code>F</code>
<code>k = dcgain(F)</code>	liefert <code>F(0)</code>
<code>pzmap(F)</code>	zeichnet PN-Diagramm
<code>[P,Z] = pzmap(F)</code>	berechnet Pole und Nullstellen von <code>F</code>
<code>[r,p,K] = residue(B,A)</code>	berechnet die Koeffizienten der Partialbruchzerlegung
<code>[B,A] = residue(r,p,K)</code>	wandelt Partialbruch- in Polynomdarstellung um
<code>r=roots(C)</code>	berechnet Nullstellen des Polynoms $c_1 x^n + c_2 x^{n-1} + \dots + c_n x + c_{n+1}$
<code>C=poly(r)</code>	berechnet Polynom- aus Nullstellendarstellung
<code>dB</code>	rechnet in Dezibel-Darstellung um

$$F(s) = \frac{B(s)}{A(s)} = \frac{r_1}{s - p_1} + \frac{r_2}{s - p_2} + \dots + \frac{r_n}{s - p_n} + K(s) \quad (\text{B.3})$$

Bemerkung 102 Die Funktion `ltiview(F)` bietet ein komfortables Interface zur Darstellung der Systemeigenschaften von `F` im Zeit- und Frequenzbereich. Die Auswahl erfolgt mit Hilfe der Menüfunktionen bzw. durch Rechtsklicken auf die leere Darstellungsfläche.

¹`tf2zp` wird im s -Bereich verwendet. Im z -Bereich und der Darstellung $(b_0 + b_1 z^{-1} + b_2 z^{-2})$ sollte `tf2zpk` verwendet werden.

B.6.1 Approximation analoger Filter

Die folgenden Funktionen liefern die Approximationen für *analoge* Filter
 Rp ... Welligkeit im Durchlassbereich, Rs ... Welligkeit im Sperrbereich, N ... Ordnung
 des Filters

- [Z,P,K] = `buttap(N)` liefert die PN-Koeffizienten für ein analoges Potenzfilter
- [Z,P,K] = `cheb1ap(N,Rp)` liefert die PN-Koeffizienten für ein analoges Tschebyscheff-Filter
- [Z,P,K] = `ellipap(N,Rp,Rs)` liefert die PN-Koeffizienten für ein analoges Cauer-Filter
- [Z,P,K] = `besselapp(N)` liefert die PN-Koeffizienten für ein analoges Bessel-Filter

B.7 Diskrete Systeme

Diskrete Systeme werden im z -Bereich in Polynom- oder PN-Form dargestellt

$$H(z) = \frac{B(z)}{A(z)} = \frac{b_1 + b_2z^{-1} + \dots + b_mz^{-m}}{a_1 + a_2z^{-1} + \dots + a_nz^{-l}} = \frac{Zae(z)}{Nen(z)} \quad (B.4)$$

$$H(z) = \frac{B(z)}{A(z)} = k \frac{(z - q_1)(z - q_2) \dots (z - q_m)}{(z - p_1)(z - p_2) \dots (z - p_l)} \quad (B.5)$$

$$a_1y[n] = b_1x[n] + b_2x[n - 1] + \dots + b_mx[n - m] - a_2y[n - 1] - \dots - a_ny[n - l] \quad (B.6)$$

- `impz(B,A)` berechnet die Impulsantwort und zeigt sie auf dem Schirm an
- `[y,t] = impz(B,A)` berechnet die Impulsantwort und weist sie `y` und `t` zu
- `stepz(B,A)` berechnet die Sprungantwort und zeigt sie auf dem Schirm an
- `y = filter(B,A,x)` berechnet Zeitanwort auf die Folge `x[n]`
- `freqz(B,A)` berechnet den Frequenzgang
- `[Z,P,k] = tf2zpk(Zae,Nen)` berechnet Pole, Nullstellen und `k` aus Polynomdarstellung²
- `[Zae,Nen] = zp2tf(Z,P,k)` berechnet die Polynomdarstellung aus Polen und, Nullstellen
- `zplane(B,A)` zeichnet PN-Diagramm % B, A ... Zeilenvektor
- `zplane(Z,P)` zeichnet PN-Diagramm % Z, P ... Spaltenvektor
- `[r,P,K] = residuez(B,A)` berechnet die Koeffizienten für die Partialbruchzerlegung
- `[B,A] = residue(r,P,K)` wandelt Partialbruch- in Polynomdarstellung um
- `C = conv(A, B)` Faltet Vektor A und B, Polynommultiplikation
- `[Q,R] = deconv(B,A)` Ent-Faltung (`B = conv(A,Q) + R`), Polynomdivisoon
- `w = fft(x)` Diskrete Fouriertransformation der Folge `x`
- `x = ifft(w)` Inverse DFT von `w`

Das Spektrum eines diskreten Signals ist periodisch fortgesetzt, die Funktion `fft` liefert die Werte von 0 bis 2π . Für die grafischen Darstellung ist es oft anschaulicher das Spektrum von $-\pi$ bis π zu erstrecken, d. h. die Komponente der Frequenz Null in die Mitte des Spektrums zu verschieben. Diese Verschiebung kann mit der Funktion `fftshift` erreicht werden.

Diskrete (und kontinuierliche) System werden häufig als Hintereinanderschaltung von Subsystemen 2. Ordnung dargestellt.

$$H(z) = g \prod_{k=1}^L H_k(z) = g \prod_{k=1}^L \frac{b_0 + b_1z^{-1} + b_2z^{-2}}{a_0 + a_1z^{-1} + a_2z^{-2}} \quad (B.7)$$

wobei die second-order sections (sos) im Matrixform dargestellt werden

$$sos = \begin{bmatrix} b_{01} & b_{11} & b_{21} & a_{01} & a_{11} & a_{21} \\ b_{02} & b_{12} & b_{22} & a_{02} & a_{12} & a_{22} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ b_{0L} & b_{1L} & b_{2L} & a_{0L} & a_{1L} & a_{2L} \end{bmatrix} \quad (B.8)$$

²`tf2zpk` wird im s -Bereich verwendet. Im z -Bereich und der Darstellung $(b_0 + b_1z^{-1} + b_2z^{-2})$ sollte `tf2zpk` verwendet werden.

Zur Umrechnung stehen folgende Funktionen zur Verfügung

<code>[sos,g] = tf2sos(Zae,Nen)</code>	wandelt Zähler- und Nennerpolynom in die SOS-Darstellung um
<code>[Zae,Nen] = sos2tf(sos,g)</code>	wandelt SOS in Zähler- und Nennerpolynom um
<code>[sos,g]=zp2sos(z,p,g)</code>	wandelt PN- in SOS-Darstellung um
<code>[z,p,g]=sos2zp(sos,g)</code>	wandelt SOS- in PN-Darstellung um

Bemerkung 103 Die Funktion `fvttool(B,A)` bietet ein komfortables Interface zur Darstellung der Systemeigenschaften von $H(z) = \frac{B(z)}{A(z)}$ im Zeit- und Frequenzbereich. Die Auswahl erfolgt mit Hilfe der Menüfunktion `Analysis`.

B.7.1 Approximation digitaler Filter

Die Parameter in den Funktionen haben folgende Bedeutung

`N` ... Ordnung des Filters,

`Wn` ... Grenzfrequenz normiert $0 < Wn < 1$, 1 entspricht der halben Abtastrate

`Wn = [W1 W2]`, liefert ein Bandpassfilter der Ordnung $2N$ mit dem Durchlassbereich $W1 < W < W2$

'type' ... 'low', 'high', 'stop' liefern Tief-, Hoch- bzw. Bandpass-Filter. Beim Bandpass muss `Wn = [W1 W2]` sein und die Ordnung ist $2N$

`Rp` ... Welligkeit im Durchlassbereich, `Rs` ... Welligkeit im Sperrbereich

FIR-Filter

`B = FIR1(N,Wn)` berechnet die FIR-Koeffizienten eines Tiefpasses

`B = FIR1(N,Wn,'high')` berechnet die FIR-Koeffizienten eines Hochpasses

`B = FIR1(N,Wn,'fenster')` FIR-Koeffizienten eines Tiefpasses mit spezifizierter Fensterfunktion

Die Fensterfunktion beeinflusst den Frequenzgang des Filters. Je steiler der Übergang vom Durchlass- in den Sperrbereich, desto geringer ist die Sperrdämpfung. Für Details siehe MATLAB Help > Search > Windows. Die Funktion `wintool` aus der Filter Design Toolbox visualisiert den Einfluss der Fensterfunktionen im Zeit- und Frequenzbereich. Die Standardeinstellung ist Hamming-Fenster.

IIR-Filter

Die folgenden Funktionen liefern die Approximationen für IIR Filter.

`[B,A] = butter(N,Wn)` liefert digitalen Potenztiefpass

`[B,A] = cheby1(N,Rp,Wn,'high')` liefert digitalen Tschebyscheff-Hochpass

`[B,A] = ellip(N,Rp,Rs,Wn)` liefert digitalen Cauer-Tiefpass

B.8 Praktischer Filterentwurf

Alle Filterfunktionen sind unter der MATLAB-Funktion `fdatool` (Filter Design & Analysis Tool) zusammengefasst. `fdatool` erlaubt die Auswahl des Filtertyps (Tief-, Hoch-, Bandpass, Bandsperre,...), der Methode (FIR- oder IIR-Filter) und der Optionen (z.B. Fenster bei FIR-Filtern), der Ordnung des Filters, der Frequenzspezifikation, der maximalen Dämpfung im Durchlassbereich, der minimalen Dämpfung im Sperrbereich.

An Analseverfahren können Amplituden- und Phasengang, Impuls- und Sprungantwort, PN-Darstellung, Filterkoeffizienten, ... durchgeführt werden.

Für die Weiterverarbeitung kann eine Filterstruktur gefunden werden, die innerhalb von Simulink simuliert werden kann. Die Filterkoeffizienten können in den MATLAB-Workspace zur weiteren Bearbeitung exportiert werden. Der Filterentwurf kann als C-Programm in ein DSP-Entwicklungssystem oder in ein FPGA-Board geladen werden.

B.9 Signalverarbeitung

`sptool` stellt eine Benützeroberfläche für Signalverarbeitungsaufgaben zur Verfügung, die aus drei Fenstern besteht.

Im Fenster **Signals** können Eingangssignale, die vom MATLAB-Workspace oder von File importiert wurden angezeigt werden. Wenn entsprechende Audiounterstützung vorhanden sind, können die Signale auch abgespielt werden.

Das Fenster **Filters** erlaubt die Anzeige und Änderung von Filtern. Die Filter können importiert oder über New selbst entworfen sein. Filter können auf Signale aus dem Signals-Menü angewandt werden und erzeugen dann ein Ausgangssignal, das im Fenster Signals gespeichert wird.

Das Fenster **Spectra** erlaubt die Berechnung und Anzeige von Spektren von Signalen aus dem Fenster Signals.

B.10 Zusammenfassung

Matlab und die *Control Systems Toolbox*, *Signal Processing Toolbox* und *Filter Design Toolbox* enthalten zahlreiche Funktionen für Analyse und Design von kontinuierlichen und diskreten LTI-Systemen. Zusätzlich zur Online-Hilfe gibt es noch eine umfangreiche Dokumentation im PDF-Format.