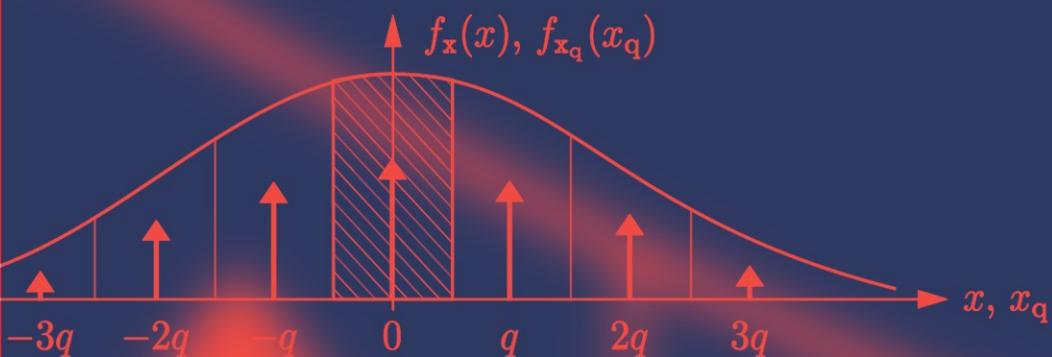


Puente León · Kiencke Messtechnik

Systemtheorie für Ingenieure
und Informatiker



9. Auflage

 Springer Vieweg

Messtechnik

Fernando Puente León • Uwe Kiencke

Messtechnik

Systemtheorie für Ingenieure und Informatiker

9., überarbeitete Auflage



Springer Vieweg

Professor Dr.-Ing. Fernando Puente León
Karlsruher Institut für Technologie (KIT)
Institut für Industrielle Informationstechnik
Karlsruhe
Deutschland

Professor Dr.-Ing. Uwe Kiencke
Karlsruher Institut für Technologie (KIT)
Karlsruhe
Deutschland

Homepage zum Buch:
<http://www.iiit.kit.edu/messtechnik>

ISBN 978-3-642-30073-8
DOI 10.1007/978-3-642-30074-5

ISBN 978-3-642-30074-5 (eBook)

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

Springer Vieweg
© Springer-Verlag Berlin Heidelberg 1995, 2001, 2005, 2008, 2011, 2012
Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung, die nicht ausdrücklich vom Urheberrechtsgesetz zugelassen ist, bedarf der vorherigen Zustimmung des Verlags. Das gilt insbesondere für Vervielfältigungen, Bearbeitungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Die Wiedergabe von Gebrauchsnamen, Handelsnamen, Warenbezeichnungen usw. in diesem Werk berechtigt auch ohne besondere Kennzeichnung nicht zu der Annahme, dass solche Namen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung als frei zu betrachten wären und daher von jedermann benutzt werden dürften.

Gedruckt auf säurefreiem und chlorkfrei gebleichtem Papier

Springer Vieweg ist eine Marke von Springer DE.
Springer DE ist Teil der Fachverlagsgruppe Springer Science+Business Media
www.springer-vieweg.de

Für Barbara

Vorwort

In der vorliegenden neunten Auflage wurden Struktur und Inhalt des Buches weitestgehend beibehalten, weil sie sich im Lehrbetrieb bewährt haben. Dennoch wurden viele Textpassagen und Beispiele überarbeitet, um die Verständlichkeit des vermittelten Stoffes zu verbessern. Ferner enthalten die Abschnitte 2.2, 3.2, 4.1, 6.1 und 7.2 inhaltliche Ergänzungen.

Dank vielen Anregungen von Hörern der gleichnamigen Vorlesung am Karlsruher Institut für Technologie sowie von Lesern des Buches konnten darüber hinaus Fehler korrigiert werden.

Mein besonderer Dank gilt Herrn Dipl.-Ing. Johannes Pallauf für viele anregende Diskussionen sowie für das sorgfältige Korrekturlesen des Manuskriptes. Schließlich möchte ich mich beim Springer-Verlag für die wiederum ausgezeichnete Zusammenarbeit bedanken.

Karlsruhe, im Sommer 2012

Fernando Puente León

Vorwort zur 8. Auflage

Dieses Lehrbuch richtet sich an Studenten der Ingenieurwissenschaften (insbesondere der Elektrotechnik, Informationstechnik, Mechatronik und des Maschinenbaus) sowie der Informatik. Es behandelt die systemtheoretischen Grundlagen der Messtechnik. Dabei werden die allen Messsystemen gemeinsamen Verfahren in den Vordergrund gestellt. Der Inhalt des Buches umfasst

- die Beschreibung des physikalischen Verhaltens von Messsystemen durch ein mathematisches Modell,
- die Verbesserung der statischen sowie der dynamischen Eigenschaften von Messsystemen,
- die Messung zufälliger Größen und stochastischer Signale,
- die rechnergestützte Messdatenerfassung und -verarbeitung sowie
- die Erfassung frequenzanaloger Signale.

Vorausgesetzt werden Kenntnisse der gebräuchlichen Integraltransformationen und der Grundlagen der Wahrscheinlichkeitsrechnung. Bei den Herleitungen von Formeln sind alle notwendigen Zwischenschritte angegeben, damit der Leser sich den Stoff auch im Selbststudium aneignen kann.

Bei der Ausarbeitung dieses Buches fand ich das Buch *Messtechnik* (Springer-Verlag, 7. Auflage, 2008) meines verehrten Vorgängers, Herrn Pro-

fessor Uwe Kiencke, vor, das die Messtechnik in der oben geschilderten, systematischen Weise hervorragend aufbereitet. Ich habe es deshalb als Grundlage herangezogen und die Struktur des Buches weitgehend beibehalten, weil sie sich im Lehrbetrieb über viele Jahre bewährt hat.

Inhaltlich wurden im Kapitel *Messsysteme und Messfehler* Messskalen und damit auch die Messung qualitativer Größen eingeführt; ferner wurden die Abschnitte über metrische Größen und Messsysteme deutlich erweitert. Das Kapitel *Zufällige Messfehler* enthält zahlreiche inhaltliche Erweiterungen und neue Beispiele. Exemplarisch seien das Bayes-Theorem, die Behandlung höherer Momente von Verteilungen sowie multivariater Verteilungen und die Eigenschaften von Schätzern genannt. Im Kapitel *Dynamisches Verhalten von Messsystemen* wurde die Gliederung aus didaktischen Überlegungen überarbeitet und es wurden neue Erkenntnisse aus dem Gebiet der Systemoptimierung berücksichtigt. In nahezu jedem Kapitel wurden neue Abbildungen und praktische Beispiele aufgenommen, die ein tieferes Verständnis der behandelten Verfahren und ihrer Einsatzgebiete vermitteln.

Ferner wurde das gesamte Buch inhaltlich und redaktionell überarbeitet und praktisch jede Abbildung neu erstellt, um eine bessere Verständlichkeit sowie eine konsistenter Verwendung der mathematischen Symbole zu erzielen. Bei dieser Gelegenheit wurde die Darstellung von Vektoren und Matrizen an die international übliche Notation angepasst, stochastische Größen wurden zu ihrer leichteren Erkennung konsistent in Schreibmaschinenschrift gesetzt.

Darüber hinaus konnten durch Hinweise von Lesern und Hörern Fehler korrigiert und Erläuterungen verbessert werden. Besonderer Dank gilt Herrn Dipl.-Ing. Marco Kruse für die Erstellung von Matlab-Übungen, für wertvolle Hinweise und Diskussionen sowie für die zweimalige akribische Korrektur des Manuskripts. Von seinem fachkundigen Einsatz und seinem didaktischen Geschick hat dieses Buch außerordentlich stark profitiert. Schließlich sei dem Springer-Verlag für die ausgezeichnete Zusammenarbeit gedankt.

Karlsruhe, im Sommer 2011

Fernando Puente León

Inhaltsverzeichnis

| | |
|--|----|
| 1 Messsysteme und Messfehler | 1 |
| 1.1 Messskalen | 3 |
| 1.2 Metrische Größen | 5 |
| 1.2.1 Einheitensystem | 6 |
| 1.2.2 Anpassung der Definitionen der Einheiten | 8 |
| 1.3 Messsysteme | 10 |
| 1.3.1 Struktur von Messsystemen | 10 |
| 1.3.2 Beschreibung von Messsystemen im Zustandsraum | 12 |
| 1.3.3 Physikalische Messkennlinie | 13 |
| 1.3.4 Messsignale als Informationsträger | 14 |
| 1.4 Messfehler | 16 |
| 1.4.1 Absoluter und relativer Fehler | 16 |
| 1.4.2 Fehlerursachen | 17 |
| 1.4.3 Spezifizierte Normalbedingungen | 20 |
| 2 Kurvenanpassung | 23 |
| 2.1 Approximation | 26 |
| 2.1.1 Approximation mit orthonormalen Funktionensystemen | 26 |
| 2.1.1.1 Approximation mit der Fourier-Reihe | 29 |
| 2.1.1.2 Approximation mit Walsh-Funktionen | 30 |
| 2.1.2 Least-Squares-Schätzer | 31 |
| 2.1.3 Regressionsanalyse | 32 |
| 2.2 Interpolation | 34 |
| 2.2.1 Polynominterpolation | 34 |
| 2.2.2 Interpolation durch Lagrange-Polynome | 36 |
| 2.2.3 Interpolation durch Newton-Polynome | 38 |
| 2.2.4 Spline-Interpolation | 41 |
| 2.2.5 Systemtheoretische Deutung der Interpolation | 46 |
| 2.3 Kennfeldinterpolation | 47 |
| 3 Stationäres Verhalten von Messsystemen | 53 |
| 3.1 Stationäre Messkennlinie und deren Fehler | 55 |
| 3.1.1 Ideale und reale Messkennlinie | 55 |
| 3.1.2 Abgleich der Messkennlinie | 56 |
| 3.1.3 Kennlinienfehler bei realer Kennlinie | 58 |
| 3.1.3.1 Relativer Kennlinienfehler | 58 |
| 3.1.3.2 Hysterese und Umkehrspanne | 59 |

| | | |
|----------|--|-----|
| 3.1.4 | Abschätzung des Kennlinienfehlers | 61 |
| 3.2 | Kennlinienfehler unter Normalbedingungen | 64 |
| 3.2.1 | Herabsetzen des Messbereichs | 64 |
| 3.2.2 | Reihenschaltung zweier nichtlinearer Glieder | 67 |
| 3.2.3 | Wahl des günstigsten Messbereichs | 71 |
| 3.2.4 | Differenzmethode | 77 |
| 3.2.5 | Gegenkopplung | 83 |
| 3.3 | Kennlinienfehler bei Abweichungen von den Normalbedingungen | 87 |
| 3.3.1 | Superponierende Störgrößen | 87 |
| 3.3.2 | Unterdrückung superponierender Störgrößen mit der Differenzmethode | 89 |
| 3.3.3 | Deformierende Störgrößen | 90 |
| 3.3.4 | Deformierende Störgrößen bei Gegenkopplung | 93 |
| 3.3.5 | Superponierende Störgrößen bei Gegenkopplung | 96 |
| 3.3.6 | Kompensation systematischer Störeinflüsse | 97 |
| 3.3.7 | Abschirmung | 98 |
| 3.3.8 | Superponierende Störgrößen in Messketten | 98 |
| 3.3.9 | Zerhackerverstärker | 100 |
| 3.4 | Rückwirkung des Messsystems | 103 |
| 4 | Zufällige Messfehler | 107 |
| 4.1 | Grundlagen der Wahrscheinlichkeitstheorie | 110 |
| 4.1.1 | Wahrscheinlichkeitsdichte | 112 |
| 4.1.2 | Wahrscheinlichkeitsdichten abgebildeter Größen | 116 |
| 4.1.3 | Momente der Statistik 1. Ordnung | 117 |
| 4.1.4 | Momente der Statistik 2. Ordnung | 120 |
| 4.1.5 | Korrelationskoeffizient | 122 |
| 4.1.6 | Charakteristische Funktion | 125 |
| 4.2 | Stichproben | 126 |
| 4.2.1 | Häufigkeitsverteilung und Histogramm | 127 |
| 4.2.2 | Stichprobenmittelwert | 128 |
| 4.2.3 | Stichprobenvarianz | 131 |
| 4.2.4 | Gesetz der großen Zahlen | 134 |
| 4.2.5 | Mittelung zur Störungsunterdrückung | 136 |
| 4.2.5.1 | Lineare Kennlinie | 136 |
| 4.2.5.2 | Nichtlineare Kennlinie | 137 |
| 4.3 | Normalverteilte Zufallsvariable | 138 |
| 4.3.1 | Normalverteilung | 138 |

| | | |
|----------|---|-----|
| 4.3.2 | Zentraler Grenzwertsatz | 141 |
| 4.3.3 | χ^2 -Verteilung | 142 |
| 4.3.4 | Student'sche t-Verteilung | 147 |
| 4.4 | Statistische Testverfahren | 148 |
| 4.4.1 | Konfidenzintervall und statistische Sicherheit | 149 |
| 4.4.1.1 | Konfidenzintervall bei bekannter Standardabweichung | 149 |
| 4.4.1.2 | Konfidenzintervall bei geschätzter Standardabweichung | 152 |
| 4.4.2 | Hypothesen und statistische Tests | 155 |
| 4.4.3 | Signifikanztest für den Stichprobenmittelwert | 156 |
| 4.4.4 | χ^2 -Anpassungstest | 157 |
| 4.5 | Qualitätssicherung | 161 |
| 4.5.1 | Beurteilung von Fertigungsprozessen | 161 |
| 4.5.2 | Bestimmung der Ausfallrate | 164 |
| 4.5.3 | Statistische Prozessüberwachung | 169 |
| 4.6 | Fehlerfortpflanzung | 173 |
| 5 | Dynamisches Verhalten von Messsystemen | 179 |
| 5.1 | Beschreibung von Messsystemen | 181 |
| 5.1.1 | Systemeigenschaften | 181 |
| 5.1.2 | Lineare, zeitinvariante Systeme (LTI-Systeme) | 182 |
| 5.1.3 | Stabilität | 184 |
| 5.2 | Empirische Analyse von Messsystemen | 186 |
| 5.2.1 | Kennwerte der Sprungantwort | 186 |
| 5.2.2 | Nichtlineares dynamisches Verhalten | 187 |
| 5.2.3 | Bestimmung des Frequenzganges | 188 |
| 5.3 | Verbesserung des dynamischen Systemverhaltens | 192 |
| 5.3.1 | Optimierung der Systemparameter | 193 |
| 5.3.2 | Änderung der Systemstruktur | 194 |
| 5.4 | Parameteroptimierung | 195 |
| 5.4.1 | Kriterium „verschwindende Momente der Impulsantwort“ | 198 |
| 5.4.2 | Kriterium „konstanter Amplitudengang für kleine Frequenzen“ | 202 |
| 5.4.3 | Kriterium „konstanter Realteil des Frequenzganges“ | 208 |
| 5.4.4 | ITAE-Kriterium | 214 |
| 5.4.4.1 | System 1. Ordnung | 214 |
| 5.4.4.2 | System 2. Ordnung | 215 |

| | | |
|----------|---|-----|
| 5.4.5 | Kriterium „quadratisches Fehlerintegral“ | 219 |
| 5.5 | Strukturänderung zur Optimierung des Zeitverhaltens | 226 |
| 5.5.1 | Kompensation des Zeitverhaltens | 226 |
| 5.5.2 | Zeitverhalten bei Gegenkopplung | 231 |
| 5.5.2.1 | P-Regler | 232 |
| 5.5.2.2 | PI-Regler | 236 |
| 6 | Messung stochastischer Signale | 241 |
| 6.1 | Stochastische Prozesse | 245 |
| 6.1.1 | Einführung | 245 |
| 6.1.2 | Wahrscheinlichkeitsverteilung und Wahrscheinlichkeitsdichte | 247 |
| 6.1.3 | Schar- und Zeitmittelwerte | 250 |
| 6.1.4 | Momente der Statistik 1. Ordnung | 250 |
| 6.1.5 | Momente der Statistik 2. Ordnung | 252 |
| 6.1.6 | Stationäre Prozesse | 254 |
| 6.1.7 | Ergodische Prozesse | 256 |
| 6.2 | Korrelationsfunktionen | 261 |
| 6.2.1 | Signalklassen | 261 |
| 6.2.1.1 | Ergodische Prozesse | 263 |
| 6.2.1.2 | Funktionenräume | 263 |
| 6.2.2 | Korrelation von Leistungssignalen | 265 |
| 6.2.3 | Korrelation von Energiesignalen | 267 |
| 6.2.4 | Eigenschaften von Auto- und Kreuzkorrelationsfunktion | 269 |
| 6.3 | Korrelationsmesstechnik | 271 |
| 6.3.1 | Messung von Korrelationsfunktionen | 271 |
| 6.3.2 | Ähnlichkeit von Signalen, Laufzeitmessung | 272 |
| 6.3.3 | Closed-loop-Korrelation | 278 |
| 6.3.4 | Polaritätskorrelation | 280 |
| 6.3.5 | Ähnlichkeit von Spektren, Dopplerfrequenzmessung . | 281 |
| 6.3.6 | Selbstähnlichkeit | 284 |
| 6.4 | Spektrale Darstellung stochastischer Signale | 285 |
| 6.4.1 | Leistungsdichtespektrum | 285 |
| 6.4.2 | Rauschen | 288 |
| 6.4.3 | Überlagerung zufälliger Störsignale | 293 |
| 6.4.4 | Übertragung stochastischer Signale durch LTI-Systeme | 295 |
| 6.5 | Systemidentifikation | 298 |
| 6.5.1 | Schätzung des Leistungsdichtespektrums | 299 |
| 6.5.2 | Systemidentifikation bei geschätzter Leistungsdichte . | 301 |

| | | |
|----------|--|-----|
| 6.5.2.1 | Quotientenbildung gemittelter Periodogramme | 301 |
| 6.5.2.2 | Quotientenbildung gemittelter Kreuzleistungsdichten | 302 |
| 6.5.3 | Dynamische Systemidentifikation | 303 |
| 6.6 | Wiener-Filter | 304 |
| 6.6.1 | Signalmodell | 305 |
| 6.6.2 | Herleitung des Wiener-Filters | 307 |
| 6.6.3 | Wiener-Filter bei linearer Verzerrung und additivem Rauschen | 308 |
| 7 | Erfassung analoger Signale | 315 |
| 7.1 | Abtastung | 318 |
| 7.1.1 | Abtasttheorem | 318 |
| 7.1.2 | Bandüberlappungsfehler (<i>Aliasing</i>) | 320 |
| 7.1.3 | Anti-Aliasing-Filter | 322 |
| 7.1.4 | Mittelwertbildung bei endlicher Abtastdauer | 325 |
| 7.1.5 | Zeitliche Abtastfehler | 328 |
| 7.2 | Quantisierung | 333 |
| 7.2.1 | Wahrscheinlichkeitsdichte der Signalwerte | 336 |
| 7.2.2 | Amplitudendichte der Fourier-Reihe | 337 |
| 7.2.3 | Quantisierungstheorem | 340 |
| 7.2.4 | Wahrscheinlichkeitsdichte des Quantisierungsfehlers . | 345 |
| 7.2.5 | Signal-Rausch-Verhältnis infolge der Quantisierung . | 347 |
| 7.2.5.1 | Sinusförmige Signale | 348 |
| 7.2.5.2 | Signale mit gleichverteilten Amplituden | 348 |
| 7.2.6 | Optimale Quantisierung | 349 |
| 7.2.7 | Minimierung des relativen Quantisierungsfehlers . | 349 |
| 7.2.8 | Dithering | 352 |
| 7.3 | Analog-Digital-Umsetzer | 357 |
| 7.3.1 | A/D-Nachlaufumsetzer | 357 |
| 7.3.2 | A/D-Umsetzer mit sukzessiver Approximation | 359 |
| 7.3.3 | Integrierende A/D-Umsetzer | 360 |
| 7.3.4 | Delta-Sigma-Umsetzer | 363 |
| 7.3.4.1 | Delta-Sigma-Modulator | 363 |
| 7.3.4.2 | Rauschformung (<i>noise shaping</i>) | 365 |
| 7.3.4.3 | Digitalfilter | 368 |
| 7.3.4.4 | Stationäres Verhalten | 369 |
| 7.3.5 | Ratiometrische Messung | 372 |
| 7.4 | Digital-Analog-Umsetzer | 372 |

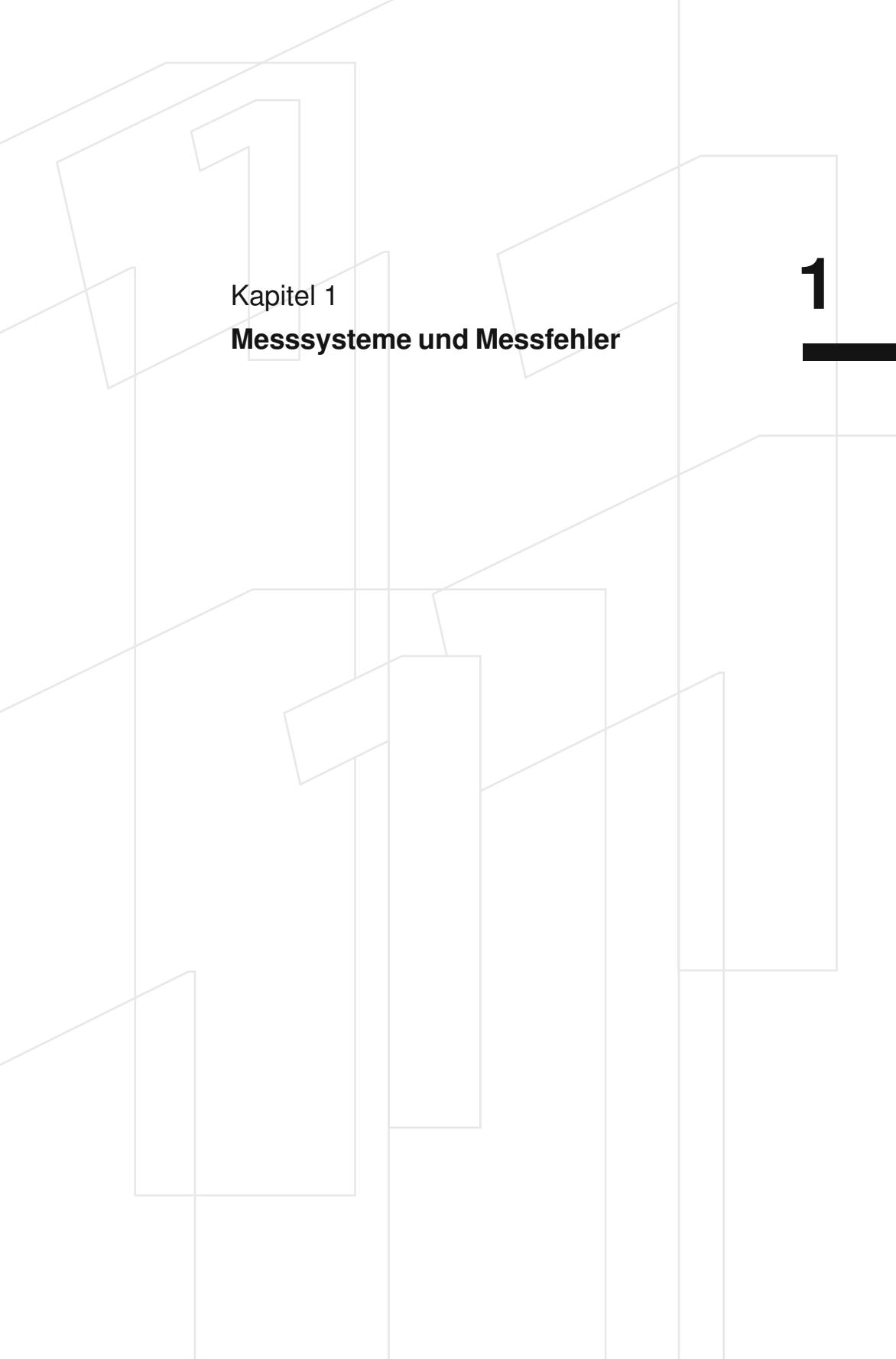
| | | |
|-----------------------------|--|-----|
| 7.4.1 | Parallele D/A-Umsetzer | 373 |
| 7.4.1.1 | D/A-Umsetzer mit dyadisch gestuften Widerständen | 373 |
| 7.4.1.2 | D/A-Umsetzer mit $R/2R$ -Kettenleiternetzwerk | 373 |
| 7.4.2 | Serielle D/A-Umsetzer | 375 |
| 8 | Erfassung frequenzanaloger Signale | 379 |
| 8.1 | Allgemeiner Frequenzbegriff | 382 |
| 8.2 | Digitale Drehzahlmessung | 389 |
| 8.2.1 | Periodendauermessung | 390 |
| 8.2.2 | Frequenzmessung | 391 |
| 8.2.3 | Maximaler Quantisierungsfehler für einen Zählvorgang | 392 |
| 8.2.3.1 | Quantisierungsfehler bei der Periodendauermessung | 393 |
| 8.2.3.2 | Quantisierungsfehler bei der Frequenzmessung | 393 |
| 8.2.4 | Mittelwertbildung bei der Drehzahlmessung | 395 |
| 8.2.5 | Abtastung bei der Drehzahlmessung | 397 |
| 8.2.6 | Quantisierung bei fortlaufenden Periodendauermessungen | 398 |
| 8.2.7 | Leistungsdichte des Quantisierungsfehlers | 402 |
| 8.2.8 | Kompensation mechanischer Fehler des Sensorrades | 405 |
| 8.2.8.1 | Stochastische Zahnflankenfehler | 405 |
| 8.2.8.2 | Schätzung der Zahnflankenfehler | 407 |
| 8.3 | Kontinuierliche Frequenzmessung | 411 |
| 8.3.1 | Phasenregelkreis | 411 |
| 8.3.2 | Frequenzregelkreis | 416 |
| 8.4 | Positions- und Richtungserkennung | 418 |
| 8.4.1 | Drehrichtungserkennung | 418 |
| 8.4.2 | Positionsbestimmung | 420 |
| A | Symbole und Tabellen | 423 |
| A.1 | Symbolverzeichnis | 425 |
| A.1.1 | Konventionen | 425 |
| A.1.2 | Operatoren | 425 |
| A.1.3 | Lateinische Symbole | 425 |
| A.1.4 | Griechische Symbole | 429 |
| A.2 | Tabellen | 430 |
| Literaturverzeichnis | | 431 |

Inhaltsverzeichnis

XV

Index

435



Kapitel 1

Messsysteme und Messfehler

1

1 Messsysteme und Messfehler

| | | |
|-------|--|----|
| 1.1 | Messskalen | 3 |
| 1.2 | Metrische Größen | 5 |
| 1.2.1 | Einheitensystem | 6 |
| 1.2.2 | Anpassung der Definitionen der Einheiten | 8 |
| 1.3 | Messsysteme | 10 |
| 1.3.1 | Struktur von Messsystemen | 10 |
| 1.3.2 | Beschreibung von Messsystemen im Zustandsraum . . | 12 |
| 1.3.3 | Physikalische Messkennlinie | 13 |
| 1.3.4 | Messsignale als Informationsträger | 14 |
| 1.4 | Messfehler | 16 |
| 1.4.1 | Absoluter und relativer Fehler | 16 |
| 1.4.2 | Fehlerursachen | 17 |
| 1.4.3 | Spezifizierte Normalbedingungen | 20 |

1 Messsysteme und Messfehler

Die Grundaufgabe der Messtechnik besteht in der Erfassung physikalischer Größen mit einer vorgegebenen Genauigkeit. Da prinzipiell jedes Messverfahren fehlerbehaftet ist, müssen die damit verbundenen Fehler abgeschätzt werden, um Aussagen über die Messgenauigkeit treffen zu können. Ziel ist es, die entstehenden Messfehler möglichst klein zu halten, wofür gegebenenfalls vorhandene Störeinflüsse kompensiert werden müssen. Dazu bedarf es allerdings in der Regel einer Systembeschreibung des Messvorgangs mit möglichst all seinen Eigenschaften und Einflussgrößen. Liegt diese Beschreibung vor, so können mit systemtheoretischen Untersuchungen die Fehlereinflüsse modelliert und durch Systemoptimierung deren Einfluss minimiert werden.

1.1 Messskalen

1.1

Im weitesten Sinne kann das **Messen** als eine Zuordnung von mathematischen Symbolen (z. B. Zahlen) zu bestimmten Merkmalen empirischer Objekte basierend auf objektiven Regeln beschrieben werden. Auf der Grundlage der bestehenden Merkmalsausprägungen lassen sich Zusammenhänge – man spricht hier von „empirischen Relationen“ – zwischen den Objekten herstellen.

Die Menge der empirischen Objekte zusammen mit den zugehörigen Relationen wird als empirisches relationales System (ERS) bezeichnet. Messen wird in diesem Kontext zu einer homomorphen Abbildung des ERS auf ein mathematisches relationales System (MRS), wobei der Homomorphismus dafür sorgt, dass die Relationen zwischen den mathematischen Symbolen die Relationen zwischen den empirischen Objekten widerspiegeln [9]. Unterschiedliche Objekte mit identischer Merkmalsausprägung werden auf dasselbe mathematische Symbol abgebildet.

Abhängig von der Aussagekraft der betrachteten Relationen zwischen den Objekten werden verschiedene **Skalentypen** (Skalenniveaus) unterschieden. Innerhalb eines Skalentyps sind gewisse Transformationen des MRS zulässig, ohne die Relationen im ERS zu beeinträchtigen. Tabelle 1.1 zeigt die wichtigsten Skalentypen zusammen mit den empirischen Relationen, die sie zu beschreiben vermögen.

- **Nominalskala:** Die Nominalskala weist den geringsten Informationsgehalt auf. Es muss lediglich sichergestellt werden, dass gleiche bzw. unterschiedliche Merkmalsausprägungen im ERS auf gleiche bzw. unter-

Tabelle 1.1. Messskalen.

| | Nominal-skala | Ordinal-skala | Intervall-skala | Verhältnis-skala |
|----------------------------|--|---|--|---|
| empirische Relationen | Äquivalenz | Äquivalenz, Ordnung | Äquivalenz, Ordnung, empirische Addition | Äquivalenz, Ordnung, empirische Addition, empirische Multiplikation |
| zulässige Transformationen | $\tilde{u} = f(u)$ mit $f(\cdot)$ bijektiv | $\tilde{u} = f(u)$ mit $f(\cdot)$ streng monoton steigend | $\tilde{u} = a u + b$ mit $a > 0$ | $\tilde{u} = a u$ mit $a > 0$ |
| Lageparameter | Modalwert | Median | arithmetischer Mittelwert | harmonischer/geometrischer Mittelwert |
| Streuungsmaße | Entropie | Quantile | Varianz | Variationskoeffizient |
| weitere Kenngrößen | Transinformation | Rangkorrelation | Korrelationskoeffizient | Variationskoeffizient, Korrelationskoeff. |
| mathematische Struktur | Menge | total geordnete Menge | affine Gerade | Körper |
| Beispiele von Merkmalen | Geschlecht, Blutgruppe | Härtegrad, Güteklassse | Temperatur, Kalenderzeit | Länge, Masse, Stromstärke |

schiedliche Symbole im MRS abgebildet werden. Als Symbole kommen neben Zahlen auch Begriffe in Frage. Zulässig sind beliebige bijektive (d. h. eineindeutige) Transformationen.

Nominale Merkmale nennt man auch *kategorial* oder *begrifflich*, da sie rein attributiven Charakter haben. Exemplarisch seien hierzu das Geschlecht und die Blutgruppe genannt.

- **Ordinalskala:** Durch eine Ordinalskala wird die empirische Ordnungsstruktur zwischen empirischen Objekten verschiedener Merkmalsausprägung wiedergegeben. Zulässig sind hier streng monoton steigende Transformationen $\tilde{u} = f(u)$.

Ordinale Merkmale nennt man auch *komparative Merkmale* oder *Rangmerkmale*, da sie die Festlegung einer Ordnung zwischen den Objekten zulassen. Beispiele hierzu sind Güteklassen sowie Härtegrade nach Mohs.

- **Intervallskala:** Hier liegt eine empirische Relation vor, die mathematisch durch Addition und Subtraktion repräsentiert werden kann. Differenzen ergeben somit einen Sinn, Verhältnisse jedoch nicht. Zulässig sind lineare Transformationen $\tilde{u} = a u + b$ mit $a > 0$. Der Nullpunkt der Skala ist willkürlich und kann über den Parameter b festgelegt werden, die Skaleneinheit kann über a frei gewählt werden.

Als Lageparameter kommt neben dem Modalwert (Modus) und dem Median auch der arithmetische Mittelwert in Betracht, als Maß für die statistische Streuung der Werte die Standardabweichung.

- **Verhältnisskala:** Bei dieser Skala liegt ein natürlicher Nullpunkt vor, während über die Skaleneinheit noch frei verfügt werden kann. Gegenüber der Intervallskala liegt zusätzlich eine empirische Relation vor, die im MRS durch Multiplikation bzw. Division repräsentiert wird.
Zulässige Transformationen haben die Form $\tilde{u} = a u$ mit $a > 0$.

Die Intervall- und die Verhältnisskala werden unter dem Oberbegriff **Kardinalskala** oder **metrische Skala** zusammengefasst. Die zugehörigen Merkmale werden als quantitative bzw. **metrische Merkmale** bezeichnet und damit von den meist als qualitativ bezeichneten nominalen oder ordinalen Merkmalen abgehoben.

1.2 Metrische Größen

1.2

Die meisten technisch relevanten **Messgrößen** sind metrischer Natur. Deren Messung wird durch die folgende Definition abgedeckt.

Definition 1.1: Messen metrischer Größen

1.1

Eine metrische Größe zu **messen** heißt, ihre Ausprägung quantitativ zu erfassen. Dafür wird die Messgröße mit einer zuvor vereinbarten **Maßeinheit** – dem **Normal** (engl. *measurement standard*) – verglichen:

$$\text{Messgröße} = \text{Zahlenwert} \cdot \text{Maßeinheit}. \quad (1.1)$$

Der Zahlenwert der Messgröße gibt dabei an, wie oft in der Messgröße die Maßeinheit enthalten ist. ◇

Diese Definition ist zwar für manche Aufgabenstellungen zu eng gefasst. Da in der Technik jedoch die Erfassung metrischer Größen überwiegt, wird im Folgenden auf die Betrachtung weiterer Definitionen bewusst verzichtet.

Für eine Messung sind zwei Voraussetzungen notwendig. Zum einen muss die zu messende Größe eindeutig definiert sein. Zum anderen muss die Einheit oder das Normal durch eine Konvention festgelegt werden. Die Festlegung einer Einheit ist im Prinzip willkürlich. Man sollte jedoch die folgenden Bedingungen beachten:

- Jede Einheit sollte praktisch sein – sie sollte also sowohl im Alltagsleben als auch in der Wissenschaft verwendbar sein.

- Sie sollte gut reproduzierbar sein.
- Das Normal sollte möglichst unveränderlich sein.

➲ 1.2.1 Einheitensystem

Für eine widerspruchsfreie Darstellung aller physikalischen Größen genügt es, aus ihnen sieben **Basisgrößen** auszuwählen, denen per Definition eine bestimmte qualitative und quantitative Eigenschaft zugeordnet wird. Alle anderen physikalischen Größen können als Verhältnis zu diesen Basisgrößen ausgedrückt werden.

Bis Mitte des 19. Jahrhunderts orientierten sich jedoch die meisten Maßeinheiten am Menschen (z. B. Elle¹ als Längenmaß). Viele Maßeinheiten waren regional unterschiedlich festgelegt (die Freiburger Elle war z. B. 54 cm lang, die Bremer Elle brachte es auf 55 cm, die Badische Elle auf 60 cm), so dass die individuell festgelegten Einheiten den internationalen Handel erschwerteten und Probleme in Forschung, Technik und Kommunikation hervorbrachten.

1790 beschloss die französische Nationalversammlung die Schaffung eines neuen **Einheitensystems**. Es sollte ausschließlich auf objektiven physikalischen Kriterien gründen und allen Nationen zugänglich sein. 1799 wurde mit dem „*Mètre des Archives*“ und dem „*Kilogramme des Archives*“ das erste metrische System vollendet. Bedingt durch politische Umstände kam es jedoch erst 1875 zum Abschluss der Meterkonvention, welche von 17 Staaten unterzeichnet wurde. Das **internationale Einheitensystem SI** (benannt nach „*le Système Internationale d'unités*“) stellt das heute gültige Einheitensystem in Deutschland dar. Die sieben in der Tabelle 1.2 aufgeführten Basisgrößen und **Basiseinheiten** des SI-Systems sind im Folgenden dargestellt [20].

- **Meter:** Das Meter ist die Länge der Strecke, die Licht im Vakuum in einem Intervall von $(1/299\,792\,458)$ Sekunden durchläuft. Die Meterdefinition weist der Lichtgeschwindigkeit c_0 einen festen Wert zu. Diese Fundamentalkonstante kann somit nicht mehr gemessen werden, sie ist jetzt exakt vorgegeben. Hieraus folgt, dass die Längeneinheit von der Zeiteinheit Sekunde abhängt.
- **Kilogramm:** Das Kilogramm ist die Einheit der Masse; es ist gleich der Masse des internationalen Kilogrammprototyps.

¹Die Elle als Maßeinheit bezeichnete ursprünglich die Länge des Unterarmes, war jedoch dann meist länger als der Abstand zwischen Ellbogen und Mittelfingerspitze.

Tabelle 1.2. Basisgrößen und Basiseinheiten.

| Gebiet | Basisgröße | Symbol | Basiseinheit | Zeichen |
|----------------|-------------|--------|--------------|---------|
| Mechanik | Länge | l | Meter | m |
| | Masse | m | Kilogramm | kg |
| | Zeit | t | Sekunde | s |
| Elektrotechnik | Stromstärke | I | Ampere | A |
| Thermodynamik | Temperatur | T | Kelvin | K |
| Optik | Lichtstärke | I_L | Candela | cd |
| Chemie | Stoffmenge | | Mol | mol |

- **Sekunde:** Die Sekunde ist das $9\,192\,631\,770$ -fache der Periodendauer der dem Übergang zwischen den beiden Hyperfeinstrukturniveaus des Grundzustandes von Atomen des Nuklids ^{133}Cs entsprechenden Strahlung.
- **Ampere:** Das Ampere ist die Stärke eines konstanten elektrischen Stromes, der, durch zwei parallele, geradlinige, unendlich lange und im Vakuum im Abstand von einem Meter voneinander angeordnete Leiter von vernachlässigbar kleinem, kreisförmigen Querschnitt fließend, zwischen diesen Leitern mit je einem Meter Leiterlänge die Kraft $2 \cdot 10^{-7}$ Newton hervorrufen würde.
- **Kelvin:** Das Kelvin, die Einheit der thermodynamischen Temperatur, ist der 273,16-te Teil der thermodynamischen Temperatur des Tripelpunktes des Wassers. Temperaturdifferenzen dürfen auch in Grad Celsius, mit dem Einheitenzeichen $^{\circ}\text{C}$, angegeben werden.
- **Mol:** Das Mol ist die Stoffmenge eines Systems, das aus ebensoviel Einzelteilchen besteht, wie Atome in 0,012 Kilogramm des Kohlenstoffnuklids ^{12}C enthalten sind. Bei Benutzung des Mol müssen die Einzelteilchen spezifiziert sein und können Atome, Moleküle, Ionen, Elektronen sowie andere Teilchen oder Gruppen solcher Teilchen genau angegebener Zusammensetzung sein.
- **Candela:** Die Candela ist die Lichtstärke in einer bestimmten Richtung einer Strahlungsquelle, die monochromatische Strahlung der Frequenz $540 \cdot 10^{12}$ Hertz aussendet und deren Strahlungsstärke in dieser Richtung $1/683$ Watt durch Steradian beträgt.

Vorteilhaft am SI-Einheitensystem ist die Tatsache, dass die Einheiten zueinander **kohärent** (d. h. aufeinander abgestimmt) sind. Abgeleitete Einheiten (wie z. B. m/s) können durch Multiplikation und Division von Basiseinheiten ermittelt werden, wobei keine Proportionalitätsfaktoren nötig sind.

Ebenfalls genormt wurden die Vorsätze der internationalen Einheiten, wobei der Zahlenfaktor sich stets auf das Dezimalsystem bezieht. Die wichtigsten genormten Vorsätze zur Umschreibung von dezimalen Vielfachen und Teilen der Einheiten sind in der Tabelle 1.3 aufgeführt. Der Vorsatz bei der jeweiligen Einheit vermeidet unhandliche Zahlenwerte und erleichtert somit auch den Umgang mit den Einheiten im Sprachgebrauch. Ebenfalls hilfreich sind die abgeleiteten SI-Einheiten, da durch ihre Festlegung die jeweilige Formelgröße nicht stets mit den kompletten Basiseinheiten aufgeführt werden muss. Eine Auswahl wichtiger abgeleiteter SI-Einheiten zeigt die Tabelle 1.4.

❸ 1.2.2 Anpassung der Definitionen der Einheiten

Die Definition der Normale ist nicht für alle Zeiten fest vorgegeben. Das Comité International des Poids et Mesures (CIPM) prüft auf regelmäßigen Sitzungen den aktuellen Entwicklungsstand und beschließt gegebenenfalls Veränderungen. Am Beispiel der Längeneinheit *Meter* sollen die Veränderungen des Normals dargestellt werden:

- **1889:** Aus 37 Prototypen wird per Losverfahren das „Urmeter“ bestimmt. Es handelt sich dabei um einen x-förmigen Stab aus Platin-Iridium, auf welchem der Abstand zweier Strichmarken bei einer Stabtemperatur von 0 °C die Längeneinheit 1 m darstellt. Das Urmeter wird mit sechs weiteren in Sèvres bei Paris aufbewahrt. Die restlichen Normale werden an die Unterzeichnerstaaten der Meterkonvention verteilt.
- **1960:** Auf der 11. Generalkonferenz wird das Meter als das 1 650 763,73-fache der Wellenlänge der von Atomen des Nuklids ^{86}Kr (Krypton) beim Übergang vom Zustand $5d_5$ zum Zustand $2p_{10}$ ausgesandten, sich im Vakuum ausbreitenden Strahlung festgelegt.
- **1983:** Auf der 17. Generalkonferenz wird das Meter in seiner bis heute gültigen Form definiert. Seither gilt als ein Meter die Länge, die das Licht im Vakuum in einem Intervall von 1/299 792 458 Sekunden zurücklegt.

Bei der Anpassung der Definitionen der Normale an den Stand der Technik versucht man einerseits, die SI-Einheiten (z. B. das Ampere) auf die sogenannten Naturkonstanten zurückzuführen (siehe Tab. 1.5). Andererseits hängen letztere vom gewählten Einheitensystem ab.

Tabelle 1.3. Genormte Vorsätze.

| Vorsatz | Zeichen | Zahlenwert | Vorsatz | Zeichen | Zahlenwert |
|---------|---------|------------|---------|---------|------------|
| Yotta | Y | 10^{24} | Yokto | y | 10^{-24} |
| Zetta | Z | 10^{21} | Zepto | z | 10^{-21} |
| Exa | E | 10^{18} | Atto | a | 10^{-18} |
| Peta | P | 10^{15} | Femto | f | 10^{-15} |
| Tera | T | 10^{12} | Piko | p | 10^{-12} |
| Giga | G | 10^9 | Nano | n | 10^{-9} |
| Mega | M | 10^6 | Mikro | μ | 10^{-6} |
| Kilo | k | 10^3 | Milli | m | 10^{-3} |
| Hekto | h | 10^2 | Zenti | c | 10^{-2} |
| Deka | da | 10^1 | Dezi | d | 10^{-1} |

Tabelle 1.4. Abgeleitete SI-Einheiten und zugehörige Größen.

| Größe und Formelzeichen | SI-Einheit | Beziehung |
|-------------------------|----------------|--|
| ebener Winkel | α | Radiant rad $1 \text{ rad} = 1 \text{ m/m}$ |
| Raumwinkel | Ω | Steradian sr $1 \text{ sr} = 1 \text{ m}^2/\text{m}^2$ |
| Frequenz | f | Hertz Hz $1 \text{ Hz} = 1/\text{s}$ |
| Kraft | F | Newton N $1 \text{ N} = 1 \text{ kg m/s}^2$ |
| Druck | p | Pascal Pa $1 \text{ Pa} = 1 \text{ N/m}^2$ |
| Energie | E | Joule J $1 \text{ J} = 1 \text{ N m} = 1 \text{ W s}$ |
| Arbeit | W | Joule J $1 \text{ J} = 1 \text{ N m} = 1 \text{ W s}$ |
| Wärmemenge | Q | Joule J $1 \text{ J} = 1 \text{ N m} = 1 \text{ W s}$ |
| Leistung | P | Watt W $1 \text{ W} = 1 \text{ J/s} = 1 \text{ N m/s}$ |
| elektrische Ladung | Q | Coulomb C $1 \text{ C} = 1 \text{ A s}$ |
| elektrische Spannung | U | Volt V $1 \text{ V} = 1 \text{ W/A}$ |
| elektrische Kapazität | C | Farad F $1 \text{ F} = 1 \text{ C/V} = \text{A s/V}$ |
| elektrischer Widerstand | R | Ohm Ω $1 \Omega = 1 \text{ W/A}^2$ |
| elektrischer Leitwert | G | Siemens S $1 \text{ S} = 1/\Omega$ |
| Induktivität | L | Henry H $1 \text{ H} = 1 \text{ Wb/A} = 1 \text{ Vs/A}$ |
| magnetischer Fluss | Φ | Weber Wb $1 \text{ Wb} = 1 \text{ Vs}$ |
| magnetische Flussdichte | B | Tesla T $1 \text{ T} = 1 \text{ Vs/m}^2$ |
| Lichtstrom | Φ | Lumen lm $1 \text{ lm} = 1 \text{ cd sr}$ |
| Beleuchtungsstärke | E _v | Lux lx $1 \text{ lx} = 1 \text{ lm/m}^2$ |
| Radioaktivität | A | Becquerel Bq $1 \text{ Bq} = 1 \text{ s}^{-1}$ |
| Energiedosis | D | Gray Gy $1 \text{ Gy} = 1 \text{ m}^2/\text{s}^2$ |
| Äquivalentdosis | H | Sievert Sv $1 \text{ Sv} = 1 \text{ m}^2/\text{s}^2$ |
| katalytische Aktivität | z | Katal kat $1 \text{ kat} = 1 \text{ mol/s}$ |

Tabelle 1.5. Naturkonstanten.

| Naturkonstante | Zeichen | Zahlenwert | Einheit |
|--------------------------------|--------------|----------------------------------|---|
| Avogadro-Konstante | N_A | $6,022\,14129 \cdot 10^{23}$ | mol^{-1} |
| Boltzmann-Konstante | k | $1,380\,6488 \cdot 10^{-23}$ | J K^{-1} |
| Elementarladung | e_0 | $1,602\,176\,565 \cdot 10^{-19}$ | As |
| elektrische Feldkonstante | ϵ_0 | $8,854\,187\,817 \cdot 10^{-12}$ | $\text{As V}^{-1} \text{m}^{-1}$ |
| Faraday-Konstante | F | 96 485,3365 | C mol^{-1} |
| Feinstrukturkonstante | α | $7,297\,352\,5698 \cdot 10^{-3}$ | |
| Gravitationskonstante | G | $6,67384 \cdot 10^{-11}$ | $\text{m}^3 \text{kg}^{-1} \text{s}^{-2}$ |
| Josephson-Konstante | K_J | $4,835\,97870 \cdot 10^{14}$ | Hz V^{-1} |
| Klitzing-Konstante | R_K | 25 812,807 443 | Ω |
| kosmologische Konstante | Λ | $2,076\,504 \cdot 10^{-43}$ | $\text{s}^2 \text{kg}^{-1} \text{m}^{-1}$ |
| Lichtgeschwindigkeit im Vakuum | c_0 | 299 792 458 | m s^{-1} |
| Loschmidt-Konstante | N_L | $2,686\,7805 \cdot 10^{25}$ | m^{-3} |
| magnetische Feldkonstante | μ_0 | $12,566\,370\,614 \cdot 10^{-7}$ | $\text{Vs A}^{-1} \text{m}^{-1}$ |
| magnetisches Flussquantum | Φ_0 | $2,067\,833\,758 \cdot 10^{-15}$ | Wb |
| Planck'sches Wirkungsquantum | h | $6,626\,06957 \cdot 10^{-34}$ | Js |
| Ruhemasse des Elektrons | m_0 | $9,109\,38291 \cdot 10^{-31}$ | kg |
| Ruhemasse des Neutrons | m_n | $1,674\,927\,351 \cdot 10^{-27}$ | kg |
| Ruhemasse des Protons | m_p | $1,672\,621\,777 \cdot 10^{-27}$ | kg |
| Rydberg-Konstante | R_∞ | 10 973 731,568 539 | m^{-1} |
| Stefan-Boltzmann-Konstante | σ | $5,670\,373 \cdot 10^{-8}$ | $\text{W m}^{-2} \text{K}^{-4}$ |
| universelle Gaskonstante | R | 8,314 4621 | $\text{J mol}^{-1} \text{K}^{-1}$ |

1.3

1.3 Messsysteme

1.3.1 Struktur von Messsystemen

Um physikalische Größen zu messen, benötigt man ein Messsystem. Je nach Art der Messaufgabe handelt es sich hierbei um Systeme unterschiedlichster Komplexität. Die einfachsten Systeme ergeben sich bei den sogenannten **direkten Messverfahren**, bei denen der gesuchte Messwert durch unmittelbaren Vergleich mit einem Bezugswert derselben Messgröße gewonnen wird.

Beispiel 1.1 (Balkenwaage): Bei einer Balkenwaage wird die unbekannte Masse m mit der bekannten Masse der Gewichtssteine verglichen. Das Messverfahren ist daher direkt. ■

Die meisten physikalischen Größen können hingegen nur **indirekt** gemessen werden. Die gesuchte Messgröße wird dabei über physikalische Zusammenhänge auf andere Größen zurückgeführt und aus diesen ermittelt.

Beispiel 1.2 (Federwaage): Die Masse m soll über die Auslenkung x einer Feder mit der Federkonstanten c bestimmt werden. Über das Kräftegleichgewicht erhält man

$$m g = c x \quad \Rightarrow \quad m = \frac{c x}{g}, \quad (1.2)$$

wobei g die Erdbeschleunigung bezeichnet. Die Messgröße ist die Masse m , abgelesen wird allerdings die Auslenkung x , die über den Zusammenhang des Kräftegleichgewichts auf die gesuchte Größe zurückgeführt wird. ■



Abbildung 1.1. Allgemeine Struktur eines Messsystems.

Anhand des in Abb. 1.1 aufgeführten Signalflussplans sollen die wichtigsten Komponenten eines Messsystems erläutert werden. Dabei müssen in einem realen Messsystem nicht zwangsläufig alle in Abb. 1.1 dargestellten Komponenten enthalten sein.

- Am Anfang der **Messkette** befindet sich der **Aufnehmer** (auch **Sensor** oder **Fühler** genannt), an dessen Eingang die zu messende Größe u anliegt. Am Ausgang liefert der Aufnehmer ein (meist elektrisches) Signal, welches von der Messgröße abhängt.
- Aufgabe des **Messumformers** ist es, das Eingangssignal in ein zur Weiterverarbeitung (z. B. Digitalisierung, Filterung, Übertragung, Speicherung) geeignetes, oft normiertes elektrisches Ausgangssignal abzubilden. Dafür kommen u. a. Messverstärker zum Einsatz.
- Die **Signalverarbeitung** hat zum Ziel, aus dem Eingangssignal den informationstragenden Parameter zu extrahieren (z. B. Amplitude oder Frequenz) und daraus das Messergebnis zu ermitteln. In modernen Messsystemen wird dafür meist das Eingangssignal digitalisiert und dessen Verarbeitung mit einem Digitalrechner oder einem Mikrocontroller realisiert.

Um das Verhalten von Messsystemen untersuchen zu können, ist zunächst eine Beschreibung des Systems erforderlich – man spricht hierbei von einer **Modellbildung**. Sie dient dazu, einen mathematischen Zusammenhang zwischen Ein- und Ausgangsgrößen am System zu ermitteln.

1.3.2 Beschreibung von Messsystemen im Zustandsraum

Zur Beschreibung des dynamischen Verhaltens eines Messsystems können neben der zu erfassenden Messgröße (Eingangsgröße) und dem angezeigten Wert (Ausgangsgröße) auch die inneren **Zustandsgrößen** des Systems berücksichtigt werden. Diese Zustandsgrößen vermitteln gewissermaßen zwischen Ein- und Ausgang und werden im Zustandsvektor $\mathbf{x}(t)$ zusammengefasst [42].

Im Zustandsraum wird das Messsystem allgemein durch die vektorwerte **Zustandsgleichung**

$$\dot{\mathbf{x}}(t) = \frac{d\mathbf{x}(t)}{dt} = \mathbf{w}(\mathbf{x}(t), u(t), \mathbf{z}(t), t) \quad (1.3)$$

und durch die **Ausgangsgleichung**

$$y(t) = F(\mathbf{x}) \quad (1.4)$$

beschrieben. Abbildung 1.2 zeigt die resultierende Darstellung des Messsystems. Dabei bezeichnet $u(t)$ die Messgröße, die als Eingangssignal fungiert; $y(t)$ ist die Ausgangsgröße, d. h. diejenige Größe, die von dem Messsystem ermittelt wird.

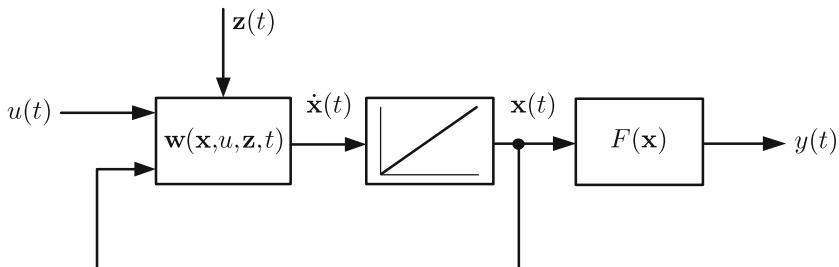


Abbildung 1.2. Zustandsraumdarstellung des Messsystems ($\mathbf{x}(t)$: Zustandsvektor des Messsystems, $u(t)$: Messgröße, $y(t)$: Ausgangsgröße, $\mathbf{z}(t)$: Störgrößenvektor).

Unter sehr allgemeinen Voraussetzungen, die bei technischen Systemen praktisch immer erfüllt sind, lässt sich der Zustandsvektor für beliebige Zeitpunkte $t > t_0$ eindeutig bestimmen, sofern der Anfangswert $\mathbf{x}(t_0)$ und die Verläufe der Eingangsgrößen im Intervall $[t_0, t]$ bekannt sind.

Neben der als skalar angenommenen Messgröße $u(t)$ können bei Messsystemen auch **Störgrößen** $\mathbf{z}(t)$ das Systemverhalten während des Betriebs ändern, weshalb sie allesamt Eingangsgrößen des Messsystems sind. Dies soll nun an zwei einfachen Beispielen gezeigt werden.

Beispiel 1.3 (Federwaage): Bei der Federwaage aus Bsp. 1.2 entspricht die Masse m der zeitlich konstanten Messgröße u . Der Zustandsvektor x enthält als einzige Komponente die Auslenkung x . Die Federkonstante c ist ein Parameter des Messsystems. Ändert sich die Federkonstante während des Betriebs, etwa aufgrund von Alterungserscheinungen, so ist c eine Störgröße z des Messsystems. Schließlich stellt der auslenkungsproportionale Schätzwert der Masse die Ausgangsgröße y dar. ■

Beispiel 1.4 (Handelswaage): Bei einer Handelswaage sind die Hebelarme Systemparameter, die sich mit der Umgebungstemperatur ändern. Man kann nun entweder die Länge der Hebelarme selbst als Störgröße z auffassen oder die Umgebungstemperatur als zusätzliche Störgröße z betrachten, deren funktionelle Wirkung auf die Hebelarme beschrieben werden muss. ■

Von einem idealen Messsystem wird verlangt, dass es unter den vielen Einflussgrößen allein die Messgröße u in die Ausgangsgröße y abbildet. In der Praxis ist das allerdings unmöglich. Am Beispiel der Handelswaage erkennt man, dass durch eine genaue Modellbildung die Wirkung vieler Störeinflüsse auf das Messergebnis mathematisch beschrieben werden kann. Das ermöglicht in manchen Fällen sogar eine Kompensation der Störungen.

Von besonderer Bedeutung ist das Verhalten eines Messsystems im stationären Zustand. In der Messtechnik beschreibt man den Zusammenhang zwischen der zu erfassenden Messgröße u und dem angezeigten Wert y im stationären Zustand mit Hilfe der Kennlinie (vgl. Kap. 3).

1.3.3 Physikalische Messkennlinie

Eine Grundaufgabe der Messtechnik ist es, **stationäre** Messgrößen zu erfassen, die sich während der Messung nicht ändern. Gesucht ist also eine stationäre Kennlinie, die sich ergibt, wenn alle Einschwingvorgänge abgeklungen sind. Im stationären Zustand des Systems gilt:

$$\dot{x} = 0. \quad (1.5)$$

Hieraus erhält man die **stationäre physikalische Messkennlinie**. Die Auflösung von (1.3) mit $\dot{x} = 0$ ergibt, dass x von der Messgröße u und vom Störgrößenvektor z abhängt. Der Zustandsvektor ist nur noch indirekt von der Zeit abhängig:

$$x = g(u, z). \quad (1.6)$$

Einsetzen von \mathbf{x} in die Ausgangsgleichung (1.4) führt zur physikalischen Messkennlinie:

$$y = F(\mathbf{x}) = F(\mathbf{g}(u, \mathbf{z})) = f(u, \mathbf{z}). \quad (1.7)$$

Für die physikalische Messkennlinie $y = f(u, \mathbf{z})$ wird im Messbereich eine stetige, streng monotone Funktion gefordert, wodurch Mehrdeutigkeiten vermieden werden. Mit beliebigem $\varepsilon > 0$ muss im gesamten Messbereich eine der folgenden Bedingungen gelten:

$$f(u + \varepsilon) > f(u) \quad \text{oder} \quad f(u + \varepsilon) < f(u). \quad (1.8)$$

❸ 1.3.4 Messsignale als Informationsträger

Innerhalb der Messkette (d. h. zwischen den Komponenten des Messsystems) wird die Information über die Messgröße u in Form von Messsignalen (meist elektrischen Spannungen und Strömen) ausgetauscht, vgl. Abb. 1.1. Die Messsignale $x(t)$ sind somit Träger des Messparameters u .

Die Messsignale selbst sind im Allgemeinen – auch bei einer konstanten Messgröße u – zeitlich veränderlich. Dabei kommen in messtechnischen Anwendungen harmonische und impulsförmige Messsignale besonders häufig vor. Bei **harmonischen Messsignalen** wird die Messgröße u entweder durch die Amplitude, die Frequenz oder die Phase dargestellt. Bei **impulsförmigen Messsignalen** können zur Beschreibung der Messgröße u. a. die Impulshöhe, die Impulsdauer und die Impulsfrequenz herangezogen werden.

Abhängig vom informationstragenden Parameter u lassen sich Messsignale wie folgt klassifizieren:

- **Amplitudenanaloge Signale** sind wertkontinuierliche Signale, bei denen die Signalamplitude proportional zur Messgröße u ist. Die Zeit kann kontinuierlich oder diskret sein. Abbildung 1.3 zeigt drei Beispiele amplitudenanaloger Messsignale. Neben dem Messgrößensignal $u(t)$ selbst zeigt die Abbildung das zeitkontinuierliche amplitudenmodulierte (AM) Signal $x_{\text{AM}}(t)$ sowie das zeitdiskrete pulsamplitudenmodulierte (PAM) Signal $x_{\text{PAM}}(n)$, wobei n den diskreten Zeitindex bezeichnet.
- **Digitale Signale** sind wert- und zeitdiskrete Signale, bei denen die Messgröße u mit Hilfe von Binärzahlen codiert wird. Als Beispiel eines digitalen Signals zeigt Abb. 1.3 ein pulscodemoduliertes (PCM) Signal $x_{\text{PCM}}(n)$, welches aus dem PAM-Signal $x_{\text{PAM}}(n)$ durch Quantisierung der Amplitudenwerte mit 3 Bit erhalten wurde (vgl. Abschn. 7.2).

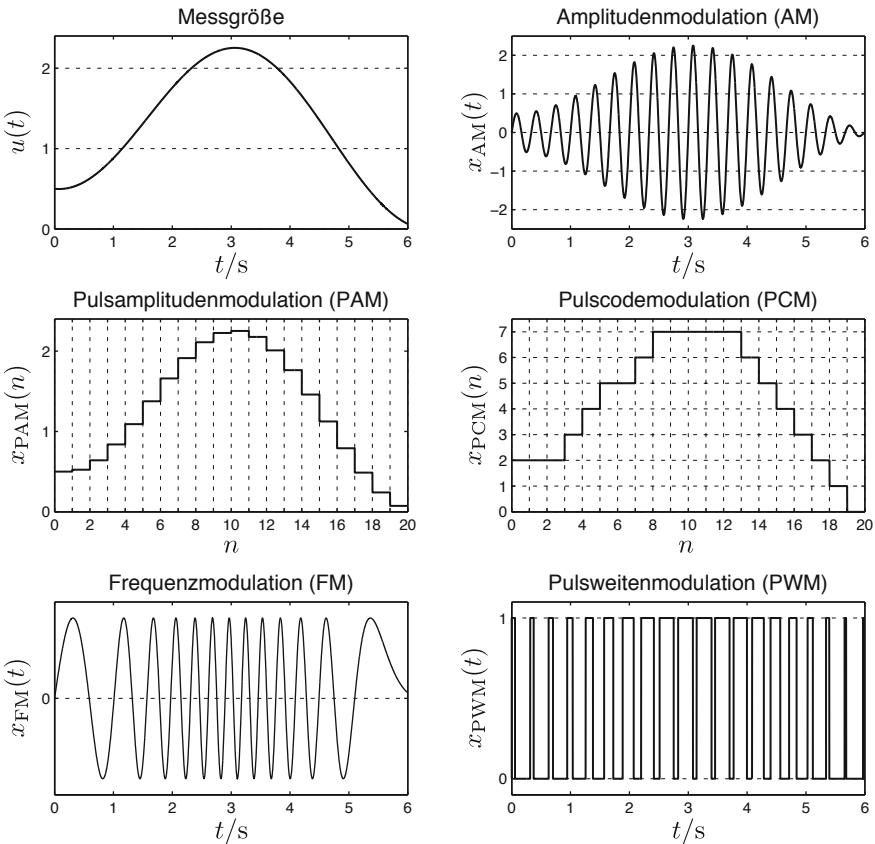


Abbildung 1.3. Informationstragende Parameter eines Messsignals $u(t)$ bei in der Messtechnik gängigen Modulationsarten: Amplitude (AM, PAM, PCM), Frequenz (FM), Zeitdauer (PWM).

- **Frequenzanaloge Signale** sind zeitkontinuierliche Signale, bei denen die Frequenz proportional zur Messgröße u ist. Die Signalwerte können kontinuierlich oder diskret sein. Exemplarisch zeigt Abb. 1.3 ein frequenzmoduliertes (FM) Signal $x_{FM}(t)$, dessen Momentanfrequenz (siehe Kap. 8) proportional zur Messgröße $u(t)$ ist.
- **Zeitanaloge Signale** sind zeitkontinuierliche impulsförmige Signale, bei denen entweder die Impulsdauer oder der Abstand zwischen aufeinanderfolgenden Impulsen proportional zur Messgröße u ist. Beim pulsweitenmodulierten (PWM) Signal $x_{PWM}(t)$ in Abb. 1.3 ist die Impulsdauer proportional zur Messgröße $u(t)$.

1.4 Messfehler

Nach der Beschreibung der physikalischen Eigenschaften von Messsystemen müssen deren **Fehler**² abgeschätzt werden. Bevor von Fehlern gesprochen werden kann, muss klar herausgestellt werden, welche Größe am Prozess als Messgröße verstanden werden soll. Bei einfachen Aufgaben wie den Abmessungen eines Werkstücks entsteht darüber keine Diskussion. Hat aber die Messgröße von Ort zu Ort verschiedene Werte, so wird man sich auf die Messung an einer oder mehreren repräsentativen Stellen einigen müssen. Bei Messung an mehreren Stellen wird man z. B. den Mittelwert bilden und diesen als Ersatz für den wahren Messwert nehmen.

Beispiel 1.5 (Temperaturmessung an Dampfturbinen): Um Überbeanspruchungen durch Wärmedehnung zu vermeiden, wird die Temperatur von Dampfturbinegehäusen überwacht. Als repräsentative Messorte werden solche ausgewählt, die bei instationären Vorgängen untereinander große Temperaturdifferenzen aufweisen. ■

Beispiel 1.6 (Heizwert von Brennstoffen): Bei der Bestimmung des Heizwertes von festen Brennstoffvorräten wird der Heizwert einer zufälligen Probe kaum interessieren. Um Fehler zu vermeiden, wird man mehrere Proben aus dem Brennstoffvorrat entnehmen und nach statistischen Methoden einen mittleren Heizwert des Vorrates schätzen (vgl. Kap. 4). ■

Zur Beurteilung einer bestimmten Messeinrichtung wird zunächst angenommen, dass die Messgröße mit einem bekannten Wert an der zu untersuchenden Messeinrichtung anliegt.

➊ 1.4.1 Absoluter und relativer Fehler

Definition 1.2: Absoluter Fehler

Der **absolute Fehler** eines Messsystems ist wie folgt definiert:

$$F = y_a - y_w . \quad (1.9)$$

²In der Messtechnik wird nach DIN 1319-1 [5] die Differenz zwischen dem Messergebnis und dem wahren Wert einer Messgröße als *Messabweichung* bezeichnet. Hier wird stattdessen der Begriff *Messfehler* verwendet.

Er beschreibt die positive oder negative Abweichung des angezeigten oder ausgegebenen Wertes y_a vom wahren Wert y_w . \diamond

Definition 1.3: Relativer Fehler

1.3

Der **relative Fehler** F_r ist eine bezogene Größe, wobei als Bezugswert in der Regel der wahre Wert y_w gewählt wird:

$$F_r = \frac{F}{y_w} = \frac{y_a - y_w}{y_w}. \quad (1.10)$$

Er ist dimensionslos und wird meist in Prozent angegeben. \diamond

Zur Abschätzung von Messfehlern ist meist nur der Betrag des Fehlers von Interesse, zur Korrektur der Fehler ist zusätzlich das Vorzeichen von Bedeutung. Es erhebt sich natürlich die Frage, wie man den Fehler bestimmen kann, wenn der wahre Wert y_w überhaupt nicht bekannt ist. Dazu gibt es im Wesentlichen zwei Möglichkeiten (Abb. 1.4).

- a) Als wahrer Wert y_w kann der von einem besonders genauen Präzisionsinstrument G_n angezeigte Messwert dienen.
- b) Mit der Messeinrichtung wird ein bekanntes Normal N (Maßverkörperung) vermessen. Der angezeigte Wert y_a wird mit dem bekannten wahren Wert des Normals y_w verglichen.

Der Fehler des Präzisionsinstruments wird z. B. durch Vermessung des Normals festgestellt. Ein Beispiel für ein (historisches) Normal ist das Urmeter.

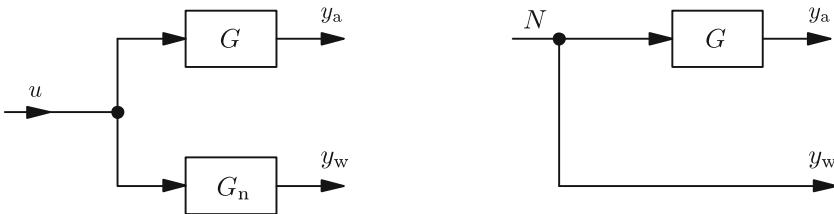


Abbildung 1.4. Möglichkeiten der Fehlerbestimmung.

1.4.2 Fehlerursachen

Jede Messung ist fehlerbehaftet. Der Grund liegt in bestimmten Eigenschaften des Messgegenstandes, in Unvollkommenheiten der Messeinrichtung

und des Messverfahrens, in wechselnden Umwelteinflüssen und in Fehlern des Beobachters. Man unterscheidet zwischen zwei Fehlerklassen:

- **Systematische Fehler:** Systematische Fehler sind dadurch charakterisiert, dass die Ursache des Fehlers und die Art der Einwirkung bekannt sind. Mit erhöhtem Aufwand im Messsystem ist deshalb eine **Kompensation** des systematischen Fehlers zumindest prinzipiell möglich.
- **Zufällige Fehler:** Zufällige (stochastische) Fehler weisen bei wiederholten Messungen unter gleichen Bedingungen verschiedene Beträge und Vorzeichen auf: die Messwerte „streuen“. Zufällige Fehler sind im Einzelnen nicht erfassbar, da ihre Ursachen teilweise unbekannt sind. Bei Messaufgaben interessiert deshalb nicht das Ergebnis einer einzelnen, zufälligen Messung, sondern z. B. der Mittelwert über viele Messungen. Ein Beispiel dafür ist die oben angeführte Bestimmung des Heizwertes von festen Brennstoffen. Andere Beispiele wären die Bestimmung der mittleren Ausfallrate von Bauelementen oder die Messung von elektrischen Spannungen, die von **Rauschen** überlagert sind. Die Messung solcher **stochastischer** Größen wird in Kap. 4 behandelt.

In dieser Einteilung steckt eine gewisse Willkür. Es kann z. B. passieren, dass ein Fehler aufgrund mangelnder Kenntnis der Messaufgabe zuerst als stochastisch deklariert wurde. Mit besserem Systemverständnis entdeckt man dann jedoch deterministische Fehlereinflüsse, so dass man nun von einem systematischen Fehler spricht.

Ungeachtet der oben aufgeführten Fehlerklassen können Fehler durch folgende Ursachen bedingt sein.

- **Vereinfachtes Modell des Messsystems:** Ein System mit verteilten Speichern werde z. B. durch konzentrierte, ideale Komponenten beschrieben. Die Messung einzelner Zustandsgrößen dieses Systems wird die tatsächlichen Abläufe nur näherungsweise erfassen. Exemplarisch sei die in Bsp. 1.5 geschilderte Temperaturmessung an Dampfturbinen erwähnt.
- **Innere Störgrößen:** Hierbei handelt es sich um Störgrößen im Messgerät selbst. Beispiele dafür sind Alterungseffekte an für die Messung wichtigen Bauteilen. Bei Drehspulinstrumenten oder Waagen ist eine Feder eingebaut, deren Eigenschaften sich im Laufe der Lebensdauer verändert, was sich in einer fehlerhaften Anzeige bemerkbar macht.
- **Äußere Störgrößen:** Der physikalische Messeffekt wird zumeist durch eine Reihe unerwünschter Einflüsse gestört. Ein Beispiel dafür ist die Temperaturabhängigkeit einer Widerstandsbrückenschaltung aus Halbleiter-Dehnungsmessstreifen zur Druckmessung. Wenn man den nicht erwünschten Einfluss isolieren und deterministisch beschreiben kann,

handelt es sich um *systematische Fehler*, die bei wiederholten Messungen gleichen Betrag und Vorzeichen aufweisen. Eine gezielte Kompensation ist möglich, was in Abschn. 3.3.6 beschrieben wird. Eine andere Art von äußereren Störgrößen sind *stochastische Einstreuungen*, die man nicht kompensieren kann. Zu ihrer Unterdrückung werden die Mittelwertbildung (Kap. 4) oder komplexere Filter (Abschn. 6.6) angewendet.

- **Beobachtungsfehler:** Der Beobachter, der eine Messung durchführt, kann als Fehlerquelle in Betracht kommen, wenn er die Anzeige falsch abliest.
- **Dynamische Fehler:** Bei vielen Messaufgaben werden zeitlich aufeinander folgende Messwerte benötigt. Ein Beispiel dafür ist der zeitliche Druckverlauf im Zylinder eines Verbrennungsmotors während des Verdichtungshubes. Das Anzeigesignal der Messeinrichtung soll der Messgröße verzögerungsfrei folgen. Abweichungen werden als dynamische Fehler bezeichnet. Eine Diskussion erfolgt in Kap. 5.
- **Rückwirkung:** Die Messeinrichtung braucht für den Messvorgang Energie oder Leistung, die dem Prozess entzogen wird. Der Wert der Messgröße mit angeschlossener Messeinrichtung unterscheidet sich vom Wert, der ohne Messeinrichtung erreicht worden wäre. Die Größe dieses Fehlers hängt davon ab, welche Messgrößenänderung der Energieaustausch im Prozess hervorruft. Dieser Fehler wird „Rückwirkung der Messeinrichtung auf den Prozess“ genannt und wird in Abschn. 3.4 behandelt.

Beispiel 1.7 (Temperaturmessung einer Flüssigkeit): Die Temperatur T einer Flüssigkeit der Wärmekapazität c_V soll mit einem Berührungsthermometer der Wärmekapazität c_F gemessen werden, das vor dem Eintauchen die Temperatur T_0 hat. Die gemessene Temperatur T_m errechnet sich aus der Energiebilanz vor und nach dem Eintauchen:

$$E_{\text{vor}} = c_V T + c_F T_0 = E_{\text{nach}} = (c_V + c_F) T_m, \quad (1.11)$$

$$c_V (T_m - T) = c_F (T_0 - T_m), \quad (1.12)$$

$$\Delta T = T_m - T = \frac{c_F}{c_V} (T_0 - T_m). \quad (1.13)$$

Der Messfehler ΔT wird klein, wenn die Wärmekapazität des Messfühlers c_F klein gegenüber der der Flüssigkeit c_V ist. ■

1.4.3 Spezifizierte Normalbedingungen

In der Spezifikation (technischen Beschreibung) eines Messsystems werden die Randbedingungen und Umwelteinflüsse festgehalten, unter denen der Hersteller einen maximalen Fehler garantiert. Dazu gehören die folgenden Angaben:

- Messbereich,
- Messgenauigkeit,
- Betriebsbedingungen,
- Einbauvorschriften,
- Energieversorgung,
- Abmessungen.

Entscheidend sind die Angaben über die Fehler, die einen Vergleich mit ähnlichen Geräten und deren Messgenauigkeit ermöglichen. Die sich im eingeschwungenen Zustand einstellenden Fehler werden als **statische Fehler** gekennzeichnet. Man kann somit zwei Klassen von statischen Fehlern unterscheiden:

- **Statische Fehler unter spezifizierten Normalbedingungen:** Die Störgrößen aus der Umgebung sind gemäß der Spezifikation konstant oder auf null zu halten:

$$\mathbf{z} = \mathbf{z}_0 . \quad (1.14)$$

- **Statische Fehler bei Abweichung von den spezifizierten Normalbedingungen:** Für jede wichtige *systematische Störgröße* ist eine definierte Abweichung von den Normalbedingungen herzustellen und die Auswirkung auf die Ausgangsgröße als Fehler festzustellen [53]:

$$\mathbf{z} = \mathbf{z}_i . \quad (1.15)$$

Im folgenden Beispiel sollen die möglichen Fehlerursachen an einer konkreten Messaufgabe dargestellt werden.

Beispiel 1.8 (Winkelgeschwindigkeitsmessung): Die Winkelgeschwindigkeit ω eines horizontal rotierenden Körpers soll über eine Beschleunigungsmessung bestimmt werden (Abb. 1.5). Die Winkelgeschwindigkeit ω wird über die Kräfte gemessen, die bei Rotation auf einen Körper wirken. Aus der Physik sind die folgenden Zusammenhänge bekannt:

$$\mathbf{a}_t = \dot{\omega} \times \mathbf{r} \quad (\text{Tangentialbeschleunigung}), \quad (1.16)$$

$$\mathbf{a}_{zp} = \omega \times \mathbf{v} \quad (\text{Zentripetalbeschleunigung}). \quad (1.17)$$

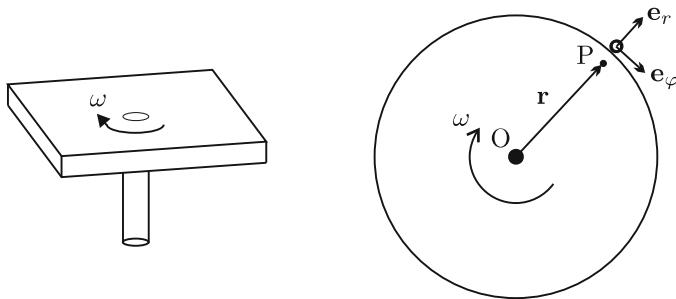


Abbildung 1.5. Winkelgeschwindigkeitsmessung.

Angewendet auf den zweidimensionalen Fall ergibt sich nach Abb. 1.5 unter Verwendung von Polarkoordinaten und den Beträgen $\omega = |\omega|$, $r = |\mathbf{r}|$ und $\dot{\omega} = |\dot{\omega}|$ eine Beschleunigung im Punkt P von

$$\mathbf{a}_P = \mathbf{a}_O + \mathbf{a}_t + \mathbf{a}_{zp} = \mathbf{a}_O + \dot{\omega} \cdot r \cdot \mathbf{e}_\varphi - \omega^2 \cdot r \cdot \mathbf{e}_r, \quad (1.18)$$

wobei \mathbf{a}_O die Führungsbeschleunigung des Aufpunktes O bezeichnet. Im Falle einer Rotation um eine feste Achse senkrecht zur Erdoberfläche kann man davon ausgehen, dass die Achse eine reine Drehbewegung ausführt und daher $\mathbf{a}_O = \mathbf{0}$ ist. Bringt man nun einen idealen Beschleunigungssensor, der nur in einer Richtung empfindlich ist, gemäß Abb. 1.6 an, so erhält man die folgende Messgleichung:

$$a_M = -\dot{\omega} \cdot r \cdot \sin \theta - \omega^2 \cdot r \cdot \cos \theta. \quad (1.19)$$

Im stationären Fall, d. h. bei konstanter Winkelgeschwindigkeit ($\dot{\omega} = 0$),

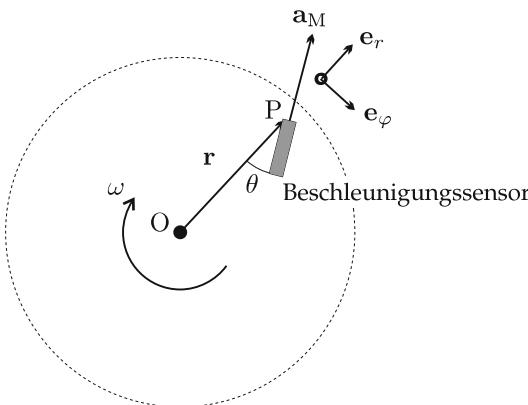


Abbildung 1.6. Beschleunigungsmessung am rotierenden Objekt.

ergibt sich daraus die folgende statische Messkennlinie:

$$a_M = -\omega^2 \cdot r \cdot \cos \theta. \quad (1.20)$$

Man erkennt sofort, dass die Sensorposition ($|r|, \theta$) das Messergebnis beeinflusst. Aber auch das Auftreten einer Führungsbeschleunigung a_O in (1.18) oder eine veränderliche Winkelgeschwindigkeit ($\dot{\omega} \neq 0$) verfälschen das Messergebnis. ■

Die stationäre Messkennlinie in Bsp. 1.8 ist quadratisch:

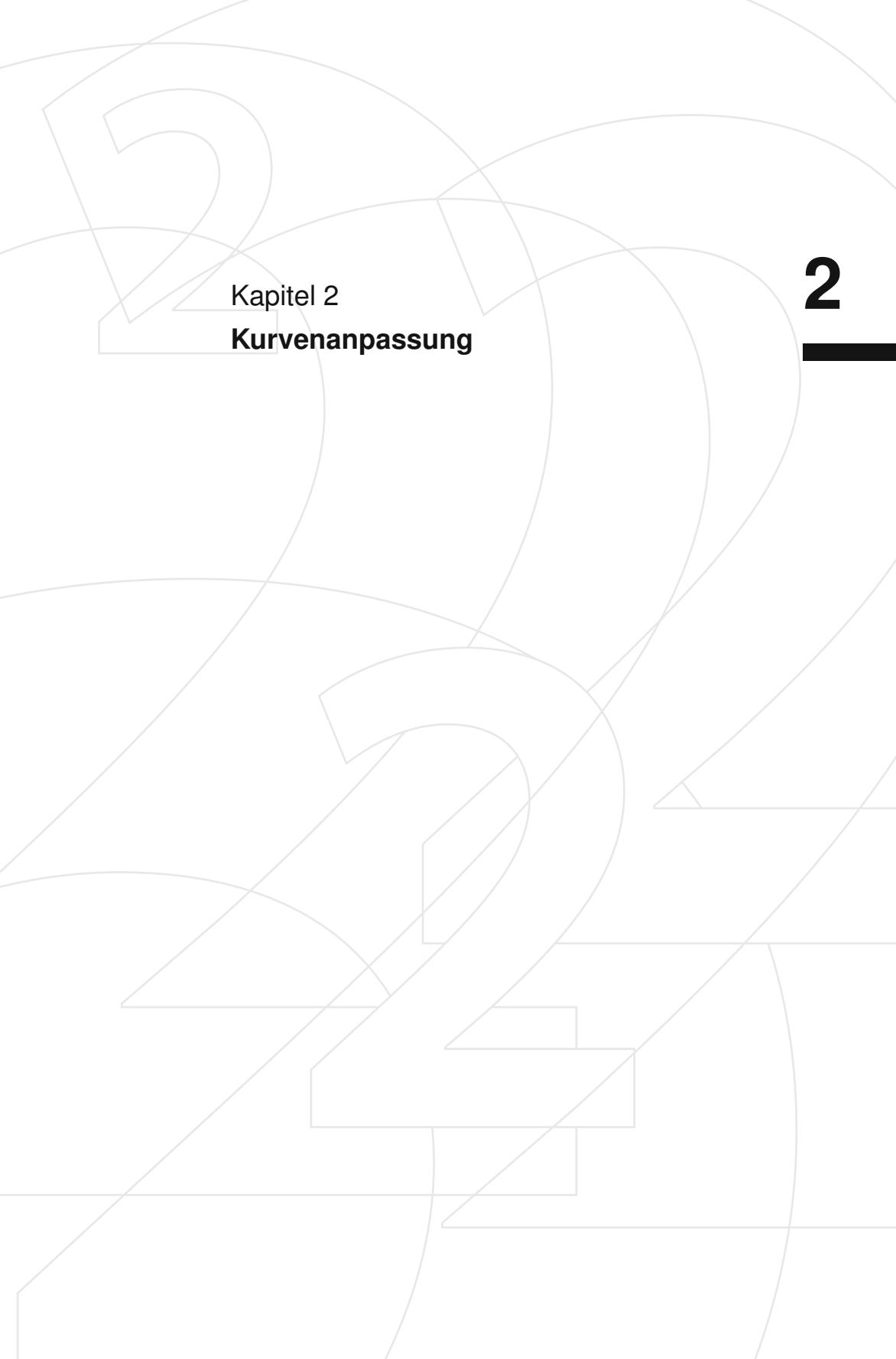
$$y = \text{const} \cdot u^2 \quad \text{mit} \quad u = \omega \quad \text{und} \quad y = a_M. \quad (1.21)$$

Die Sensorposition $|r|, \theta$ ist eine innere Störgröße des Messsystems. Als Normalbedingung ist dann beispielsweise die konstruktiv vorgegebene Sensorposition $|r_0|, \theta_0$ spezifiziert.

Die Führungsbeschleunigung a_O ist eine äußere Störgröße, die je nach Anwendung systematischer Natur (Achse steht nicht senkrecht zur Erdoberfläche) oder stochastischer Natur (etwa bei unbekannten Vibrationen der Messanordnung) sein kann.

Statt einer quadratischen stationären Kennlinie wäre eine lineare Kennlinie wünschenswert. Sie bietet den Vorteil einer konstanten Empfindlichkeit im gesamten Messbereich. Methoden zur Linearisierung der Messkennlinie werden in Kap. 3 behandelt.

Beim Übergang auf eine neue stationäre Winkelgeschwindigkeit ω treten dynamische Fehler auf, die proportional zu $\dot{\omega}$ sind. Ändert sich ω sprungförmig, so ist die Messung erst nach einer gewissen Zeit brauchbar, wenn die dynamischen Fehlerterme weitgehend abgeklungen sind. Das dynamische Verhalten von Messsystemen wird in Kap. 5 behandelt.



Kapitel 2

Kurvenanpassung

2

2 Kurvenanpassung

| | | |
|---------|--|----|
| 2.1 | Approximation | 26 |
| 2.1.1 | Approximation mit orthonormalen Funktionensystemen | 26 |
| 2.1.1.1 | Approximation mit der Fourier-Reihe | 29 |
| 2.1.1.2 | Approximation mit Walsh-Funktionen | 30 |
| 2.1.2 | Least-Squares-Schätzer | 31 |
| 2.1.3 | Regressionsanalyse | 32 |
| 2.2 | Interpolation | 34 |
| 2.2.1 | Polynominterpolation | 34 |
| 2.2.2 | Interpolation durch Lagrange-Polynome | 36 |
| 2.2.3 | Interpolation durch Newton-Polynome | 38 |
| 2.2.4 | Spline-Interpolation | 41 |
| 2.2.5 | Systemtheoretische Deutung der Interpolation | 46 |
| 2.3 | Kennfeldinterpolation | 47 |

2 Kurvenanpassung

Die analytische Darstellung einer Messkennlinie erfordert eine Modellbildung des Systems. Da das zugrundeliegende Modell in der Praxis oft unbekannt ist, liegt die stationäre Messkennlinie häufig nicht in analytischer Form, sondern nur als Menge von n Messpunkten $(u_k, y_k), k \in \{0, \dots, n-1\}$, vor. Gesucht wird nun eine analytische Darstellung der Kennlinie, welche die gemessenen Punkte in geeigneter Weise nachbildet. Dadurch können für beliebige Zwischenwerte u die zugehörigen Werte y angegeben werden. Des Weiteren kann die auf diese Weise ermittelte Messkennlinie mit den in Kap. 3 besprochenen Methoden genauer untersucht und optimiert werden.

Bei der Konstruktion einer analytischen Kennlinie aus Messpunkten können zwei grundsätzlich verschiedene Ansätze verfolgt werden (vgl. Abb. 2.1).

- **Interpolation:** Liegen nur wenige Messwerte ohne überlagerte Störungen vor, so wird man verlangen, dass die analytische Kennlinie exakt durch alle Messpunkte verläuft. Verwendet man beispielsweise Polynome $p(u)$ zur Interpolation, so erhält man bei n Messpunkten Polynome vom Grad $\deg\{p(u)\} \leq n - 1$. Man erkennt sofort, dass eine Interpolation nur für kleine n sinnvoll ist. Für eine große Anzahl von Messwerten wird die Interpolationsfunktion sehr schnell unhandlich und weist ein stark oszillierendes Verhalten auf. Das Interpolationsproblem wird im Abschn. 2.2 behandelt.
- **Approximation:** Liegen dagegen sehr viele Messwerte vor oder sind diesen Messwerten Störungen überlagert, so ist die Interpolation ein unpraktischer Ansatz. Man sucht vielmehr *einfache* Funktionen, welche die Menge der Messpunkte so nachbilden, dass der Fehler zwischen den Messpunkten und der analytischen Funktion möglichst klein wird. Als Stichwort sei hier die Regressionsanalyse genannt (Abschn. 2.1.3). Will man für beliebige Zwischenwerte u die zugehörigen Werte y angeben, so wird die Kennlinie als Approximation in einer endlichen Reihe analytischer Funktionen $\varphi_i(u)$ dargestellt:

$$\hat{y}(u) = \sum_{i=0}^{m-1} a_i \varphi_i(u) \quad (2.1)$$

bzw. mit $\hat{y}_k = \hat{y}(u_k)$

$$\hat{y}_k = \sum_{i=0}^{m-1} a_i \varphi_i(u_k), \quad k \in \{0, \dots, n-1\}. \quad (2.2)$$

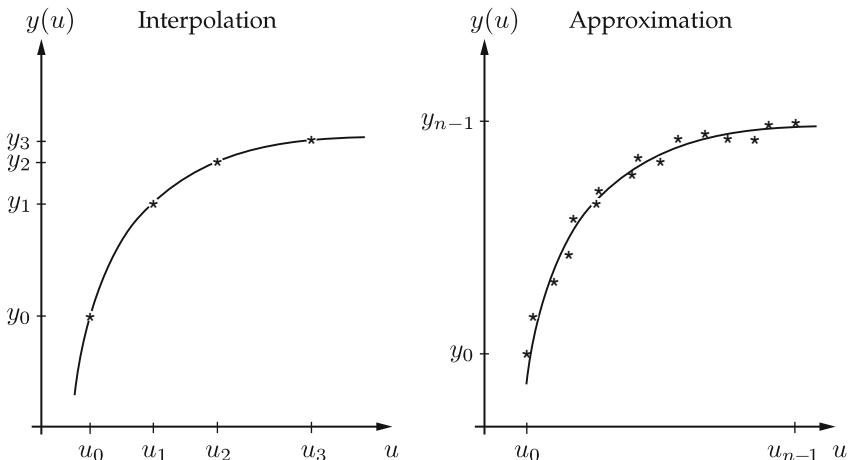


Abbildung 2.1. Kennlinie in Form von n Messpunkten und Ergebnis der Kurvenanpassung.

Die Koeffizienten a_i werden dann über die Minimierung eines Gütemaßes Q bestimmt, wobei hierfür üblicherweise die Summe der Approximationsfehlerquadrate herangezogen wird:

$$Q = \sum_{k=0}^{n-1} \left(y_k - \sum_{i=0}^{m-1} a_i \varphi_i(u_k) \right)^2 \rightarrow \min. \quad (2.3)$$

Der Vorteil dieser Vorgehensweise liegt in der Tatsache, dass man bereits mit einer begrenzten Anzahl m einfacher Basisfunktionen $\varphi_i(u)$ die Kennlinie nachbilden kann, wobei im Allgemeinen $m \ll n$ gilt. Die so gewonnene analytische Kennlinie verläuft allerdings nicht exakt durch die gemessenen Punkte.

2.1

2.1 Approximation

2.1.1 Approximation mit orthonormalen Funktionensystemen

Zunächst stellt sich die Frage nach den Vorteilen orthonormaler Funktionensysteme zur Signaldarstellung. Zur Verdeutlichung sei an die Vektorrechnung erinnert. Im dreidimensionalen Raum \mathbb{R}^3 wird jeder Vektor durch seine Komponenten in x -, y - und z -Richtung repräsentiert. Beispielsweise sei der folgende Vektor betrachtet:

$$\mathbf{a} = (a_0, a_1, a_2)^T. \quad (2.4)$$

Mittels der zueinander orthogonalen Einheitsvektoren

$$\mathbf{e}_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{e}_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad (2.5)$$

lässt sich der Vektor \mathbf{a} als Linearkombination der Vektoren \mathbf{e}_i darstellen:

$$\mathbf{a} = a_0 \mathbf{e}_0 + a_1 \mathbf{e}_1 + a_2 \mathbf{e}_2 = \sum_{i=0}^2 a_i \mathbf{e}_i. \quad (2.6)$$

Die Menge $\{\mathbf{e}_0, \mathbf{e}_1, \mathbf{e}_2\}$ bildet dabei eine **orthonormale Basis** des Vektorraumes \mathbb{R}^3 . Sie hat die besonderen Eigenschaften, dass das Innenprodukt zwischen zwei verschiedenen Basisvektoren verschwindet – die Vektoren sind also zueinander orthogonal [42] – und dass die Norm der Basisvektoren gleich eins ist:

$$\langle \mathbf{e}_i, \mathbf{e}_j \rangle = \sum_{k=0}^{n-1} e_{i,k} \cdot e_{j,k} = \delta_i^j, \quad (2.7)$$

wobei

$$\delta_i^j = \begin{cases} 1 & \text{für } i = j \\ 0 & \text{für } i \neq j \end{cases} \quad (2.8)$$

das **Kronecker-Delta** bezeichnet. Ein großer Vorteil orthonormaler Basissysteme ist die Tatsache, dass bei Hinzunahme einer neuen Dimension (aufgespannt z. B. durch den Basisvektor \mathbf{e}_2) lediglich die entsprechende Komponente a_2 zur Darstellung des resultierenden Vektors bestimmt werden muss, *ohne* dass sich dabei eine Änderung der bisherigen Komponenten (im vorliegenden Beispiel a_0 und a_1) ergäbe.

Bei der Approximation von Messkennlinien macht man sich genau diese Eigenschaft orthonormaler Basissysteme zu Nutze. Man verwendet zur Approximation Funktionensysteme mit der Eigenschaft

$$\langle \varphi_i, \varphi_j \rangle = \sum_{k=0}^{n-1} \varphi_i(u_k) \cdot \varphi_j^*(u_k) = \delta_i^j. \quad (2.9)$$

Die Funktionswerte $\varphi_i(u_k)$ an den Stützstellen u_k liefern orthonormale, n -dimensionale Vektoren. Im Folgenden wird angenommen, dass die Stützstellen äquidistant über dem gesamten Orthogonalitätsintervall des Funktionensystems verteilt sind. Zur Bestimmung der Koeffizienten wird das

Gütemaß (2.3)

$$Q = \sum_{k=0}^{n-1} \left(y_k - \sum_{i=0}^{m-1} a_i \varphi_i(u_k) \right) \left(y_k - \sum_{j=0}^{m-1} a_j \varphi_j(u_k) \right)^* \quad (2.10)$$

mit Hilfe der Kettenregel nach den Koeffizienten a_j abgeleitet:

$$\frac{\partial Q}{\partial a_j} = -2 \sum_{k=0}^{n-1} \left(y_k - \sum_{i=0}^{m-1} a_i \varphi_i(u_k) \right) \varphi_j^*(u_k) \quad (2.11)$$

$$= -2 \left(\sum_{k=0}^{n-1} y_k \varphi_j^*(u_k) - \sum_{k=0}^{n-1} \sum_{i=0}^{m-1} a_i \varphi_i(u_k) \varphi_j^*(u_k) \right) \stackrel{!}{=} 0, \quad (2.12)$$

$$\sum_{k=0}^{n-1} y_k \varphi_j^*(u_k) - \sum_{i=0}^{m-1} a_i \underbrace{\sum_{k=0}^{n-1} \varphi_i(u_k) \varphi_j^*(u_k)}_{\delta_i^j} = 0 \quad (2.13)$$

$$\Rightarrow a_j = \sum_{k=0}^{n-1} y_k \varphi_j^*(u_k). \quad (2.14)$$

Man erkennt sofort den Nutzen orthogonaler Funktionen. Die Koeffizienten a_j zur Darstellung der Kennlinie hängen nur von der zugehörigen Basisfunktion $\varphi_j(u_k)$ ab. Werden weitere Funktionen φ_i zur Approximation herangezogen, so bleiben die bisher berechneten Koeffizienten unverändert. Der quadratische Fehler zwischen den Messpunkten und der approximierten Kennlinie berechnet sich zu

$$Q = \sum_{k=0}^{n-1} \left(y_k - \sum_{i=0}^{m-1} a_i \varphi_i(u_k) \right) \left(y_k - \sum_{j=0}^{m-1} a_j \varphi_j(u_k) \right)^* \quad (2.15)$$

$$\begin{aligned} &= \sum_{k=0}^{n-1} y_k^2 - \sum_{i=0}^{m-1} a_i \underbrace{\sum_{k=0}^{n-1} \varphi_i(u_k) y_k^*}_{a_i^*} - \sum_{j=0}^{m-1} a_j^* \underbrace{\sum_{k=0}^{n-1} \varphi_j^*(u_k) y_k}_{a_j} \\ &\quad + \sum_{i=0}^{m-1} \sum_{j=0}^{m-1} a_i a_j^* \underbrace{\sum_{k=0}^{n-1} \varphi_i(u_k) \varphi_j^*(u_k)}_{\delta_i^j}, \end{aligned} \quad (2.16)$$

$$Q = \sum_{k=0}^{n-1} y_k^2 - \sum_{i=0}^{m-1} |a_i|^2. \quad (2.17)$$

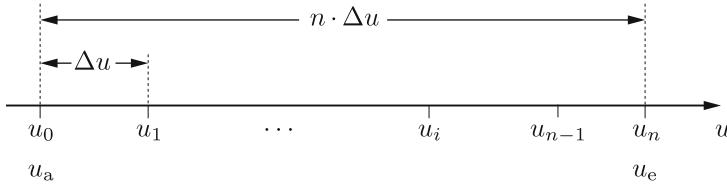


Abbildung 2.2. Stützstellenabstände einer gemessenen Kennlinie.

Mit wachsendem Grad m der Funktionenreihe wird der Approximationsfehler geringer. Mit $Q \geq 0$ folgt die bekannte Bessel'sche Ungleichung [42].

② 2.1.1.1 Approximation mit der Fourier-Reihe

Nun stellt sich die Frage, welche Funktionensysteme die Orthogonalitätsbedingung (2.9) erfüllen. Am bekanntesten sind die Funktionen der **Fourier-Reihe**:

$$F_i(u) = \frac{1}{\sqrt{n}} \cdot \exp\left(j2\pi i \frac{u - u_a}{u_e - u_a}\right). \quad (2.18)$$

Diese Funktionen bilden im Messbereich $[u_a, u_e]$ bei n äquidistanten Stützstellen im Abstand Δu (vgl. Abb. 2.2) ein orthonormales Funktionensystem:

$$\langle F_i(u_k), F_j(u_k) \rangle = \frac{1}{n} \sum_{k=0}^{n-1} \exp\left(j2\pi i \frac{u_k - u_a}{u_e - u_a}\right) \exp\left(-j2\pi j \frac{u_k - u_a}{u_e - u_a}\right). \quad (2.19)$$

Mit dem Stützstellenabstand Δu und der Intervallbreite $(u_e - u_a) = n \cdot \Delta u$ gilt

$$u_k = k \cdot \Delta u + u_a \quad \Rightarrow \quad \frac{u_k - u_a}{u_e - u_a} = \frac{k}{n}. \quad (2.20)$$

Damit lässt sich das Innenprodukt schreiben und die Orthogonalität zeigen:

$$\langle F_i(u_k), F_j(u_k) \rangle = \frac{1}{n} \sum_{k=0}^{n-1} \exp\left(j2\pi(i-j)\frac{k}{n}\right) = \delta_i^j. \quad (2.21)$$

Zur Veranschaulichung von (2.21) hilft ein Zeigerdiagramm in der komplexen Ebene (Abb. 2.3). Für $i \neq j$ ergibt sich ein geschlossener Polygonzug, d. h. die Summe der Zeiger verschwindet. Nur für $i = j$ ergibt sich ein Wert ungleich null.

Die Approximation einer Messkennlinie mit den Funktionen aus (2.18) entspricht gerade der Fourier-Reihe bei periodischen Funktionen.

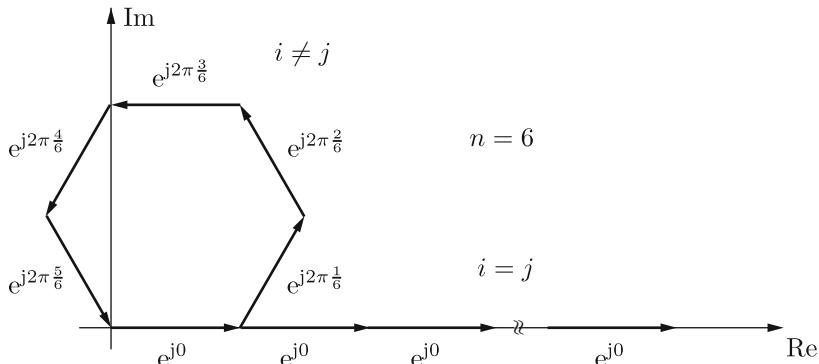


Abbildung 2.3. Veranschaulichung der Orthogonalität der Fourier-Funktionen.

② 2.1.1.2 Approximation mit Walsh-Funktionen

Ein Nachteil der Verwendung der Fourier-Reihe ist die notwendige Rechnung mit komplexen Exponentialfunktionen. Geradezu ideal für die Implementierung im Rechner ist hingegen das orthonormale System der **Walsh-Funktionen** $\text{wal}(i, u)$ geeignet. Sie sind im Intervall $[0; 1]$ definiert und nehmen lediglich die Funktionswerte $+1$ und -1 an.

Abbildung 2.4 zeigt einige Funktionen dieses Funktionensystems. Von ihrer Orthogonalität überzeugt man sich leicht durch Summenbildung über äquidistant verteilte Stützstellen. Die Berechnung der Koeffizienten a_j nach

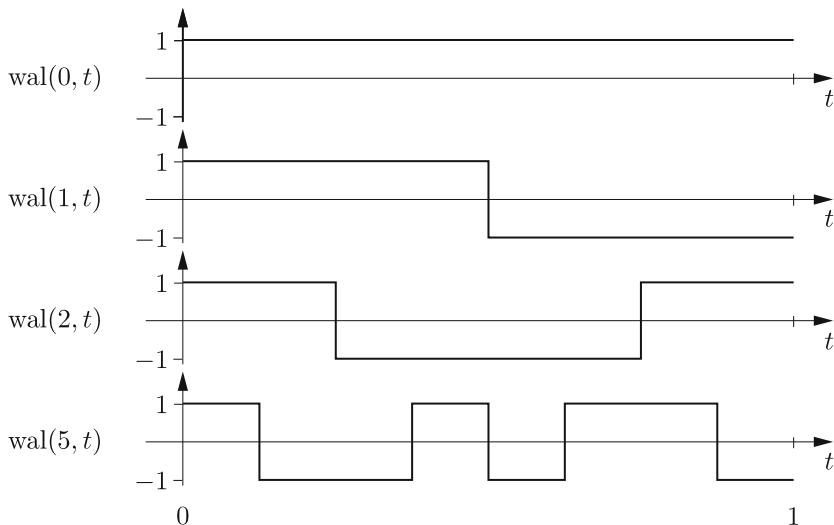


Abbildung 2.4. Walsh-Funktionen.

(2.14) reduziert sich bei diesem Basissystem auf eine einfache Summe über die Funktionswerte y_i . Für das Rechnen mit Walsh-Funktionen und die Erzeugung beliebiger Basisfunktionen $\text{wal}(i, u)$ sei auf die entsprechende Fachliteratur verwiesen [14].

An dieser Stelle soll noch auf eine Eigenschaft hingewiesen werden. Wie man aus Abb. 2.4 erkennt, sind die i Nulldurchgänge der Walsh-Funktionen nicht gleichmäßig über das Intervall verteilt. Um dennoch einen Frequenzbegriff wie bei Sinus- und Cosinusfunktionen zu erhalten, kann man die Häufigkeit der Nulldurchgänge im Intervall heranziehen. Damit gelangt man zur verallgemeinerten Frequenz, die in Kap. 8 eingeführt wird.

2.1.2 Least-Squares-Schätzer

In vielen Anwendungen ist eine Approximation mit nicht orthogonalen Basisfunktionen $\varphi_i(u)$ gewünscht. Zur Minimierung der quadratischen Summe der Approximationsfehler (2.3) kann der **Least-Squares-Schätzer** (kurz: LS-Schätzer) herangezogen werden. Die Summe der Approximationsfehlerquadrate lautet in Vektorschreibweise

$$Q = \sum_{k=0}^{n-1} (y_k - \hat{y}_k)^2 = (\mathbf{y} - \hat{\mathbf{y}})^T (\mathbf{y} - \hat{\mathbf{y}}). \quad (2.22)$$

Es wird folgender Approximationsansatz für n Messpunkte verwendet:

$$\hat{\mathbf{y}} = \begin{bmatrix} \hat{y}_0 \\ \vdots \\ \hat{y}_{n-1} \end{bmatrix} = \begin{bmatrix} \varphi_0(u_0) & \cdots & \varphi_{m-1}(u_0) \\ \vdots & \ddots & \vdots \\ \varphi_0(u_{n-1}) & \cdots & \varphi_{m-1}(u_{n-1}) \end{bmatrix} \cdot \begin{bmatrix} a_0 \\ \vdots \\ a_{m-1} \end{bmatrix} = \Phi \mathbf{a}. \quad (2.23)$$

Einsetzen von (2.23) in (2.22) ergibt

$$Q = (\mathbf{y} - \Phi \mathbf{a})^T (\mathbf{y} - \Phi \mathbf{a}) = \mathbf{y}^T \mathbf{y} - 2\mathbf{a}^T \Phi^T \mathbf{y} + \mathbf{a}^T \Phi^T \Phi \mathbf{a}. \quad (2.24)$$

Zur Bestimmung von \mathbf{a} wird das Gütemaß Q minimiert:

$$\frac{dQ}{d\mathbf{a}} = -2\Phi^T \mathbf{y} + 2\Phi^T \Phi \mathbf{a} = \mathbf{0} \quad (2.25)$$

$$\Rightarrow \mathbf{a} = (\Phi^T \Phi)^{-1} \Phi^T \mathbf{y}. \quad (2.26)$$

Der gesuchte Parametervektor \mathbf{a} berechnet sich somit aus dem Produkt der nach E. H. Moore und R. Penrose benannten Pseudoinversen $(\Phi^T \Phi)^{-1} \Phi^T$ von Φ und dem Messpunktevektor \mathbf{y} .

Der LS-Schätzer ist von großer praktischer Bedeutung. Er wird oft benutzt, um aus stark verrauschten Messungen Kennlinienfunktionen zu be-

stimmen, was auch als Regressionsanalyse bezeichnet wird (Abschn. 2.1.3). Der LS-Schätzer kann damit einerseits als ein **Optimalfilter** angesehen werden (Abschn. 6.6), welches Vorwissen über die herauszufilternde Funktion explizit einbezieht – z. B. dass sie als Polynom 2. Grades darstellbar ist. Dieses Vorwissen bezeichnet man als **Signalmodell** und bestimmt den Aufbau der Matrix Φ bzw. die Wahl der Basisfunktionen $\varphi_i(u)$.

Andererseits wird ein funktionaler Zusammenhang zwischen Abszissen- und Ordinatenvariablen hergestellt. Damit ist das Verfahren auch zur **Extrapolation** geeignet, etwa um zukünftige Funktionswerte eines Zeitsignals zu vorhersagen [27]. Eine Anwendung hierzu findet der LS-Schätzer in der statistischen Prozessüberwachung, die in Abschn. 4.5.3 behandelt wird. Ein weiteres Anwendungsgebiet des LS-Schätzers ist die **Parameterschätzung** (Abschn. 8.2.8).

2.1.3 Regressionsanalyse

In der Statistik dient die **Regressionsanalyse** allgemein dazu, einen funktionalen Zusammenhang zwischen Variablen herzustellen. Praktisch werden hierfür häufig Polynome zusammen mit der in Abschn. 2.1.2 behandelten Methode der kleinsten Quadrate verwendet, aber es können genauso auch andere Modelle und Optimierungsziele zum Einsatz kommen. Im Folgenden wird nur auf den Sonderfall der **linearen Regression** eingegangen.

Eine in der Praxis häufig auftretende Aufgabe ist die Suche nach einer Geraden durch eine Menge von Messpunkten (Abb. 2.5). Die Gerade habe die Form

$$\hat{y}(u) = a_1 u + a_0 . \quad (2.27)$$

Die unbekannten Parameter a_1 und a_0 werden durch Minimierung der Fehlerquadrate gemäß (2.3) bestimmt. Mit der Gütefunktion

$$Q = \sum_{k=0}^{n-1} (y_k - a_1 u_k - a_0)^2 \quad (2.28)$$

können die Parameter durch Differentiation bestimmt werden:

$$\frac{dQ}{da_1} = -2 \sum_{k=0}^{n-1} u_k (y_k - a_1 u_k - a_0) , \quad (2.29)$$

$$\frac{dQ}{da_0} = -2 \sum_{k=0}^{n-1} (y_k - a_1 u_k - a_0) . \quad (2.30)$$

Durch Nullsetzen der beiden Ableitungen gelangt man schließlich zu folgendem Gleichungssystem

$$a_1 \sum_{k=0}^{n-1} u_k^2 + a_0 \sum_{k=0}^{n-1} u_k = \sum_{k=0}^{n-1} u_k y_k , \quad (2.31)$$

$$a_1 \sum_{k=0}^{n-1} u_k + n a_0 = \sum_{k=0}^{n-1} y_k \quad (2.32)$$

und erhält durch Auflösen die gesuchten Parameter der Regressionsgeraden:

$$a_0 = \frac{1}{n} \sum_{k=0}^{n-1} y_k - a_1 \cdot \frac{1}{n} \sum_{k=0}^{n-1} u_k , \quad (2.33)$$

$$a_1 = \frac{n \cdot \sum_{k=0}^{n-1} u_k y_k - \sum_{k=0}^{n-1} u_k \sum_{k=0}^{n-1} y_k}{n \cdot \sum_{k=0}^{n-1} u_k^2 - \left(\sum_{k=0}^{n-1} u_k \right)^2} . \quad (2.34)$$

Die Berechnung der Parameter mit dem LS-Schätzer (2.26) führt auf das gleiche Ergebnis. Das zugehörige Signalmodell lautet für eine Geradengleichung:

$$\hat{\mathbf{y}} = \begin{bmatrix} \hat{y}_0 \\ \vdots \\ \hat{y}_{n-1} \end{bmatrix} = \begin{bmatrix} 1 & u_0 \\ \vdots & \vdots \\ 1 & u_{n-1} \end{bmatrix} \cdot \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \Phi \mathbf{a} . \quad (2.35)$$

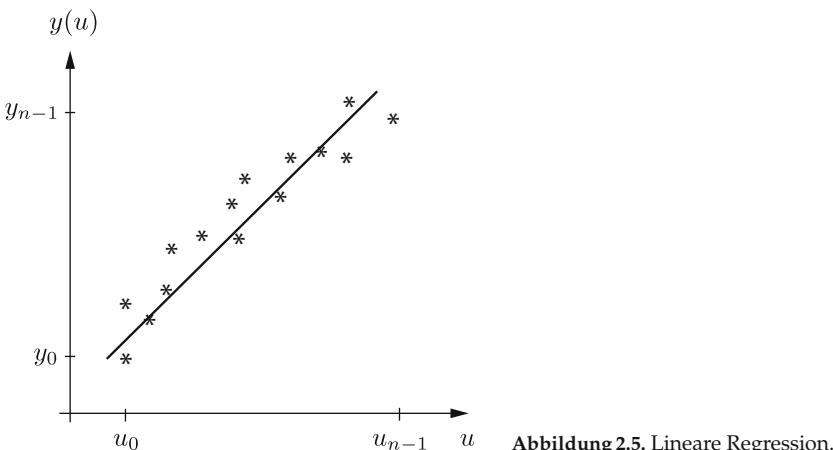


Abbildung 2.5. Lineare Regression.

Die Übereinstimmung von (2.33) und (2.34) mit dem LS-Schätzer möge der Leser als Übung durchführen. Es können natürlich auch Polynome höheren Grades für die Regression verwendet werden.

Im Gegensatz zur Approximation mit orthonormalen Funktionen müssen bei der Regressionsanalyse alle Koeffizienten a_i neu berechnet werden, wenn die Ordnung des Regressionspolynoms erhöht wird.

Abschließend soll noch auf einen Zusammenhang hingewiesen werden. Die Ergebnisgleichungen (2.33) und (2.34) der linearen Regression enthalten Terme der Form

$$\frac{1}{n} \sum_{k=0}^{n-1} u_k \quad \text{bzw.} \quad n \cdot \sum_{k=0}^{n-1} u_k^2 - \left(\sum_{k=0}^{n-1} u_k \right)^2. \quad (2.36)$$

Wie später in Kap. 4 gezeigt wird, handelt es sich hierbei um den Mittelwert und die mit dem Faktor $(n - 1)/n$ gewichtete Varianz der Größe u .

2.2

2.2 Interpolation

Ist eine experimentelle Kennlinie nur durch wenige Punkten gegeben, so wird man oft verlangen, dass die zu ermittelnde Funktion die Werte (u_k, y_k) exakt wiedergibt und dass sie zwischen diesen Stützstellen einen glatten Verlauf aufweist. Diese Aufgabenstellung führt zum klassischen **Interpolationsproblem**.

2.2.1 Polynominterpolation

Zur Kennliniendarstellung wird meist ein **Polynomansatz** in der Messgröße u gewählt:

$$\hat{y}(u) = \sum_{i=0}^{n-1} a_i u^i = \mathbf{a}^T \mathbf{p} \quad (2.37)$$

mit $\mathbf{a} = (a_0, \dots, a_{n-1})^T$ und $\mathbf{p} = (1, u, \dots, u^{n-1})^T$. Zur Bestimmung der n Koeffizienten a_i stehen n Gleichungen in den Stützstellen zur Verfügung:

$$y_k(u_k) = \sum_{i=0}^{n-1} a_i u_k^i, \quad k \in \{0, \dots, n - 1\}. \quad (2.38)$$

Die n Gleichungen lassen sich in Matrixschreibweise angeben:

$$\mathbf{y} = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_{n-1} \end{pmatrix} = \begin{pmatrix} 1 & u_0 & u_0^2 & \cdots & u_0^{n-1} \\ 1 & u_1 & u_1^2 & \cdots & u_1^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & u_{n-1} & u_{n-1}^2 & \cdots & u_{n-1}^{n-1} \end{pmatrix} \cdot \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \end{pmatrix} = \mathbf{V} \cdot \mathbf{a}, \quad (2.39)$$

wobei \mathbf{y} den Messvektor und \mathbf{V} die **Vandermonde-Matrix** bezeichnen. Um die Gleichung zu lösen, muss die Vandermonde-Matrix invertiert werden. Dies setzt voraus, dass ihre Determinante $|\mathbf{V}|$ ungleich null ist, was der Fall ist, wenn die Stützstellen u_i paarweise verschieden sind.

Die Matrix \mathbf{V} hat den Rang $r = n$. Ihre Determinante lässt sich geschlossen berechnen, indem man sie rekursiv in Unterdeterminanten auflöst. Dazu multipliziert man jeweils die vorherige Spalte mit u_{n-r} und subtrahiert sie von der gerade betrachteten Spalte:

$$r = n : \quad (2.40)$$

$$|\mathbf{V}| = \begin{vmatrix} 1 & 0 & \cdots & 0 \\ 1 & u_1 - u_0 & \cdots & u_1^{n-1} - u_1^{n-2} \cdot u_0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & u_{n-1} - u_0 & \cdots & u_{n-1}^{n-1} - u_{n-1}^{n-2} \cdot u_0 \end{vmatrix} \quad \begin{array}{l} \text{2. Spalte} - 1. \text{ Spalte} \times u_0 \\ \text{3. Spalte} - 2. \text{ Spalte} \times u_0 \\ \vdots \\ n. \text{ Sp.} - (n-1). \text{ Sp.} \times u_0 \end{array}.$$

Die Determinante wird nach der ersten Zeile entwickelt, bei der nur das erste Element ungleich null ist. Die Terme $(u_j - u_0)$ werden aus den Zeilen herausgezogen ($j = 1, \dots, n-1$):

$$r = n-1 : \quad (2.41)$$

$$|\mathbf{V}| = (u_1 - u_0) (u_2 - u_0) \cdots (u_{n-1} - u_0) \begin{vmatrix} 1 & u_1 & u_1^2 & \cdots & u_1^{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & u_{n-1} & u_{n-1}^2 & \cdots & u_{n-1}^{n-2} \end{vmatrix}.$$

Das Verfahren wird nun mit $r = n-1$ und Multiplikation der Spalten mit u_1 wiederholt usw. Für die Determinante erhält man schließlich:

$$\begin{aligned} |\mathbf{V}| &= (u_1 - u_0) (u_2 - u_0) (u_3 - u_0) \cdots (u_{n-1} - u_0) \times \\ &\quad (u_2 - u_1) (u_3 - u_1) \cdots (u_{n-1} - u_1) \times \\ &\quad (u_3 - u_2) \cdots (u_{n-1} - u_2) \times \quad (2.42) \\ &\quad \vdots \\ &\quad (u_{n-1} - u_{n-2}). \end{aligned}$$

Die Determinante berechnet sich somit aus dem Produkt aller möglicher Differenzen zwischen den Stützstellen. Die Inversion von \mathbf{V} ist genau dann gut konditioniert (d. h. numerisch gut berechenbar), wenn es kein sehr eng beieinander liegendes Stützstellenpaar gibt. Die Koeffizienten berechnen sich dann mittels

$$\mathbf{a} = \mathbf{V}^{-1} \mathbf{y}, \quad (2.43)$$

wodurch die Interpolationsgleichung (2.37) folgende Form annimmt:

$$\hat{y} = \mathbf{a}^T \mathbf{p} = \mathbf{p}^T \mathbf{a} = \mathbf{p}^T \mathbf{V}^{-1} \mathbf{y}. \quad (2.44)$$

❷ 2.2.2 Interpolation durch Lagrange-Polynome

Der Aufwand zur Lösung des Gleichungssystems (2.39) ist vergleichsweise groß. Anstelle des Polynomansatzes (2.37) arbeitet man daher in der Praxis oft mit **Lagrange-Polynomen**

$$\hat{y} = \sum_{i=0}^{n-1} y_i L_i(u) \quad (2.45)$$

oder, in Vektorschreibweise,

$$\hat{y} = \mathbf{L}^T \mathbf{y}. \quad (2.46)$$

Dank einer komplexeren Basis bestehend aus echten Polynomen $L_i(u)$ anstelle der in (2.37) gewählten Monome u_i werden Lagrange-Polynome direkt mit den Messwerten y_i gewichtet. Dadurch ist keine Berechnung weiterer Koeffizienten erforderlich.

Aus einem Vergleich von (2.44) und (2.46) sieht man, dass sich die Lagrange-Polynome gerade aus der invertierten Vandermonde-Matrix ergeben:

$$\mathbf{L} = (\mathbf{V}^{-1})^T \mathbf{p}. \quad (2.47)$$

Die Lagrange-Polynome sind damit

$$L_i(u) = \frac{(u - u_0) \cdots (u - u_{i-1})(u - u_{i+1}) \cdots (u - u_{n-1})}{(u_i - u_0) \cdots (u_i - u_{i-1})(u_i - u_{i+1}) \cdots (u_i - u_{n-1})}. \quad (2.48)$$

Aus ihrer Eigenschaft $L_i(u_j) = \delta_i^j$ folgt mit (2.45):

$$\hat{y}(u_j) = \sum_{i=0}^{n-1} y_i L_i(u_j) = y_j. \quad (2.49)$$

Die Stützstellen eines Lagrange-Polynoms werden somit exakt interpoliert.

Beispiel 2.1 (Lagrange-Interpolation mit drei Stützstellen): Eine Messkennlinie sei durch drei äquidistante Stützstellen u_a , u_m und u_e gegeben (Abb. 2.6):

$$u_m - u_a = u_e - u_m = h. \quad (2.50)$$

Die Lagrange-Polynome $L_i(u)$ sind damit

$$L_0(u) = \frac{(u - u_m)(u - u_e)}{(u_a - u_m)(u_a - u_e)} = \frac{1}{2h^2} (u - u_m)(u - u_e), \quad (2.51)$$

$$L_1(u) = \frac{(u - u_a)(u - u_e)}{(u_m - u_a)(u_m - u_e)} = -\frac{1}{h^2} (u - u_a)(u - u_e), \quad (2.52)$$

$$L_2(u) = \frac{(u - u_a)(u - u_m)}{(u_e - u_a)(u_e - u_m)} = \frac{1}{2h^2} (u - u_a)(u - u_m). \quad (2.53)$$

Einsetzen der $L_i(u)$ in (2.45) ergibt die Interpolationsgleichung:

$$\hat{y} = \sum_{i=0}^2 y_i L_i(u) \quad (2.54)$$

$$= \frac{y_a}{2h^2} (u - u_m)(u - u_e) - \frac{y_m}{h^2} (u - u_a)(u - u_e) + \frac{y_e}{2h^2} (u - u_a)(u - u_m). \quad (2.55)$$

Mit $u_a = 0$, $y_a = 0$ sowie $u_m = h$, $u_e = 2h$ erhält man

$$\hat{y} = u \left(\frac{2y_m}{h} - \frac{y_e}{2h} \right) + u^2 \left(\frac{y_e}{2h^2} - \frac{y_m}{h^2} \right). \quad (2.56)$$

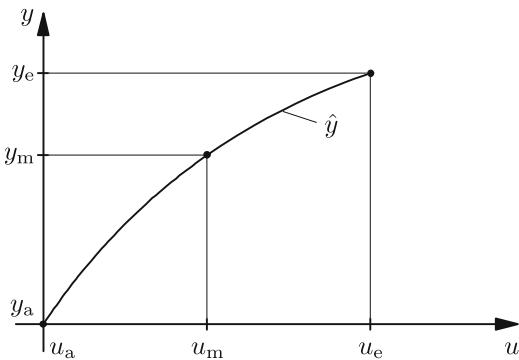


Abbildung 2.6. Skizze der zu interpolierenden Messkurve.

Liegen die drei Stützstellen auf einer Geraden, d. h. ist $y_m = y_e/2$, so wird mit

$$\hat{y} = u \cdot \frac{y_e}{2h} \quad (2.57)$$

die resultierende Kurve linear. ■

2.2.3 Interpolation durch Newton-Polynome

Für die Interpolation der Punkte (u_i, y_i) der Messkennlinie durch **Newton-Polynome** wird der folgende Ansatz formuliert:

$$\begin{aligned} \hat{y} = & a_0 + a_1(u - u_0) + a_2(u - u_0)(u - u_1) + \dots \\ & + a_{n-1}(u - u_0)(u - u_1) \cdots (u - u_{n-2}) . \end{aligned} \quad (2.58)$$

Die Koeffizienten werden rekursiv aus den Interpolationsbedingungen in den Stützstellen berechnet:

$$y_0 = a_0 , \quad (2.59)$$

$$y_1 = a_0 + a_1(u_1 - u_0) , \quad (2.60)$$

⋮

$$\begin{aligned} y_{n-1} = & a_0 + a_1(u_{n-1} - u_0) + a_2(u_{n-1} - u_0)(u_{n-1} - u_1) + \dots \\ & + a_{n-1}(u_{n-1} - u_0)(u_{n-1} - u_1) \cdots (u_{n-1} - u_{n-2}) . \end{aligned} \quad (2.61)$$

Damit weisen Newton-Polynome eine leichtere Erweiterbarkeit beim Hinzufügen von Stützstellen als Lagrange-Polynome auf. Außerdem sind die Basispolynome einfacher. Zu ihrer einfachen Berechnung werden **Differenzen** von Funktionen eingeführt.

Definition 2.1: Differenzen einer Funktion

Die **erste Differenz** einer diskreten Funktion y_i ist wie folgt definiert:

$$\Delta y_i = y_{i+1} - y_i , \quad (2.62)$$

wobei Δ als **Differenzenoperator** bezeichnet wird. Für den Differenzenoperator gelten das Distributivgesetz sowie das Kommutativgesetz in Bezug auf eine Konstante. Damit lassen sich auch **höhere Differenzen**

$$\Delta^j y_i = \Delta^{j-1}(\Delta y_i) = \Delta^{j-1} y_{i+1} - \Delta^{j-1} y_i \quad (2.63)$$

rekursiv definieren. ◇

Exemplarisch soll die **zweite Differenz** von y_i berechnet werden. Mit (2.63) und (2.62) ergibt sich für $j = 2$:

$$\Delta^2 y_i = \Delta(\Delta y_i) = \Delta(y_{i+1} - y_i) = \Delta y_{i+1} - \Delta y_i \quad (2.64)$$

$$= y_{i+2} - 2y_{i+1} + y_i . \quad (2.65)$$

Satz 2.1: Differenzenoperator

2.1

Für den Differenzenoperator Δ gilt die folgende Rechenregel:

$$y_i = (1 + \Delta)^i y_0 . \quad (2.66)$$

◊

Beweis 2.1 (Differenzenoperator): Der Beweis erfolgt durch vollständige Induktion. Der Induktionsanfang ($i = 1$) lautet:

$$y_1 = (1 + \Delta) y_0 = y_0 + \Delta y_0 = y_0 + y_1 - y_0 = y_1 . \quad (2.67)$$

Für den Induktionsschluss von i auf $(i + 1)$ erhält man:

$$\begin{aligned} y_{i+1} &= (1 + \Delta)^{i+1} y_0 = (1 + \Delta) \underbrace{(1 + \Delta)^i y_0}_{y_i} \\ &= (1 + \Delta) y_i = y_i + y_{i+1} - y_i = y_{i+1} . \end{aligned} \quad (2.68)$$

◆

Durch Ausmultiplizieren der rechten Seite von (2.66) mit Hilfe der allgemeinen binomischen Formel ergibt sich:

$$\begin{aligned} y_i &= y_0 + \frac{i}{1!} \Delta y_0 + \frac{i(i-1)}{2!} \Delta^2 y_0 + \dots \\ &\quad + \frac{i(i-1)\cdots(i-j+1)}{j!} \Delta^j y_0 + \dots + \frac{i!}{i!} \Delta^i y_0 . \end{aligned} \quad (2.69)$$

Aus dem Ansatz (2.59)–(2.61) folgt bei konstantem Stützstellenabstand h :

$$y_0 = a_0 , \quad (2.70)$$

$$y_1 = a_0 + a_1 h , \quad (2.71)$$

⋮

$$y_i = a_0 + a_1 ih + \dots + a_j i(i-1)\cdots(i-j+1) h^j + \dots + a_i i! h^i , \quad (2.72)$$

⋮

$$y_{n-1} = a_0 + a_1 (n-1) h + a_2 (n-1)(n-2) h^2 + \dots + a_{n-1} (n-1)! h^{n-1} . \quad (2.73)$$

Vergleicht man (2.72) mit den Differenzen aus (2.69), so erhält man für die Koeffizienten vom Grad j jeweils

$$a_j \cdot i(i-1)\cdots(i-j+1) \cdot h^j = i(i-1)\cdots(i-j+1) \cdot \Delta^j y_0 / j!. \quad (2.74)$$

Die Koeffizienten des Newton-Polynoms lassen sich damit aus den Differenzen wie folgt berechnen:

$$a_j = \frac{\Delta^j y_0}{j! h^j}. \quad (2.75)$$

Mit den Koeffizienten wird die Newton'sche Interpolationsformel für *äquidistante* Stützstellen zu

$$\begin{aligned} \hat{y} = y_0 + \frac{\Delta y_0}{h}(u - u_0) + \frac{\Delta^2 y_0}{2h^2}(u - u_0)(u - u_1) + \cdots \\ + \frac{\Delta^{n-1} y_0}{(n-1)! h^{n-1}} (u - u_0) \cdots (u - u_{n-2}). \end{aligned} \quad (2.76)$$

Die höheren Differenzen lassen sich auf einfache Weise mit dem folgenden **Differenzenschema** durch fortlaufende Subtraktion berechnen:

| u | y | Δy | $\Delta^2 y$ | $\Delta^3 y$ | $\Delta^4 y$ |
|------|-------|--------------|----------------|----------------|----------------|
| 0 | y_0 | | | | |
| h | | Δy_0 | | | |
| | y_1 | | $\Delta^2 y_0$ | | |
| | | Δy_1 | | $\Delta^3 y_0$ | |
| $2h$ | y_2 | | $\Delta^2 y_1$ | | $\Delta^4 y_0$ |
| | | Δy_2 | | $\Delta^3 y_1$ | |
| $3h$ | y_3 | | $\Delta^2 y_2$ | | |
| | | Δy_3 | | | |
| $4h$ | y_4 | | | | |

(2.77)

Bei nicht äquidistanten Stützstellen lässt sich das Differenzenschema verallgemeinern, indem man die Funktionsdifferenzen durch den Stützstellenabstand teilt.

Beispiel 2.2 (Newton-Interpolation mit drei Stützstellen): Gegeben sei wieder eine Kennlinie mit drei äquidistanten Stützstellen u_a , u_m und u_e mit:

$$u_a = 0, \quad u_m = h, \quad u_e = 2h, \quad y_a = 0. \quad (2.78)$$

Aus dem Differenzenschema

| u | y | Δy | $\Delta^2 y$ |
|------|--|------------|--------------|
| 0 | y_a | | |
| h | $y_m \quad \begin{array}{l} \nearrow y_m - y_a \\ \nearrow y_e - 2y_m + y_a \\ \nearrow y_e - y_m \end{array}$ | | |
| $2h$ | y_e | | |

(2.79)

ergibt sich folgende Interpolationsfunktion:

$$\hat{y} = \frac{y_m}{h} u + \frac{y_e - 2y_m}{2h^2} u(u - h). \quad (2.80)$$

Das Ergebnis stimmt mit dem Ergebnis der Lagrange-Interpolation (2.56) überein. Für $y_e = 2y_m$ verschwindet ebenfalls das quadratische Glied. ■

2.2.4 Spline-Interpolation

Wie eingangs erwähnt, ist die Polynominterpolation bei einer großen Anzahl von Messpunkten ungeeignet, weil man dann Interpolationspolynome hohen Grades erhält, die ein stark oszillierendes Verhalten aufweisen. Eine Alternative besteht darin, den Teilintervallen verschiedene Polynome niedrigen Grades zuzuordnen. Das einfachste Verfahren verbindet hierbei benachbarte Messpunkte linear (Abb. 2.7). Bei n Messpunkten besteht die Interpolationsfunktion aus $n - 1$ Geradenstücken in den Intervallen

$$[u_i, u_{i+1}], \quad i \in \{0, \dots, n - 2\}. \quad (2.81)$$

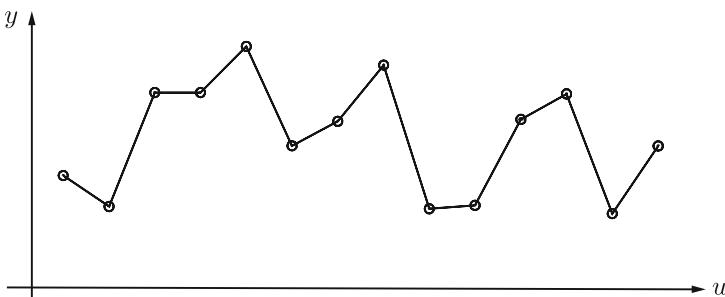


Abbildung 2.7. Lineares Interpolieren zwischen den Stützstellen.

Nachteilig an dieser Methode ist die Tatsache, dass bereits die erste Ableitung der Interpolationsfunktion nicht mehr stetig ist. Zur Kennlinieninterpolation ist solch ein Ansatz nicht geeignet, weil man sich für die erste Ableitung der Messkennlinie, die der Empfindlichkeit entspricht (vgl. Ab-

schn. 3.1), eine stetige Funktion wünscht. Abhilfe schafft die Interpolation mit **Splines**.

Spline-Funktionen gehen auf das mechanische Modell von dünnen Latten (engl. *splines*) zurück, wie sie im Schiffbau oder in der Konstruktion eingesetzt werden. Das Modell geht davon aus, dass durch die gegebenen Stützstellen eine dünne, homogene Latte gelegt sei. In den Stützstellen sei die Latte so gelagert, dass dort keine äußeren Kräfte einwirken. Die so entstehende Biegelinie der Latte soll die Lösung $s(u)$ der Interpolationsaufgabe sein. Die aufgrund der Biegekraft in der Latte gespeicherte Energie lässt sich, abgesehen von physikalischen und geometrischen Konstanten, als das Integral über das Quadrat der **Krümmung** $\kappa(u) \approx s''(u)$ beschreiben¹:

$$E = \frac{1}{2} \int_{u_0}^{u_{n-1}} (s''(u))^2 \, du. \quad (2.83)$$

Der stabile Arbeitspunkt eines Systems stellt sich dann ein, wenn die Energie des Systems ihr Minimum annimmt. Auf das Modell übertragen wird genau die Biegelinie $s(u)$ angenommen, die (2.83) minimiert:

$$\int_{u_0}^{u_{n-1}} (s''(u))^2 \, du \rightarrow \min. \quad (2.84)$$

Die gesuchte Spline-Funktion $s(u)$ minimiert also obige Gleichung unter der Voraussetzung, dass $s(u)$ mindestens einmal stetig differenzierbar ist. Aus diesen Bedingungen ergeben sich nach Lösung von (2.84) mit Hilfe der Variationsrechnung [12, 31] folgende Eigenschaften der Spline-Interpolierenden:

$$s_i(u_i) = y_i, \quad i \in \{0, 1, \dots, n - 1\}, \quad (2.85)$$

$$s''_i(u_i + 0) = s''_{i-1}(u_i - 0), \quad i \in \{1, 2, \dots, n - 2\}, \quad (2.86)$$

$$s''_0(u_0) = s''_{n-2}(u_{n-1}), \quad (2.87)$$

$$s'''(u) = 0, \quad u \neq u_0, \dots, u_{n-1}. \quad (2.88)$$

¹Streng genommen ist die Krümmung durch

$$\kappa(u) = \frac{s''(u)}{(1 + s'(u)^2)^{3/2}} \quad (2.82)$$

gegeben, weshalb die Näherung $\kappa(u) \approx s''(u)$ lediglich bei kleinen Steigungen $s'(u)^2 \ll 1$ zulässig ist. Dennoch liefert (2.83) auch bei Verletzung dieser Annahme meist eine gute Näherung der Biegeenergie der Latte.

Gleichung (2.85) ist gerade die Interpolationsbedingung. Aus (2.86) folgt, dass nicht nur die erste, sondern auch die zweite Ableitung der Interpolationsfunktion stetig ist. Wegen (2.88) ist $s(u)$ in jedem Teilintervall ein kubisches Polynom. Die interpolierende Spline-Funktion setzt sich also stückweise aus Polynomen dritten Grades zusammen (Abb. 2.8). Man bezeichnet daher $s(u)$ auch als kubischen Spline. Natürlich können auch Splines höherer Ordnung durch Erweiterung von (2.84) auf höhere Ableitungen erzeugt werden. Diese Erweiterung ist allerdings dann rein mathematischer Natur, ohne physikalische Motivation.

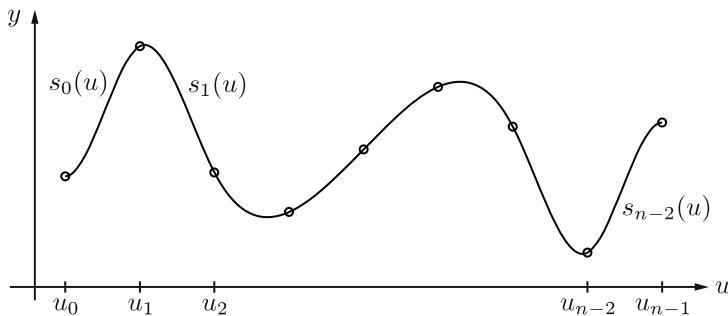


Abbildung 2.8. Interpolation mit kubischen Splines.

④ Berechnung der kubischen Spline-Interpolierenden

Mit den Eigenschaften in (2.85) bis (2.88) kann nun die gesuchte Interpolationsfunktion $s(u)$ konstruiert werden. Für jedes Teilintervall $[u_i, u_{i+1}]$ der Länge

$$h_i = u_{i+1} - u_i \quad (2.89)$$

wählt man als Ansatz ein allgemeines Polynom 3. Grades:

$$s_i(u) = a_i (u - u_i)^3 + b_i (u - u_i)^2 + c_i (u - u_i) + d_i. \quad (2.90)$$

Für den Funktionswert und die ersten beiden Ableitungen an den Enden des Intervalls – d. h. an den Stützstellen – erhält man:

$$s_i(u_i) = d_i = y_i, \quad (2.91)$$

$$s_i(u_{i+1}) = a_i h_i^3 + b_i h_i^2 + c_i h_i + d_i = y_{i+1}, \quad (2.92)$$

$$s'_i(u_i) = c_i, \quad (2.93)$$

$$s'_i(u_{i+1}) = 3 a_i h_i^2 + 2 b_i h_i + c_i, \quad (2.94)$$

$$s_i''(u_i) = 2 b_i = y_i'', \quad (2.95)$$

$$s_i''(u_{i+1}) = 6 a_i h_i + 2 b_i = y_{i+1}'', \quad (2.96)$$

Die unbekannten Parameter a_i, b_i, c_i und d_i können damit durch die gegebenen Stützpunkte y_i und y_{i+1} sowie die noch unbekannten zweiten Ableitungen y_i'' und y_{i+1}'' ausgedrückt werden:

$$a_i = \frac{1}{6h_i} (y_{i+1}'' - y_i''), \quad (2.97)$$

$$b_i = \frac{1}{2} y_i'', \quad (2.98)$$

$$c_i = \frac{1}{h_i} (y_{i+1} - y_i) - \frac{1}{6} h_i (y_{i+1}'' + 2y_i''), \quad (2.99)$$

$$d_i = y_i. \quad (2.100)$$

Die unbekannten zweiten Ableitungen sollen nun aus der Bedingung für die Stetigkeit der ersten Ableitung an den inneren Stützstellen berechnet werden. Für die erste Ableitung am Intervallende u_{i+1} ergibt sich nach Einsetzen der Parameter a_i, b_i, c_i

$$s_i'(u_{i+1}) = \frac{1}{h_i} (y_{i+1} - y_i) + \frac{h_i}{6} (2y_{i+1}'' + y_i''). \quad (2.101)$$

Die Forderung nach Stetigkeit der ersten Ableitung an den inneren Stützstellen ergibt

$$s_i'(u_{i+1}) = s_{i+1}'(u_{i+1}) \quad (2.102)$$

mit

$$s_{i+1}'(u_{i+1}) = c_{i+1} = \frac{1}{h_{i+1}} (y_{i+2} - y_{i+1}) - \frac{h_{i+1}}{6} (y_{i+2}'' + 2y_{i+1}''). \quad (2.103)$$

Setzt man beide Gleichungen gleich und ordnet nach den unbekannten zweiten Ableitungen, so erhält man

$$\begin{aligned} h_i y_i'' + 2(h_i + h_{i+1}) y_{i+1}'' + h_{i+1} y_{i+2}'' \\ = \frac{6}{h_{i+1}} (y_{i+2} - y_{i+1}) - \frac{6}{h_i} (y_{i+1} - y_i). \end{aligned} \quad (2.104)$$

Berücksichtigt man (2.87), also $y_0'' = y_{n-1}'' = 0$, so resultieren $(n-2)$ lineare Gleichungen für die unbekannten zweiten Ableitungen $y_1'', y_2'', \dots, y_{n-2}''$:

$$\begin{aligned}
& \begin{bmatrix} 2(h_0+h_1) & h_1 & 0 & \cdots & 0 \\ h_1 & 2(h_1+h_2) & h_2 & 0 & \cdots & 0 \\ \vdots & & & \ddots & & \vdots \\ 0 & \cdots & 0 & 0 & h_{n-3} & 2(h_{n-3}+h_{n-2}) \end{bmatrix} \begin{bmatrix} y''_1 \\ y''_2 \\ \vdots \\ y''_{n-2} \end{bmatrix} = \\
& = \begin{bmatrix} \frac{6}{h_1}(y_2 - y_1) - \frac{6}{h_0}(y_1 - y_0) \\ \frac{6}{h_2}(y_3 - y_2) - \frac{6}{h_1}(y_2 - y_1) \\ \vdots \\ \frac{6}{h_{n-2}}(y_{n-1} - y_{n-2}) - \frac{6}{h_{n-3}}(y_{n-2} - y_{n-3}) \end{bmatrix}. \tag{2.105}
\end{aligned}$$

Der Rechenweg zur Bestimmung der kubischen Spline-Interpolierenden $s(u)$ ist nun naheliegend. Aus den gegebenen Stützstellen (u_i, y_i) , $i \in \{0, \dots, n-1\}$, werden die Längen der Intervalle h_i bestimmt und das Gleichungssystem (2.105) zur Bestimmung der y''_i aufgestellt. Nach dessen Lösung werden nach (2.97)–(2.100) die Koeffizienten a_i, b_i, c_i und d_i der zum Teilintervall $[u_i, u_{i+1}]$ gehörenden kubischen Polynome $s_i(u)$ bestimmt.

Beispiel 2.3 (Interpolation einer Sinusfunktion): Im Folgenden soll die Funktion $\sin(\frac{\pi}{2}u)$ im Intervall $[-2; 2]$ sowohl durch Lagrange-Polynome als auch durch Spline-Funktionen interpoliert werden. Die Stützstellen sind

$$\begin{aligned}
\mathbf{u} &= [-2, -1, 0, 1, 2], \\
\mathbf{y} &= [0, -1, 0, 1, 0].
\end{aligned} \tag{2.106}$$

Die Lagrange-Interpolation liefert das Polynom

$$\hat{y}_L = -\frac{1}{3}u^3 + \frac{4}{3}u, \quad -1 \leq u < 1. \tag{2.107}$$

Mit (2.105) erhält man für die zweiten Ableitungen der Spline-Interpolation

$$\mathbf{y}'' = [0, +3, 0, -3, 0] \tag{2.108}$$

und gelangt damit zu folgender Interpolationsfunktion:

$$\hat{y}_s(u) = s(u) = \begin{cases} \frac{1}{2}u^3 + 3u^2 + \frac{9}{2}u + 1 & \text{für } u < -1, \\ -\frac{1}{2}u^3 + \frac{3}{2}u & \text{für } -1 \leq u < 1, \\ \frac{1}{2}u^3 - 3u^2 + \frac{9}{2}u - 1 & \text{für } u \geq 1. \end{cases} \tag{2.109}$$

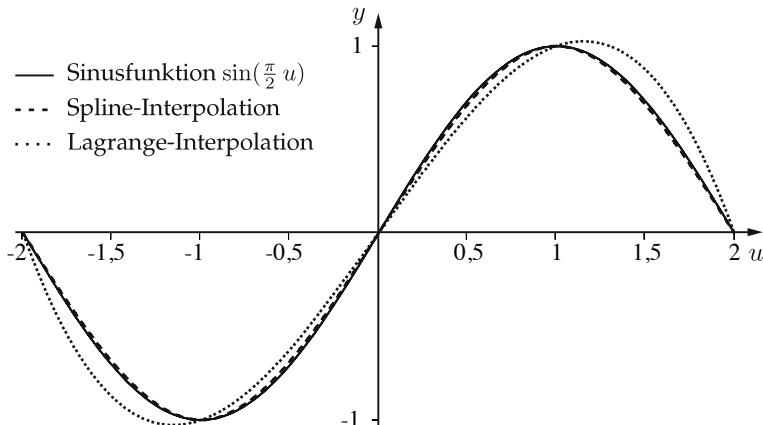


Abbildung 2.9. Vergleich von Spline- und Lagrange-Interpolation.

Abbildung 2.9 zeigt das Interpolationsergebnis. Man erkennt, dass die Spline-Interpolation in diesem Fall das bessere Ergebnis liefert. ■

2.2.5 Systemtheoretische Deutung der Interpolation

Bei äquidistanten Stützstellen lässt sich die zu interpolierende Funktion durch eine Multiplikation einer kontinuierlichen Funktion $g(x)$ mit einer zweidimensionalen Impulsreihe ausdrücken [42]:

$$y_*(u) = y(u) \cdot \sum_{n=-\infty}^{\infty} \delta(u - n \Delta u), \quad (2.110)$$

wobei Δu den Stützstellenabstand bezeichnet. Durch die Multiplikation mit der Impulsreihe verschwindet die Abtastfunktion $y_*(u)$ für $u \neq n \Delta u$, $n \in \mathbb{Z}$. Bestimmte Interpolationsverfahren können als Faltung des diskreten Signals $y_*(u)$ mit einem linearen, verschiebungsinvarianten Interpolationsfilter mit der **Impulsantwort** $i(u)$ dargestellt werden:

$$\hat{y}(u) = \left[y(u) \sum_{n=-\infty}^{\infty} \delta(u - n \Delta u) \right] * i(u) \quad (2.111)$$

$$\hat{Y}(f) = \left[\frac{1}{\Delta u} \sum_{k=-\infty}^{\infty} Y\left(f - \frac{k}{\Delta u}\right) \right] \cdot I(f). \quad (2.112)$$

Die Faltung des Abtastsignals $y_*(u)$ mit der Impulsantwort $i(u)$ geht dann im Frequenzbereich in eine Multiplikation des periodisch fortgesetz-

Tabelle 2.1. Vergleich von Interpolationsarten.

| Interpolationsart | Ordnung | $i(u)$ | $I(f)$ |
|--------------------------------|----------|--|--------------------------------------|
| Nächster-Nachbar-Interpolation | 0 | $\text{rect}\left(\frac{u}{\Delta u}\right)$ | $\Delta u \text{sinc}(f \Delta u)$ |
| Lineare Interpolation | 1 | $\Lambda\left(\frac{u}{\Delta u}\right)$ | $\Delta u \text{sinc}^2(f \Delta u)$ |
| Ideale Interpolation | ∞ | $\text{sinc}\left(\frac{u}{\Delta u}\right)$ | $\Delta u \text{rect}(f \Delta u)$ |

ten Spektrums mit der Übertragungsfunktion $I(f)$ des Interpolationsfilters über. Weist das Interpolationsfilter $I(f)$ eine geeignete Tiefpasscharakteristik auf, so kann durch die Multiplikation das kontinuierliche Signal rekonstruiert werden (vgl. Abschn. 7.1). Tabelle 2.1 zeigt eine Übersicht einfacher Interpolationsarten zusammen mit ihrer jeweiligen Impulsantwort und Übertragungsfunktion. Dabei sind die Rechteck-, die Dreieck- und die Sinc-Funktion wie folgt definiert:

$$r_B(x) = \text{rect}\left(\frac{x}{B}\right) = \begin{cases} 1 & \text{für } |x| < \frac{B}{2} \\ 0 & \text{sonst} \end{cases}, \quad (2.113)$$

$$\Lambda\left(\frac{x}{B}\right) = \text{rect}\left(\frac{x}{B}\right) * \text{rect}\left(\frac{x}{B}\right) = \begin{cases} 1 - \left|\frac{x}{B}\right| & \text{für } |x| < B \\ 0 & \text{sonst} \end{cases}, \quad (2.114)$$

$$\text{sinc}(x) = \frac{\sin(\pi x)}{\pi x}. \quad (2.115)$$

2.3 Kennfeldinterpolation

2.3

In den Abschnitten 2.1 und 2.2 wurde eine analytische Kennliniendarstellung aus Messpunkten bestimmt. Dies geschah bei der Auslegung des Messsystems – und damit *vor* der eigentlichen Messung. Wenn die physikalische Kennlinie des Messsystems eine nichtlineare Funktion einer unabhängigen Messgröße u und einer messbaren systematischen Störgröße z ist, kann man die Ausgangsgröße y *in jedem Messvorgang* durch eine **zweidimensionale Interpolation** berechnen. Dabei wird der bekannte Einfluss der Störgröße z auf die Ausgangsgröße kompensiert. Dazu werden die zuvor bestimmten Kennfeldwerte y in den äquidistanten Stützstellen (u_i, z_j) messtechnisch erfasst und abgespeichert. Die Zwischenwerte in den Stützstellenintervallen werden für feiner aufgelöste Eingangsgrößen interpoliert.

Zur Herleitung der Interpolationsvorschrift auf der Grundlage von Polynomen muss (2.37) auf zwei Dimensionen erweitert werden:

$$\hat{y}(u, z) = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} a_{ij} u^i z^j. \quad (2.116)$$

Alternativ darf bei Polynomen 1. Grades die Ausgangsgröße y an der Stelle (u_i, z_j) in eine Taylor-Reihe entwickelt werden, die nach den linearen Gliedern abgebrochen wird. Zur besseren Übersicht wird die folgende verkürzte Schreibweise verwendet:

$$\left. \frac{\partial f(u, z)}{\partial u} \right|_{u_i, z_j} \rightarrow \frac{\partial f}{\partial u}(u_i, z_j). \quad (2.117)$$

Es ergibt sich

$$y(u, z) = f(u_i + \Delta u, z_j + \Delta z) \quad (2.118)$$

$$\begin{aligned} &\approx f(u_i, z_j) + \frac{\partial f}{\partial u}(u_i, z_j) \Delta u + \frac{\partial f}{\partial z}(u_i, z_j) \Delta z \\ &+ \frac{\partial^2 f}{\partial u \partial z}(u_i, z_j) \Delta u \Delta z. \end{aligned} \quad (2.119)$$

Durch Approximation der Ableitungen durch Differenzenquotienten folgt:

$$\begin{aligned} y(u, z) &\approx y(u_i, z_j) + \frac{\Delta y(u_i)}{\Delta u_i} \Delta u + \frac{\Delta y(z_j)}{\Delta z_j} \Delta z \\ &+ \frac{\Delta^2 y(u_i, z_j)}{\Delta u_i \Delta z_j} \Delta u \Delta z. \end{aligned} \quad (2.120)$$

Beispiel 2.4 (Bilineare Interpolation): Die Näherung (2.120) entspricht der sogenannten **bilinearen Interpolation**. Durch Einführung der Abkürzungen $y_{kl} = y(u_{i+k}, z_{j+l})$ lassen sich die Differenzen in (2.120) wie folgt schreiben:

$$\Delta y(u_i) = y_{10} - y_{00}, \quad \Delta y(z_j) = y_{01} - y_{00}, \quad (2.121)$$

$$\Delta^2 y(u_i, z_j) = y_{11} - y_{10} - y_{01} + y_{00}. \quad (2.122)$$

Bei normierten Stützstellenweiten $\Delta u_i = \Delta z_j = 1$ ergibt sich schließlich:

$$\begin{aligned} y(u, z) &\approx y_{00} + (y_{10} - y_{00}) \Delta u + (y_{01} - y_{00}) \Delta z \\ &+ (y_{11} - y_{10} - y_{01} + y_{00}) \Delta u \Delta z \end{aligned} \quad (2.123)$$

$$\begin{aligned}
 &= y_{00} (1 - \Delta u) (1 - \Delta z) + y_{10} \Delta u (1 - \Delta z) \\
 &\quad + y_{01} (1 - \Delta u) \Delta z + y_{11} \Delta u \Delta z.
 \end{aligned} \tag{2.124}$$

Abbildung 2.10 veranschaulicht die bilineare Interpolation. ■

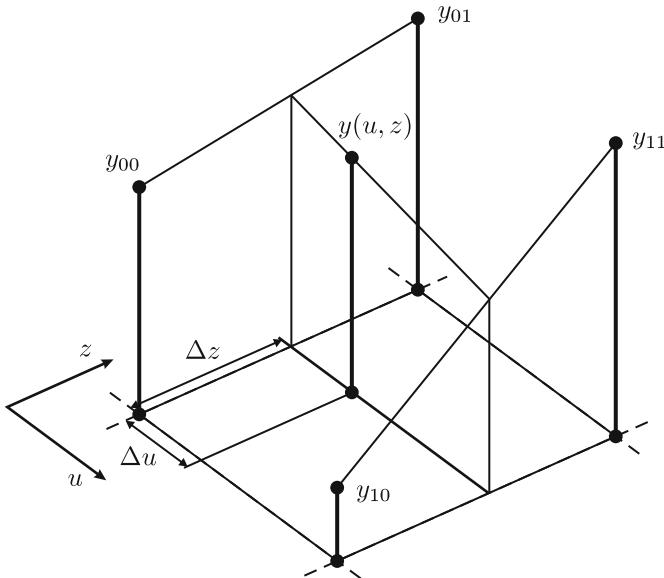


Abbildung 2.10. Bilineare Interpolation.

④ Rechnergestützte Kennfeldberechnung

Neben der allgemeinen Darstellung der Kennfeldinterpolation interessiert ferner eine günstige Darstellung für die Implementierung auf einem Rechner. Hierbei müssen häufig nicht beliebige Zwischenwerte interpoliert werden, sondern man beschränkt sich auf ein festes Raster der Breite q_u bzw. q_z zwischen den Stützstellen. Mit einer Rasterung des Intervalls in Zweierpotenzen,

$$\Delta u_i = u_{i+1} - u_i = 2^r q_u, \quad \Delta z_j = z_{j+1} - z_j = 2^r q_z, \tag{2.125}$$

und der Annahme, dass die Intervallbreiten beider unabhängiger Variablen u und z die gleiche Auflösung 2^r bei vorgegebener Quantisierung q_u, q_z be-

sitzen, lässt sich die Approximation (2.120) folgendermaßen darstellen:

$$\begin{aligned} y(u, z) &\approx y(u_i, z_j) \\ &+ \frac{y(u_{i+1}, z_j) - y(u_i, z_j)}{2^r} \frac{\Delta u}{q_u} + \frac{y(u_i, z_{j+1}) - y(u_i, z_j)}{2^r} \frac{\Delta z}{q_z} \\ &+ \frac{y(u_{i+1}, z_{j+1}) - y(u_i, z_{j+1}) - y(u_{i+1}, z_j) + y(u_i, z_j)}{2^{2r}} \frac{\Delta u \Delta z}{q_u q_z}. \end{aligned} \quad (2.126)$$

Durch Ordnen der Summanden nach den Stützstellenelementen ergibt sich

$$\begin{aligned} y(u, z) &= y(u_i, z_j) \left[1 - \frac{1}{2^r} \frac{\Delta u}{q_u} - \frac{1}{2^r} \frac{\Delta z}{q_z} + \frac{1}{2^{2r}} \cdot \frac{\Delta u \Delta z}{q_u q_z} \right] \\ &+ y(u_{i+1}, z_j) \left[\frac{1}{2^r} \frac{\Delta u}{q_u} - \frac{1}{2^{2r}} \frac{\Delta u \Delta z}{q_u q_z} \right] \\ &+ y(u_i, z_{j+1}) \left[\frac{1}{2^r} \frac{\Delta z}{q_z} - \frac{1}{2^{2r}} \frac{\Delta u \Delta z}{q_u q_z} \right] + y(u_{i+1}, z_{j+1}) \frac{1}{2^{2r}} \frac{\Delta u \Delta z}{q_u q_z}. \end{aligned} \quad (2.127)$$

Ausklemmen des Faktors 2^{2r} führt auf

$$\begin{aligned} y(u, z) &= \frac{1}{2^{2r}} \left[y(u_i, z_j) \left(2^r - \frac{\Delta u}{q_u} \right) \left(2^r - \frac{\Delta z}{q_z} \right) + y(u_{i+1}, z_{j+1}) \frac{\Delta u \Delta z}{q_u q_z} \right. \\ &\quad \left. + y(u_{i+1}, z_j) \frac{\Delta u}{q_u} \left(2^r - \frac{\Delta z}{q_z} \right) + y(u_i, z_{j+1}) \left(2^r - \frac{\Delta u}{q_u} \right) \frac{\Delta z}{q_z} \right]. \end{aligned} \quad (2.128)$$

In Summenschreibweise ergibt sich die Form

$$y(u, z) = \frac{1}{2^{2r}} \cdot \sum_{m=0}^1 \sum_{n=0}^1 k_{m,n} \cdot y(u_{i+m}, z_{j+n}). \quad (2.129)$$

Man erhält als Interpolationsfunktion $y(u, z)$ eine gewichtete Mittelwertbildung, bei der die Funktionsamplituden $y(u_{i+m}, z_{j+n})$ an den Ecken des Intervallquadrats mit dem Produkt

$$k_{m,n} = \left((m-1) 2^r + (-1)^m \frac{\Delta u}{q_u} \right) \left((n-1) 2^r + (-1)^n \frac{\Delta z}{q_z} \right) \quad (2.130)$$

der gegenüberliegenden Teilintervalle gewichtet werden (siehe Abb. 2.11).

Die Funktionsamplitude $y(u_i, z_{j+1})$ wird beispielsweise mit

$$k_{0,1} = \left(-2^r + \frac{\Delta u}{q_u} \right) \left(-\frac{\Delta z}{q_z} \right) = \frac{(2^r q_u - \Delta u)}{q_u} \cdot \frac{\Delta z}{q_z} \quad (2.131)$$

gewichtet, d. h. mit dem Produkt der Teilintervalle $(u_{i+1} - \Delta u) \cdot \Delta z$. Je näher die Interpolationsstelle (u, z) an eine der vorab gemessenen Stützstellen

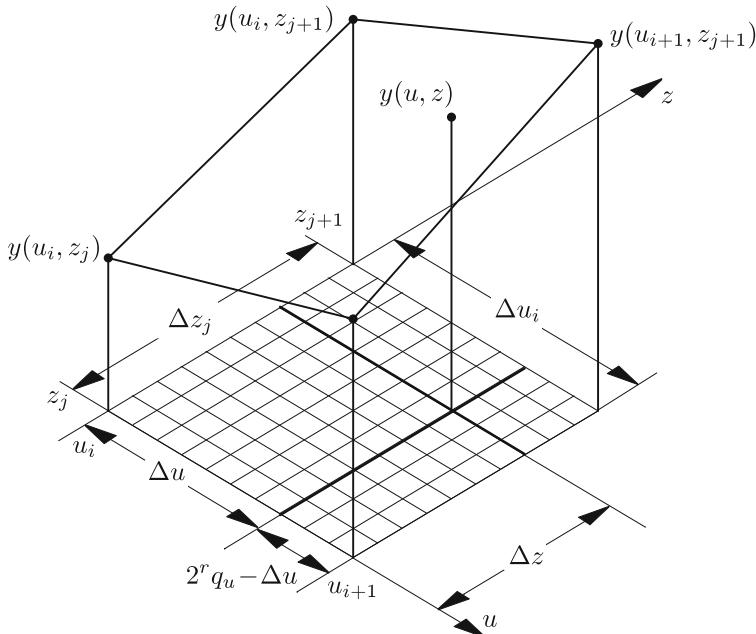


Abbildung 2.11. Interpolation in einem Intervall.

an den Intervallgrenzen \$(u_{i+m}, z_{j+n})\$ heranrückt, desto stärker geht diese in das Interpolationsergebnis ein. Wegen der linearen Interpolation im Intervall weist das Interpolationskennfeld an den Intervallgrenzen Unstetigkeiten in der Steigung auf. Die Summe der Koeffizienten ist begrenzt auf

$$\left| \sum_{m=0}^1 \sum_{n=0}^1 k_{m,n} \right| = \left| 2^{2r} - 2^r \frac{\Delta u}{q_u} - 2^r \frac{\Delta z}{q_z} + 2 \frac{\Delta u \Delta z}{q_u q_z} + 2^r \frac{\Delta u}{q_u} - \frac{\Delta u \Delta z}{q_u q_z} + 2^r \frac{\Delta z}{q_z} - \frac{\Delta u \Delta z}{q_u q_z} \right| = 2^{2r}. \quad (2.132)$$

3

Kapitel 3

Stationäres Verhalten von Messsystemen

3 Stationäres Verhalten von Messsystemen

| | | |
|---------|--|-----|
| 3.1 | Stationäre Messkennlinie und deren Fehler | 55 |
| 3.1.1 | Ideale und reale Messkennlinie | 55 |
| 3.1.2 | Abgleich der Messkennlinie | 56 |
| 3.1.3 | Kennlinienfehler bei realer Kennlinie | 58 |
| 3.1.3.1 | Relativer Kennlinienfehler | 58 |
| 3.1.3.2 | Hysterese und Umkehrspanne | 59 |
| 3.1.4 | Abschätzung des Kennlinienfehlers | 61 |
| 3.2 | Kennlinienfehler unter Normalbedingungen | 64 |
| 3.2.1 | Herabsetzen des Messbereichs | 64 |
| 3.2.2 | Reihenschaltung zweier nichtlinearer Glieder | 67 |
| 3.2.3 | Wahl des günstigsten Messbereichs | 71 |
| 3.2.4 | Differenzmethode | 77 |
| 3.2.5 | Gegenkopplung | 83 |
| 3.3 | Kennlinienfehler bei Abweichungen von den Normalbedingungen | 87 |
| 3.3.1 | Superponierende Störgrößen | 87 |
| 3.3.2 | Unterdrückung superponierender Störgrößen mit der Differenzmethode | 89 |
| 3.3.3 | Deformierende Störgrößen | 90 |
| 3.3.4 | Deformierende Störgrößen bei Gegenkopplung | 93 |
| 3.3.5 | Superponierende Störgrößen bei Gegenkopplung | 96 |
| 3.3.6 | Kompensation systematischer Störeinflüsse | 97 |
| 3.3.7 | Abschirmung | 98 |
| 3.3.8 | Superponierende Störgrößen in Messketten | 98 |
| 3.3.9 | Zerhackerverstärker | 100 |
| 3.4 | Rückwirkung des Messsystems | 103 |

3 Stationäres Verhalten von Messsystemen

In diesem Kapitel wird das stationäre Verhalten von Messsystemen untersucht. Im stationären Zustand sind alle Einschwingvorgänge des Messsystems abgeklungen und es stellt sich eine stabile Anzeige am Messsystem ein. Das Verhalten des Messsystems wird nun alleine durch die stationäre Messkennlinie bestimmt.

3.1 Stationäre Messkennlinie und deren Fehler

3.1

3.1.1 Ideale und reale Messkennlinie

Die stationäre Messkennlinie $y(u)$ beschreibt den funktionalen Zusammenhang zwischen den Messwerten u und den Anzeigewerten y im stationären Zustand, wobei zunächst der Einfluss von Störungen vernachlässigt werden soll. Die wesentlichen Begriffe werden anhand von Abb. 3.1 definiert.

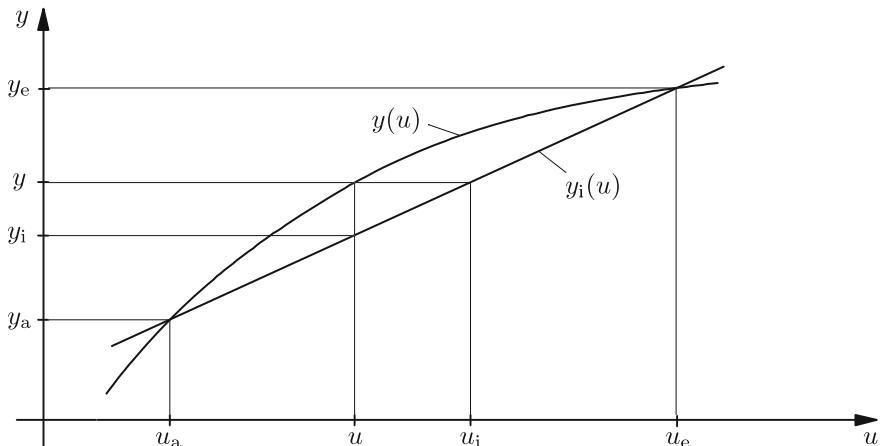
- **Messbereich:** Der Messbereich $[u_a, u_e]$ wird vom Messanfang u_a und vom Messende u_e begrenzt. Dessen Breite beträgt $u_e - u_a$.
- **Anzeigebereich:** Analog wird der Anzeigebereich $[y_a, y_e]$ vom Anzeigeanfang $y_a = y(u_a)$ und vom Anzeigeende $y_e = y(u_e)$ begrenzt. Die Breite des Anzeigebereiches beträgt $y_e - y_a$.
- **Empfindlichkeit:** Die Empfindlichkeit (engl. *sensitivity*) $S(u)$ ist die Steigung der Messkennlinie $y(u)$:

$$S(u) = \frac{\partial y(u)}{\partial u} = \frac{\partial f(u)}{\partial u}. \quad (3.1)$$

Die Empfindlichkeit ist im Allgemeinen eine Funktion der Messgröße u . Diese Abhängigkeit von u charakterisiert das *nichtlineare* Verhalten der realen Kennlinie.

- **Ideale Kennlinie:** Zur Vereinfachung werden Messsysteme häufig durch eine lineare Kennlinie $y_i(u)$ beschrieben, die den Messanfang und das Messende durch eine Gerade miteinander verbindet. Man nennt diese Kennlinie auch ideale Kennlinie. Die ideale Kennlinie besitzt über dem gesamten Messbereich $[u_a, u_e]$ eine konstante Steigung

$$S_i = \frac{y_e - y_a}{u_e - u_a}, \quad (3.2)$$

Abbildung 3.1. Reale Messkennlinie $y(u)$ und ideale Messkennlinie $y_i(u)$.

die man als **ideale Empfindlichkeit** bezeichnet. Damit kann die ideale Messkennlinie angegeben werden:

$$y_i(u) = S_i \cdot (u - u_a) + y_a . \quad (3.3)$$

Gleich große Änderungen der Messgröße gehen mit der gleichen Empfindlichkeit in die Messwertanzeige ein. Diese Eigenschaft wird gerade durch eine lineare Kennlinie gewährleistet. Die Fehler, die durch Verwendung der idealen Kennlinie statt der realen, nichtlinearen Kennlinie entstehen, werden in Abschn. 3.1.3 beschrieben.

- Der **ideale Messwert** u_i resultiert aus dem tatsächlichen Anzeigewert y bei Annahme einer idealen Kennlinie mit konstanter Steigung S_i . Entsprechend ist die ideale Messspanne $u_i - u_a$ abhängig von y .
- Der **ideale Anzeigewert** y_i folgt aus dem tatsächlichen Messwert u bei Annahme einer idealen Kennlinie mit konstanter Steigung S_i . Somit ist die ideale Anzeigespanne $y_i - y_a$ eine Funktion von u .

3.1.2 Abgleich der Messkennlinie

Vor dem Abgleich werden die Normalbedingungen des Messsystems gemäß seiner Spezifikation eingestellt. Äußere Störgrößen werden dabei auf den spezifizierten Werten konstant gehalten oder durch Abschirmung bzw. Filterung ferngehalten. Der Einstellvorgang, der das Messsystem gemäß seiner technischen Beschreibung für die vorgesehene Messaufgabe tauglich macht, ist der **Abgleich** oder die **Justierung**. Unter *Abgleichen* oder *Justieren* versteht man einen physikalischen Eingriff in das Gerät oder seine Maß-

verkörperung mit dem Ziel, den *Messbereich auf den vorgesehenen Bereich der Ausgabeeinrichtung, des Ausgangssignals oder der Anzeige abzubilden*.

Beispiel 3.1 (Justierung eines Temperaturmessgerätes): Eine digitale Temperaturanzeige soll im Messbereich zwischen 20°C und 100°C auf einen Anzeigebereich von 0 bis 100 % gebracht werden. Die Anzeigegröße ist eine elektrische Spannung U zwischen 2 und 300 mV. Eine vorhandene Skala eines Zeigerinstrumentes soll auf möglichst kleine Fehler justiert werden, d. h. bei $U = 2 \text{ mV}$ soll der Zeiger im Anfangspunkt der Skala (20°C) und bei $U = 300 \text{ mV}$ im Endpunkt der Skala (100°C) stehen. ■

Messgeräte haben im Allgemeinen zwei Möglichkeiten für den Abgleich: die **additive Verschiebung** und die **multiplikative Drehung** der Kennlinie. Zwei Verfahren sind bei der Justierung gebräuchlich (siehe Abb. 3.2):

- **Fixpunktjustierung:** Nach der Fixpunktjustierung geht die Kennlinie durch den Anfangspunkt (u_a, y_a) und durch den Endpunkt (u_e, y_e) hindurch. Der Messbereich $[u_a, u_e]$ wird auf den Anzeigebereich $[y_a, y_e]$ abgebildet. Im Messanfang und Messende ist damit der Fehler null.
- **Toleranzbandjustierung:** Die Toleranzbandjustierung entsteht durch eine zusätzliche additive Verschiebung der Fixpunktjustierung. Ziel ist es, den maximalen Fehler im Messbereich möglichst klein zu gestalten. Der maximale Fehler wird im Vergleich zur Fixpunktjustierung auf die Hälfte reduziert. Die Kennlinie geht dann allerdings nicht mehr zwangsläufig durch den Anfangs- und den Endpunkt.

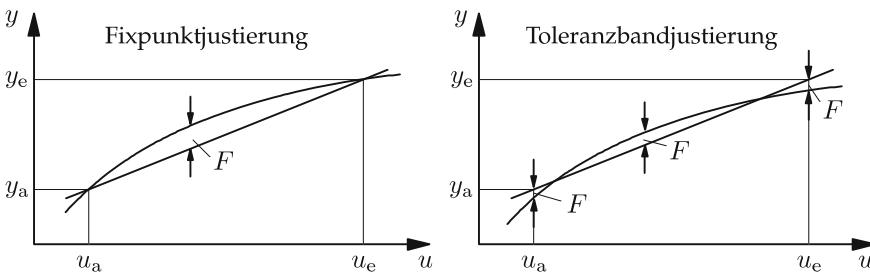


Abbildung 3.2. Fixpunkt- und Toleranzbandjustierung.

In der Praxis wird der Fixpunktjustierung der Vorzug gegeben, weil die Justierung erheblich weniger arbeitsaufwendig ist. Der Abgleich unter Normalbedingungen bedeutet, dass er für den festen, spezifizierten Störgrößenvek-

tor \mathbf{z}_0 vorgenommen wird. Die physikalische Messkennlinie

$$y = f(u, \mathbf{z}) \quad (3.4)$$

wird bei Fixpunktjustierung so justiert, dass die resultierende Kennlinie durch den gewünschten Anfangspunkt

$$y_a \stackrel{!}{=} f(u_a, \mathbf{z}_0) \quad (3.5)$$

geht. Der **Nullpunktfehler** (Offsetfehler)

$$e(\mathbf{z}) = f(u_a, \mathbf{z}) - y_a \quad (3.6)$$

ist unter Normalbedingungen, d. h. für den Störgrößenvektor $\mathbf{z} = \mathbf{z}_0$, aufgrund der Justierung null:

$$e(\mathbf{z}_0) = f(u_a, \mathbf{z}_0) - y_a = 0. \quad (3.7)$$

Vom Justieren oder Abgleichen zu unterscheiden ist das **Eichen**. Das Eichen ist eine gesetzliche Maßnahme, die durch eine Prüfbehörde erfolgt.

❸ 3.1.3 Kennlinienfehler bei realer Kennlinie

Nach der Justierung eines Messsystems werden alle systematischen Fehler als **Kennlinienfehler** interpretiert. Zu diesen Fehlern gehören sowohl die unerwünschten Nichtlinearitäten, die als Abweichung von der idealen Messkennlinie auftreten, als auch der Einfluss von Störgrößen $\mathbf{z} \neq \mathbf{z}_0$. Abschnitt 3.2 widmet sich der Reduktion von Fehlern infolge von Nichtlinearitäten, während in Abschn. 3.3 die Unterdrückung von Störgrößeneinflüssen diskutiert wird.

Zu einem Anzeigewert des Messgerätes y gehört der wahre Messwert u aufgrund der physikalischen Gesetzmäßigkeiten. Die ideale Kennlinie des Messgerätes ordnet aber dem Anzeigewert y einen fehlerhaften Messwert u_i zu. Umgekehrt ordnet die ideale Kennlinie dem wahren Messwert u einen fehlerhaften Anzeigewert y_i zu.

④ 3.1.3.1 Relativer Kennlinienfehler

Bezieht man die Messgröße u auf den Messanfang u_a (Unterdrückung des Anfangswertes), so wird der **relative Kennlinienfehler** nach (1.10) zu

$$F_r = \frac{(u_i - u_a) - (u - u_a)}{u - u_a} = \frac{u_i - u}{u - u_a}. \quad (3.8)$$

Mit Hilfe der idealen Kennlinie (Abb. 3.1)

$$y - y_a = S_i \cdot (u_i - u_a), \quad (3.9)$$

$$y_i - y_a = S_i \cdot (u - u_a), \quad (3.10)$$

$$y - y_i = S_i \cdot (u_i - u) \quad (3.11)$$

kann der relative Kennlinienfehler mittels der Anzeigegrößen angegeben werden:

$$F_r = \frac{y - y_i}{S_i} \cdot \frac{S_i}{y_i - y_a} = \frac{y - y_i}{y_i - y_a}. \quad (3.12)$$

Hierbei handelt es sich um den relativen Kennlinienfehler *bezogen auf die Anzeigespanne*. Neben dieser Variante sind weitere Definitionen des relativen Kennlinienfehlers gängig, bei welchen andere Bezugswerte gewählt werden:

- relativer Kennlinienfehler bezogen auf den Anzeigebereich:

$$F_{rA} = \frac{y - y_i}{y_e - y_a}, \quad (3.13)$$

- relativer Kennlinienfehler bezogen auf den Sollwert y_i :

$$F_{rS} = \frac{y - y_i}{y_i} \quad \text{sowie} \quad (3.14)$$

- relativer Kennlinienfehler bezogen auf den Endwert y_e :

$$F_{rE} = \frac{y - y_i}{y_e}. \quad (3.15)$$

④ 3.1.3.2 Hysterese und Umkehrspanne

Steigert man die Messgröße u langsam und beobachtet dabei die Ausgangsgröße y , und nimmt dann die Messgröße wieder langsam auf den alten Wert zurück, wird man – insbesondere bei Geräten mit feinwerktechnischen Teilen – oftmals feststellen, dass der alte Wert der Ausgangsgröße nicht mehr erreicht wird. Durch den Messvorgang werden im Messgerät langsam veränderliche Vorgänge ausgelöst, die noch nachwirken. Es werden je nach Durchführung des Versuchs zwei Kenngrößen bestimmt.

- **Hysterese:** Zur Ermittlung der Hysterese wird die Messgröße langsam vom Messanfang u_a bis zum Messende u_e gesteigert und wieder langsam auf den Messanfang zurückgenommen (Abb. 3.3 links). Die größte dabei auftretende Differenz H der Messgröße zwischen den richtungsabhängigen Kennlinien wird auf den Messbereich bezogen und als Hysterese h

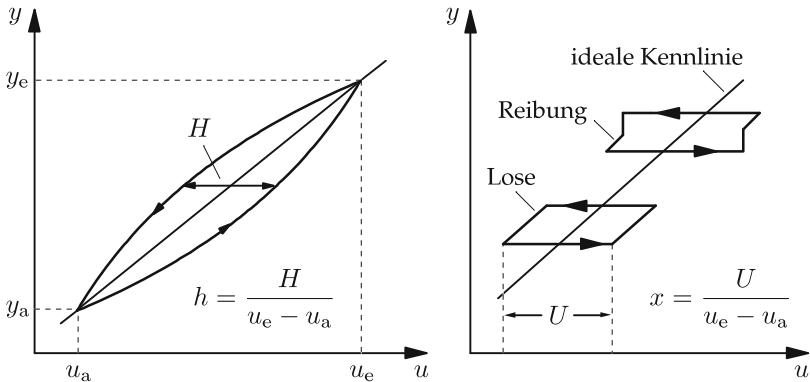


Abbildung 3.3. Hysterese (links) und Umkehrspanne (rechts).

angegeben:

$$h = \frac{H}{u_e - u_a}. \quad (3.16)$$

- **Umkehrspanne:** Die Umkehrspanne wird ähnlich wie die Hysterese ermittelt. Die Messgröße wird aber dazu nur um einige wenige Prozent geändert (Abb. 3.3 rechts). Der größte Fehler U zwischen Auf- und Abwärtsgang wird auf den Messbereich bezogen und als Umkehrspanne x angegeben:

$$x = \frac{U}{u_e - u_a}. \quad (3.17)$$

Beide Erscheinungen können zahlreiche Ursachen haben. Eine Umkehrspanne ist immer dann vorhanden, wenn im Messsystem *Hemmungen* existieren, die verhindern, dass bei einer stationären Messgröße der Gleichgewichtszustand eingenommen wird. Solche Hemmungen sind z. B. mechanische **Reibung** und **Lose**. In Abb. 3.3 sind rechts Umkehrspannen durch Reibung und Lose eingezeichnet. Bei Reibung ändert sich die Anzeigegröße y sprunghaft, wenn die größere Haftreibung überwunden wird.

Existiert eine Umkehrspanne, wird man auch eine Hysterese messen. Das gleiche gilt jedoch nicht umgekehrt. Oft wird Hysterese durch innere Störgrößen hervorgerufen, wie z. B. durch die elastische Nachwirkung von Messfedern. Die Größe der Hystereseschleife ist in solchen Fällen abhängig von der Größe und der Dauer der Belastung, sie verschwindet aber für kleine Änderungen der Messgröße.

3.1.4 Abschätzung des Kennlinienfehlers

Als Grundlage für eine quantitative Bewertung von Maßnahmen zur Fehlerunterdrückung in den Abschnitten 3.2 und 3.3 soll nun der relative Kennlinienfehler abgeschätzt werden. Nach (3.12) resultiert dieser aus der normierten Differenz von physikalischer und idealer Kennlinie:

$$F_r = \frac{y - y_i}{y_i - y_a} . \quad (3.18)$$

Bei dieser Betrachtung werden sowohl äußere Störgrößen \mathbf{z} als auch Nichtlinearitäten der Kennlinie erfasst.

Mit der Empfindlichkeit $S(u, \mathbf{z})$ aus (3.1) – die bei Berücksichtigung von Störungen ebenfalls vom Störgrößenvektor \mathbf{z} abhängig ist – lässt sich die reale, stationäre Kennlinie im Messbereich $[u_a, u_e]$ durch Integration berechnen:

$$y = f(u_a, \mathbf{z}) + \int_{u_a}^u S(u, \mathbf{z}) du . \quad (3.19)$$

Subtrahiert man y_a von beiden Seiten in (3.19), so folgt mit dem Nullpunktfehler (3.6)

$$y - y_a = \int_{u_a}^u S(u, \mathbf{z}) du + \underbrace{f(u_a, \mathbf{z}) - y_a}_{e(\mathbf{z})} = \int_{u_a}^u S(u, \mathbf{z}) du + e(\mathbf{z}) , \quad (3.20)$$

vgl. Abb. 3.4. Der Nullpunktfehler $e(\mathbf{z})$ im Messanfang u_a ist – bedingt durch den Abgleich (3.7) für den Störgrößenvektor \mathbf{z}_0 – gerade null. Die

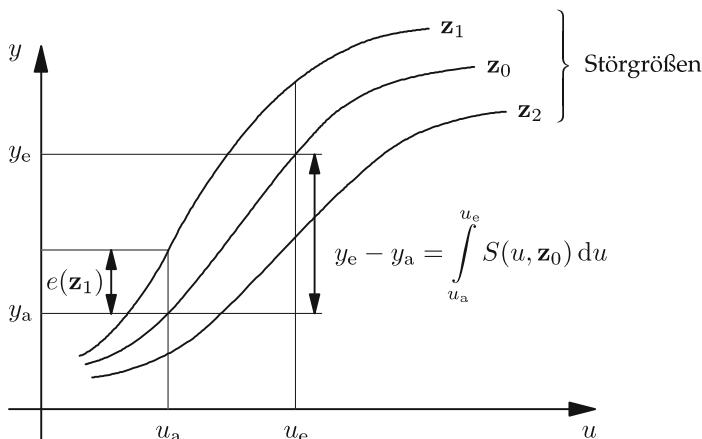


Abbildung 3.4. Physikalische Messkennlinie für unterschiedliche Störgrößenvektoren.

mittlere Empfindlichkeit $\bar{S}(u, u_a, \mathbf{z}_0)$ der realen nichtlinearen Kennlinie des Messgerätes erhält man aus dem Ansatz

$$y - y_a = \int_{u_a}^u S(u, \mathbf{z}_0) du \stackrel{!}{=} \bar{S}(u, u_a, \mathbf{z}_0) \cdot (u - u_a). \quad (3.21)$$

Auflösen nach $\bar{S}(u, u_a, \mathbf{z}_0)$ ergibt:

$$\bar{S}(u, u_a, \mathbf{z}_0) = \frac{1}{u - u_a} \int_{u_a}^u S(u, \mathbf{z}_0) du = \frac{y - y_a}{u - u_a}. \quad (3.22)$$

Für das Messende $u = u_e$ stimmt diese mittlere Empfindlichkeit mit der idealen Empfindlichkeit überein (Abb. 3.5):

$$\bar{S}(u_e, u_a, \mathbf{z}_0) = \frac{1}{u_e - u_a} \int_{u_a}^{u_e} S(u, \mathbf{z}_0) du = S_i. \quad (3.23)$$

Andernfalls lautet die **Empfindlichkeitsdifferenz**:

$$\Delta \bar{S}(u, \mathbf{z}_0) = \bar{S}(u, u_a, \mathbf{z}_0) - S_i. \quad (3.24)$$

Sie hängt von der jeweiligen Position des Messwertes im Messbereich und von der Wahl des Messanfangs ab. Der relative Kennlinienfehler ist nach

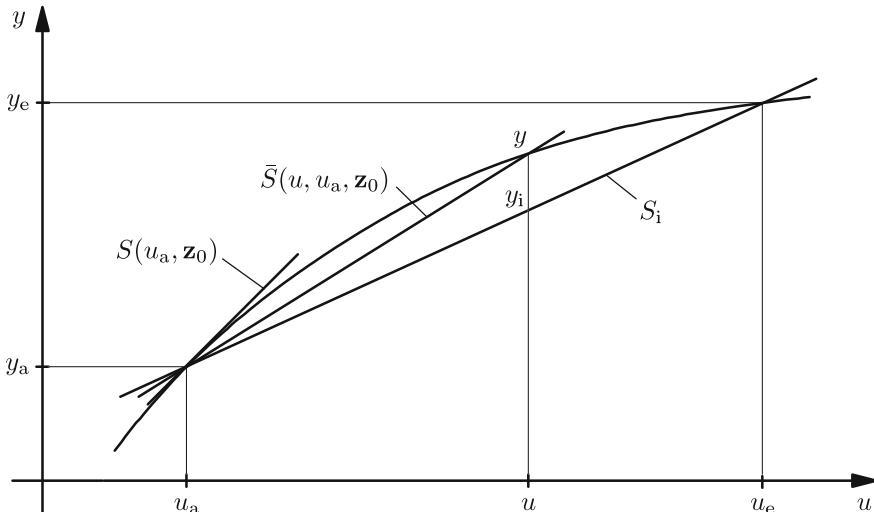


Abbildung 3.5. Veranschaulichung der mittleren Empfindlichkeit $\bar{S}(u, u_a, \mathbf{z}_0)$ unter spezifizierten Normalbedingungen.

(3.12)

$$F_r = \frac{y - y_i}{y_i - y_a} = \frac{(y - y_a) - (y_i - y_a)}{y_i - y_a} \quad (3.25)$$

$$= \frac{(\bar{S}(u, u_a, \mathbf{z}_0) - S_i) \cdot (u - u_a)}{S_i \cdot (u - u_a)} = \frac{\Delta \bar{S}(u, \mathbf{z}_0)}{S_i}, \quad (3.26)$$

wobei für die Umformung (3.26) der Zusammenhang (3.21) verwendet wurde. Aus (3.25) erhält man

$$y - y_a = (y_i - y_a) \cdot (1 + F_r) \quad (3.27)$$

$$= S_i \cdot (u - u_a) \cdot (1 + F_r). \quad (3.28)$$

Nach (3.26) ist der relative Kennlinienfehler F_r am Messanfang $u = u_a$ gerade gleich der Empfindlichkeitsdifferenz $\Delta \bar{S}(u_a, \mathbf{z}_0)$, bezogen auf die ideale Empfindlichkeit S_i . Am Messanfang ist die physikalische Kennlinie y gleich der idealen Kennlinie y_i , d. h. die Differenz ist null. Der dennoch von null abweichende Kennlinienfehler F_r am Messanfang liegt am Bezug auf die Anzeigespanne, die dort ebenfalls null ist. Die Kennlinie (1.7) wird zur Abschätzung des Fehlers in eine Taylor-Reihe um den Messanfang $u = u_a$ entwickelt:

$$\begin{aligned} y &= f(u_a, \mathbf{z}_0) + \left. \frac{\partial f(u, \mathbf{z}_0)}{\partial u} \right|_{u=u_a} (u - u_a) + \frac{1}{2!} \left. \frac{\partial^2 f(u, \mathbf{z}_0)}{\partial u^2} \right|_{u=u_a} (u - u_a)^2 \\ &\quad + \cdots + \frac{1}{(\nu+1)!} \left. \frac{\partial^{(\nu+1)} f(u, \mathbf{z}_0)}{\partial u^{(\nu+1)}} \right|_{u=u_a} (u - u_a)^{\nu+1} + \cdots. \end{aligned} \quad (3.29)$$

Durch Einsetzen der Empfindlichkeit $S(u, \mathbf{z}) = \frac{\partial f(u, \mathbf{z})}{\partial u}$ erhält man:

$$\begin{aligned} y &= f(u_a, \mathbf{z}_0) + S(u_a, \mathbf{z}_0) \cdot (u - u_a) + \frac{1}{2!} S'(u_a, \mathbf{z}_0) \cdot (u - u_a)^2 + \cdots \\ &\quad + \frac{1}{(\nu+1)!} S^{(\nu)}(u_a, \mathbf{z}_0) \cdot (u - u_a)^{\nu+1} + \cdots \end{aligned} \quad (3.30)$$

oder, durch Subtraktion von y_a und Ausklammern von $(u - u_a)$:

$$\begin{aligned} y - y_a &= (u - u_a) \left\{ S(u_a, \mathbf{z}_0) + \frac{1}{2!} S'(u_a, \mathbf{z}_0) \cdot (u - u_a) + \cdots \right. \\ &\quad \left. + \underbrace{\frac{1}{(\nu+1)!} S^{(\nu)}(u_a, \mathbf{z}_0) \cdot (u - u_a)^\nu}_{\bar{S}(u, u_a, \mathbf{z}_0)} + \cdots \right\}. \end{aligned} \quad (3.31)$$

Durch Vergleich mit der mittleren Empfindlichkeit der realen Kennlinie in (3.21) wird diese näherungsweise zu

$$\begin{aligned}\bar{S}(u, u_a, \mathbf{z}_0) &\approx S(u_a, \mathbf{z}_0) + \frac{1}{2!} S'(u_a, \mathbf{z}_0) \cdot (u - u_a) + \dots \\ &\quad + \frac{1}{(\nu + 1)!} S^{(\nu)}(u_a, \mathbf{z}_0) \cdot (u - u_a)^\nu.\end{aligned}\quad (3.32)$$

Schließlich werden in (3.32) die Glieder zweiter und höherer Ordnung vernachlässigt:

$$\bar{S}(u, u_a, \mathbf{z}_0) \approx S(u_a, \mathbf{z}_0) + \frac{1}{2!} S'(u_a, \mathbf{z}_0) \cdot (u - u_a). \quad (3.33)$$

Der relative Kennlinienfehler F_r ist dann nach (3.26)

$$F_r \approx \frac{S(u_a, \mathbf{z}_0) + \frac{1}{2} S'(u_a, \mathbf{z}_0) \cdot (u - u_a) - S_i}{S_i} \quad (3.34)$$

$$\approx \frac{\Delta \bar{S}(u_a, \mathbf{z}_0)}{S_i} + \frac{1}{2} \frac{S'(u_a, \mathbf{z}_0)}{S_i} (u - u_a). \quad (3.35)$$

Über die Definition der mittleren Empfindlichkeit \bar{S} und über deren approximativen Berechnung mittels der Taylor-Reihenentwicklung lässt sich jetzt der relative Kennlinienfehler F_r gemäß (3.34) in Abhängigkeit der Empfindlichkeit und deren Ableitungen angeben.

3.2

3.2 Kennlinienfehler unter Normalbedingungen

In diesem Abschnitt wird das Messsystem unter den spezifizierten Normalbedingungen betrieben. Der Störgrößenvektor hat dabei konstant den Wert \mathbf{z}_0 . Das Gerät ist so justiert, dass der Nullpunktfehler $e(\mathbf{z}_0) = 0$ ist. Die auftretenden Messfehler werden als Abweichung der Messkennlinie von der idealen Messkennlinie interpretiert. Durch die im Folgenden geschilderten Verfahren sollen die Kennlinie der idealen Grundform angenähert werden und damit die Messfehler reduziert werden. Ziel ist also eine möglichst lineare Kennlinie mit konstanter Empfindlichkeit $S = S_i = \text{const.}$

3.2.1 Herabsetzen des Messbereichs

Abbildung 3.6 zeigt das Blockschaltbild der Anordnung zur Herabsetzung des Messbereichs. Das Glied mit nichtlinearer Kennlinie hat die Empfindlichkeit $S(u)$. Angestrebt wird eine lineare Kennlinie mit $S = \text{const.}$ über den Messbereich hinweg. Die Linearisierung soll mit einem *vorgeschalteten* linearen Übertragungsglied der Empfindlichkeit $S_0 \ll 1$ und einem *nachgeschalteten*

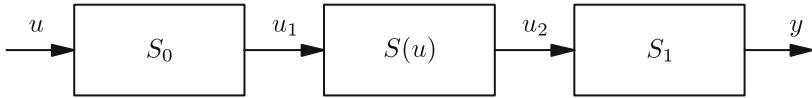


Abbildung 3.6. Herabsetzen des Messbereichs.

schalteten linearen Übertragungsglied der Empfindlichkeit $S_1 \gg 1$ erreicht werden.

Mit den Bezeichnungen aus Abb. 3.6, (3.3) und (3.28) wird

$$u_1 - u_{1a} = S_0 \cdot (u - u_a), \quad (3.36)$$

$$u_2 - u_{2a} = S_i \cdot (1 + F_r) (u_1 - u_{1a}), \quad (3.37)$$

$$y - y_a = S_1 \cdot (u_2 - u_{2a}). \quad (3.38)$$

Daraus erhält man die Kennlinie der gesamten Kette zu

$$y - y_a = S_0 \cdot S_1 \cdot S_i \cdot (1 + F_r) (u - u_a). \quad (3.39)$$

Wegen $S_0 \ll 1$ wird das nichtlineare Glied am Anfang u_{1a} des Messbereiches betrieben. Der relative Fehler des nichtlinearen Gliedes lautet nach (3.35)

$$\begin{aligned} F_r &\approx \frac{\Delta \bar{S}(u_{1a}, z_0)}{S_i} + \frac{1}{2} \frac{S'(u_{1a}, z_0)}{S_i} (u_1 - u_{1a}) \\ &= \frac{S(u_{1a}, z_0)}{S_i} - 1 + \underbrace{\frac{1}{2} \frac{S'(u_{1a}, z_0)}{S_i} S_0 (u - u_a)}_{\approx 0}, \end{aligned} \quad (3.40)$$

wobei der zweite Term wegen $S_0 \ll 1$ vernachlässigt werden kann. Die Ausgangsgröße in (3.39) wird damit

$$y - y_a \approx S_0 \cdot S_1 \cdot S(u_{1a}, z_0) (u - u_a). \quad (3.41)$$

Sie hängt nur noch von der konstanten Empfindlichkeit $S(u_{1a}, z_0)$ im Messanfang des nichtlinearen Gliedes ab. Zur Kompensation des kleinen S_0 muss S_1 entsprechend groß gewählt werden, damit die Bedingung

$$S_0 \cdot S_1 = 1 \quad (3.42)$$

eingehalten wird. Die Messkennlinie wird auf diese Weise linearisiert. Diese Vorgehensweise entspricht einer Linearisierung um einen Arbeitspunkt (hier $u_1 = u_{1a}$). Dabei stellt das Glied mit der Empfindlichkeit $S_0 \ll 1$ si-

cher, dass sich die Eingangsgröße des nichtlinearen Systems u_1 nur wenig vom Arbeitspunkt weg bewegt.

Beispiel 3.2 (Wegmessung mit Dehnungsmessstreifen): Dehnungsmessstreifen (DMS) ermöglichen die Messung von Dehnungen, Stauchungen, Torsionen und anderen mechanischen Beanspruchungen im μm -Bereich. Bei mechanischen Beanspruchungen ändern sich die Länge ℓ und der Querschnitt $A = \pi d^2/4$ eines Leiters (Abb. 3.7) und somit ebenfalls der elektrische Widerstand R :

$$R \propto \frac{\ell}{A}, \quad (3.43)$$

wobei d den Durchmesser des Leiters bezeichnet. Mit Hilfe von einem DMS soll eine **Wegmessung** mit einem Messbereich von 1 mm realisiert werden. Dies kann mit einer Anordnung nach Abb. 3.8 links gelöst werden. Eine Blattfeder formt den für einen DMS zu großen Weg x in eine kleine Dehnung $\varepsilon = \Delta\ell/\ell$ um. Diesen kleinen Ausschlag $\varepsilon \cdot \ell$ wandelt der DMS nahezu linear in ein elektrisches Signal um, das in einem Verstärker auf das gewünschte Niveau gebracht wird.

Insgesamt erfolgt die Linearisierung in drei Schritten (siehe Abb. 3.8 rechts):

- Umsetzung des Weges x in eine kleine Längenänderung $\Delta\ell/\ell$ mittels der Biegefeder: $S_0 \ll 1$.
 - Umsetzung der Längenänderung $\Delta\ell/\ell$ in eine Widerstandsänderung $\Delta R/R$. Dabei wird im linearen Bereich der DMS-Kennlinie Abb. 3.9 gearbeitet. Im Hysteresebereich findet dagegen keine elastische Verformung mehr statt.
 - Umsetzung von $\Delta R/R$ in eine Spannung $\Delta U/U$ mit anschließender Verstärkung: $S_1 \gg 1$.
-

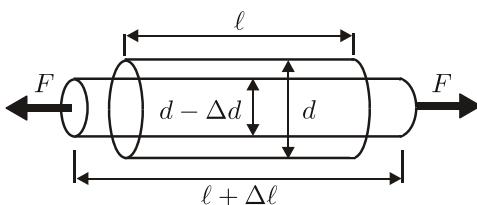


Abbildung 3.7. Dehnungsmessstreifen.

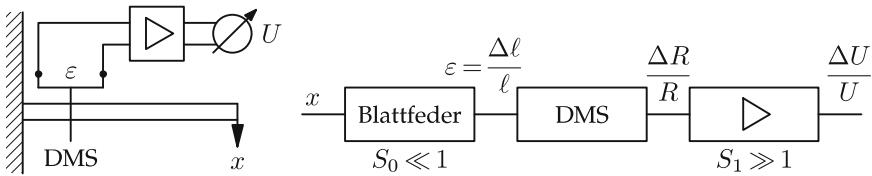


Abbildung 3.8. Beispiel für das Herabsetzen der Messspanne.

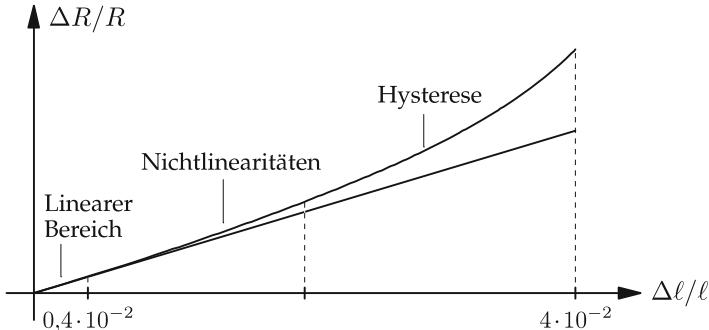


Abbildung 3.9. Kennlinie eines Dehnungsmessstreifens (DMS).

3.2.2 Reihenschaltung zweier nichtlinearer Glieder

Hier werden die Bedingungen untersucht, die eine Reihenschaltung zweier nichtlinearer Glieder mit den Empfindlichkeiten $S_1(u_1)$ und $S_2(u_2)$ in einer Messkette gemäß Abb. 3.10 erfüllen muss, damit eine lineare Gesamtkennlinie resultiert.

Die Kennlinien und deren Empfindlichkeiten sind:

$$u_2 = f_1(u_1), \quad y = f_2(u_2), \quad (3.44)$$

$$\frac{\partial u_2}{\partial u_1} = \frac{\partial f_1(u_1)}{\partial u_1} = S_1(u_1), \quad \frac{\partial y}{\partial u_2} = \frac{\partial f_2(u_2)}{\partial u_2} = S_2(u_2). \quad (3.45)$$

Für die Gesamtkennlinie $y(u_1)$ erhält man die Empfindlichkeit:

$$\frac{\partial y}{\partial u_1} = \frac{\partial y}{\partial u_2} \cdot \frac{\partial u_2}{\partial u_1} = S_1(u_1) S_2(u_2). \quad (3.46)$$

Für eine lineare Gesamtkennlinie muss folgender Zusammenhang gelten:

$$S = S_1(u_1) S_2(u_2) \stackrel{!}{=} \text{const.} \quad (3.47)$$

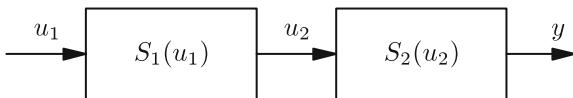


Abbildung 3.10. Reihenschaltung zweier nichtlinearer Glieder.

Formal erhält man die exakte Lösung, wenn die Kennlinie $y = f_2(u_2)$ des zweiten Gliedes der Umkehrfunktion $y = f_1^{-1}(u_2)$ der Kennlinie des ersten Gliedes entspricht. Zur Approximation dieses Verhaltens wird (3.47) in Abhängigkeit von der Eingangsgröße u_1 in eine Taylor-Reihe um den Arbeitspunkt entwickelt. Als Linearitätsbedingung werden die ersten Koeffizienten der Taylor-Reihe zu null gesetzt (konstante Steigung, konstante Krümmung, konstante Krümmungsänderung usw.):

$$\frac{dS}{du_1} = S'_1 S_2 + S_1 S'_2 \frac{du_2}{du_1} = S'_1 S_2 + S_1^2 S'_2 = 0, \quad (3.48)$$

$$\frac{d^2S}{du_1^2} = S''_1 S_2 + 3S_1 S'_1 S'_2 + S_1^3 S''_2 = 0, \quad (3.49)$$

$$\frac{d^3S}{du_1^3} = S'''_1 S_2 + 4S_1 S''_1 S'_2 + 3S_1^2 S'_2 + 6S_1^2 S'_1 S''_2 + S_1^4 S'''_1 = 0. \quad (3.50)$$

Die Gleichungen sind mit steigenden Ansprüchen an die Linearisierung der Reihe nach zu erfüllen. Die Reihenfolge der nichtlinearen Glieder in der Messkette ist dabei für die Linearisierung wesentlich, da die Beziehungen nicht symmetrisch in S_1 und S_2 sind. Anhand von (3.48) erkennt man etwa, dass die Steigung S'_2 das zu S'_1 entgegengesetzte Vorzeichen haben muss:

$$S'_2 = -\frac{S_2}{S_1^2} S'_1. \quad (3.51)$$

Beispiel 3.3 (Widerstandsthermometer in Brückenschaltung): Gegeben sei ein temperaturabhängiger Widerstand mit einer Kennlinie nach Abb. 3.11. Die Kennlinie kann wie folgt beschrieben werden:

$$\Delta R_T = f_1(T) = R_T - R_{T_0} = a(T - T_0) + b(T - T_0)^2. \quad (3.52)$$

Zur Widerstandsmessung wird eine **Brückenschaltung** im Abgleichverfahren verwendet (Abb. 3.12). Hierin ist in Reihe zum Temperaturwider-

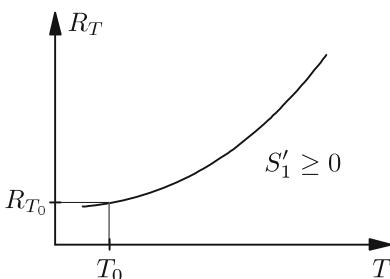


Abbildung 3.11. Kennlinie eines Widerstandsthermometers.

stand R_T ein konstanter, temperaturunabhängiger Justierwiderstand R_i geschaltet, so dass mit

$$R = R_{T_0} + R_i \quad (3.53)$$

die Brücke im Messanfang $\Delta R_T = R_T - R_{T_0} = 0$ auf $U = 0$ abgeglichen ist. Die Brückenausgangsspannung ist

$$U = f_2(\Delta R_T) = U_T - U_R = U_0 \left[\frac{R_T + R_i}{R_T + R_i + R} - \frac{R}{R + R} \right] \quad (3.54)$$

$$= U_0 \left[\frac{R_{T_0} + \Delta R_T + R_i}{R_{T_0} + \Delta R_T + R_i + R} - \frac{1}{2} \right] \quad (3.55)$$

$$= U_0 \left[\frac{\Delta R_T + R}{\Delta R_T + 2R} - \frac{1/2 \Delta R_T + R}{\Delta R_T + 2R} \right] = U_0 \frac{\Delta R_T}{4R} \cdot \frac{1}{1 + \frac{\Delta R_T}{2R}}. \quad (3.56)$$

Beide Kennlinien $\Delta R_T = f_1(T)$ und $U = f_2(\Delta R_T)$ sind nichtlinear. Die Gesamtanordnung soll jedoch ein lineares Verhalten aufweisen. Die Empfindlichkeit S_1 erhält man aus der Widerstandskennlinie $\Delta R_T = f_1(T)$ (3.52) durch Differentiation:

$$S_1(T = T_0) = \frac{d(\Delta R_T)}{dT} \Big|_{T=T_0} = a, \quad (3.57)$$

$$S'_1(T = T_0) = \frac{d^2(\Delta R_T)}{dT^2} \Big|_{T=T_0} = 2b. \quad (3.58)$$

Die Empfindlichkeit S_2 folgt aus der Kennlinie (3.56) der Brückenschaltung $U = f_2(\Delta R_T)$, wobei am Arbeitspunkt $T = T_0$ nach (3.52) $\Delta R_T = 0$

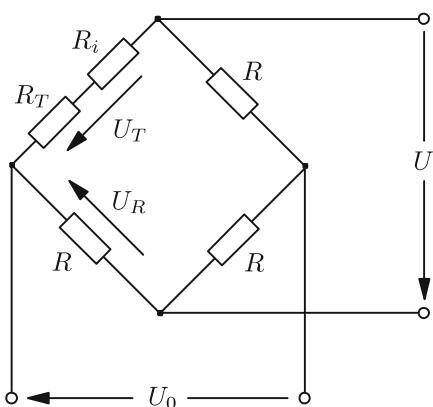


Abbildung 3.12. Brückenschaltung zur Linearisierung.

gilt:

$$S_2(\Delta R_T = 0) = \left. \frac{dU}{d\Delta R_T} \right|_{\Delta R_T=0} = \frac{U_0}{4R}, \quad (3.59)$$

$$S'_2(\Delta R_T = 0) = \left. \frac{d^2U}{d\Delta R_T^2} \right|_{\Delta R_T=0} = -\frac{U_0}{4R^2}. \quad (3.60)$$

Durch Einsetzen von S_1 , S'_1 , S_2 und S'_2 in die Bedingung (3.48) für eine konstante Steigung erhält man:

$$2b \frac{U_0}{4R} - a^2 \frac{U_0}{4R^2} = 0 \quad \Rightarrow \quad R = \frac{a^2}{2b}, \quad b > 0. \quad (3.61)$$

Der Justierwiderstand wird damit

$$R_i = R - R_{T_0} = \frac{a^2}{2b} - R_{T_0}, \quad b > 0. \quad (3.62)$$

Die Beziehung für den Justierwiderstand R_i zeigt, dass bei dieser Schaltung der Koeffizient b *positiv* sein muss und weiter, dass der Quotient $a^2/2b \geq R_{T_0}$ sein muss. Für die weitverbreiteten Platin-Widerstandsthermometer ist beispielsweise b negativ, weshalb eine Linearisierung damit nicht erreicht werden kann.

Nickel-Thermometer: Bei einem Nickel-Thermometer mit einem Widerstand von

$$R_{T_0} = 100 \Omega \quad \text{bei} \quad T_0 = 0^\circ\text{C} \quad (3.63)$$

und einem Messbereich von 100°C sind die beiden Koeffizienten

$$a = 0,574 \Omega/\text{ }^\circ\text{C} \quad \text{und} \quad b = 0,0007 \Omega/(\text{ }^\circ\text{C})^2. \quad (3.64)$$

Der Koeffizient b ist positiv. Der Justierwiderstand wird damit

$$R_i = a^2/2b - R_{T_0} = 135,34 \Omega. \quad (3.65)$$

Im Folgenden sollen die resultierenden Kennlinien des Nickel-Widerstandsthermometers ohne und mit Linearisierung miteinander verglichen werden. Eine Approximation durch Newton-Polynome ergibt:

a) ohne Linearisierung ($R_i = 0$):

$$R = R_i + R_{T_0} = R_{T_0} = 100 \Omega \quad (3.66)$$

$$\rightarrow \quad \frac{U}{U_0} = 1,87 \cdot 10^{-3} \frac{T}{\text{ }^\circ\text{C}} \left(1 - 1,7 \cdot 10^{-2} \frac{T}{\text{ }^\circ\text{C}} \right). \quad (3.67)$$

b) mit Linearisierung ($R_i = 135,34 \Omega$):

$$R = R_i + R_{T_0} = \frac{a^2}{2b} = 235,34 \Omega \quad (3.68)$$

$$\rightarrow \frac{U}{U_0} = 0,61 \cdot 10^{-3} \frac{T}{^\circ\text{C}} \left(1 - 2,0 \cdot 10^{-4} \frac{T}{^\circ\text{C}} \right). \quad (3.69)$$

Die Nichtlinearität wurde damit um zwei Größenordnungen reduziert. Dafür wurde allerdings die Empfindlichkeit auf ein Drittel verringert. ■

3.2.3 Wahl des günstigsten Messbereichs

Bei Messsystemen stehen oft Kennlinien mit einem großen Bereich der Eingangsgröße zur Verfügung. Für den gewünschten kleineren Messbereich ist nun daraus ein möglichst linearer Teil mit hoher Empfindlichkeit auszuwählen. Bei höheren Ansprüchen ist für die Wahl des Arbeitsbereiches der einfache Blick auf die Kennlinie unzureichend. Besser ist es, die *Empfindlichkeit* in Abhängigkeit von der Eingangsgröße aufzutragen – oder, wenn die Kennlinie nicht analytisch gegeben ist, auch die Differenzen erster Ordnung Δy_i nach (2.62) oder die Steigung $\Delta y_i/h$ mit h als dem Stützstellenabstand aufzutragen (Abb. 3.13). Im Intervall $[u_a, u_e]$ sei die Empfindlichkeit S_i der physikalischen Messkennlinie ausreichend hoch und wenig veränderlich. Dieses Intervall der Breite $d = u_e - u_a$ wird daher als Arbeitsbereich des Messsystems gewählt.

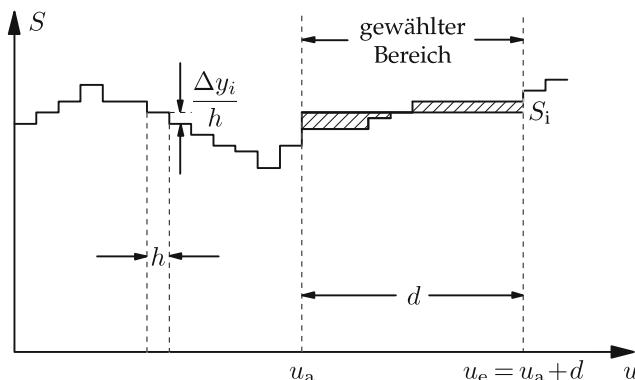


Abbildung 3.13. Wahl des günstigsten Messbereichs.

Entsprechend der Fixpunktjustierung wird die mittlere Empfindlichkeit in diesem Bereich als ideale Empfindlichkeit S_i gewählt; vgl. (3.2). Neben der grafischen Wahl des günstigsten Messbereichs ist auch eine analytische Berechnung möglich. Dazu wird als Gütemaß die quadratische Abweichung

der Empfindlichkeit von der idealen Empfindlichkeit S_i über dem Messbereich integriert:

$$Q = \int_{u_a}^{u_a+d} (S(u) - S_i)^2 du . \quad (3.70)$$

Gesucht ist der Messanfang u_a , der mit Rücksicht auf die Nebenbedingung

$$S_i = \frac{1}{d} \int_{u_a}^{u_a+d} S(u) du \quad (3.71)$$

das Gütemaß Q minimiert. Durch Multiplikation des Binoms kann (3.70) wie folgt geschrieben werden:

$$Q = \int_{u_a}^{u_a+d} S^2(u) du - 2S_i \underbrace{\int_{u_a}^{u_a+d} S(u) du}_{S_i d} + S_i^2 d \quad \text{bzw.} \quad (3.72)$$

$$Q = \int_{u_a}^{u_a+d} S^2(u) du - S_i^2 d . \quad (3.73)$$

Die Differentiation von (3.73) nach u_a ergibt die notwendige Bedingung

$$\frac{\partial Q}{\partial u_a} = S^2(u_a + d) - S^2(u_a) - 2S_i (S(u_a + d) - S(u_a)) \quad (3.74)$$

$$= \underbrace{(S(u_a + d) - S(u_a))}_I \cdot \underbrace{(S(u_a + d) + S(u_a) - 2S_i)}_{II} \stackrel{!}{=} 0 . \quad (3.75)$$

Daraus erhält man zwei Bedingungen, von denen eine zu erfüllen ist.

– **Kriterium I: Arbeitsbereich um einen Wendepunkt**

Die erste Bedingung aus (3.75) lautet:

$$S(u_a + d) - S(u_a) = y'(u_a + d) - y'(u_a) \stackrel{!}{=} 0 . \quad (3.76)$$

Demnach ist der Arbeitspunkt möglichst um einen Wendepunkt der Kennlinie zu legen (Abb. 3.14). Geht man von einem Wendepunkt der Kennlinie im Messbereich aus, so nimmt auf beiden Seiten des Wendepunktes – unabhängig von der Richtung steigender oder fallender Eingangsgröße u – die Steigung $S(u)$ immer entweder zu oder ab. Man findet also meistens zwei Punkte u_a und $u_a + d$, für die sich (3.76) erfüllen lässt. Die diese Endpunkte verbindende lineare Kennlinie konstanter

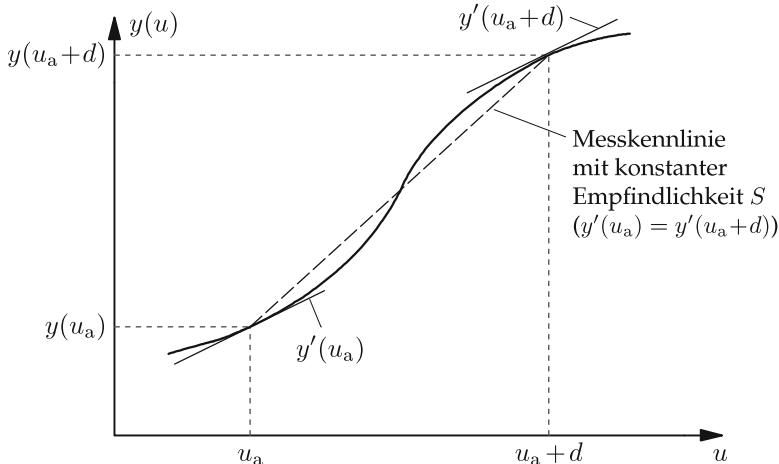


Abbildung 3.14. Wahl des besten Messbereiches nach dem Kriterium I „Arbeitsbereich um einen Wendepunkt“.

Empfindlichkeit verläuft allerdings nicht unbedingt durch den Wendepunkt.

– Kriterium II: Arbeitsbereich ohne Wendepunkt

Die zweite Bedingung aus (3.75) lautet:

$$\frac{1}{2} (S(u_a + d) + S(u_a)) - S_i = \frac{1}{2} (y'(u_a + d) + y'(u_a)) - S_i \stackrel{!}{=} 0. \quad (3.77)$$

Der Arbeitsbereich ist demnach so zu wählen, dass der arithmetische Mittelwert der Steigungen am Anfangspunkt u_a und am Endpunkt $u_a + d$ gerade der mittleren Steigung S_i entspricht (Abb. 3.15). Die lineare Kennlinie, die sich aus der Mittelung beider Steigungen ergibt, verläuft im Allgemeinen nicht durch den Endpunkt, d. h. sie entspricht nicht der idealen Kennlinie, die aus der Fixpunktjustierung resultiert.

Will man neben der Linearität auch eine möglichst hohe Empfindlichkeit erreichen, so kann anstelle des Gütemaßes Q aus (3.70) alternativ das Gütemaß

$$R = \frac{1}{S_i^2 d} \int_{u_a}^{u_a + d} (S(u) - S_i)^2 du = \frac{1}{S_i^2 d} \int_{u_a}^{u_a + d} S^2(u) du - 1 \quad (3.78)$$

herangezogen werden. Die Ableitung von R nach u_a ergibt

$$\underbrace{(S(u_a + d) - S(u_a))}_I \left[S_i d (S(u_a + d) + S(u_a)) - 2 \int_{u_a}^{u_a + d} S^2(u) du \right] = 0. \quad (3.79)$$

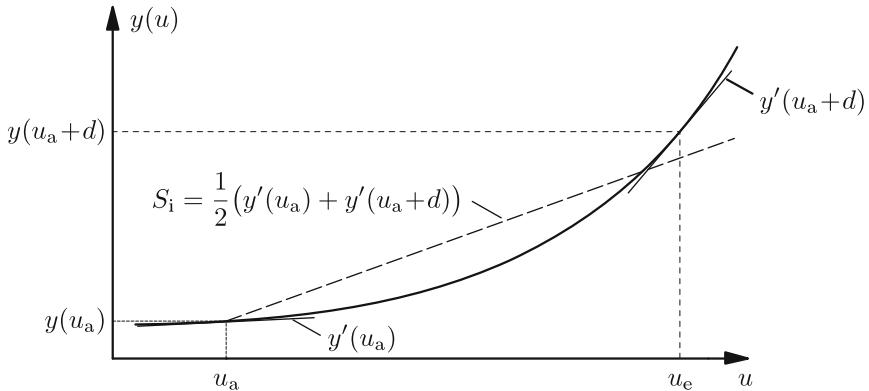


Abbildung 3.15. Wahl des besten Messbereiches nach dem Kriterium II ohne Wendepunkt.

Der erste Faktor entspricht gerade dem ersten Term in (3.76), also dem Kriterium I. Hat die physikalische Kennlinie im ausgewählten Messbereich einen Wendepunkt, so hat auch das Gütemaß R ein Extremum. Durch Abschätzung der Empfindlichkeiten an zwei Messwerten kann man entscheiden, ob es sich um ein Maximum handelt.

Beispiel 3.4 (Kurbeltrieb): Als Beispiel für das analytische Verfahren wird das Problem der Umformung eines Drehwinkels ϕ mit dem Hub $\pi/2$ in eine lineare Hubbewegung mit Hilfe zweier Hebel behandelt (Abb. 3.16). In Abhängigkeit von der Anbringung wählt man einen Teil der Kennlinie als Messbereich aus. Mit den Bezeichnungen des Bildes erhält man für den Ausschlag y :

$$y = r \sin \phi + \sqrt{l^2 - r^2 \cos^2 \phi}. \quad (3.80)$$

Erwünscht ist ein möglichst linearer Zusammenhang zwischen y und ϕ .

Das Verhältnis der Hebellängen r/l sei erheblich kleiner als 1. Die Gleichung vereinfacht sich mit der Näherung $\sqrt{1-x} \approx 1-x/2$ für $x \ll 1$ zu

$$\frac{y}{l} = \frac{r}{l} \sin \phi + 1 - \frac{r^2}{2l^2} \cos^2 \phi. \quad (3.81)$$

Für die Empfindlichkeit der Anordnung gilt

$$S(\phi) = \frac{d(y/l)}{d\phi} = \frac{r}{l} \cos \phi + \frac{r^2}{l^2} \sin \phi \cos \phi. \quad (3.82)$$

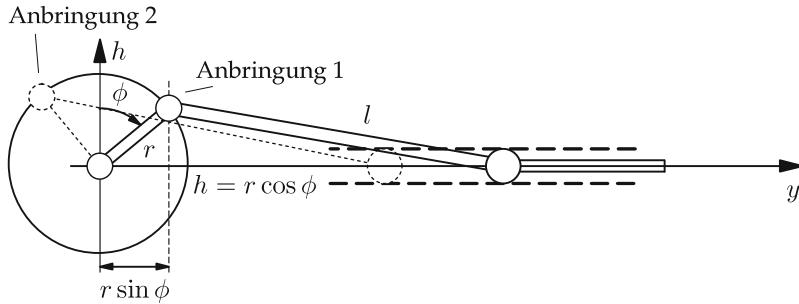


Abbildung 3.16. Umformung eines Drehwinkels ϕ in einen Hub y mit einem Kurbeltrieb.

Im Beispiel gelte

$$u_a = \phi_a, \quad d = \frac{\pi}{2}, \quad \sin\left(\phi_a + \frac{\pi}{2}\right) = \cos \phi_a, \quad (3.83)$$

$$\cos\left(\phi_a + \frac{\pi}{2}\right) = -\sin \phi_a. \quad (3.84)$$

a) Durch Anwendung des **Kriteriums I** (3.76) erhält man:

$$-\frac{r}{l} \sin \phi_a - \frac{r^2}{l^2} \sin \phi_a \cos \phi_a = \frac{r}{l} \cos \phi_a + \frac{r^2}{l^2} \cos \phi_a \sin \phi_a \quad (3.85)$$

$$\Leftrightarrow \quad \frac{2r}{l} = -\left(\frac{1}{\cos \phi_a} + \frac{1}{\sin \phi_a}\right). \quad (3.86)$$

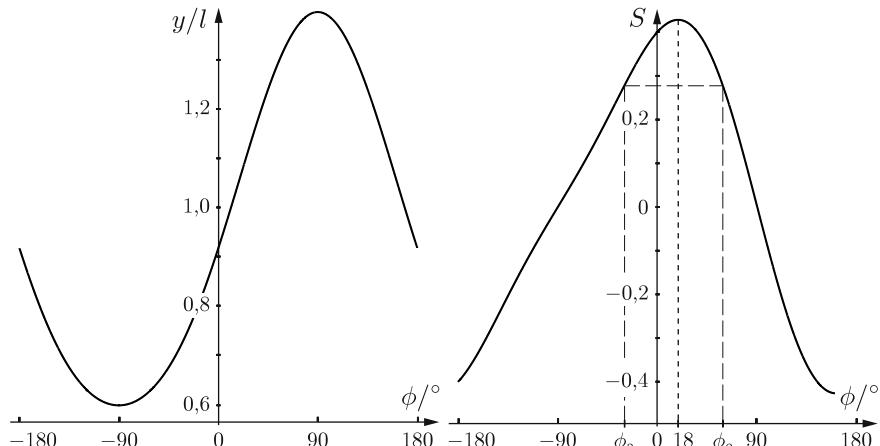
Die Gleichung lässt sich analytisch nicht lösen. Für ein Hebelverhältnis von $r/l = 0,4$ ergibt sich durch numerische oder grafische Lösung ein Messbereich von $\phi_a \approx -30,7^\circ$ bis $\phi_e \approx 59,3^\circ$. Die Kennlinie und die Empfindlichkeit sind in Abb. 3.17 gezeichnet. Man erkennt, dass die Kennlinie einen Wendepunkt bei $\phi \approx 18^\circ$ besitzt, und kann hieraus schließen, dass das Kriterium I tatsächlich anwendbar ist. In Abb. 3.18 ist die ideale Kennlinie eingezeichnet als S_{ia} .

b) Das **Kriterium II** (3.77) ist erfüllt, wenn folgende Gleichung gilt:

$$\frac{r}{2l} (\cos \phi_a - \sin \phi_a) = S_i. \quad (3.87)$$

Die ideale Empfindlichkeit erhält man aus der Forderung nach einer Fixpunktjustierung zu

$$S_i = \frac{y(\phi_a + d) - y(\phi_a)}{d}. \quad (3.88)$$

Abbildung 3.17. Kennlinie und Empfindlichkeit $S(\phi)$ beim Kurbeltrieb.

Durch Einsetzen von (3.88) in (3.87) ergibt sich:

$$\begin{aligned} \frac{r}{2l} (\cos \phi_a - \sin \phi_a) &= \frac{2r}{\pi l} (\cos \phi_a - \sin \phi_a + \frac{r}{2l} (\cos^2 \phi_a - \sin^2 \phi_a)) \\ \Leftrightarrow (\cos \phi_a - \sin \phi_a) \left(\frac{r}{2l} - \frac{2r}{\pi l} - \frac{r^2}{\pi l^2} (\cos \phi_a + \sin \phi_a) \right) &= 0. \quad (3.89) \end{aligned}$$

Somit ist diese Bedingung erfüllt, wenn entweder

$$\cos \phi_a - \sin \phi_a = 0 \quad (3.90)$$

oder

$$\cos \phi_a + \sin \phi_a = \left(\frac{\pi}{4} - 1 \right) \frac{2l}{r} \quad (3.91)$$

gilt. Beide Gleichungen sind wieder nur mit Näherungsverfahren lösbar. Für die erste Gleichung (3.90) ergibt sich als Lösung ein Bereich von $\phi_a \approx -52^\circ$ bis $\phi_e \approx 38^\circ$, der in Abb. 3.18 als S_{ib} eingezeichnet ist. Die Lösung der zweiten Gleichung (3.91) führt mit $\phi_a \approx 45^\circ$ und $\phi_e \approx 135^\circ$ zu einer Maximierung des Gütemaßes Q , weshalb in diesem Fall keine Lösung erhalten wird. Man erkennt in Abb. 3.17, dass ein Wendepunkt (Maximum des Empfindlichkeitsverlaufs) bei $\phi \approx 18^\circ$ vorliegt. Durch Vergleich der beiden Lösungen wird man sich in diesem Fall für das Kriterium I entscheiden, da es den besseren Arbeitsbereich liefert. Abbildung 3.18 lässt ebenfalls erkennen, dass bei der Kennlinie S_{ia} der Linearisierungsfehler Q kleiner als bei der Kennlinie S_{ib} ist. ■

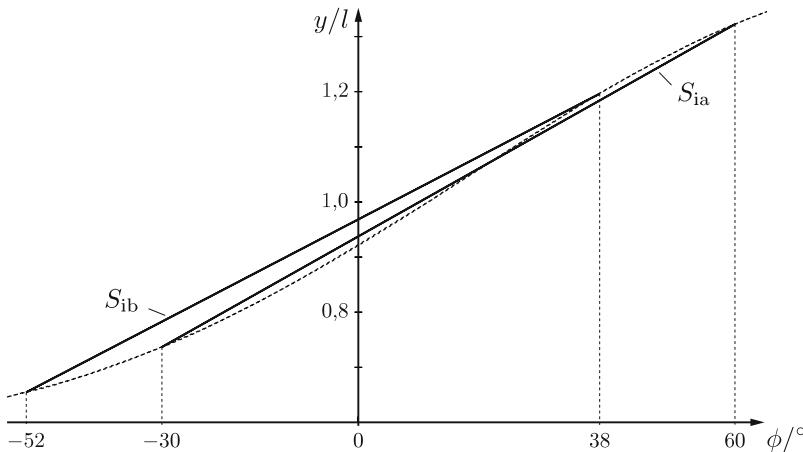


Abbildung 3.18. Ausschnittsvergrößerung von Abb. 3.17 (links) mit den beiden möglichen idealen Kennlinien.

3.2.4 Differenzmethode

Bei der **Differenzmethode** werden zwei gleichartige Teilsysteme mittels einer Parallelschaltung verknüpft. Dadurch werden eine Linearisierung der Kennlinie und eine Steigerung der Empfindlichkeit erzielt. Abbildung 3.19 zeigt das Prinzip. Die Messgröße $u = u_0 + \Delta u$ wird um den Arbeitspunkt u_0 betrachtet. Die Teilsysteme – typischerweise gleichartige Sensoren – sind derart angeordnet, dass die Messgrößenabweichung Δu auf beide gegensinnig wirkt. Dies kann erreicht werden, indem man das Eingangssignal u wie in Abb. 3.19 bearbeitet oder, wie in Bsp. 3.5 gezeigt, indem die Signale $u_0 + \Delta u$ und $u_0 - \Delta u$ direkt abgegriffen werden. Schließlich wird durch die Parallelschaltung die Differenz der Ausgangsgrößen y_1 und y_2 gebildet.

Die Wirkweise der Differenzmethode lässt sich am besten durch Zerlegung einer Funktion $f(\Delta u)$ in einen geraden und einen ungeraden Anteil

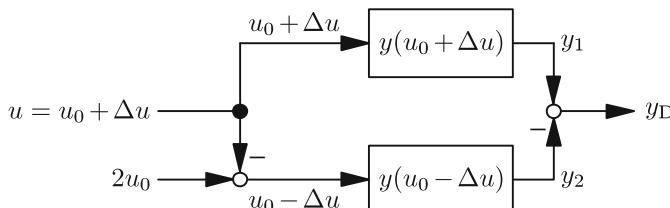


Abbildung 3.19. Parallelschaltung gleichartiger Messsysteme.

$f_g(\Delta u)$ bzw. $f_u(\Delta u)$ darstellen:

$$f(\Delta u) = f_g(\Delta u) + f_u(\Delta u), \quad (3.92)$$

wobei als Argumente lediglich Abweichungen $\Delta u = u - u_0$ vom Arbeitspunkt u_0 betrachtet werden sollen. Mit

$$f_g(-\Delta u) = f_g(\Delta u) \quad \text{und} \quad f_u(-\Delta u) = -f_u(\Delta u) \quad (3.93)$$

erhält man den geraden Funktionsteil als Summe

$$f(\Delta u) + f(-\Delta u) = 2f_g(\Delta u) \quad (3.94)$$

und den ungeraden Funktionsteil als Differenz

$$f(\Delta u) - f(-\Delta u) = 2f_u(\Delta u). \quad (3.95)$$

Der gerade Anteil besteht ausschließlich aus unerwünschten Nichtlinearitäten. Wird der Arbeitspunkt u_0 in der Mitte des Messbereichs gewählt und die Kennlinie durch Differenzbildung als ungerader Funktionsteil gebildet,

$$y_D = f(\Delta u) - f(-\Delta u) = 2f_u(\Delta u), \quad (3.96)$$

so wird der Einfluss der geraden Funktionsteile auf die Differenzkennlinie unterdrückt. Den Linearisierungseffekt erkennt man in Abb. 3.20, wo für eine Funktion $f(\Delta u)$ die Differenzkennlinie $2f_u(\Delta u)$ konstruiert ist.

Entwickelt man die Funktionen y_1 und y_2 nach (3.30) in eine Taylor-Reihe um den Arbeitspunkt u_0 ,

$$y_1 = y(u_0 + \Delta u) \quad (3.97)$$

$$= y(u_0) + S(u_0) \Delta u \left(1 + \frac{S'(u_0)}{S(u_0)} \frac{\Delta u}{2!} + \frac{S''(u_0)}{S(u_0)} \frac{\Delta u^2}{3!} + \dots \right),$$

$$y_2 = y(u_0 - \Delta u) \quad (3.98)$$

$$= y(u_0) + S(u_0) (-\Delta u) \left(1 + \frac{S'(u_0)}{S(u_0)} \frac{-\Delta u}{2!} + \frac{S''(u_0)}{S(u_0)} \frac{(-\Delta u)^2}{3!} + \dots \right),$$

so erhält man mit $y_D = y_1 - y_2$ die Differenzkennlinie

$$y_D = 2S(u_0) \Delta u \left(1 + \frac{S''(u_0) \Delta u^2}{S(u_0) \cdot 3!} + \dots + \frac{S^{(2v)}(u_0) \Delta u^{2v}}{S(u_0) \cdot (2v+1)!} + \dots \right). \quad (3.99)$$

Die Krümmung der Messkennlinie wird durch den Term

$$\frac{1}{2} S'(u_0) (\Delta u)^2 \quad (3.100)$$

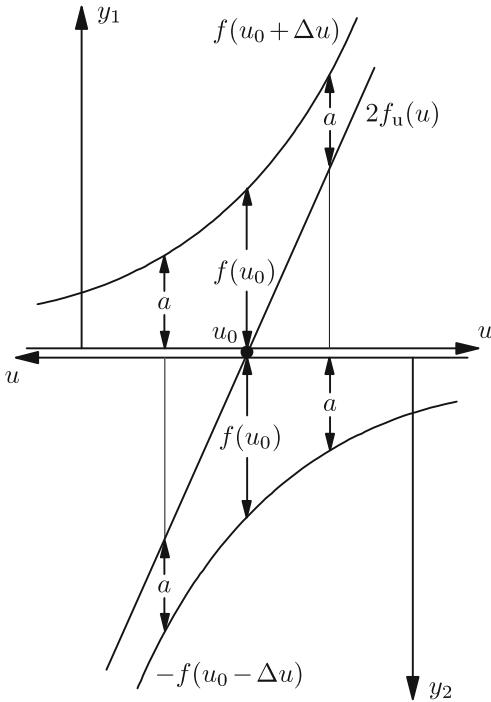


Abbildung 3.20. Grafische Konstruktion der Kennlinie in Differenzschaltung.

der Taylor-Reihe (3.30) repräsentiert. Als gerade Funktion fällt er bei der Differenzbildung heraus – genauso wie alle sonstigen geradzahligen Terme. Dadurch erklärt sich die linearisierende Wirkung der Differenzmethode.

Der relative Fehler der ursprünglichen Kennlinie ist nach (3.34)

$$F_r = \frac{y - y_i}{y_i} \approx \frac{S(u_0, z_0) + \frac{1}{2} S'(u_0, z_0) \Delta u - S_i}{S_i}. \quad (3.101)$$

Die ideale Kennlinie der Differenzanordnung

$$y_{Di} = 2S_i \cdot \Delta u \quad (3.102)$$

hat die doppelte Empfindlichkeit, so dass man entsprechend (3.34) für den relativen Fehler der Differenzkennlinie aus (3.99) folgenden Ausdruck erhält:

$$F_{rD} = \frac{y_D - y_{Di}}{y_{Di}} \approx \frac{2S(u_0, z_0) + \frac{1}{3} S''(u_0, z_0) \Delta u^2 - 2S_i}{2S_i}. \quad (3.103)$$

Die Empfindlichkeit der ursprünglichen Kennlinie im Arbeitspunkt u_0 ist gerade gleich der halben Empfindlichkeit der linearisierten Differenzkenn-

linie:

$$S(u_0, \mathbf{z}_0) = \frac{1}{2} (2S_i) = S_i . \quad (3.104)$$

Der relative Fehler der Differenzkennlinie (3.103) wird damit zu

$$F_{rD} \approx \frac{1}{6} \frac{S''(u_0, \mathbf{z}_0)}{S_i} \Delta u^2 . \quad (3.105)$$

Er ist proportional zur zweiten Ableitung der Empfindlichkeit der ursprünglichen Kennlinie im Arbeitspunkt und wächst quadratisch nach beiden Seiten an. Diesen Betrachtungen liegen Taylor-Entwicklungen zugrunde, die nach dem 3. Term abgebrochen wurden. Diese Näherung ist bei den meisten physikalischen Messkennlinien ausreichend. Im Arbeitspunkt u_0 ist der Kennlinienfehler null – diese Aussage gilt exakt. Das Verhältnis des Fehlers der Differenzkennlinie zu dem der ursprünglichen Kennlinie ist

$$\frac{F_{rD}}{F_r} \approx \frac{1}{3} \frac{S''(u_0, \mathbf{z}_0)}{S'(u_0, \mathbf{z}_0)} \cdot \Delta u . \quad (3.106)$$

Beispiel 3.5 (Kapazitiver Beschleunigungssensor): Mit einer mikromechanischen Technologie wird eine kammartige Struktur geätzt, die beweglich an vier Punkten aufgehängt ist (Abb. 3.21). Die Struktur bildet ein **Feder-Masse-Dämpfer-System**, das durch die Beschleunigung a um Δd ausgelenkt wird:

$$m \Delta \ddot{d} + \beta \Delta \dot{d} + c \Delta d = m a = F_m , \quad (3.107)$$

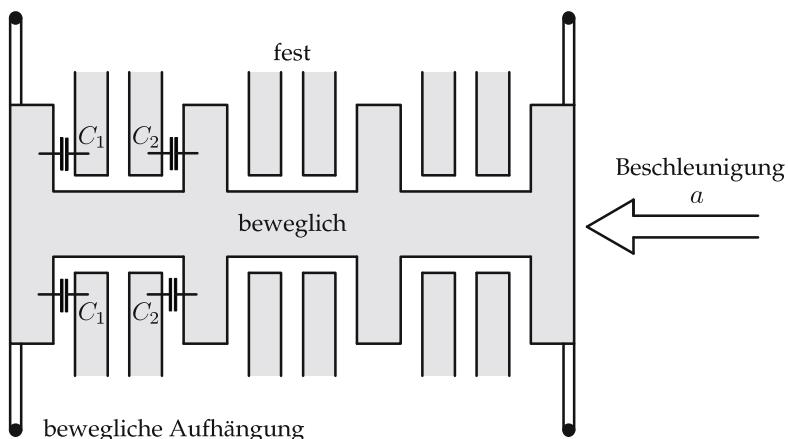


Abbildung 3.21. Kapazitiver Beschleunigungssensor.

wobei m die Masse, β die Dämpfungskonstante und c die Federkonstante bezeichnen. Im stationären Zustand ($\Delta d = \text{const.}$) verschwinden die zeitlichen Ableitungen:

$$F_m = m a = c \Delta d, \quad (3.108)$$

woraus man die Auslenkung

$$\Delta d = \frac{m}{c} a = \frac{1}{\omega_0^2} a \quad (3.109)$$

erhält. Aufgrund der hohen Eigenfrequenz ω_0 sind die Auslenkungen Δd klein. Zwischen den Kämmen liegen jeweils zwei ortsfeste Strukturen, welche gegen den Kamm die Kapazitäten C_1 und C_2 bilden. Bei Auslenkung um Δd nach links wird C_2 größer und gleichzeitig C_1 kleiner:

$$\frac{1}{C_1} = \frac{1}{\varepsilon A} \cdot (d + \Delta d) = \frac{d}{\varepsilon A} \left(1 + \frac{\Delta d}{d} \right) = \frac{1}{C_0} \left(1 + \frac{1}{d\omega_0^2} a \right), \quad (3.110)$$

$$\frac{1}{C_2} = \frac{1}{\varepsilon A} \cdot (d - \Delta d) = \frac{d}{\varepsilon A} \left(1 - \frac{\Delta d}{d} \right) = \frac{1}{C_0} \left(1 - \frac{1}{d\omega_0^2} a \right). \quad (3.111)$$

Die Kapazitätsänderungen an C_1 und C_2 können mittels Spannungsteiler erfasst werden, die von der Wechselspannung u_\sim gespeist werden (Abb. 3.22). Mit Hilfe einer Maschenanalyse ergibt sich

$$u_{1\sim} = \frac{1/C_1}{1/C_1 + 1/C_0} u_\sim = \frac{1 + \Delta d/d}{2 + \Delta d/d} u_\sim, \quad (3.112)$$

$$u_{2\sim} = \frac{1/C_2}{1/C_2 + 1/C_0} u_\sim = \frac{1 - \Delta d/d}{2 - \Delta d/d} u_\sim. \quad (3.113)$$

Daraus kann die Differenzspannung

$$\Delta u_\sim = u_{1\sim} - u_{2\sim} = \frac{2 \frac{\Delta d}{d}}{4 - (\frac{\Delta d}{d})^2} u_\sim \quad (3.114)$$

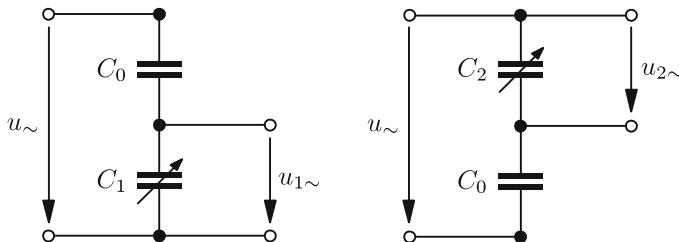


Abbildung 3.22. Schaltbild zur Beschleunigungsmessung.

gebildet werden. Für kleine Auslenkungen $\Delta d/d \ll 1$ gilt die Näherung

$$\Delta u_{\sim} \approx \frac{1}{2} \frac{\Delta d}{d} u_{\sim}. \quad (3.115)$$

Somit besteht näherungsweise ein linearer Zusammenhang zwischen der Differenzspannung Δu_{\sim} und der Auslenkung Δd .

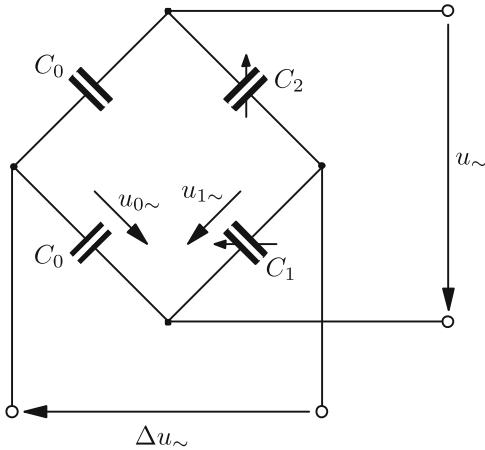


Abbildung 3.23. Brückenschaltung der Messkondensatoren.

Die separaten Spannungsteiler nach Abb. 3.22 mit nachfolgender Differenzbildung sind allerdings für praktische Anwendungen weniger geeignet. Daher verwendet man eine **Wechselstrombrücke** nach Abb. 3.23, in der zwei veränderliche Kapazitäten C_1 und C_2 in Serie geschaltet sind. Mit ihr kommt man zu einem entsprechenden Ergebnis:

$$\Delta u_{\sim} = u_{1\sim} - u_{0\sim} = \frac{C_1^{-1}}{C_1^{-1} + C_2^{-1}} u_{\sim} - \frac{1}{2} u_{\sim} = \frac{1}{2} u_{\sim} \frac{\Delta d}{d}. \quad (3.116)$$

Die Differenzspannung Δu_{\sim} bildet eine lineare Messkennlinie in Abhängigkeit von Δd . Nach einer Gleichrichtung erhält man im stationären Zustand die Ausgangsspannung als

$$|\Delta u_{\sim}| = \frac{1}{2} |u_{\sim}| \frac{\Delta d}{d} = \frac{1}{2} |u_{\sim}| \frac{m}{cd} a = \frac{1}{2} |u_{\sim}| \frac{1}{\omega_0^2 d} a = m S_i a. \quad (3.117)$$

Die Empfindlichkeit S_i beschreibt den Zusammenhang zwischen der Beschleunigungskraft F_m auf die Masse des beweglichen Kammes und der Ausgangsspannung $|\Delta u_{\sim}|$:

$$S_i = \frac{|\Delta u_{\sim}|}{F_m} = \frac{|u_{\sim}|}{2F_m} \frac{\Delta d}{d} = \frac{|u_{\sim}|}{2d} \frac{1}{c}. \quad (3.118)$$

Ist die Kreisfrequenz der Brückenspeisespannung u_{\sim} höher als die Eigenfrequenz ω_0 des Feder-Masse-Dämpfer-Systems, so führt die Spannung u_{\sim} zu keinen nennenswerten Ausschlägen Δd . ■

3.2.5 Gegenkopplung

Die **Gegenkopplung** ist eine wirkungsvolle Methode, um eine nahezu ideale Kennlinie zu erhalten. Ihr Kennzeichen ist der geschlossene Kreis (Abb. 3.24). Bei der Gegenkopplung wird die Messgröße u mit einer vom Ausgangssignal y abgeleiteten Größe $K(y)$ verglichen und ein Abgleich durchgeführt, bis die Differenz v gleich null ist. Man spricht von einem Kompressionsverfahren. Voraussetzung für die Anwendung der Gegenkopplung ist die Existenz eines Übertragungs- bzw. Messgliedes $K(y)$ im Kompressionszweig, das ein Signal in der gleichen physikalischen Größe wie die Messgröße u liefert. Nur dann ist ein Vergleich mit der Messgröße möglich.

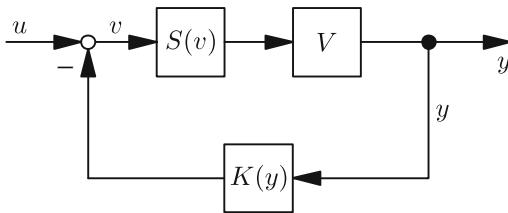


Abbildung 3.24. Strukturbild der Gegenkopplung.

Ist $f(v)$ die physikalische Messkennlinie, so kann die Empfindlichkeit

$$S(v) = \frac{\partial f(v)}{\partial v} \quad (3.119)$$

wegen des kleinen Messbereiches $v_e - v_a$ (kleine Regelabweichung im stationären Zustand) als konstant angenommen werden:

$$S(v_e) = S(v_a) = S_i > 0. \quad (3.120)$$

Aus dem Strukturbild Abb. 3.24 erhält man die beiden Beziehungen

$$y = S_i V \cdot v, \quad (3.121)$$

$$v = u - K(y). \quad (3.122)$$

Daraus resultieren folgende Gleichungen:

$$v = \frac{1}{S_i V} y = u - K(y), \quad (3.123)$$

$$K(y) = u - \frac{1}{S_i V} y. \quad (3.124)$$

Für große Verstärkungen $V \gg 1$ wird die letzte Gleichung zu

$$K(y) \approx u, \quad (3.125)$$

d. h. für monoton steigende Kennlinien $K(y)$ geht die Kennlinie der gesamten Anordnung in die inverse Rückkopplungsfunktion über:

$$y = K^{-1}(u). \quad (3.126)$$

Somit wird die Messkennlinie des geschlossenen Kreises bei ausreichend hoher Verstärkung V ausschließlich vom Rückkopplungszweig bestimmt. Falls eine Auftrennung von $K(y)$ in $K' \cdot y$ möglich ist, ergibt sich mit (3.124)

$$K(y) = K' \cdot y = u - \frac{1}{S_i V} y \quad (3.127)$$

$$\Leftrightarrow K' \left(1 + \frac{1}{K' S_i V} \right) y = u \quad (3.128)$$

und durch Auflösen nach y :

$$y = \frac{u}{K'} \cdot \frac{1}{1 + \frac{1}{K' S_i V}}. \quad (3.129)$$

Mit der Näherung $1/(1+x) \approx 1-x$ für $|x| \ll 1$, was einer nach dem ersten Glied abgebrochenen Potenzreihenentwicklung entspricht, erhält man

$$y \approx \frac{u}{K'} \cdot \left(1 - \frac{1}{K' S_i V} \right). \quad (3.130)$$

Mit der idealen Kennlinie

$$y_i = \frac{1}{K'} \cdot u \quad (3.131)$$

wird der relative Kennlinienfehler der gesamten Gegenkopplungsanordnung zu

$$F_{rG} = \frac{y - y_i}{y_i} \approx -\frac{1}{K' S_i V}. \quad (3.132)$$

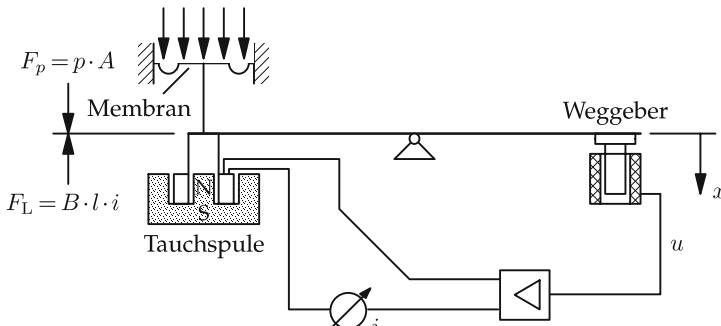
Bei ausreichend hoher Verstärkung V wird ein Linearisierungseffekt der Kennlinie wie im Abschn. 3.2.1 durch Herabsetzen des Messbereichs erreicht. Es muss jedoch darauf hingewiesen werden, dass in vielen Fällen geeignete Gegenkopplungsglieder $K(y)$ fehlen, welche die Ausgangsgröße y auf die Messgröße u abbilden. So ist z. B. kein Übertragungsglied $K(y)$ bekannt, das eine elektrische Ausgangsgröße unmittelbar in eine Tempe-

ratur abbildet und so einen gegengekoppelten Temperatur-Messumformer ermöglicht.

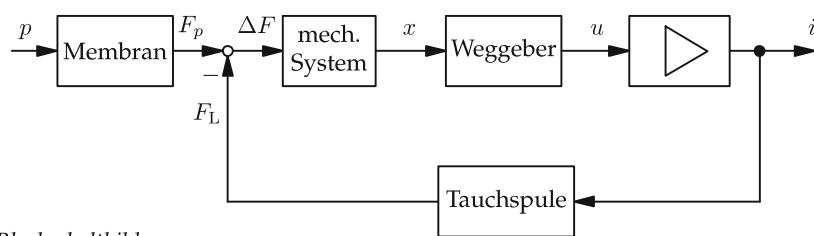
Ein Nachteil der Gegenkopplung ist die durch den Regelkreis hinzukommende Dynamik. Bei zu hoher Verstärkung V kann das System sehr lange Einschwingzeiten besitzen oder sogar instabil werden (Abschn. 5.5.2).

Beispiel 3.6 (Druck-Messumformer): Ein Beispiel für die Gegenkopplung ist der **Druck-Messumformer** in Abb. 3.25. Die Gegenkopplung wird nicht unmittelbar als Druckvergleich, sondern als Kräftevergleich durchgeführt, weil es nur dafür ein geeignetes Übertragungsglied gibt. Für $V \gg 1$ und damit $\Delta F \approx 0$ erhält man aus dem Kräftegleichgewicht $F_p = F_L$ von der Druckkraft $F_p = p A$ und der Lorentz-Kraft $F_L = B l i$ als Ausgangsgröße die Stromstärke i in Abhängigkeit von der Messspannung u :

$$i = \frac{A}{B l} p, \quad (3.133)$$



Physikalisches Ersatzbild des Messformers



Blockschaltbild

Abbildung 3.25. Druck-Messumformer mit Rückkopplung über Kräftevergleich.

wobei A die Fläche der Membran, B die magnetische Flussdichte und l die Leiterlänge bezeichnen. ■

Die Gegenkopplung wird ebenfalls verwendet, um nichtlineare Kennlinien zu realisieren. Umfasst der Messbereich mehrere Größenordnungen, wie z. B. bei Teilchenraten in der Kerntechnik, so verwendet man logarithmische Verstärker. Die Messkennlinie wird dann durch Kompensationsglieder $K(y)$ im Gegenkopplungszweig mit exponentieller Charakteristik realisiert (z. B. Diodenkennlinie).

Beispiel 3.7 (Durchflussmessung): Bei der **Durchflussmessung nach dem Wirkdruckverfahren** ist der Wirkdruck

$$\Delta p = \frac{\rho}{2} \left(\frac{1}{A_2^2} - \frac{1}{A_1^2} \right) q_V^2 \quad (3.134)$$

proportional zum Quadrat des Volumendurchflusses q_V , wobei A_1 und A_2 die Querschnitte vor bzw. nach der Rohrverengung und ρ die Dichte des Fluids bezeichnen; vgl. Abb. 3.26 [8]. Verleiht man der Messkennlinie durch Gegenkopplung insgesamt eine Wurzelcharakteristik, so erhält man gerade den gesuchten Durchfluss:

$$y = \sqrt{\frac{2}{\rho (A_2^{-2} - A_1^{-2})}} \cdot \sqrt{\Delta p} = q_V \quad (3.135)$$

Die Wurzelcharakteristik wird durch eine quadratische Kennlinie

$$K(y) = K \cdot y^2 \quad (3.136)$$

im Rückkopplungszweig erzeugt. ■

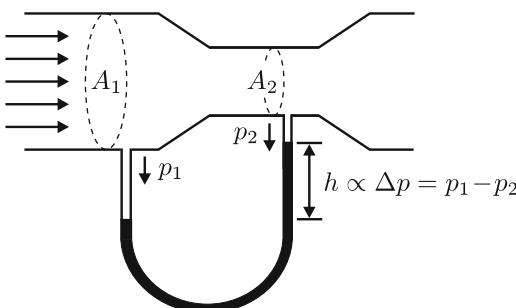


Abbildung 3.26. Durchflussmessung nach dem Wirkdruckverfahren.

Die Messkennlinien des gegengekoppelten Kreises beruhen in diesem Kapitel auf rein stationären Beziehungen. Bei Stabilitätsproblemen muss man auf die Methoden der Regelungstechnik zurückgreifen [11].

3.3 Kennlinienfehler bei Abweichungen von den Normalbedingungen

3.3

Im vorherigen Abschnitt wurden die stationären Eigenschaften der Messkennlinie bei den spezifizierten Normalbedingungen behandelt und Verbesserungen der Kennlinie vorgenommen. Weichen die Betriebsbedingungen von den Normalbedingungen ab, so ändert sich der Störgrößenvektor von \mathbf{z}_0 auf \mathbf{z} , was in der Regel eine Änderung des Kennlinienabgleichs mit sich bringt. Beispiele für sich ändernde Störgrößen sind:

- Temperatur, Temperaturgradient,
- Feuchte,
- mechanische Erschütterungen, Stöße,
- Aussteuerung des Messsystems über den Messbereich hinaus sowie
- Änderung von Hilfsenergien.

Nach ihrer Wirkung auf die Messkennlinie lassen sich **superponierende** (additive) und **deformierende** (multiplikative) **Störgrößen** unterscheiden. Im Folgenden werden Maßnahmen zu deren Unterdrückung behandelt.

3.3.1 Superponierende Störgrößen

Abbildung 3.27 zeigt exemplarisch einen superponierenden Fehler. Er ist dadurch gekennzeichnet, dass er über den gesamten Messbereich konstant ist. Der Nullpunktfehler $e(\mathbf{z})$ ändert sich abhängig vom Störgrößenvektor \mathbf{z} :

$$e(\mathbf{z}) = \underbrace{e(\mathbf{z}_0)}_{=0} + \Delta e(\mathbf{z}). \quad (3.137)$$

Die Empfindlichkeit der Kennlinie ändert sich dagegen bei rein superponierenden Fehlern nicht:

$$\Delta S(u, \mathbf{z}) = 0. \quad (3.138)$$

Die physikalische Messkennlinie wird dann mit $y_{a0} = y(u_a, \mathbf{z}_0)$:

$$y = y_{a0} + \int_{u_a}^u S(u, \mathbf{z}_0) du + \Delta e(\mathbf{z}). \quad (3.139)$$

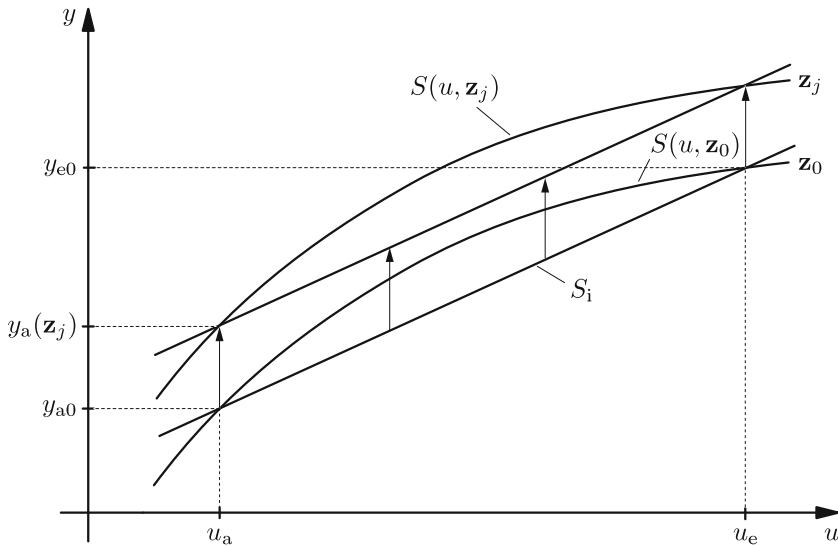


Abbildung 3.27. Superponierender Fehler durch Störgrößenänderung.

Die Änderung des Nullpunktfehlers wird durch eine Taylor-Reihe 1. Ordnung im Störgrößenabgleichpunkt \mathbf{z}_0 abgeschätzt:

$$\Delta e(\mathbf{z}) = y(u_a, \mathbf{z}) - y_{a0} \approx \sum_j \left. \frac{\partial y(u_a, \mathbf{z})}{\partial z_j} \right|_{\mathbf{z}=\mathbf{z}_0} \Delta z_j . \quad (3.140)$$

Der superponierende Fehler bezogen auf die Anzeigespanne ist dann

$$F_{r,sup} = \frac{\Delta e(\mathbf{z})}{y(u, \mathbf{z}_0) - y_{a0}} = \frac{\sum_j \left. \frac{\partial y(u_a, \mathbf{z})}{\partial z_j} \right|_{\mathbf{z}=\mathbf{z}_0} \Delta z_j}{\int_{u_a}^u S(u, \mathbf{z}_0) du} . \quad (3.141)$$

Für eine vereinfachte Abschätzung des Fehlers wird im Folgenden angenommen, dass die Empfindlichkeit im Messbereich näherungsweise konstant sei:

$$S(u, \mathbf{z}_0) \approx S_i , \quad (3.142)$$

$$\left. \frac{\partial y(u_a, \mathbf{z})}{\partial z_j} \right|_{\mathbf{z}=\mathbf{z}_0} \approx \frac{y_a(z_j) - y_{a0}}{\Delta z_j} . \quad (3.143)$$

Damit wird der superponierende Kennlinienfehler (3.141) zu

$$F_{r,sup} \approx \frac{1}{S_i(u - u_a)} \sum_j (y_a(z_j) - y_{a0}). \quad (3.144)$$

Weil $F_{r,sup}$ am Messanfang u_a gegen unendlich strebt, verwendet man auch den superponierenden Fehler bezogen auf den Anzeigebereich:

$$F_{rA,sup} \approx \sum_j \frac{y_a(z_j) - y_{a0}}{S_i(u_e - u_a)} = \sum_j \frac{y_a(z_j) - y_{a0}}{y_{e0} - y_{a0}}. \quad (3.145)$$

Bei gleichzeitigem Vorhandensein von superponierendem und deformierendem Fehler wird der superponierende Fehler am Messanfang betrachtet.

3.3.2 Unterdrückung superponierender Störgrößen mit der Differenzmethode

Zur Eliminierung superponierender Störgrößen wird die Differenzanordnung nach Abb. 3.19 herangezogen. Anhand von (3.96) erkennt man, dass in die Differenzkennlinie nur der ungerade Teil eingeht, wenn gleichartige Teilsysteme parallel geschaltet werden. Superponierende Störgrößen gehören zum geraden Teil der Messkennlinie und werden daher bei der Differenzmethode unterdrückt. Die Kennlinien der Einzelmesssysteme sind allgemein analog zu Abb. 3.19 durch

$$y_1 = y(u_0 + \Delta u) + z_1 \quad \text{und} \quad y_2 = y(u_0 - \Delta u) + z_2 \quad (3.146)$$

gegeben, wobei hier im Gegensatz zu Abschn. 3.2.4 die superponierenden Störungen z_1 bzw. z_2 berücksichtigt wurden. Für die Differenzkennlinie ergibt sich mit (3.99):

$$y_D = y_1 - y_2 \approx 2S(u_0) \Delta u \left(1 + \frac{S''(u_0) \Delta u^2}{S(u_0) \cdot 3!} + \dots \right) + z_1 - z_2. \quad (3.147)$$

Bei Abbruch der Taylor-Reihe nach dem linearen Glied erhält man daraus die Näherung:

$$y_D \approx 2S(u_0) \Delta u \left(1 + \frac{z_1 - z_2}{2S(u_0) \Delta u} \right). \quad (3.148)$$

Der superponierende Kennlinienfehler der Differenzanordnung

$$F_{r,sup} = \frac{z_1 - z_2}{2S(u_0) \Delta u} \quad (3.149)$$

ist bei gleichartig einwirkenden Störgrößen $z_1 = z_2$ gerade null. Dieser Eigenschaft und dem Linearisierungseffekt auf die Kennlinie (Abschn. 3.2.4)

verdankt das Differenzverfahren seine weite Verbreitung in der Messtechnik. Eine superponierende Störgröße z_j (Nullpunktverschiebung) wirkt sich gleichmäßig auf beide Kennlinien aus. Dieser Hub fällt gemäß (3.148) bei der Differenzbildung wieder heraus.

Beispiel 3.8 (Absolutdruckmesser): Beim **Absolutdruckmesser** aus Abb. 3.28 wird der Druck in einer Membrandose gemessen, die eine zur Differenz $p_1 - p_0$ zwischen Mess- und Luftdruck proportionale Kraft erzeugt. Nimmt man an, dass der Luftdruck p_0 konstant bleibt, so kann der durch den Messdruck erzeugte Ausschlag auch auf einer Skala, die in Absolutdruckwerten geteilt ist, angegeben werden. Ändert sich hingegen der Luftdruck p_0 gegenüber den Normalbedingungen, bewirkt diese additive Störgröße Δp_0 einen Fehler, der durch ein zweites Messwerk kompensiert werden kann. Die zweite Dose ist evakuiert ($p = 0$). Sie erfährt eine Messkraft, die allein dem Umgebungsluftdruck $p_0 + \Delta p_0$ entspricht. Im Messgerät kommt insgesamt die Differenz der Messkräfte der Dosen 1 und 2 zur Anzeige. ■

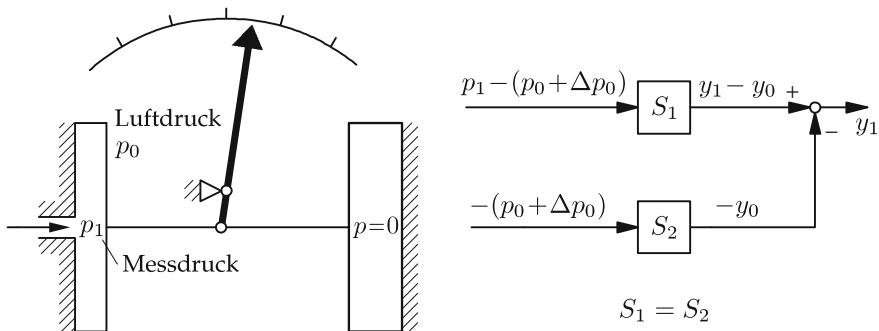


Abbildung 3.28. Absolutdruckmesser als Beispiel für die Differenzmethode.

3.3.3 Deformierende Störgrößen

Die Empfindlichkeit der physikalischen Messkennlinie ändere sich in Abhängigkeit vom Störgrößenvektor \mathbf{z} (Abb. 3.29):

$$S(u, \mathbf{z}) = S(u, \mathbf{z}_0) + \Delta S(u, \mathbf{z}). \quad (3.150)$$

Der Nullpunktfehler ändert sich bei rein deformierenden Fehlern nicht:

$$\Delta e(\mathbf{z}) \stackrel{!}{=} 0. \quad (3.151)$$

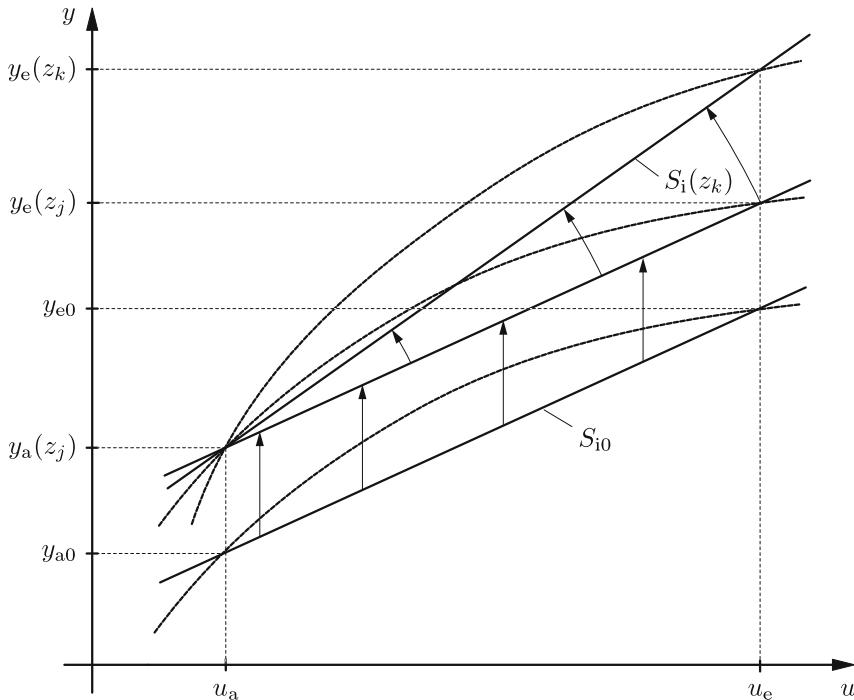


Abbildung 3.29. Einfluss eines deformierenden Fehlers auf die Messkennlinie.

Die physikalische Messkennlinie wird damit

$$y = y_a + \int_{u_a}^u S(u, \mathbf{z}_0) du + \int_{u_a}^u \Delta S(u, \mathbf{z}) du \quad (3.152)$$

$$= y(u, \mathbf{z}_0) + \int_{u_a}^u \Delta S(u, \mathbf{z}) du. \quad (3.153)$$

Die Empfindlichkeitsänderung wird durch eine Taylor-Reihe im Störgrößenabgleichpunkt \mathbf{z}_0 abgeschätzt:

$$\Delta S(u, \mathbf{z}) \approx \sum_k \frac{\partial S(u, \mathbf{z}_0)}{\partial z_k} \Delta z_k. \quad (3.154)$$

Der absolute Kennlinienfehler ergibt sich durch Einsetzen dieser Näherung in (3.153) zu

$$\Delta y = y - y(u, \mathbf{z}_0) = \sum_k \int_{u_a}^u \frac{\partial S(u, \mathbf{z}_0)}{\partial z_k} \Delta z_k du. \quad (3.155)$$

Der relative deformierende Kennlinienfehler wird damit

$$F_{r,\text{def}} = \frac{y - y(u, \mathbf{z}_0)}{y(u, \mathbf{z}_0) - y_a} = \frac{\sum_k \int_{u_a}^u \frac{\partial S(u, \mathbf{z}_0)}{\partial z_k} \Delta z_k \, du}{\int_{u_a}^u S(u, \mathbf{z}_0) \, du}. \quad (3.156)$$

Für eine vereinfachte Abschätzung des Fehlers wird im Folgenden angenommen, dass zur Berechnung der Anzeigespanne im Nenner von (3.156) mit konstanten Empfindlichkeiten im Messbereich gerechnet werden kann:

$$S(u, \mathbf{z}_0) \approx S_{i0}. \quad (3.157)$$

Des Weiteren seien die Empfindlichkeitsänderungen durch die Störgrößen näherungsweise

$$\frac{\partial S_i(u, \mathbf{z}_0)}{\partial z_k} \approx \frac{S_i(z_k) - S_{i0}}{\Delta z_k} = \frac{\Delta S_i(z_k)}{\Delta z_k}. \quad (3.158)$$

Damit erhält man für den deformierenden Kennlinienfehler (3.156):

$$F_{r,\text{def}} = \frac{1}{S_{i0}} \sum_k \Delta S_i(z_k). \quad (3.159)$$

Mit diesen Abschätzungen kann nun der Einfluss einer einzelnen deformierenden Störgröße z_k auf die Steigung der Messkennlinie angegeben werden. Dabei wird für den Anzeigewert nicht der Abgleichpunkt \mathbf{z}_0 herangezogen, sondern der Wert, der sich bei Berücksichtigung einer *superponierenden* Störung z_j ergibt. Mit dem Anfangswert $y_a(z_j)$ erhält man den Kennlinienfehler

$$F_{r,\text{def}} = \frac{S_i(z_k) - S_{i0}}{S_{i0}} = \frac{\Delta S_i(z_k)}{S_{i0}} \quad (3.160)$$

$$= \frac{[y_e(z_k) - y_a(z_j)] - [y_e(z_j) - y_a(z_j)]}{y_e(z_j) - y_a(z_j)} \quad (3.161)$$

$$= \frac{[y_e(z_k) - y_a(z_j)] - (y_{e0} - y_{a0})}{y_{e0} - y_{a0}}, \quad (3.162)$$

wobei die Identität $y_e(z_j) - y_a(z_j) = y_{e0} - y_{a0}$ durch die superponierende Natur des Fehlers z_j zu erklären ist.

Wirken je ein superponierender und ein deformierender Fehler gleichzeitig auf die Messkennlinie ein, so können beide Fehler überlagert werden.

Die Taylor-Reihe im Störgrößenabgleichpunkt \mathbf{z}_0 lautet dann allgemein

$$y = y(u, \mathbf{z}_0) + \frac{\partial y(u, \mathbf{z}_0)}{\partial z_j} \Delta z_j + \frac{\partial y(u, \mathbf{z}_0)}{\partial z_k} \Delta z_k \quad (3.163)$$

und für lineare Kennlinien mit (3.153) und (3.162)

$$y = y_{a0} + S_{i0} (u - u_a) + \Delta e(z_j) + \frac{\Delta S_i(z_k)}{S_{i0}} S_{i0} (u - u_a) \quad (3.164)$$

$$\begin{aligned} &= \underbrace{y_{a0} + S_{i0} (u - u_a)}_{y(u, \mathbf{z}_0)} + [y_a(z_j) - y_{a0}] \\ &\quad + \frac{[y_e(z_k) - y_a(z_j)] - (y_{e0} - y_{a0})}{y_{e0} - y_{a0}} S_{i0} (u - u_a). \end{aligned} \quad (3.165)$$

Der relative Kennlinienfehler bezogen auf die Anzeigespanne ist damit

$$F_r = \frac{y - y(u, \mathbf{z}_0)}{y(u, \mathbf{z}_0) - y_{a0}} \quad (3.166)$$

$$= \frac{y_a(z_j) - y_{a0}}{S_{i0} (u - u_a)} + \frac{y_e(z_k) - y_a(z_j) - (y_{e0} - y_{a0})}{y_{e0} - y_{a0}}. \quad (3.167)$$

Im Messanfang geht dieser Fehler wieder gegen unendlich. Der relative Fehler bezogen auf den Anzeigebereich ist

$$F_{rA} = \frac{y - y(u, \mathbf{z}_0)}{y_{e0} - y_{a0}} = \frac{y_e(z_k) - y_{e0}}{y_{e0} - y_{a0}}. \quad (3.168)$$

3.3.4 Deformierende Störgrößen bei Gegenkopplung

Im Folgenden soll der Einfluss einer deformierenden Störgröße z_k auf die Steigung S_i der Messkennlinie durch eine Gegenkopplung reduziert werden. Die Messkennlinie des gegengekoppelten Messsystems Abb. 3.24 ist nach (3.129)

$$y = \frac{u}{K'} \frac{1}{1 + \frac{1}{K' S_i V}}. \quad (3.169)$$

Die vorhin als konstant angenommene Empfindlichkeit S_i des Messgliedes ohne Rückführung ändere sich im Messbereich aufgrund einer deformierenden Störgröße z_k :

$$S_i(z_k) = S_{i0} + \Delta S_i(z_k). \quad (3.170)$$

Durch Differentiation der Messkennlinie (3.129) nach S_i mittels der Quotientenregel folgt die Steigung der Kennlinie bedingt durch Änderungen von

S_i :

$$\frac{\partial y}{\partial S_i} = u \frac{V}{(1 + K' \cdot S_i V)^2}. \quad (3.171)$$

Damit wird, zur Abschätzung des Kennlinienfehlers, die Ausgangsgröße y in eine Taylor-Reihe um S_{i0} herum entwickelt:

$$\begin{aligned} y &\approx y(S_{i0}) + \left. \frac{\partial y}{\partial S_i} \right|_{S_{i0}} \Delta S_i(z_k) + \dots \\ &\approx u \frac{S_{i0} V}{1 + K' \cdot S_{i0} V} \left(1 + \frac{1}{1 + K' \cdot S_{i0} V} \cdot \frac{\Delta S_i(z_k)}{S_{i0}} \right). \end{aligned} \quad (3.172)$$

Der relative Fehler bei Gegenkopplung ist dann:

$$F_{rG,def} = \frac{y - y(S_{i0})}{y(S_{i0})} = \frac{1}{1 + K' \cdot S_{i0} V} \cdot \frac{\Delta S_i(z_k)}{S_{i0}}. \quad (3.173)$$

Bildet man das Verhältnis des relativen Fehlers des gegengekoppelten Systems (3.173) zum nicht gegengekoppelten Messsystem in (3.160), so erhält man

$$\frac{F_{rG,def}}{F_{r,def}} = \frac{1}{1 + K' S_{i0} V}. \quad (3.174)$$

Durch die Gegenkopplung wird der relative Kennlinienfehler aufgrund einer deformierenden Störgröße z_k um die Verstärkung V des offenen Kreises reduziert.

Beispiel 3.9 (Kapazitiver Beschleunigungsmesser): Auf das Beispiel 3.5 des kapazitiven Beschleunigungsmessers in Abschn. 3.2.4 soll eine Gegenkopplung angewendet werden. Die Empfindlichkeit

$$S_{i0} = \frac{|u_{\sim}|}{2dc} \quad (3.175)$$

hängt von der Wechselspannung $|u_{\sim}|$ ab, mit der die kapazitive Brücke gespeist wird. Änderungen von $|u_{\sim}|$ gehen unmittelbar als deformierende Störgröße in die Ausgangsspannung $|\Delta u_{\sim}|$ ein. Durch eine Gegenkopplung gemäß Abb. 3.30 mit dem nichtlinearen Übertragungsglied

$$u_C = \sqrt{V |\Delta u_{\sim}|} = \sqrt{V \frac{1}{2} |u_{\sim}| \frac{\Delta d}{d}} \quad (3.176)$$

soll der Einfluss der Spannung $|u_{\sim}|$ auf die Ausgangsgröße $|\Delta u_{\sim}|$ verringert werden. Die resultierende Gesamtübertragungsfunktion K' der Rückführung soll linear sein und bewirken, dass durch Rückkopplung

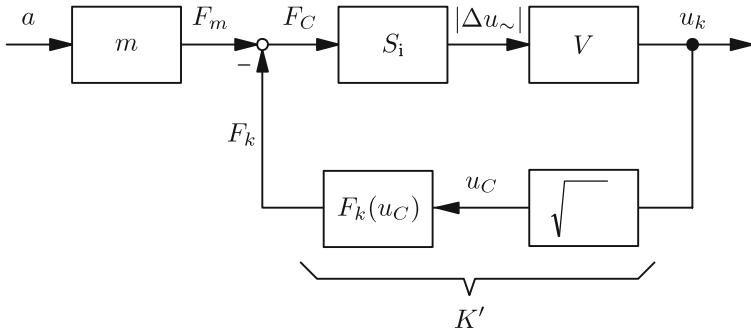


Abbildung 3.30. Kapazitiver Beschleunigungssensor mit Gegenkopplung.

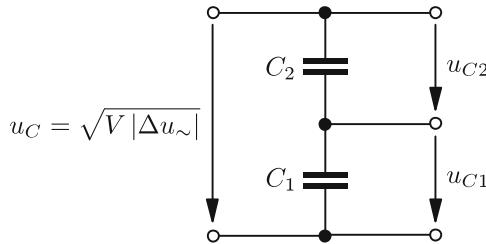


Abbildung 3.31. Spannungsrückkopplung.

der verstärkten Ausgangsspannung u_k auf die Kammstruktur eine kompensierende Anziehungs Kraft F_k ausgeübt wird, die diese in die Nulllage zurückholt und auf diese Weise die Beschleunigungskraft F_m ausgleicht. Dafür wird die zurückgeführte Spannung u_C gemäß Abb. 3.31 auf beide Kondensatoren C_1 und C_2 verteilt:

$$u_{C1} = \frac{1}{2} \left(1 + \frac{\Delta d}{d} \right) u_C = \frac{1}{2} \left(1 + \frac{\Delta d}{d} \right) \sqrt{V \frac{1}{2} \frac{\Delta d}{d} |u_\sim|}, \quad (3.177)$$

$$u_{C2} = \frac{1}{2} \left(1 - \frac{\Delta d}{d} \right) u_C = \frac{1}{2} \left(1 - \frac{\Delta d}{d} \right) \sqrt{V \frac{1}{2} \frac{\Delta d}{d} |u_\sim|}. \quad (3.178)$$

Die Energie auf den Kondensatoren

$$E_{1,2} = \frac{1}{2} C_{1,2} u_{C1,2}^2 = \frac{C_0 V}{16} |u_\sim| \left(1 \pm \frac{\Delta d}{d} \right) \frac{\Delta d}{d} \quad (3.179)$$

wird nach Δd differenziert, was zu den folgenden Kräften führt:

$$F_{k1,2} = \frac{C_0 V}{8} |u_\sim| \left(\frac{1}{2d} \pm \frac{\Delta d}{d^2} \right). \quad (3.180)$$

Deren Differenz F_k zieht die bewegten Kondensatorplatten in Richtung der Nulllage zurück:

$$F_k = F_{k1} - F_{k2} = \frac{C_0 V}{4d} |u_{\sim}| \frac{\Delta d}{d} \quad (3.181)$$

$$= \frac{C_0 V}{2d} |\Delta u_{\sim}| = \frac{C_0}{2d} u_k = K' \cdot u_k. \quad (3.182)$$

Die Gesamtübertragungsfunktion der Rückführung entspricht also einem Proportionalglied mit dem Verstärkungsfaktor

$$K' = \frac{C_0}{2d}. \quad (3.183)$$

Bildet man das Kräftegleichgewicht der rückgekoppelten Anordnung, so ergibt sich

$$F_m = F_C + F_k, \quad (3.184)$$

$$m a = c \Delta d + \frac{C_0}{2d} u_k \quad (3.185)$$

$$= \frac{c 2d}{|u_{\sim}|} \frac{u_k}{V} + \frac{C_0}{2d} u_k. \quad (3.186)$$

Bei einer großen Verstärkung V und somit einer Auslenkung von $\Delta d \approx 0$ ist $F_m \approx F_k$, wodurch die Differenzkraft F_C zu null wird. Die Empfindlichkeit des gegengekoppelten Beschleunigungssensors ist gerade gleich der inversen Übertragungskonstante der Rückführung:

$$u_k = \frac{1}{K'} F_m = \frac{2 m d}{C_0} a. \quad (3.187)$$

Die Schwankungen der Speisewechselspannung $|u_{\sim}|$ und der Federsteifigkeit c gehen nur noch um den Verstärkungsfaktor V reduziert in das Messergebnis ein. Die Masse des bewegten Kammes m , der Abstand in der Ruhelage d und die Kapazität C_0 können mit den heute verfügbaren lithographischen Prozessen sehr präzise eingehalten werden. Beim gegengekoppelten Beschleunigungssensor kann auf eine übermäßig genaue Stabilisierung der Speisespannung $|u_{\sim}|$ verzichtet werden. ■

3.3.5 Superponierende Störgrößen bei Gegenkopplung

Dem Messsystem seien additive Störungen z überlagert. Es soll nun untersucht werden, ob derartige Störeinflüsse durch eine Gegenkopplung reduziert werden können (Abb. 3.32).

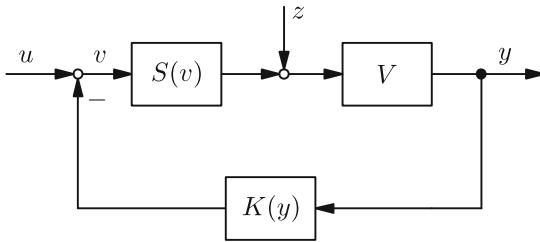


Abbildung 3.32. Superponierende Störgrößen bei Gegenkopplung.

Bei großer Verstärkung $V \gg 1$ ist die Kennlinie des gegengekoppelten Systems nach (3.126)

$$y = K^{-1} \left(u + \frac{z}{S_i} \right). \quad (3.188)$$

Superponierende Störgrößen lassen sich demnach *nicht* grundsätzlich durch eine Gegenkopplung eliminieren. Sie bleiben im Wesentlichen erhalten.

3.3.6 Kompensation systematischer Störeinflüsse

Das im Folgenden vorgestellte Verfahren kompensiert den Einfluss einer einzigen dominanten Störgröße z . Dazu muss der Einfluss dieser Störgröße auf die Messgröße y a priori bekannt sein.

Eine Möglichkeit zur Korrektur der Ausgangsgröße y in Abhängigkeit von der Störgröße z ist die in Abschn. 2.3 behandelte Kennfeldinterpolation. Während bislang allein die physikalische Kennlinie $y = f(u)$ Grundlage der Messkennlinie war, wird nun die physikalische Abhängigkeit der Ausgangsgröße

$$y = f(u, z) \quad (3.189)$$

vom Messsignal u und *zusätzlich* von der dominanten Störgröße z gleichzeitig vorab vermessen und in einem Kennfeld abgespeichert. Bei der tatsächlichen Messung wird die Störgröße z neben der eigentlichen Messgröße u erfasst und im gespeicherten Kennfeld die aktuelle Ausgangsgröße nach (2.120) interpoliert:

$$\begin{aligned} y(u, z) \approx & y(u_i, z_j) + \frac{\Delta y(u_i)}{\Delta u_i} \Delta u + \frac{\Delta y(z_j)}{\Delta z_j} \Delta z \\ & + \frac{\Delta^2 y(u_i, z_j)}{\Delta u_i \Delta z_j} \Delta u \Delta z. \end{aligned} \quad (3.190)$$

Das Verfahren entspricht einer fortlaufenden Adaption der Ausgangsgröße an die sich ändernde dominante Störgröße. Es können reproduzierbare

Messergebnisse hoher Genauigkeit erzielt werden, wenn der Einfluss der Störgröße konstant bleibt, wenn es sich also um eine systematische Störgröße handelt.

3.3.7 Abschirmung

Abschirmung zielt in der Messtechnik darauf ab, Störgrößen vom Messsystem fernzuhalten. Beispiele sind etwa die Thermostatisierung (d. h. Temperaturregelung) von Baugruppen, um Temperaturfehler zu unterdrücken, oder eine luftdichte Verpackung, um einen Einfluss der Luftfeuchte zu vermeiden. Auch elektrische und magnetische Abschirmungen gehören hierzu. Diese Maßnahme wird zwar oft eingesetzt, sie bringt jedoch häufig einen beträchtlichen Mehraufwand mit sich.

3.3.8 Superponierende Störgrößen in Messketten

Das Signal $u(t)$ wird in einer Messkette in einer Richtung übertragen. An den Schnittstellen seien jeweils Störungen z_j überlagert (Abb. 3.33). Die Kennlinie des j -ten Gliedes sei durch

$$y_j = S_j y_{j-1} + z_j \quad (3.191)$$

gegeben. Die Kennlinie der gesamten Anordnung erhält man, beginnend mit $y_0 = u$, durch wiederholte Elimination der y_j :

$$y_n = (S_n S_{n-1} \cdots S_1) u + (S_n S_{n-1} \cdots S_2) z_1 + \cdots + S_n z_{n-1} + z_n .$$

Im *absoluten Fehler* F des Ausgangssignals summieren sich die Störungen z_j der einzelnen Glieder jeweils multipliziert mit der Verstärkung vom $(j+1)$ -ten bis zum letzten Glied der Kette:

$$F = \sum_{j=1}^{n-1} \left(\prod_{k=j}^{n-1} S_{k+1} \right) z_j + z_n . \quad (3.192)$$

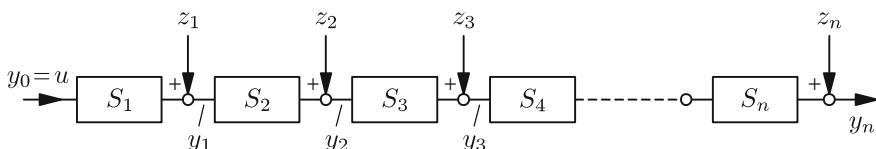


Abbildung 3.33. Superponierende Fehler in einer Messkette.

Der Beitrag eines Übertragungsgliedes zum Fehler ist somit abhängig von seiner Position in der Kette! Mit der idealen Kennlinie

$$y = u \cdot \prod_{j=1}^n S_j \quad (3.193)$$

erhält man für den relativen Fehler bezogen auf den Anzegebereich $y(u = d)$:

$$F_{rA} = \frac{F}{y} = \frac{1}{d} \sum_{j=1}^n \frac{z_j}{\prod_{k=1}^j S_k} = \frac{z_1}{S_1 d} + \frac{z_2}{S_1 S_2 d} + \dots + \frac{z_n}{S_1 S_2 \dots S_n d} \quad (3.194)$$

$$= \sum_{j=1}^n F_{rAj}. \quad (3.195)$$

Der *relative Fehler* F_{rA} einer Messkette infolge superponierender Störgrößen setzt sich nach (3.195) additiv aus den relativen Fehlern der einzelnen Glieder zusammen.

Bei einem mehrstufigen Verstärker ist immer $S_j > 1$. Der superponierende Fehler wird wirksam herabgesetzt, wenn die erste Stufe eine möglichst hohe Verstärkung $S_1 \gg S_j, j > 1$, aufweist. Oft lässt sich allerdings die Verstärkung der ersten Stufe nicht beliebig erhöhen. Dann ist darauf zu achten, dass in der ersten Stufe hochwertige Verstärker zum Einsatz kommen, die nur minimal von superponierenden Störgrößen (z. B. Drift) beeinflusst werden. Diese Regel ist beim Entwurf von Gleichspannungsverstärkern zu beachten. In der 1. Stufe wird oftmals ein großer Aufwand getrieben, um den Einfluss superponierender Störgrößen herabzusetzen, indem etwa durch eine Differenzschaltung (Abschn. 3.2.4), durch Thermostatisierung (Abschn. 3.3.7) oder durch einen „Zerhacker“ (Abschn. 3.3.9) der Fehler der ersten Stufe herabgesetzt wird.

Beispiel 3.10 (Zweistufiger Verstärker): Zum Bau eines zweistufigen Spannungsverstärkers stehen zwei Stufen mit gleicher Verstärkung $S_a = S_b = 30$ zur Verfügung. Bei der Stufe a ist mit einer Nullpunkt drift von $0,5 \text{ mV}$ zu rechnen, bei der Stufe b mit einer von 1 mV . Mit (3.192) erhält man für die Anordnung a–b einen Fehler von $F = (30 \cdot 0,5 + 1) \text{ mV} = 16 \text{ mV}$, bei der Anordnung b–a einen Fehler von $F = (30 \cdot 1 + 0,5) \text{ mV} = 30,5 \text{ mV}$. ■

3.3.9 Zerhackerverstärker

Der **Zerhackerverstärker** (engl. *chopper amplifier*) ist ein Modulationsverstärker zur driftfreien Verstärkung kleiner Gleichspannungen. Dadurch lässt sich insbesondere der Einfluss temperaturabhängiger Offsetspannungen am Verstärkereingang, wie sie bei realen Operationsverstärkern vorkommen, unterdrücken. Zu diesem Zweck wird das niederfrequente Eingangssignal u durch **Modulation** zu höheren Frequenzen hin verschoben.

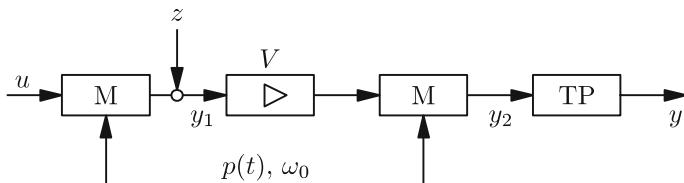


Abbildung 3.34. Synchrone Zerhackung und Gleichrichtung.

Abbildung 3.34 zeigt die Wirkungskette des Zerhackerverstärkers. Die Modulatoren M werden zeitgleich angesteuert: sie arbeiten synchron zueinander. Die erste Modulation erfolgt durch Multiplikation des Eingangssignals mit einem mittelwertfreien Rechtecksignal $p(t)$ (Abb. 3.35). Die Störung $z(t)$ (z. B. der Offset des Operationsverstärkers) wird dabei nicht moduliert:

$$y_1(t) = u(t) \cdot p(t) + z(t). \quad (3.196)$$

Die Umschaltfunktion $p(t)$ habe die Frequenz $\omega_0 = 2\pi/T$ (Abb. 3.35).

Die nochmalige Modulation des verstärkten Signals y_1 mit der Umschaltfunktion $p(t)$ gemäß Abb. 3.34 ergibt mit (3.196)

$$y_2(t) = V \cdot y_1(t) \cdot p(t) \quad (3.197)$$

$$= V \underbrace{u(t) \cdot p^2(t)}_{=1} + V \cdot z(t) p(t) = V \cdot u(t) + V \cdot z(t) p(t). \quad (3.198)$$

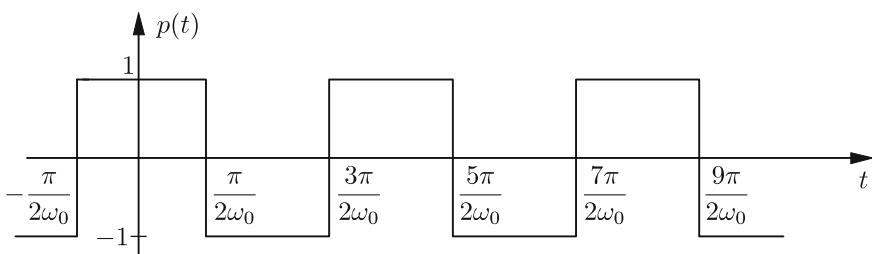


Abbildung 3.35. Umschaltfunktion $p(t)$.

Das Eingangssignal $u(t)$ wird unverzerrt übertragen, während die Störung $z(t)$ mit der mittelwertfreien Umschaltfunktion $p(t)$ moduliert bleibt. Die Störung kann deshalb mit einem einfachen Tiefpass weggefiltert werden, auch wenn der Erwartungswert $E\{z(t)\} \neq 0$ ist, d.h. wenn die Störung einen konstanten superponierenden Anteil enthält. Die Umschaltfunktion $p(t)$ erzeugt kein eigenes superponierendes Störsignal, da über einer Periode das Integral verschwindet:

$$\int_0^T p(t) dt = 0. \quad (3.199)$$

Der Übersichtlichkeit halber soll im Folgenden nur mit der Grundschwingung der Umschaltfunktion $p(t)$ gerechnet werden. Die höherfrequenten Anteile der Fourier-Reihe von $p(t)$ und der Faktor $4/\pi$ werden hingegen vernachlässigt:

$$p(t) \approx \cos(\omega_0 t). \quad (3.200)$$

Modulation: Die Multiplikation des Eingangssignals $u(t)$ mit der Umschaltfunktion $p(t)$ entspricht einer Faltung von $U(\omega)$ mit $P(\omega)$ im Frequenzbereich. Wegen des harmonischen Charakters von (3.200) erscheint das Spektrum $U(\omega)$ des Eingangssignals an den Frequenzen der Modulationsschwingung $\pm\omega_0$. Die überlagerte Störung $z(t)$ greift am Eingang des Verstärkers an (Abb. 3.36):

$$y_1(t) = u(t) \cdot \cos(\omega_0 t) + z(t) \quad (3.201)$$

$$Y_1(\omega) = \frac{1}{2} U(\omega - \omega_0) + \frac{1}{2} U(\omega + \omega_0) + Z(\omega). \quad (3.202)$$

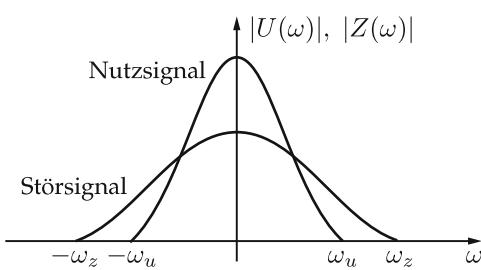


Abbildung 3.36. Bandbreite von Nutzsignal und Störsignal.

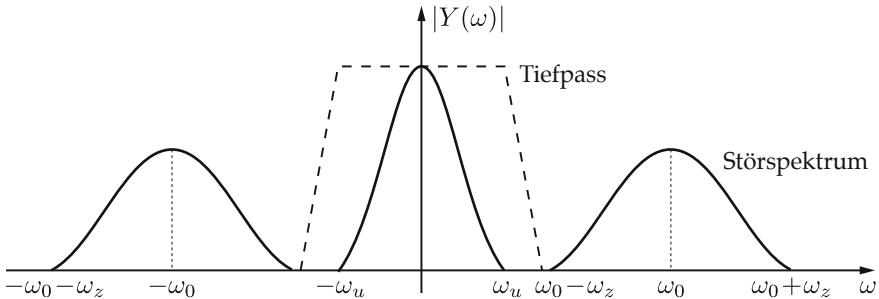


Abbildung 3.37. Nutzsignal- und Störspektrum beim Zerhackerverstärker.

Synchrongleichrichtung: Die Synchrongleichrichtung entspricht einer nochmaligen Modulation mit $p(t)$, die synchron zur ersten Modulation erfolgt. Unter Berücksichtigung der Verstärkung V erhält man:

$$y_2(t) = V u(t) \cos^2(\omega_0 t) + V z(t) \cos(\omega_0 t). \quad (3.203)$$

Mit der Beziehung $\cos^2 x = \frac{1}{2} + \frac{1}{2} \cos(2x)$ kann man das Signal schreiben als

$$y_2(t) = \frac{V}{2} u(t) + \frac{V}{2} u(t) \cos(2\omega_0 t) + V z(t) \cos(\omega_0 t). \quad (3.204)$$

Im Spektralbereich bedeutet das eine weitere Verschiebung des Störspektrums $Z(\omega)$ um $\pm 2\omega_0$ (Abb. 3.37) und ein zusätzliches Auftreten des Nutzsignals bei $\pm 2\omega_0$:

$$Y_2(\omega) = \frac{V}{2} U(\omega) + \frac{V}{4} U(\omega \pm 2\omega_0) + \frac{V}{2} Z(\omega \pm 2\omega_0). \quad (3.205)$$

Tiefpass: Aus dem Signal $Y_2(\omega)$ in (3.205) müssen jetzt noch die hochfrequenten Anteile entfernt werden. Dies geschieht mit einem Tiefpass (TP) mit der Filtercharakteristik

$$|T(\omega)| = \begin{cases} 1 & \text{für } |\omega| \leq |\omega_u| \\ 0 & \text{für } |\omega| > |\omega_0| - |\omega_z| \end{cases}. \quad (3.206)$$

Hierin bezeichnen ω_u die Bandbreite des Eingangssignals und ω_z die der Störung (Abb. 3.36). Die hochfrequenten Störungen werden damit aus dem Ausgangssignal $Y_2(\omega)$ herausgefiltert:

$$Y(\omega) = Y_2(\omega) \cdot T(\omega) = \frac{V}{2} \cdot U(\omega). \quad (3.207)$$

Um spektrale Überlappungen zu vermeiden, muss die Umschaltfrequenz ω_0 sehr viel größer sein als die Summe der Bandbreiten des Nutzsignals

und des Störsignals:

$$|\omega_0| > |\omega_u| + |\omega_z|. \quad (3.208)$$

Durch die Zerhackung wird die nutzbare Signalbandbreite ω_u eingeschränkt. Zerhackerverstärker werden insbesondere verwendet, wenn ein Offset in der Verstärkereingangsstufe kompensiert werden soll.

3.4 Rückwirkung des Messsystems

3.4

Beim Abgriff einer Messgröße durch das Messsystem findet ein Energiefluss zwischen Prozess und Messsystem statt, der die ursprüngliche Messgröße verfälschen kann. Die daraus resultierenden Fehler sollen im Folgenden untersucht werden.

An die vorgesehene Messstelle werde eine Schnittstelle gelegt. Die Schnittstelle teilt den Prozess in Teil 1 und Teil 2 auf (Abb. 3.38). Es wird mit den verallgemeinerten, konjugierten Variablen Kraft F und Fluss Q gerechnet. Beispiele dafür sind Spannung und Strom oder Federkraft und Auslenkungsgeschwindigkeit. Die Messgröße sei die verallgemeinerte Kraft F_1 .

Für die Flussvariablen ohne Messsystem gilt

$$Q_{10} = Q_{20} \quad (3.209)$$

bzw. mit Messsystem entsprechend

$$Q_1 = Q_2 + Q_m \quad \text{oder} \quad Q_{10} + \Delta Q_1 = Q_{20} + \Delta Q_2 + Q_m. \quad (3.210)$$

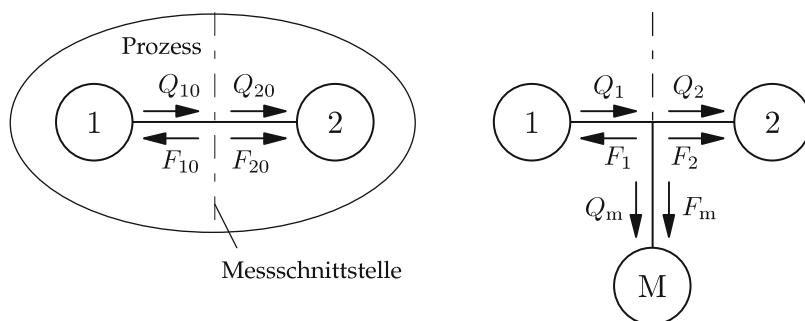


Abbildung 3.38. Prozess und Messsystem.

Für die Kraftvariablen ohne und mit Messsystem gilt:

$$F_{10} = -F_{20} \quad \text{bzw.} \quad F_1 = -F_2 = -F_m . \quad (3.211)$$

Die physikalischen Gesetzmäßigkeiten ergeben allgemeine Zusammenhänge zwischen Kraft und Flussgrößen der Form

$$Q_i = f_i(F_1, F_2, \dots) . \quad (3.212)$$

Kleine Abweichungen der Flussgrößen ΔQ_i lassen sich dann näherungsweise durch das Differential $\frac{\partial Q_i}{\partial F_i} \Delta F_i$ berechnen. Für das Messsystem wird ein linearer Zusammenhang $F_m = Q_m W_m$ zwischen den konjugierten Größen angenommen. Die Bezeichnung W_m des Proportionalitätsfaktors soll dabei auf den Innenwiderstand des Messgerätes hinweisen. Gleichung (3.210) wird damit

$$\frac{\partial Q_1}{\partial F_1} \Delta F_1 = \frac{\partial Q_2}{\partial F_2} \Delta F_2 + \frac{F_m}{W_m} \quad (3.213)$$

oder mit (3.211)

$$\frac{\partial Q_1}{\partial F_1} \Delta F_1 = -\frac{\partial Q_2}{\partial F_2} \Delta F_1 - \frac{F_1}{W_m} , \quad (3.214)$$

woraus man

$$\frac{\Delta F_1}{F_1} = -\frac{1}{W_m} \cdot \left(\frac{\partial Q_1}{\partial F_1} + \frac{\partial Q_2}{\partial F_2} \right)^{-1} \quad (3.215)$$

erhält. Durch den endlichen Energie- oder Leistungsverbrauch des Messsystems ändert sich die Messgröße F_1 um ΔF_1 . Ist die Messgröße der verallgemeinerte Fluss, so lassen sich analoge Beziehungen herleiten.

Beispiel 3.11 (Spannungsmessung an einem Spannungsteiler): An die Schnittstelle des Spannungsteilers in Abb. 3.39 wird parallel zu R_a eine Messeinrichtung mit dem Innenwiderstand R_m gelegt. Es entsprechen sich

$$F_1 = U_1 , \quad Q_1 = I_1 , \quad W_m = R_m , \quad (3.216)$$

$$\frac{\partial Q_1}{\partial F_1} = \frac{1}{R_i} , \quad \frac{\partial Q_2}{\partial F_2} = \frac{1}{R_a} , \quad (3.217)$$

was mit (3.215)

$$\frac{\Delta U_1}{U_1} = -\frac{1}{R_m} \cdot \left(\frac{1}{R_i} + \frac{1}{R_a} \right)^{-1} \quad (3.218)$$

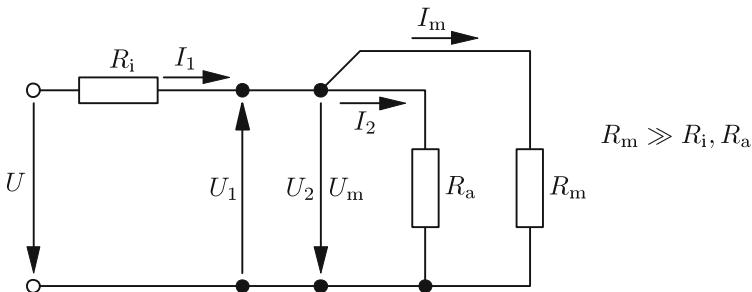


Abbildung 3.39. Spannungsteiler.

ergibt. Der Fehler geht gegen null, wenn $R_m \rightarrow \infty$ geht. Eine genaue Ableitung ergibt im Nenner noch einen zusätzlichen Term $1/R_m$. Die Entwicklung von (3.215) nach den Kräften stellt lediglich eine Näherung dar, weil die höheren Glieder der Taylor-Reihe vernachlässigt wurden.

Um eine Rückwirkung des Messsystems zu verhindern, muss bei einer **Spannungsmessung** die Spannung hochohmig abgegriffen werden. Bei einer **Strommessung** sind die Ströme dagegen niederohmig zu messen, um die Stromquelle minimal zu belasten. ■

Beispiel 3.12 (Längenmesstaster): Eine vergleichbare Abschätzung des Fehlers erhält man, wenn eine Reihenschaltung vorliegt. Ein Werkstück der Länge x soll mit einem Längenmesstaster gemessen werden, der die Federcharakteristik $F_m = c_m x_m$ hat (Abb. 3.40). Das Werkstück habe die Federkonstante c mit $c \gg c_m$. Das System umfasst die beiden Baugruppen Werkstück und Längenmesstaster. Die Messgröße ist die Zustandsvariable x . Der stabile Arbeitspunkt ist erreicht, wenn beide Längen gleich sind:

$$x = x_m. \quad (3.219)$$

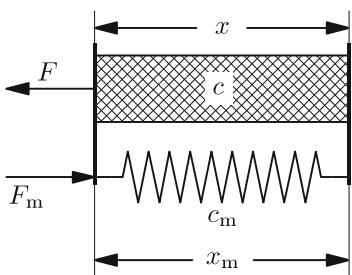


Abbildung 3.40. Längenmesstaster.

Die beiden Federkräfte sind dann ebenfalls betragsmäßig gleich groß:

$$F = -F_m \quad \text{oder} \quad \underbrace{F_0}_{=0} + \Delta F = -F_m . \quad (3.220)$$

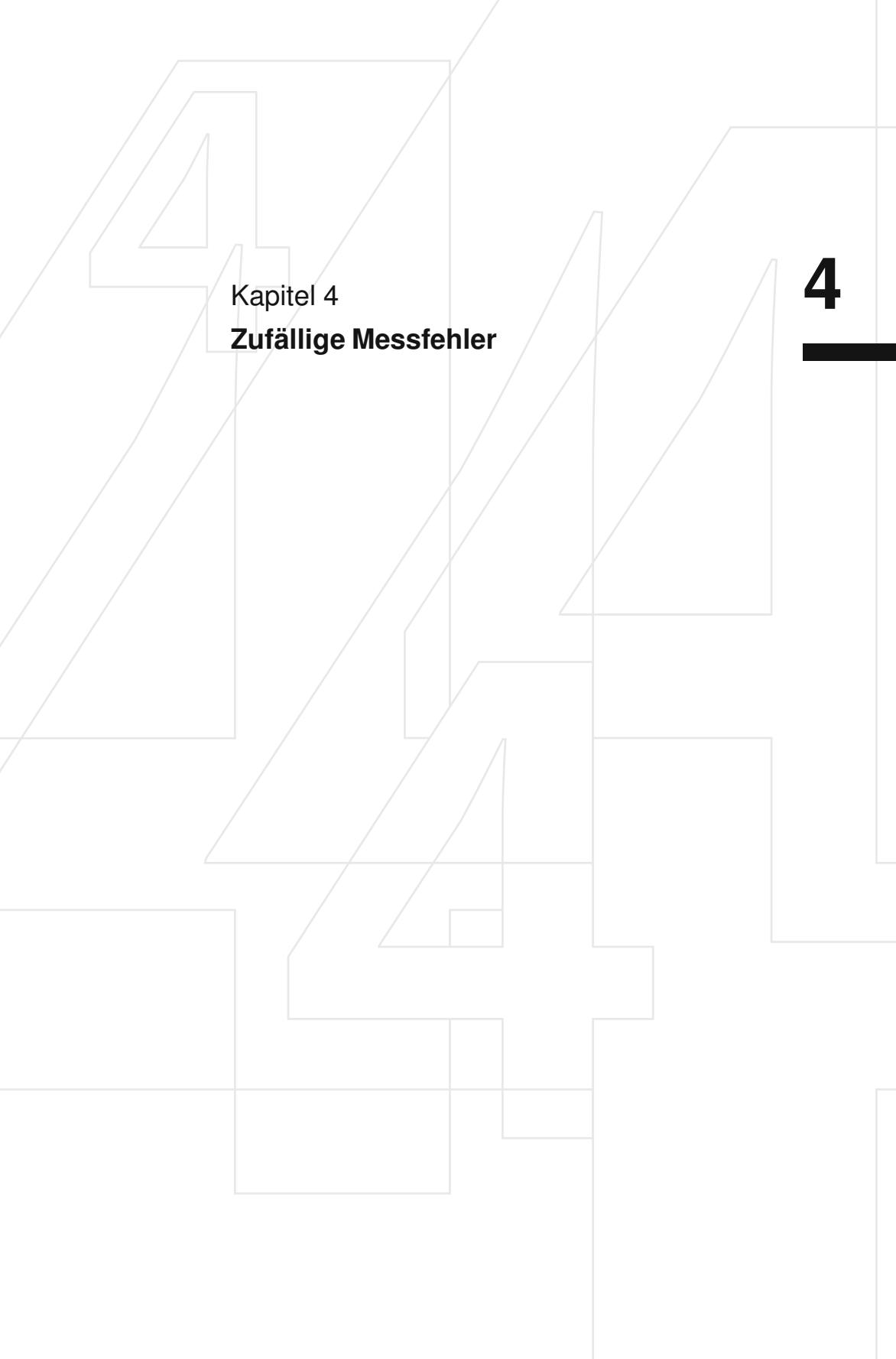
Ohne Anlegen des Längenmesstasters ist die Kraft im Werkstück $F_0 = 0$. Die Entwicklung der Kraftänderung ΔF nach dem Weg Δx ergibt

$$\Delta F = -F_m \quad \text{oder} \quad \frac{\partial F}{\partial x} \Delta x = -c_m x , \quad (3.221)$$

woraus man mit $\partial F / \partial x = c$ die relative Wegänderung erhält:

$$\frac{\Delta x}{x} = -\frac{c_m}{\partial F / \partial x} = -\frac{c_m}{c} . \quad (3.222)$$

Für $c_m \ll c$ ist die Wegänderung erwartungsgemäß klein. ■



Kapitel 4

Zufällige Messfehler

4

4

4 Zufällige Messfehler

| | | |
|---------|---|-----|
| 4.1 | Grundlagen der Wahrscheinlichkeitstheorie | 110 |
| 4.1.1 | Wahrscheinlichkeitsdichte | 112 |
| 4.1.2 | Wahrscheinlichkeitsdichten abgebildeter Größen | 116 |
| 4.1.3 | Momente der Statistik 1. Ordnung | 117 |
| 4.1.4 | Momente der Statistik 2. Ordnung | 120 |
| 4.1.5 | Korrelationskoeffizient | 122 |
| 4.1.6 | Charakteristische Funktion | 125 |
| 4.2 | Stichproben | 126 |
| 4.2.1 | Häufigkeitsverteilung und Histogramm | 127 |
| 4.2.2 | Stichprobenmittelwert | 128 |
| 4.2.3 | Stichprobenvarianz | 131 |
| 4.2.4 | Gesetz der großen Zahlen | 134 |
| 4.2.5 | Mittelung zur Störungsunterdrückung | 136 |
| 4.2.5.1 | Lineare Kennlinie | 136 |
| 4.2.5.2 | Nichtlineare Kennlinie | 137 |
| 4.3 | Normalverteilte Zufallsvariable | 138 |
| 4.3.1 | Normalverteilung | 138 |
| 4.3.2 | Zentraler Grenzwertsatz | 141 |
| 4.3.3 | χ^2 -Verteilung | 142 |
| 4.3.4 | Student'sche t-Verteilung | 147 |
| 4.4 | Statistische Testverfahren | 148 |
| 4.4.1 | Konfidenzintervall und statistische Sicherheit | 149 |
| 4.4.1.1 | Konfidenzintervall bei bekannter Standardabweichung | 149 |
| 4.4.1.2 | Konfidenzintervall bei geschätzter Standardabweichung | 152 |
| 4.4.2 | Hypothesen und statistische Tests | 155 |
| 4.4.3 | Signifikanztest für den Stichprobenmittelwert | 156 |
| 4.4.4 | χ^2 -Anpassungstest | 157 |
| 4.5 | Qualitätssicherung | 161 |
| 4.5.1 | Beurteilung von Fertigungsprozessen | 161 |
| 4.5.2 | Bestimmung der Ausfallrate | 164 |
| 4.5.3 | Statistische Prozessüberwachung | 169 |
| 4.6 | Fehlerfortpflanzung | 173 |

4 Zufällige Messfehler

Messfehler lassen sich aufgrund von Versuchen in systematische und zufällige Fehler einteilen. Erhält man bei wiederholten Versuchen das gleiche Ergebnis, spricht man von systematischen, bei voneinander im Betrag und Vorzeichen abweichenden Ergebnissen dagegen von zufälligen Fehlern. Die Fehler einer Messreihe können abhängig vom Standpunkt des Beobachters oder von den Versuchsbedingungen zu den zufälligen oder zu den systematischen Fehlern zählen. Dies macht man sich am besten an einem Beispiel klar.

Beispiel 4.1 (Zufällige Fehler durch ungenaue Modellkenntnis): Ein Spannungsmesser werde an ein Spannungsnormale angeschlossen und im Labor werden über den Tag verteilt mehrere Messungen durchgeführt. Die Fehler der Messungen unterscheiden sich im Betrag und im Vorzeichen. Sie sind also den zufälligen Fehlern zuzurechnen. Eine eingehende Untersuchung lässt vermuten, dass ein Zusammenhang mit Temperaturänderungen besteht. Die Versuche werden im Temperaturschrank mit einstellbaren Temperaturen wiederholt, ebenso werden die Messungen im Labor bei gleichzeitiger Registrierung der Temperatur noch einmal durchgeführt. Es zeigt sich, dass die Fehler eindeutig von der Raumtemperatur abhängen. Es sind jetzt systematische Fehler. ■

Das Beispiel lässt etwas Grundsätzliches erkennen. Mit verfeinerten Versuchsbedingungen und besserer Systemkenntnis werden immer mehr zufällige Fehler zu systematischen. Selbst das Standardbeispiel der Wahrscheinlichkeitsrechnung für ein zufälliges Ereignis – das Würfeln – braucht grundsätzlich nicht zufällig zu sein. Sind die Geschwindigkeit, die Richtung, der Drehimpuls des Würfels beim Wurf bekannt, ließe sich das Ergebnis eigentlich nach den analytischen Gesetzen der Mechanik berechnen. Nach den Erkenntnissen der Chaostheorie führen allerdings winzige Abweichungen in den Anfangsbedingungen bereits zu völlig unterschiedlichen Systemzuständen, so dass die abweichenden Ergebnisse doch als zufällige Fehler gedeutet werden. Diese Art der Fehler kann mit den Methoden der Wahrscheinlichkeitsrechnung und der Statistik untersucht und beschrieben werden.

4.1 Grundlagen der Wahrscheinlichkeitstheorie

In diesem Abschnitt werden grundlegende Zusammenhänge der Wahrscheinlichkeitstheorie behandelt, die für eine stochastische Beschreibung von Messsystemen erforderlich sind. Für eine eingehende Darstellung anderer Aspekte der Wahrscheinlichkeitstheorie sei auf die zahlreichen hervorragenden Lehrbücher verwiesen [7, 17, 23, 37].

Der Ausgang eines Messsystems ist abhängig vom zeitlichen Verlauf aller Eingangssignale (vgl. Abb. 1.2). Dabei werden Eingangs- und Ausgangsgrößen gewöhnlich **deterministisch** beschrieben, also als Funktionen der Signalamplitude über der Zeit. Allerdings ist der Verlauf der Störgrößen $z(t)$ nicht genau bekannt, weshalb diese meist statistisch beschrieben werden. Anders als bei deterministischen Signalen, die im Zeitbereich oder im Frequenzbereich beschrieben werden [42], werden **zufällige** Signale (insbesondere Störsignale und störungsbehaftete Messsignale) etwas unverbindlicher im sogenannten „Amplitudenbereich“ dargestellt.

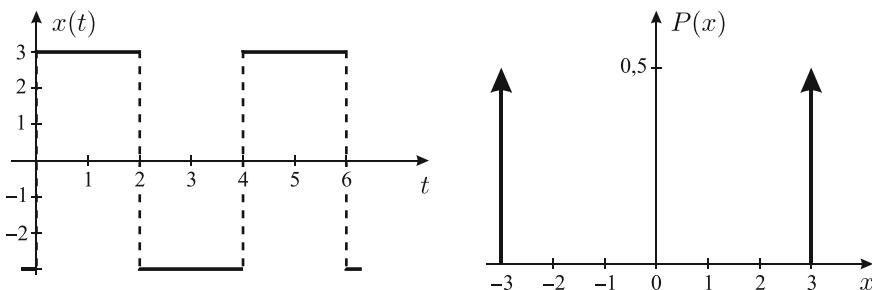


Abbildung 4.1. Rechtecksignal im Zeitbereich (links) und im Amplitudenbereich (rechts).

Exemplarisch zeigt Abb. 4.1 den zeitlichen Verlauf eines Rechtecksignals und die zugehörige Beschreibung im Amplitudenbereich. Dabei werden den Amplitudewerten $x(t)$ Wahrscheinlichkeiten $P(x)$ zugeordnet. Dass die Darstellung im Amplitudenbereich unverbindlicher und daher mit einem Informationsverlust verbunden ist, wird dadurch deutlich, dass aus den Amplitudewahrscheinlichkeiten im rechten Diagramm kein Rückschluss auf den zeitlichen Verlauf des Signals möglich ist.

Allgemein lassen sich diese einzelnen Amplituden- oder Signalwerte als Ergebnisse eines **Zufallsexperimentes** interpretieren, die mitunter auch diskret (z. B. beim Würfeln) oder nominal skaliert (beim Münzwerfen sind es etwa „Kopf“ oder „Zahl“) sein können. Um eine einheitliche Handhabung derartiger Größen im Amplitudenbereich zu ermöglichen, ist eine Abbildung der **Ergebnismenge** des Zufallsexperimentes – d. h. der Menge der sogenannten **Elementarereignisse** ξ_i – auf eine geeignete Wertemenge er-

forderlich (meist handelt es sich hierbei um die Menge der reellen Zahlen). Diese Abbildung wird von einer sogenannten **Zufallsvariablen** x vorgenommen. Mit ihr ordnet man der Ergebnismenge eines Zufallsexperimentes reellwertige Zahlen zu.

Definition 4.1: Zufallsvariable

Jede auf der Ergebnismenge eines Zufallsexperimentes definierte reelle Funktion wird als Zufallsvariable bezeichnet. Ist x das Symbol einer Zufallsvariablen, so bezeichnet man die reelle Zahl, die dem Elementarereignis ξ durch x zugeordnet wird, mit $x(\xi)$; siehe Abb. 4.2. ◇

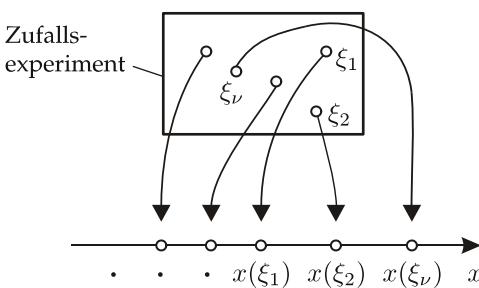


Abbildung 4.2. Zufallsexperiment, Elementarereignis und Zufallsvariable.

Der Begriff Zufallsvariable (engl. *random variable*) hat sich zwar etabliert, ist jedoch leider irreführend. Wie in Abb. 4.2 veranschaulicht, handelt es sich bei $x(\xi)$ um keine Variable, sondern um eine wohldefinierte Funktion, welche die Elementarereignisse ξ_i auf die reellen Werte $x(\xi_i)$ abbildet.¹ Der Zufall spielt sich alleine beim Zufallsexperiment ab.

Es gibt diskrete Zufallsvariable, die bei Zufallsexperimenten mit abzählbaren Elementarereignissen auftreten (z. B. das Würfeln oder das Werfen einer Münze). Kontinuierliche Zufallsvariable sind immer mit Experimenten verbunden, bei denen die Elementarereignisse nicht abzählbar sind (z. B. die Messung einer metrischen Größe). Dies ist in der Messtechnik der häufigste Fall.

Beispiel 4.2 (Diskrete Zufallsvariable): In einem Würfelexperiment wird ein Würfel zweimal geworfen; die Ergebnisse seien die Elementarereignisse

¹Zur übersichtlicheren Darstellung wird bei Zufallsvariablen $x(\xi)$ oftmals die Abhängigkeit vom Elementarereignis ξ unterdrückt. Zur Erinnerung, dass es sich weiterhin um Funktionen handelt, wird hier x in Schreibmaschinenschrift gesetzt. Die Werte $x(\xi_i)$, die x annehmen kann, sind allerdings deterministisch und werden kursiv gesetzt.

ξ_1 und ξ_2 . Die Zufallsvariable x bezeichne die Summe der Augenzahlen:

$$x(\xi_1, \xi_2) = \xi_1 + \xi_2 \quad \text{für} \quad \xi_1, \xi_2 \in \{1, 2, \dots, 6\}. \quad (4.1)$$

x ist also eine diskrete Zufallsvariable, welche die Werte $2, 3, \dots, 12$ annehmen kann. ■

Beispiel 4.3 (Kontinuierliche Zufallsvariable): Eine Spannungsquelle mit einer Nennspannung $U_0 = 5 \text{ V}$ wird vermessen. Die gemessenen Werte schwanken zufällig im Intervall $4,9 \text{ V} \leq u \leq 5,1 \text{ V}$. Als Zufallsvariable wird die Abweichung $x = u - U_0$ zur Nennspannung U_0 betrachtet. Diese ist kontinuierlich und nimmt Werte im Intervall $-0,1 \text{ V} \leq x \leq 0,1 \text{ V}$ an. ■

Die folgenden Abschnitte widmen sich, sofern nichts anderes angegeben wird, den kontinuierlichen Zufallsvariablen.

4.1.1 Wahrscheinlichkeitsdichte

4.2

Definition 4.2: Wahrscheinlichkeitsverteilung

Die Wahrscheinlichkeitsverteilung (kurz: Verteilung)

$$F_x(x) = P(x \leq x) \quad (4.2)$$

einer Zufallsvariablen x gibt die Wahrscheinlichkeit P an, mit welcher der Funktionswert von x kleiner oder höchstens gleich x ist. Die Wahrscheinlichkeitsverteilung besitzt die folgenden Eigenschaften:

$$\lim_{x \rightarrow -\infty} F_x(x) = 0, \quad (4.3)$$

$$\lim_{x \rightarrow \infty} F_x(x) = 1. \quad (4.4)$$

Ferner ist $F_x(x)$ eine monoton steigende Funktion, nimmt also mit wachsendem x nirgends ab. ◇

Anstelle der Verteilungsfunktion wird meist die Wahrscheinlichkeitsdichtefunktion $f_x(x)$ (auch Verteilungsdichtefunktion genannt) verwendet, die dieselbe Information enthält.

Definition 4.3: Wahrscheinlichkeitsdichte

Die Wahrscheinlichkeitsdichte (kurz: Dichte) $f_x(x)$ einer Zufallsvariablen x ist wie folgt definiert:

$$f_x(x) = \frac{dF_x(x)}{dx} \quad \text{mit} \quad F_x(x) = \int_{-\infty}^x f_x(u) du. \quad (4.5)$$

Die Wahrscheinlichkeitsdichte besitzt folgende Eigenschaften:

$$f_x(x) \geq 0, \quad (4.6)$$

$$\int_{-\infty}^{\infty} f_x(x) dx = 1, \quad (4.7)$$

$$\int_a^b f_x(x) dx = P(a < x \leq b) \quad (4.8)$$

mit $a, b \in \mathbb{R}$. \diamond

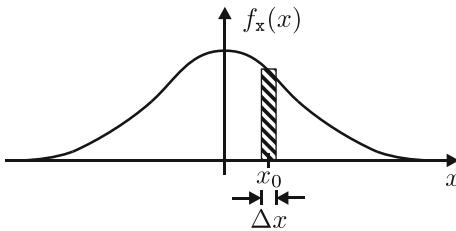


Abbildung 4.3. Interpretation der Wahrscheinlichkeitsdichte.

Abbildung 4.3 zeigt exemplarisch die Wahrscheinlichkeitsdichte $f_x(x)$ einer Zufallsvariablen x . Die Wahrscheinlichkeitsdichte $f_x(x)$ beschreibt die Wahrscheinlichkeit pro Umgebungsbreite, dass x in einer schmalen Umgebung der Breite Δx um x liegt:

$$f_x(x) \approx \frac{P(x - \frac{\Delta x}{2} < x \leq x + \frac{\Delta x}{2})}{\Delta x}. \quad (4.9)$$

Bei diskreten Zufallsvariablen enthält die Wahrscheinlichkeitsdichte Dirac-Impulse.

Der Einfachheit halber wird die Zufallsvariable als Index oft weggelassen, wenn dies zu keinen Verwechslungen führt:

$$f_x(x) = f(x) \quad \text{bzw.} \quad F_x(x) = F(x). \quad (4.10)$$

Definiert man über derselben Ergebnismenge mehrere Zufallsvariable, so kann man für diese gemeinsame Verteilungen und Dichten angeben. Die Darstellung soll auf den häufigen Fall von zwei Zufallsvariablen x, y beschränkt werden. Eine Verallgemeinerung ist jedoch leicht möglich.

4.4

Definition 4.4: Verbundwahrscheinlichkeitsverteilung

Die Verbundwahrscheinlichkeitsverteilung oder gemeinsame Wahrscheinlichkeitsverteilung

$$F_{xy}(x, y) = P(x \leq x \cap y \leq y) \quad (4.11)$$

zweier Zufallsvariabler x, y gibt die Wahrscheinlichkeit P an, mit welcher der Funktionswert von x kleiner oder höchstens gleich x ist und der Funktionswert von y kleiner oder höchstens gleich y ist. \diamond

Meist wird anstelle der Verbundverteilungsfunktion die Verbundverteilungsdichtefunktion verwendet, die dieselbe Information enthält.

4.5

Definition 4.5: Verbundwahrscheinlichkeitsdichte

Die Verbundwahrscheinlichkeitsdichte (oder auch: gemeinsame Wahrscheinlichkeitsdichte) zweier Zufallsvariabler x und y ist

$$f_{xy}(x, y) = \frac{\partial^2 F_{xy}(x, y)}{\partial x \partial y}. \quad (4.12)$$

Wenn keine Verwechslungsgefahr besteht, kann hier analog zu (4.10) die Kurzschriftweise

$$f_{xy}(x, y) = f(x, y) \quad \text{bzw.} \quad F_{xy}(x, y) = F(x, y) \quad (4.13)$$

verwendet werden. \diamond

Aus der gemeinsamen Dichte kann die gemeinsame Verteilung durch Integration berechnet werden:

$$F_{xy}(x, y) = \int_{-\infty}^x \int_{-\infty}^y f_{xy}(u, v) du dv. \quad (4.14)$$

Aus der gemeinsamen Wahrscheinlichkeitsdichte lässt sich ebenfalls die sogenannte **Randdichte** einer einzelnen Zufallsvariablen berechnen, indem die andere Zufallsvariable herausintegriert wird.

Definition 4.6: Randdichte

Ist die Wahrscheinlichkeitsdichte $f_{xy}(x, y)$ zweier Zufallsvariablen x, y gegeben, so berechnen sich die Randdichten der einzelnen Zufallsvariablen durch sogenannte Marginalisierung zu

$$f_x(x) = \int_{-\infty}^{\infty} f_{xy}(x, y) dy \quad \text{bzw.} \quad f_y(y) = \int_{-\infty}^{\infty} f_{xy}(x, y) dx. \quad (4.15)$$

Der umgekehrte Weg ist im Allgemeinen nicht möglich. Nur wenn x und y statistisch unabhängig sind, lässt sich aus den Randdichten die Verbundwahrscheinlichkeitsdichte berechnen. ◇

Definition 4.7: Statistische Unabhängigkeit

Zwei Zufallsvariable x und y sind statistisch unabhängig, wenn

$$F_{xy}(x, y) = F_x(x) \cdot F_y(y) \quad \text{bzw.} \quad f_{xy}(x, y) = f_x(x) \cdot f_y(y) \quad (4.16)$$

gilt. ◇

Statistische Unabhängigkeit ist eine Eigenschaft, die experimentell höchstens näherungsweise nachgewiesen werden kann. Bei der Formulierung eines Modells für ein Messsystem kann sie meist nur als Voraussetzung angenommen werden. Der große Vorteil der statistischen Unabhängigkeit liegt in der wesentlich vereinfachten Modellanalyse.

Hilfreich ist auch die Definition einer bedingten Wahrscheinlichkeitsdichte, aus der unmittelbar das **Bayes-Theorem** resultiert.

Definition 4.8: Bedingte Wahrscheinlichkeitsdichte

Die bedingte Wahrscheinlichkeitsdichte

$$f_{x|y}(x|y) = \frac{f_{xy}(x, y)}{f_y(y)} \quad (4.17)$$

ist die Wahrscheinlichkeitsdichte der Zufallsvariablen x unter der Bedingung, dass das Ereignis $y = y$ eingetreten ist. ◇

Sind x und y statistisch unabhängig, so hängt das Auftreten von $x = x$ in keiner Weise von der Bedingung $y = y$ ab. Die bedingte Wahrscheinlichkeit wird dann

$$f_{x|y}(x|y) = \frac{f_{xy}(x, y)}{f_y(y)} = \frac{f_x(x) \cdot f_y(y)}{f_y(y)} = f_x(x). \quad (4.18)$$

Satz 4.1: Bayes-Theorem

Aus der Def. 4.8 folgt unmittelbar:

$$f_{xy}(x, y) = f_{x|y}(x|y) \cdot f_y(y) = f_{y|x}(y|x) \cdot f_x(x). \quad (4.19)$$

Die rechte Seite dieser Gleichung entspricht dem Bayes-Theorem, das einen Zusammenhang zwischen den bedingten Wahrscheinlichkeiten $f_{x|y}(x|y)$ und $f_{y|x}(y|x)$ herstellt und welches von enormer Bedeutung in der Statistik, Schätztheorie und Mustererkennung ist. \diamond

② 4.1.2 Wahrscheinlichkeitsdichten abgebildeter Größen

Eine Zufallsvariable x kann durch eine Funktion $g(x)$ auf eine andere Zufallsvariable abgebildet werden. Dabei interessiert man sich für die Wahrscheinlichkeitsdichte der neuen Zufallsvariablen. Die Berechnung erfolgt nach folgendem Satz.

Satz 4.2: Wahrscheinlichkeitsdichten transformierter Variabler

Wird eine Zufallsvariable x mit der Wahrscheinlichkeitsdichte $f_x(x)$ durch eine Funktion

$$y = g(x) \quad (4.20)$$

in eine neue Zufallsvariable transformiert, und existiert eine Umkehrfunktion mit n Lösungen

$$x_i = g^{-1}(y), \quad i \in \{1, \dots, n\}, \quad (4.21)$$

so gilt für die Wahrscheinlichkeitsdichte von y

$$f_y(y) = \sum_{i=1}^n f_x(x_i) \left| \frac{dg(x)}{dx} \right|_{x=x_i}^{-1}. \quad (4.22)$$

\diamond

Der Zusammenhang (4.22) soll nun grafisch veranschaulicht werden; der ausführliche Beweis kann in [37] nachgelesen werden. Aus Abb. 4.4 entnimmt man für die Wahrscheinlichkeit im Intervall $y \leq y \leq y + dy$

$$f_y(y) dy = f_x(x_1) |dx_1| + f_x(x_2) |dx_2| + \dots. \quad (4.23)$$

Mit

$$dx_i = \frac{dy}{g'(x_i)} \quad (4.24)$$

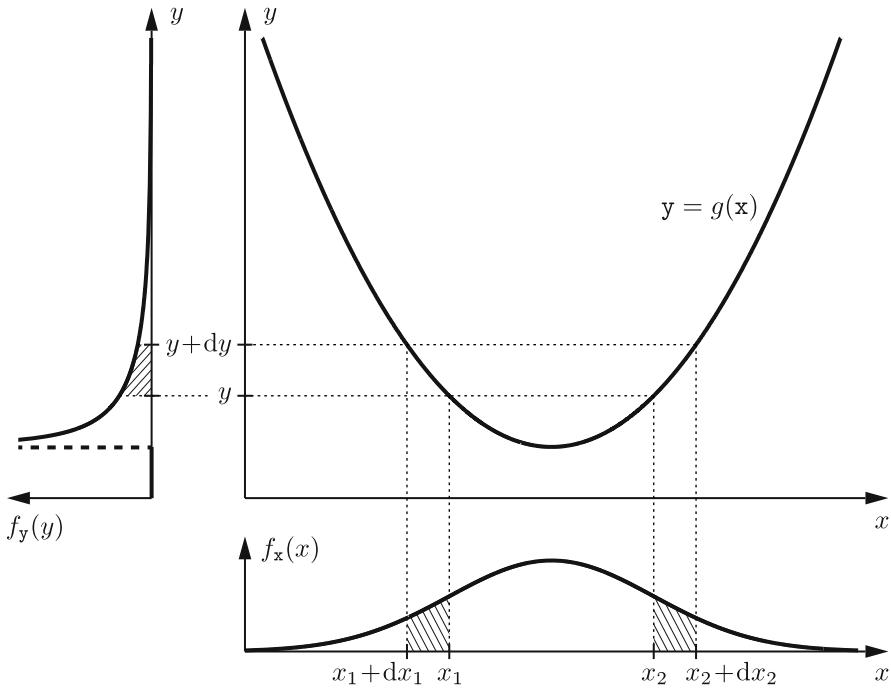


Abbildung 4.4. Zur Wahrscheinlichkeitsdichte abgebildeter Größen.

folgt sofort die Behauptung (4.22).

4.1.3 Momente der Statistik 1. Ordnung

Bisher wurden Zufallsvariable x anhand ihrer Wahrscheinlichkeitsverteilung $F_x(x)$ bzw. ihrer Dichtefunktion $f_x(x)$ charakterisiert. Zur kompakten Beschreibung von Zufallsvariablen werden Kenngrößen wie z. B. der Mittelwert oder die Varianz herangezogen. Diese lassen sich über den **Erwartungswert** definieren, der im Folgenden für eine Statistik 1. Ordnung eingeführt wird.

Definition 4.9: Statistik n -ter Ordnung

4.9

Werden bei einer statistischen Betrachtung n Zufallsvariable x_1, \dots, x_n berücksichtigt, so spricht man von einer **Statistik n -ter Ordnung**. Diese wird durch die Verbundwahrscheinlichkeitsdichte

$$f_{x_1 \dots x_n}(x_1, \dots, x_n) \quad (4.25)$$

vollständig beschrieben. \diamond

4.10**Definition 4.10: Erwartungswert**

Der Erwartungswert einer Funktion $g(x)$ einer Zufallsvariablen x mit der Dichte $f_x(x)$ ist als

$$E\{g(x)\} = \int_{-\infty}^{\infty} g(x) \cdot f_x(x) dx \quad (4.26)$$

definiert. \diamond

Der Erwartungswert hat als linearer Operator folgende Eigenschaften:

$$E\{a g(x)\} = a \cdot E\{g(x)\}, \quad (4.27)$$

$$E\{g(x) + h(x)\} = E\{g(x)\} + E\{h(x)\}. \quad (4.28)$$

Setzt man für die Funktion $g(x)$ Potenzen x^n ein, so erhält man die **Momente**.

4.11**Definition 4.11: Moment**

Das m -te Moment einer Zufallsvariablen x ist definiert zu

$$\mu_{x,m} = E\{x^m\} = \int_{-\infty}^{\infty} x^m f_x(x) dx. \quad (4.29)$$

Das erste Moment

$$\mu_{x,1} = \mu_x = E\{x\} = \int_{-\infty}^{\infty} x f_x(x) dx \quad (4.30)$$

ist der **Mittelwert** oder auch Schwerpunkt von x . Es ist ein Lageparameter von x und beschreibt, wo sich die Zufallsgröße im Mittel befindet. \diamond

Grundsätzlich ist streng zwischen der Ordnung n einer Statistik und der Ordnung m eines Momentes zu unterscheiden. Während die Ordnung einer Statistik die Anzahl betrachteter Zufallsvariabler bezeichnet, beschreibt die Ordnung eines Momentes den Grad m der Potenz in (4.29).

4.12**Definition 4.12: Zentrales Moment**

Das m -te **zentrale Moment** einer Zufallsvariablen x ist zu

$$E\{(x - E\{x\})^m\} = \int_{-\infty}^{\infty} (x - E\{x\})^m f_x(x) dx \quad (4.31)$$

definiert. \diamond

Das zweite zentrale Moment ist die **Varianz** σ_x^2 . Ihre Wurzel σ_x wird als **Standardabweichung** bezeichnet. Als Streuungsmaß von x gibt die Varianz einen Anhalt über die Breite der Dichtefunktion. Die schraffierten Flächen in Abb. 4.5 stellen die Varianzen der Zufallsvariablen x_1 und x_2 dar.

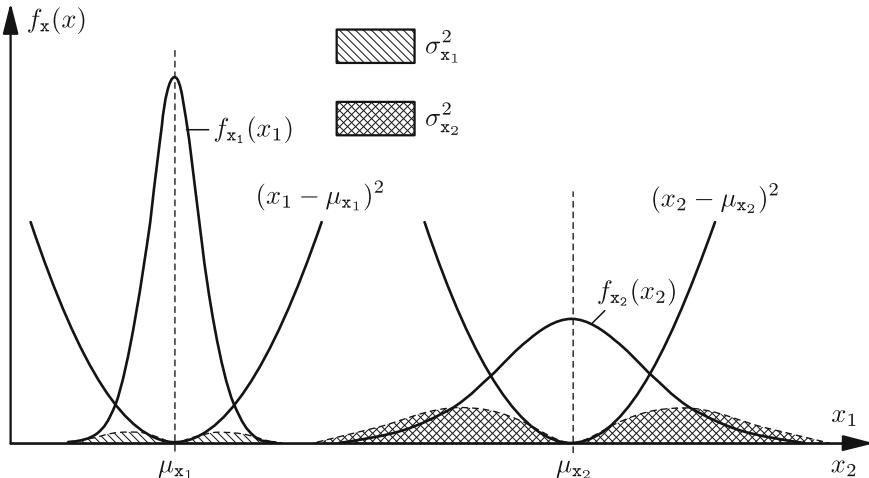


Abbildung 4.5. Mittelwerte und Varianzen zweier stochastischer Variablen x_1 und x_2 .

Häufige Verwendung finden auch die **höheren Momente**. Die **Schiefe**

$$\varrho_x = \frac{E\{(x - E\{x\})^3\}}{\sigma_x^3} \quad (4.32)$$

ist das auf die dritte Potenz der Standardabweichung bezogene dritte zentrale Moment und stellt ein Maß für die Asymmetrie einer Verteilung dar. Bei $\varrho_x < 0$ heißt eine Verteilung linksschief, bei $\varrho_x > 0$ rechtsschief. Für symmetrische Verteilungen gilt zwar $\varrho_x = 0$, allerdings darf aus $\varrho_x = 0$ nicht automatisch auf eine symmetrische Verteilung geschlossen werden.

Der **Exzess**

$$\varepsilon_x = \frac{E\{(x - E\{x\})^4\}}{\sigma_x^4} - 3. \quad (4.33)$$

basiert auf dem vierten zentralen Moment und ist ein Maß für die Abweichung einer unimodalen (d. h. eingipfligen) Verteilung von der Normalverteilung (vgl. Abschn. 4.3). Durch die Subtraktion des rechten Terms wird der Exzess der Normalverteilung zu null (vgl. Abb. 4.6). Bei $\varepsilon_x < 0$ heißt eine

Verteilung flachgipflig oder platykurtisch, bei $\varepsilon_x > 0$ heißt sie steilgipflig oder leptokurtisch.

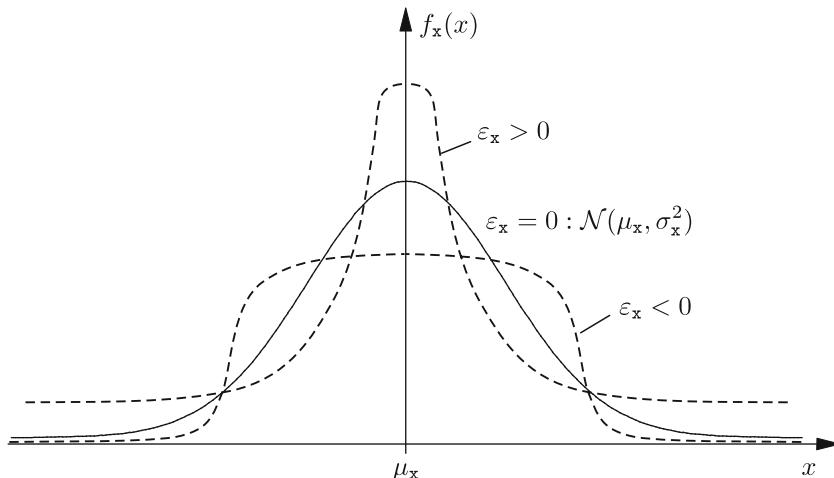


Abbildung 4.6. Exzess ε_x zur Beurteilung der Verteilung von Zufallsvariablen.

4.1.4 Momente der Statistik 2. Ordnung

Der Erwartungswert-Operator lässt sich ebenfalls auf das Produkt mehrerer Zufallsvariabler anwenden. Allgemein kann man auch hier Funktionen der Zufallsgrößen – insbesondere deren Potenzen – zulassen.

4.13

Definition 4.13: Gemeinsames Moment

Das gemeinsame Moment zweier Zufallsvariabler ist wie folgt definiert:

$$\mu_{xy,km} = E\{x^k y^m\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^k y^m f_{xy}(x, y) dx dy, \quad (4.34)$$

wobei die Summe $k + m$ die Ordnung des Momentes bezeichnet. \diamond

In der Anwendung beschränkt man sich in der Regel auf das einfache Produkt $E\{xy\}$, welches für $k = m = 1$ resultiert. Hierbei verwendet man meist das zweite zentrale Moment der Statistik 2. Ordnung, welches als **Kovarianz** bezeichnet wird.

Definition 4.14: Kovarianz

Die Kovarianz zweier Zufallsvariabler x und y ist durch

$$C_{xy} = E\{(x - \mu_x)(y - \mu_y)\} \quad (4.35)$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_x)(y - \mu_y) f_{xy}(x, y) dx dy \quad (4.36)$$

gegeben. \diamond

Die Kovarianz sagt etwas über die Korrelation zwischen zwei Zufallsgrößen aus, also über ihre lineare statistische Abhängigkeit.

Definition 4.15: Unkorrelierte Größen

Zwei Zufallsvariable x und y sind unkorreliert, wenn für sie

$$E\{xy\} = E\{x\} \cdot E\{y\} \quad \text{bzw.} \quad C_{xy} = 0 \quad (4.37)$$

gilt. Beide Aussagen sind nach (4.35) äquivalent. \diamond

Insbesondere gilt für unkorrelierte Zufallsvariable x_i und x_j :

$$C_{x_i x_j} = \sigma_x^2 \delta_{ij} = \begin{cases} 0 & \text{für } i \neq j, \\ \sigma_x^2 & \text{für } i = j. \end{cases} \quad (4.38)$$

Die statistische Unabhängigkeit schließt immer die Unkorreliertheit ein. Die Umkehrung gilt nur, falls beide Zufallsvariable normalverteilt sind, da hier die höheren Momente der Statistik 1. Ordnung nur vom ersten und zweiten Moment abhängig sind (Abschn. 4.3). In allen anderen Fällen können zwei Zufallsvariable zwar unkorreliert, aber statistisch abhängig sein.

Beispiel 4.4 (Unkorreliertheit bei statistischer Abhängigkeit): Gegeben seien zwei Zufallsvariable x und y , deren Verbundwahrscheinlichkeitsdichte $f_{xy}(x, y)$ aus der Addition vierer gleicher unimodaler Verteilungen mit verschiedenen Mittelwerten resultiert (Abb. 4.7 links). Die Zufallsgrößen x und y sind statistisch abhängig, da sich $f_{xy}(x, y)$ nicht als Produkt der Randdichten $f_x(x)$ und $f_y(y)$ darstellen lässt (vgl. Abb. 4.7 rechts):

$$f_{xy}(x, y) \neq f_x(x) \cdot f_y(y). \quad (4.39)$$

Da aber

$$C_{xy} = E\{(x - \mu_x)(y - \mu_y)\} = E\{xy\} = 0 \quad (4.40)$$

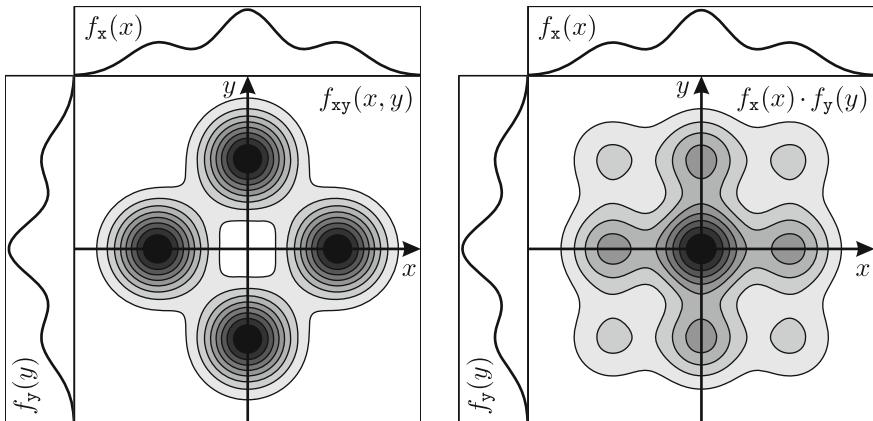


Abbildung 4.7. Höhenlinienplot der Verbunddichte $f_{xy}(x, y)$ zweier unkorrelierter, statistisch abhängiger Zufallsgrößen (links) und durch Marginalisierung berechnete Randdichten $f_x(x)$ und $f_y(y)$. Die Verbunddichte ist nicht als Produkt $f_x(x) \cdot f_y(y)$ der Randdichten darstellbar (rechts).

gilt, sind die Zufallsgrößen x und y unkorreliert. ■

4.1.5 Korrelationskoeffizient

Die Kovarianz C_{xy} sagt zwar etwas über die lineare Abhängigkeit statistischer Größen aus, ist allerdings als Vergleichsmaß nicht geeignet, da sie nicht invariant gegenüber multiplikativen Skalierungen der Größen ist. Man führt daher einen **Korrelationskoeffizienten** als Maß für die statistische Abhängigkeit von Zufallsgrößen ein.

4.16

Definition 4.16: Korrelationskoeffizient

Der Korrelationskoeffizient ρ_{xy} zwischen den Größen x und y ist definiert zu

$$\rho_{xy} = \frac{C_{xy}}{\sigma_x \sigma_y} = \frac{E\{(x - \mu_x)(y - \mu_y)\}}{\sqrt{E\{(x - \mu_x)^2\} E\{(y - \mu_y)^2\}}} . \quad (4.41)$$

Der Wertebereich erstreckt sich auf

$$-1 \leq \rho_{xy} \leq 1 , \quad (4.42)$$

wobei ρ_{xy} bei starrer Bindung von x und y den Wert $+1$ oder -1 und bei unkorrelierten Größen den Wert 0 annimmt. ◇

Beweis 4.1 (Wertebereich des Korrelationskoeffizienten): Für den Beweis ist es hilfreich, die Zufallsgrößen x und y als verallgemeinerte Vektoren in einem unitären Raum zu interpretieren. Für sie werden das Innenprodukt

$$\langle x, y \rangle = E\{(x - \mu_x)(y - \mu_y)\} \quad (4.43)$$

durch die Kovarianz und die Norm

$$\|x\| = \sqrt{\langle x, x \rangle} = \sqrt{E\{(x - \mu_x)^2\}} \quad (4.44)$$

durch die Standardabweichung definiert. Mit der Schwarz'schen Ungleichung

$$|\langle x, y \rangle| \leq \|x\| \cdot \|y\| \quad (4.45)$$

kann man nun die Kovarianzfunktion abschätzen:

$$|E\{(x - \mu_x)(y - \mu_y)\}| \leq \sqrt{E\{(x - \mu_x)^2\}} \sqrt{E\{(y - \mu_y)^2\}}, \quad (4.46)$$

$$|C_{xy}| \leq \sigma_x \cdot \sigma_y. \quad (4.47)$$

Mit dem Korrelationskoeffizienten ρ_{xy} gilt gerade das Gleichheitszeichen:

$$C_{xy} = \rho_{xy} \cdot \sigma_x \sigma_y \quad \text{mit} \quad |\rho_{xy}| \leq 1. \quad (4.48)$$

1. Fall: Starre lineare Bindung zwischen den Größen:

$$y = kx + a, \quad k, a \in \mathbb{R}. \quad (4.49)$$

Der Korrelationskoeffizient ist dann

$$\rho_{xy} = \frac{E\{(x - \mu_x) \cdot k \cdot (y - \mu_y)\}}{\sqrt{E\{(x - \mu_x)^2\} \cdot k^2 \cdot E\{(y - \mu_y)^2\}}} = \pm 1. \quad (4.50)$$

Für die Kovarianz gilt in diesem Fall

$$C_{xy} = \sigma_x \sigma_y. \quad (4.51)$$

2. Fall: Unkorrelierte oder statistisch unabhängige Größen:

$$C_{xy} = 0 \quad \Rightarrow \quad \rho_{xy} = 0. \quad (4.52)$$

Damit ist die letzte Aussage von Def. 4.16 bewiesen. ♦

Tabelle 4.1. Messwertreihe.

| | | | | | | | | | | | | |
|-------|-----|------|------|-----|-----|------|------|-----|-----|------|-----|-----|
| x_i | 0,8 | 1,3 | 2,1 | 2,8 | 3,4 | 4,9 | 5,5 | 6,6 | 7,2 | 8,1 | 9,4 | 9,6 |
| y_i | 0,3 | 0,75 | 1,15 | 1,2 | 1,8 | 2,35 | 2,65 | 3,5 | 3,5 | 4,15 | 4,6 | 4,9 |

Beispiel 4.5 (Korrelation von Messwerten): In der Tabelle 4.1 ist eine Messreihe von $n = 12$ Wertepaaren x_i, y_i dargestellt, die Realisierungen der Zufallsvariablen x und y sind. Es soll der Korrelationskoeffizient zwischen x und y berechnet werden. Da nur einzelne Messwerte vorliegen, müssen die statistischen Kennwerte gemäß Abschn. 4.2 geschätzt werden. Die Erwartungswerte können durch die Stichprobenmittelwerte (Abschn. 4.2.2) geschätzt werden:

$$\hat{\mu}_x = \frac{1}{n} \sum_{i=1}^n x_i = 5,14, \quad \hat{\mu}_y = \frac{1}{n} \sum_{i=1}^n y_i = 2,57. \quad (4.53)$$

Die Kovarianz kann durch die Stichprobenkovarianz geschätzt werden:

$$C_{xy} \approx \frac{1}{n-1} \sum_{i=1}^n (x_i - \hat{\mu}_x)(y_i - \hat{\mu}_y) = 4,8. \quad (4.54)$$

Analog erhält man für die Standardabweichungen (Abschn. 4.2.3):

$$\sigma_x \approx \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \hat{\mu}_x)^2} = 3,08, \quad (4.55)$$

$$\sigma_y \approx \sqrt{\frac{1}{n-1} \sum_{i=1}^n (y_i - \hat{\mu}_y)^2} = 1,56. \quad (4.56)$$

Daraus berechnet sich der Korrelationskoeffizient zu

$$\rho_{xy} = \frac{C_{xy}}{\sigma_x \sigma_y} \approx 0,997. \quad (4.57)$$

Die Wertepaare sind stark voneinander abhängig, was auch aus ihrer grafischen Darstellung (Abb. 4.8) ersichtlich ist. ■

Ein hoher Korrelationskoeffizient ρ_{xy} sagt lediglich etwas über die lineare statistische Abhängigkeit zwischen den Größen x und y aus. Einen kausalen Zusammenhang kann allerdings daraus *nicht* abgeleitet werden, was folgendes Beispiel verdeutlichen soll.

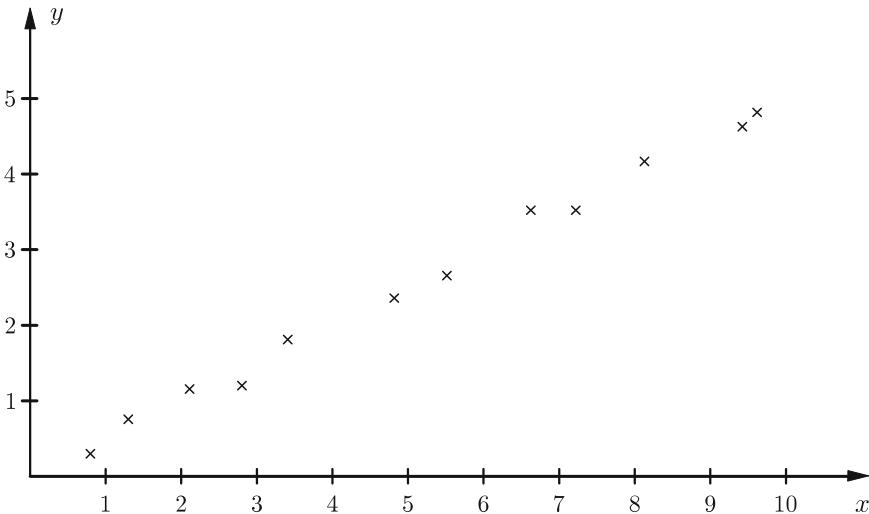


Abbildung 4.8. Messwertreihe aus Tab. 4.1.

Beispiel 4.6 (Korrelation und kausaler Zusammenhang): Zwischen der Anzahl x der Geburten pro Monat und der Zahl y der sich niedergelassenen Störche im gleichen Monat bestehe über das ganze Jahr eine statistische Abhängigkeit. Der Korrelationskoeffizient liege z. B. im Intervall $0,5 \leq \rho_{xy} \leq 1$. Dann darf daraus nicht der kausale Zusammenhang geschlossen werden, die Störche seien der Grund für die Geburten. ■

4.1.6 Charakteristische Funktion

Definition 4.17: Charakteristische Funktion

Die **charakteristische Funktion** $\Phi_x(f)$ einer Zufallsvariablen x ist durch den Erwartungswert

$$\Phi_x(f) = E\{e^{i2\pi f x}\} = \int_{-\infty}^{\infty} f_x(x) e^{i2\pi f x} dx = \mathcal{F}^{-1}\{f_x(x)\}. \quad (4.58)$$

definiert. ◇

Offensichtlich entspricht (4.58) der inversen Fourier-Transformierten der Wahrscheinlichkeitsdichte $f_x(x)$. Da aber $f_x(x)$ stets reell ist, kann (4.58) ebenfalls als die komplexe konjugierte Fourier-Transformierte von $f_x(x)$ aufgefasst werden, weshalb f als die mit x korrespondierende Frequenz inter-

pretiert werden kann. Insbesondere gilt:

$$|\Phi_x(f)| = |\mathcal{F}\{f_x(x)\}|. \quad (4.59)$$

Wegen der Normierung und der Nichtnegativität des Wahrscheinlichkeitsmaßes [42] gelten für die charakteristische Funktion die folgenden Zusammenhänge:

$$\Phi_x(0) = 1, \quad |\Phi_x(f)| \leq 1. \quad (4.60)$$

Für die charakteristische Funktion gibt es zwei wichtige Anwendungsbereiche:

- **Berechnung der Momente:** Wahrscheinlichkeitsdichten können – zumindest näherungsweise – durch die Angabe weniger Momente beschrieben werden. Die Momente können leicht aus der charakteristischen Funktion gewonnen werden. Durch k -fache Differentiation von (4.58) nach der Frequenz f ergibt sich:

$$\frac{d^k \Phi_x(f)}{df^k} = \int_{-\infty}^{\infty} (j2\pi x)^k \cdot f_x(x) e^{j2\pi fx} dx. \quad (4.61)$$

Für $f = 0$ lässt sich daraus das k -te Moment der Zufallsvariablen x wie folgt berechnen:

$$\mu_{x,k} = E\{x^k\} = \int_{-\infty}^{\infty} x^k f_x(x) dx = \frac{1}{(j2\pi)^k} \left. \frac{d^k \Phi_x(f)}{df^k} \right|_{f=0}. \quad (4.62)$$

- **Addition von Zufallsvariablen:** Werden n unabhängige Zufallsvariable x_i addiert, erhält man die Dichte der Summe $x = \sum_{i=1}^n x_i$ durch Faltung der einzelnen Wahrscheinlichkeitsdichten. Die Faltung entspricht einer Multiplikation im Frequenzbereich, so dass man für die charakteristische Funktion eines Summensignals folgenden Ausdruck erhält:

$$\Phi_x(f) = E\{e^{j2\pi f(\sum_i x_i)}\} = E\{e^{j2\pi f x_1}\} \cdot \dots \cdot E\{e^{j2\pi f x_n}\} \quad (4.63)$$

$$= \prod_{i=1}^n \Phi_{x_i}(f). \quad (4.64)$$

4.2 Stichproben

In der Praxis sind die Größen aus der Wahrscheinlichkeitstheorie meist nicht bekannt. Vor allem die Wahrscheinlichkeitsdichte $f_x(x)$, aus der sich Kenn-

größen wie der Mittelwert μ_x und die Varianz σ_x^2 berechnen lassen, ist nicht gegeben. Man muss sich daher mit einer Stichprobe behelfen. Eine Stichprobe ist ein Zufallsexperiment, bei dem n Messwerte $x_i, i \in \{1, \dots, n\}$, aus einer Grundgesamtheit zur weiteren statistischen Analyse herangezogen werden. Mit den Werten x_i versucht man, Schätzwerte für die zugrundeliegende Wahrscheinlichkeitsdichte, den Mittelwert und die Varianz der Grundgesamtheit zu ermitteln.

4.2.1 Häufigkeitsverteilung und Histogramm

Liegt die Wahrscheinlichkeitsdichte $f_x(x)$ einer Messgröße x nicht vor, so kann man diese anhand einer repräsentativen Stichprobe schätzen. Das Ergebnis der Schätzung ist eine empirische **Häufigkeitsverteilung**, die tabellarisch oder grafisch in Form eines sogenannten **Histogramms** (Abb. 4.9) angegeben werden kann. Zur Darstellung des Histogramms müssen die Elemente x_i der Stichprobe nach Größenklassen der Breite Δx sortiert werden,

$$\nu \cdot \Delta x \leq x_i < (\nu + 1)\Delta x, \quad (4.65)$$

wobei ν den Klassenindex bezeichnet. Von den n Stichprobenelementen werden diejenigen n_ν der Klasse ν zugeordnet, deren Werte x_i im Intervall (4.65) liegen. Die relative Häufigkeit n_ν/n der Messwerte in der Klasse ν

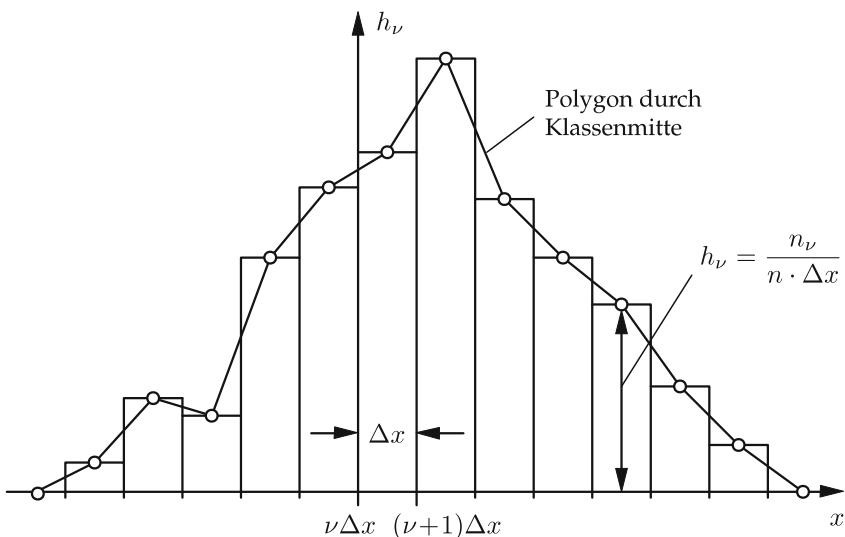


Abbildung 4.9. Beispiel für ein Histogramm.

bezogen auf die Klassenbreite Δx ergibt die Häufigkeitsverteilung

$$h_\nu = \frac{n_\nu}{n \cdot \Delta x}, \quad (4.66)$$

die durch den Bezug unabhängig von der Klassenbreite wird. Die Gesamtzahl aller Messwerte ist

$$n = \sum_{\nu=1}^m n_\nu. \quad (4.67)$$

Der in Klassen einzuteilende Bereich von x sollte alle Messwerte umfassen. Die Klassenbreite Δx ist so zu wählen, dass der Polygonzug durch die Klassenmitte einigermaßen „glatt“ ist (Abb. 4.9). Für normalverteilte Zufallsgrößen (Abschn. 4.3) gestattet die folgende Formel eine optimale Wahl der Klassenbreite im Sinne des mittleren quadratischen Fehlers [44]:

$$\Delta x = \frac{3,49 s_x}{\sqrt[3]{n}}, \quad (4.68)$$

wobei s_x die Standardabweichung der Stichprobe bezeichnet (Abschn. 4.2.3).

Für jedes Histogramm ist die Fläche A zwischen Kurve und Abszisse gleich eins. Die Treppenkurve in Abb. 4.9 ist flächengleich mit dem Polygonzug:

$$A = \sum_{\nu=1}^m h_\nu \Delta x = \sum_{\nu=1}^m \frac{n_\nu}{n \cdot \Delta x} \Delta x = \frac{1}{n} \sum_{\nu=1}^m n_\nu = 1. \quad (4.69)$$

Sind einem konstanten Signal keine Störungen überlagert, dann fallen alle Messwerte der Stichprobe in eine Klasse. Bei beträchtlichen Schwankungen der Messwerte wird das Histogramm breiter und flacher. Die Breite des Histogramms ist also ein Maß für die Streubreite.

4.2.2 Stichprobenmittelwert

Ist die Wahrscheinlichkeitsverteilung einer Zufallsvariablen nicht bekannt, so kann der Mittelwert nicht nach (4.30) berechnet werden. In der Praxis wird man vielmehr eine Stichprobe mit einer begrenzten Anzahl von Messwerten haben. Daraus kann man den **Stichprobenmittelwert** berechnen.

Definition 4.18: Stichprobenmittelwert

Der Stichprobenmittelwert aus n Werten x_i berechnet sich zu

$$\hat{x} = \frac{1}{n} \sum_{i=1}^n x_i . \quad (4.70)$$

Der Stichprobenmittelwert \hat{x} ist ein Schätzwert² des wahren Mittelwertes μ_x und ist damit selbst wieder eine stochastische Größe. \diamond

Häufig stellt sich die Frage, wie gut eine Schätzfunktion (auch: **Schätzer**) überhaupt ist. Zur Bewertung von Schätzern werden folgende Kriterien herangezogen: **Erwartungstreue**, **Konsistenz** und **Effizienz**. Dabei wird im Folgenden davon ausgegangen, dass die zu schätzende Größe deterministisch und konstant ist.

Definition 4.19: Erwartungstreue

Einen Schätzer \hat{y} nennt man **erwartungstreu**, wenn bei wiederholten Stichproben der wahre Wert y_w im Mittel richtig geschätzt wird:

$$E\{\hat{y}\} = y_w . \quad (4.71)$$

Die Differenz zwischen dem Erwartungswert des Schätzers $E\{\hat{y}\}$ und dem wahren Wert y_w ist der systematische Fehler (engl. *bias*). Erwartungstreue Schätzer weisen somit keinen systematischen Fehler auf. \diamond

Definition 4.20: Konsistenz

Ein Schätzer \hat{y} heißt **konsistent**, wenn mit wachsendem Stichprobenumfang n der wahre Wert y_w mit Sicherheit ermittelt wird,

$$\lim_{n \rightarrow \infty} \hat{y}_n = y_w , \quad (4.72)$$

und damit die Varianz des Schätzers gegen null geht:

$$\lim_{n \rightarrow \infty} \sigma_{\hat{y}}^2 = 0 . \quad (4.73)$$

Konsistente Schätzer werden somit mit wachsendem Stichprobenumfang n immer genauer. \diamond

²Schätzwerte werden oft mit einem Dach (^) gekennzeichnet.

Definition 4.21: Effizienz

Ein Schätzer \hat{y} heißt **effizient** (auch: wirksam), wenn er aus allen erwartungstreuen Schätzern die kleinste Varianz hat. \diamond

Nun soll der Stichprobenmittelwert (4.70) auf Erwartungstreue und Konsistenz hin überprüft werden. Der Erwartungswert ist

$$E\{\hat{x}\} = E\left\{\frac{1}{n} \sum_{i=1}^n x_i\right\} = \frac{1}{n} \sum_{i=1}^n \underbrace{E\{x_i\}}_{\mu_x} = \frac{1}{n} n \mu_x = \mu_x. \quad (4.74)$$

Der Erwartungswert des Stichprobenmittelwertes \hat{x} ist gerade der wahre Mittelwert μ_x . Die Schätzung \hat{x} von μ_x ist somit erwartungstreu. Die Varianz des Stichprobenmittelwertes ist

$$\sigma_{\hat{x}}^2 = E\{(\hat{x} - \mu_x)^2\} \quad (4.75)$$

$$= E\left\{\left[\left(\frac{1}{n} \sum_{i=1}^n x_i\right) - \mu_x\right]^2\right\} = E\left\{\left[\frac{1}{n} \sum_{i=1}^n (x_i - \mu_x)\right]^2\right\} \quad (4.76)$$

$$= \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n E\{(x_i - \mu_x)(x_j - \mu_x)\} \quad (4.77)$$

$$= \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n C_{x_i x_j}. \quad (4.78)$$

Für die Kovarianz $C_{x_i x_j}$ lassen sich wieder die beiden Extremfälle der starren Bindung und der statistischen Unabhängigkeit unterscheiden.

1. Starre Bindung zwischen x_i und x_j :

Für gleiche Varianzen von x_i und x_j gilt nach (4.51)

$$C_{x_i x_j} = \sigma_{x_i} \sigma_{x_j} = \sigma_x^2. \quad (4.79)$$

Damit erhält man

$$\sigma_{\hat{x}}^2 = \frac{1}{n^2} n^2 \sigma_x^2 \quad \Rightarrow \quad \sigma_{\hat{x}}^2 = \sigma_x^2. \quad (4.80)$$

Bei starrer Bindung der Messwerte ist die Varianz des Stichprobenmittelwertes gleich der Varianz der Messwerte. Mehrere Messwerte enthalten somit nicht mehr Information als ein einziger Messwert. Die Schätzung ist in diesem Fall *nicht konsistent*.

2. Statistisch unabhängige Messwerte x_i und x_j für $i \neq j$:

Nach (4.38) gilt:

$$C_{x_i x_j} = \sigma_x^2 \delta_{ij}. \quad (4.81)$$

Die Varianz des Stichprobenmittelwertes wird damit

$$\sigma_{\hat{x}}^2 = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \sigma_x^2 \delta_{ij} = \frac{1}{n^2} \sum_{i=1}^n \sigma_x^2 = \frac{\sigma_x^2}{n}. \quad (4.82)$$

Für statistisch unabhängige Messgrößen nimmt die Varianz des Stichprobenmittelwertes $\sigma_{\hat{x}}^2$ mit wachsendem Stichprobenumfang n gegen null ab. Der Stichprobenmittelwert \hat{x} strebt dann gegen den wahren Mittelwert μ_x . Die Schätzung ist *konsistent*.

Die Abnahme der Standardabweichung eines Schätzers mit $1/\sqrt{n}$ gemäß (4.82) ist typisch für die meisten praktisch relevanten Aufgabenstellungen.

4.2.3 Stichprobenvarianz

Bei den meisten messtechnischen Aufgaben ist die Wahrscheinlichkeitsdichte $f_x(x)$ nicht bekannt. Es liegen lediglich einzelne Messwerte x_i und der Stichprobenmittelwert \hat{x} vor. Die unbekannte Varianz σ_x^2 wird durch die **Stichprobenvarianz** $s_x^2 = \hat{\sigma}_x^2$ geschätzt.

Definition 4.22: Stichprobenvarianz

4.22

Die Stichprobenvarianz ist wie folgt definiert:

$$s_x^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \hat{x})^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i^2 - 2x_i \hat{x} + \hat{x}^2) \quad (4.83)$$

$$= \frac{1}{n-1} \sum_{i=1}^n x_i^2 - 2\hat{x} \frac{1}{n-1} \sum_{i=1}^n x_i + \frac{1}{n-1} n \hat{x}^2 \quad (4.84)$$

$$= \frac{1}{n-1} \sum_{i=1}^n x_i^2 - \frac{n}{n-1} \hat{x}^2. \quad (4.85)$$

Ihre Wurzel s_x wird Standardabweichung der Stichprobe genannt. ◇

Im Folgenden soll der Zusammenhang zwischen der Stichprobenvarianz s_x^2 und der wahren Varianz σ_x^2 der Verteilung betrachtet werden. Die Stichpro-

Stichprobenvarianz s_x^2 selbst ist wieder eine Zufallsgröße. Ihr Erwartungswert ist

$$E\{s_x^2\} = E\left\{\frac{1}{n-1} \sum_{i=1}^n (x_i - \hat{x})^2\right\} \quad (4.86)$$

$$= \frac{1}{n-1} E\left\{\sum_{i=1}^n ((x_i - \mu_x) - (\hat{x} - \mu_x))^2\right\} \quad (4.87)$$

$$\begin{aligned} &= \frac{1}{n-1} E\left\{\sum_{i=1}^n (x_i - \mu_x)^2 - 2 \sum_{i=1}^n (x_i - \mu_x)(\hat{x} - \mu_x) + \right. \\ &\quad \left. + n(\hat{x} - \mu_x)^2\right\}. \end{aligned} \quad (4.88)$$

Formt man den Stichprobenmittelwert (4.70) um zu

$$n(\hat{x} - \mu_x) = \sum_{i=1}^n (x_i - \mu_x) \quad (4.89)$$

und setzt dies in die obige Gleichung ein, so erhält man

$$\begin{aligned} E\{s_x^2\} &= \frac{1}{n-1} \left[\underbrace{\sum_{i=1}^n E\{(x_i - \mu_x)^2\}}_{\sigma_x^2} - 2n \underbrace{E\{(\hat{x} - \mu_x)^2\}}_{\sigma_{\hat{x}}^2} \right. \\ &\quad \left. + n \underbrace{E\{(\hat{x} - \mu_x)^2\}}_{\sigma_{\hat{x}}^2} \right]. \end{aligned} \quad (4.90)$$

Mit (4.77) folgt schließlich der Erwartungswert der Stichprobenvarianz

$$E\{s_x^2\} = \frac{n}{n-1} (\sigma_x^2 - \sigma_{\hat{x}}^2). \quad (4.91)$$

Man kann wieder die beiden Extremfälle der Messwertstatistik unterscheiden.

1. Starre Bindung zwischen x_i und x_j :

Nach (4.80) gilt $\sigma_{\hat{x}}^2 = \sigma_x^2$ und damit

$$E\{s_x^2\} = 0. \quad (4.92)$$

Der Erwartungswert der Stichprobenvarianz ist bei starrer Bindung der Messwerte null. Das liegt daran, dass der Stichprobenmittelwert \hat{x} die gleiche Varianz wie die Messwerte selbst aufweist. Daher ist in diesem Fall die Stichprobenvarianz als Schätzung der wahren Varianz unbrauchbar.

2. Statistisch unabhängige Messwerte x_i und x_j für $i \neq j$:

Nach (4.82) gilt $s_x^2 = \sigma_x^2/n$ und damit

$$E\{s_x^2\} = \frac{n}{n-1} \sigma_x^2 \left(1 - \frac{1}{n}\right) = \sigma_x^2. \quad (4.93)$$

Für statistisch unabhängige Messwerte ist die Stichprobenvarianz s_x^2 eine erwartungstreue Schätzung für die Varianz σ_x^2 der Verteilung.

Damit die Schätzung der Stichprobenvarianz erwartungstreue ist, wurde diese mit dem Faktor $1/(n-1)$ anstelle von $1/n$ versehen. Bei Einzelmessungen ($n=1$) kann allerdings keine Stichprobenvarianz ermittelt werden. Bei der Messung abhängiger Werte ist die Stichprobenvarianz s_x^2 kleiner oder höchstens gleich der Varianz der Wahrscheinlichkeitsverteilung σ_x^2 .

Beispiel 4.7 (Abweichung der Stichprobenvarianz): Ein Messgerät zeige im Prüffeld eine sehr geringe Stichprobenvarianz. Beim eigentlichen Einsatz in einer verfahrenstechnischen Anlage liege dagegen die Stichprobenvarianz deutlich höher. Dies kann als Indiz für einen stochastischen Fehler gewertet werden, der nur in der Anlage und nicht im Prüffeld auftritt. Die Messwerte im Prüffeld sind dann weniger voneinander unabhängig als die Messwerte in der Anlage. Die im Prüffeld ermittelte Stichprobenvarianz s_x^2 schätzt die Varianz σ_x^2 der Messwerte im Prozess zu klein. ■

Manchmal sind auch die höheren Momente von Interesse. Die Schiefe (4.32) kann aus einer Stichprobe wie folgt erwartungstreue geschätzt werden:

$$\hat{\varrho}_x = \frac{1}{s_x^3} \cdot \frac{n}{(n-1)(n-2)} \sum_{i=1}^n (x_i - \hat{x})^3. \quad (4.94)$$

Sie ist ein Maß für die Asymmetrie der Messwerteverteilung und wird bei symmetrischen Verteilungen zu null.

Der Exzess (4.33) ist ein Maß für die Wölbung der Messwerteverteilung. Dessen erwartungstreuer Schätzwert lautet:

$$\hat{\varepsilon}_x = \frac{1}{s_x^4} \cdot \frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum_{i=1}^n (x_i - \hat{x})^4 - \frac{3(n-1)^2}{(n-2)(n-3)}. \quad (4.95)$$

② Numerische Berechnung von Mittelwert und Varianz

Bei der numerischen Berechnung des Mittelwertes und der Varianz ist es oft vorteilhaft, anstelle einer Addition vieler großer Zahlen x_i mit den Abwei-

chungen Δx_i von einem Näherungswert x_0 zu rechnen:

$$\Delta x_i = x_i - x_0 . \quad (4.96)$$

Der Stichprobenmittelwert (4.70) wird dann

$$\hat{x} = x_0 + \frac{1}{n} \sum_{i=1}^n \Delta x_i \quad \Rightarrow \quad \Delta \hat{x} = \hat{x} - x_0 = \frac{1}{n} \sum_{i=1}^n \Delta x_i . \quad (4.97)$$

Die Stichprobenvarianz (4.83) wird durch Ausmultiplizieren des Binoms und Einsetzen der Ausdrücke für x_i und \hat{x} zu

$$s_x^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \hat{x})^2 \quad (4.98)$$

$$= \frac{1}{n-1} \left(\sum_{i=1}^n x_i^2 - n \hat{x}^2 \right) \quad (4.99)$$

$$= \frac{1}{n-1} \left(\sum_{i=1}^n (x_0 + \Delta x_i)^2 - n (x_0 + \Delta \hat{x})^2 \right) \quad (4.100)$$

$$= \frac{1}{n-1} \left(nx_0^2 + 2x_0 \underbrace{\sum_{i=1}^n \Delta x_i}_{n \Delta \hat{x}} + \sum_{i=1}^n (\Delta x_i)^2 - nx_0^2 - 2nx_0 \Delta \hat{x} - n(\Delta \hat{x})^2 \right) \quad (4.101)$$

$$= \frac{1}{n-1} \left(\sum_{i=1}^n (\Delta x_i)^2 - n(\Delta \hat{x})^2 \right) . \quad (4.102)$$

Dadurch wird die Stichprobenvarianz auf die Quadratsumme der Abweichungen Δx_i und den Mittelwert der Abweichungen $\Delta \hat{x}$ zurückgeführt.

4.2.4 Gesetz der großen Zahlen

Die Wahrscheinlichkeitsdichten können nur selten aus den Versuchsbedingungen hergeleitet werden. Öfter kann zumindest aus dem Histogramm der Typ der Verteilung angegeben werden. Die für die Verteilung wichtigen Parameter müssen dann aus Stichproben experimentell bestimmt werden. Die Ergebnisse der Wahrscheinlichkeitsrechnung, hergeleitet mit Hilfe der Mengen- und Maßtheorie, gelten streng. Die Verbindung dieser Theorie zu Messergebnissen aus Stichproben geschieht über verschiedene Grenzwertsätze, z. B. über das **Bernoulli'sche Gesetz der großen Zahlen** [7, 50].

Es sei x eine Zufallsvariable mit der Wahrscheinlichkeitsdichte $f_x(x)$. Die Wahrscheinlichkeit für das Ereignis $|x - \mu_x| \geq \varepsilon$ ist

$$P(|x - \mu_x| \geq \varepsilon) = \int_{-\infty}^{\mu_x - \varepsilon} f_x(x) dx + \int_{\mu_x + \varepsilon}^{\infty} f_x(x) dx. \quad (4.103)$$

Mit $|x - \mu_x| \geq \varepsilon$ ist

$$\frac{(x - \mu_x)^2}{\varepsilon^2} \geq 1, \quad (4.104)$$

so dass man damit die Wahrscheinlichkeit abschätzen kann:

$$\begin{aligned} P(|x - \mu_x| \geq \varepsilon) &\leq \int_{-\infty}^{\mu_x - \varepsilon} \frac{(x - \mu_x)^2}{\varepsilon^2} f_x(x) dx + \int_{\mu_x + \varepsilon}^{\infty} \frac{(x - \mu_x)^2}{\varepsilon^2} f_x(x) dx \\ &\leq \frac{1}{\varepsilon^2} \int_{-\infty}^{\infty} (x - \mu_x)^2 f_x(x) dx = \frac{\sigma_x^2}{\varepsilon^2}. \end{aligned} \quad (4.105)$$

Die Ungleichung (4.105) ist die **Tschebyscheff'sche Ungleichung**. Sie besagt, dass für eine Zufallsvariable x mit endlicher Varianz σ_x^2 die Realisierungen x mit einer gewissen Wahrscheinlichkeit um den Erwartungswert μ_x liegen. Diese Ungleichung kann nun auf den Stichprobenmittelwert \hat{x} als Zufallsvariable mit der Varianz (4.82) angewandt werden:

$$P(|\hat{x} - \mu_x| \geq \varepsilon) \leq \frac{\sigma_{\hat{x}}^2}{\varepsilon^2} = \frac{\sigma_x^2}{n \cdot \varepsilon^2}. \quad (4.106)$$

Mit größer werdendem Stichprobenumfang n strebt die Wahrscheinlichkeit $P(|\hat{x} - \mu_x| \geq \varepsilon)$ gegen null, dass die Schätzung \hat{x} um mehr als die beliebig kleine Schranke ε vom wahren Mittelwert μ_x abweicht. Die Versuchsergebnisse aus großen Stichproben nähern sich also den Ergebnissen der Wahrscheinlichkeitsrechnung.

Einen entsprechenden Zusammenhang kann man zwischen der Häufigkeitsverteilung einer Stichprobe $h(x)$ und der Wahrscheinlichkeitsdichte $f_x(x)$ herstellen. Dazu werden Indikatorvariablen $J_{\nu i}$ definiert, die beschreiben, ob ein Ereignis x_i einer bestimmten Klasse ν angehört:

$$J_{\nu i} = \begin{cases} 1 & \text{für } \nu \Delta x \leq x_i < (\nu + 1) \Delta x \\ 0 & \text{sonst} \end{cases}. \quad (4.107)$$

Die Ereignisse seien statistisch unabhängig. Die Wahrscheinlichkeit dafür, dass ein Ereignis der Klasse ν angehört, wird nach (4.66) durch die relative

Häufigkeit geschätzt und lässt sich als Stichprobenmittelwert der Indikatorvariablen $J_{\nu i}$ darstellen:

$$\Delta x h_\nu = \frac{n_\nu}{n} = \frac{1}{n} \sum_{i=1}^n J_{\nu i}. \quad (4.108)$$

Durch Erwartungswertsbildung erhält man die Wahrscheinlichkeit $f_x(x_\nu) \Delta x$:

$$E\{\Delta x h_\nu\} = E\left\{\frac{1}{n} \sum_{i=1}^n J_{\nu i}\right\} = \frac{1}{n} \sum_{i=1}^n \underbrace{E\{J_{\nu i}\}}_{f_x(x_\nu) \Delta x} = f_x(x_\nu) \Delta x. \quad (4.109)$$

Mit dem Mittelwertsatz der Integralrechnung lässt sich zeigen, dass im Intervall $[\nu \Delta x, (\nu + 1) \Delta x]$ ein x_ν existiert, das obige Gleichung erfüllt.

Der Schätzer für die Häufigkeitsverteilung $h(x)$ hat die gleiche Struktur wie der Schätzer für den Stichprobenmittelwert. Für die Varianz der Häufigkeitsverteilung $h(x)$ erhält man daher bei unabhängigen Ereignissen nach (4.82):

$$E\{(h(x) - f_x(x))^2\} = \frac{\sigma_J^2}{n}. \quad (4.110)$$

Einsetzen von $h(x)$ in die Tschebyscheff'sche Ungleichung (4.105) ergibt das **Bernoulli'sche Gesetz der großen Zahlen**:

$$P(|h(x) - f_x(x)| \geq \varepsilon) \leq \frac{1}{\varepsilon^2} E\{(h(x) - f_x(x))^2\} = \frac{\sigma_J^2}{n \varepsilon^2}. \quad (4.111)$$

Mit wachsendem Stichprobenumfang n geht also die Häufigkeitsverteilung $h(x)$ in die Wahrscheinlichkeitsdichte $f_x(x)$ über.

② 4.2.5 Mittelung zur Störungsunterdrückung

Vielfach sind einer deterministischen Messgröße u zufällige Störungen e additiv überlagert. Die Unterdrückung derartiger Störungen kann durch eine **Mittelung** von n Messwerten y_i nach Def. 4.18 erfolgen. Um diesen Ansatz zu untersuchen, soll unterschien werden, ob die Kennlinie des Messsystems linear ist oder nicht.

③ 4.2.5.1 Lineare Kennlinie

Zunächst sei angenommen, dass das Messsystem eine *lineare Kennlinie* mit der Empfindlichkeit S_i aufweist. Der Anzeigewert lautet dann:

$$y = S_i(u + e). \quad (4.112)$$

Um das Ergebnis der Mittelung zu untersuchen, wird der Erwartungswert von y gebildet:

$$\mu_y = E\{y\} = E\{S_i(u + e)\} = S_i(u + \mu_e). \quad (4.113)$$

Es können zwei Fälle unterschieden werden:

1. Die Störung e ist mittelwertfrei:

$$E\{e\} = \mu_e = 0. \quad (4.114)$$

Der Mittelwert des Ausgangssignals μ_y entspricht dem idealen Anzeigewert $S_i u$. Durch die Mittelung wird also das Störsignal e unterdrückt. Die Mittelung ist die einfachste Methode zur Bestimmung des wahren Wertes der Messgröße, wenn diese von mittelwertfreien Störungen überlagert ist.

2. Die Störung habe einen endlichen Mittelwert:

$$E\{e\} = \mu_e \neq 0. \quad (4.115)$$

Der Mittelwert der Störung $S_i \mu_e$ am Ausgang des Messsystems kann als deterministische additive Störung (systematischer Fehler) vom Ausgangssignal subtrahiert werden:

$$\tilde{y} = y - S_i \mu_e = S_i(u + e - \mu_e). \quad (4.116)$$

Die in \tilde{y} verbleibende Störung $S_i(e - \mu_e)$ ist dann wieder mittelwertfrei, so dass eine Mittelung zur Störungsunterdrückung herangezogen werden kann.

Bei linearen Messkennlinien können also durch eine Mittelung überlagerte mittelwertfreie Störungen aus dem Ausgangssignal herausgefiltert werden. Dies ist allerdings bei *nichtlinearen Messkennlinien* oder bei deformierenden (multiplikativen) Störgrößen nicht der Fall.

④ 4.2.5.2 Nichtlineare Kennlinie

In Abb. 4.10 ist das Eingangssignal u des Messsystems von einer mittelwertfreien Störung e überlagert. Das Messsystem habe nun eine gekrümmte Kennlinie $f(u)$ mit der Empfindlichkeit S , die um den Arbeitspunkt u_0 herum in eine Taylor-Reihe entwickelt wird:

$$\Delta y = S(u_0)(\Delta u + e) \cdot \left[1 + \frac{1}{2} \frac{S'(u_0)}{S(u_0)} (\Delta u + e) + \dots \right]. \quad (4.117)$$

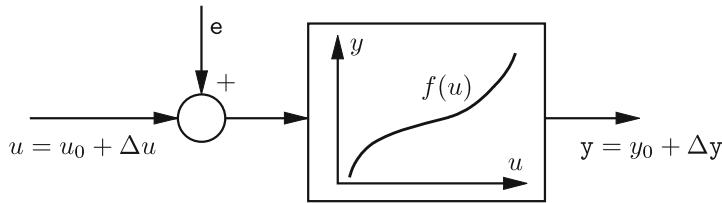


Abbildung 4.10. Störung bei nichtlinearer Kennlinie.

Bei Unkorreliertheit von Δu und e ergibt sich für den Erwartungswert der Ausgangsgröße Δy die folgende Näherung:

$$\mu_{\Delta y} = E\{\Delta y\} \approx S(u_0) [\Delta u + E\{e\}] + \frac{1}{2} S'(u_0) [\Delta u^2 + E\{e^2\}] \quad (4.118)$$

$$= S(u_0) \cdot \Delta u + \frac{1}{2} S'(u_0) \cdot (\Delta u^2 + \sigma_e^2). \quad (4.119)$$

Obwohl die Störung e mittelwertfrei ist, weicht der Mittelwert der Ausgangsgröße $\mu_{\Delta y}$ bei gekrümmter Kennlinie vom idealen Anzeigewert ab. Auch im Arbeitspunkt u_0 ($\Delta u = 0$) ist der Erwartungswert der Messabweichung $\mu_{\Delta y} \neq 0$. Die Abweichung ist proportional zur Kennlinienkrümmung $S'(u_0)$ im Arbeitspunkt und zur Varianz σ_e^2 des Störsignals. Möchte man Störungen des Messsignals durch eine Mittelwertbildung unterdrücken, so muss zuvor die Messkennlinie linearisiert werden.

4.3

4.3 Normalverteilte Zufallsvariable

In der praktischen Anwendung spielen normalverteilte Zufallsvariable eine große Rolle. So wird bei unbekannter Wahrscheinlichkeitsdichte in vielen Fällen berechtigterweise eine **Normalverteilung** angenommen. Der Grund liegt im zentralen Grenzwertsatz (Abschn. 4.3.2). Davor sollen einige wichtige Eigenschaften normalverteilter Zufallsvariabler behandelt werden.

4.3.1 Normalverteilung

Im Folgenden wird eine Zufallsvariable x mit dem Mittelwert μ_x und der Varianz σ_x^2 betrachtet.

Definition 4.23: Normalverteilung

Eine Zufallsvariable $x \sim \mathcal{N}(\mu_x, \sigma_x^2)$ mit der Wahrscheinlichkeitsdichte³

$$f_x(x) = \mathcal{N}(\mu_x, \sigma_x^2) = \frac{1}{\sigma_x \sqrt{2\pi}} \exp\left(-\frac{(x - \mu_x)^2}{2\sigma_x^2}\right) \quad (4.120)$$

heißt normalverteilt. Eine normalverteilte Zufallsgröße wird durch die zwei Momente Mittelwert μ_x und Varianz σ_x^2 vollständig charakterisiert. \diamond

Die Momente der Normalverteilung lassen sich über die charakteristische Funktion (4.58) berechnen. Für die Normalverteilung (4.120) lautet diese:

$$\Phi_x(f) = \exp\left(j2\pi f \mu_x - \frac{1}{2}(2\pi f \sigma_x)^2\right). \quad (4.121)$$

Der Mittelwert ergibt sich nach (4.62) zu:

$$E\{x\} = \frac{1}{(j2\pi)} \left. \frac{d\Phi_x(f)}{df} \right|_{f=0} = \mu_x. \quad (4.122)$$

Für das zweite Moment erhält man analog:

$$E\{x^2\} = \frac{1}{(j2\pi)^2} \left. \frac{d^2\Phi_x(f)}{df^2} \right|_{f=0} = \sigma_x^2 + \mu_x^2. \quad (4.123)$$

Alle höheren Momente lassen sich auf die Parameter μ_x und σ_x zurückführen. Daher ist die Normalverteilung alleine durch den Mittelwert μ_x und die Varianz σ_x^2 (oder Standardabweichung σ_x) der Zufallsvariablen x bestimmt.

Eine auf den Mittelwert $\mu_x = 0$ und die Varianz $\sigma_x^2 = 1$ normierte Normalverteilung nennt man **Standardnormalverteilung**

$$f_x(x) = \mathcal{N}(0, 1) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right). \quad (4.124)$$

Wichtig ist der Zusammenhang, dass jede lineare Transformation einer normalverteilten Zufallsvariablen x ,

$$z = a x + b \quad \text{mit} \quad a, b \in \mathbb{R}, \quad (4.125)$$

³Man sagt: „ x ist verteilt nach $\mathcal{N}(\mu_x, \sigma_x^2)$ “ und schreibt $x \sim \mathcal{N}(\mu_x, \sigma_x^2)$.

wieder eine normalverteilte Zufallsvariable z ergibt. Mit der Transformation⁴

$$z = \frac{x - \mu_x}{\sigma_x} \quad (4.126)$$

kann daher eine Normalverteilung $\mathcal{N}(\mu_x, \sigma_x^2)$ in eine Standardnormalverteilung $\mathcal{N}(0, 1)$ überführt werden. Dies ist in der Praxis häufig nötig, da das Integral der Dichtefunktion (4.120) nicht geschlossen dargestellt werden kann. Zur Berechnung des Integrals werden Tabellen oder Softwareprogramme verwendet, bei denen meist eine Standardnormalverteilung vorausgesetzt wird.

4.24

Definition 4.24: Mehrdimensionale Normalverteilung

Eine mehrdimensionale Zufallsvariable $\mathbf{x} = (x_1, \dots, x_d)^T \in \mathbb{R}^d$ mit der Wahrscheinlichkeitsdichte

$$f_{\mathbf{x}}(\mathbf{x}) = \mathcal{N}(\boldsymbol{\mu}_{\mathbf{x}}, \boldsymbol{\Sigma}_{\mathbf{x}}) \quad (4.127)$$

$$= \frac{1}{(2\pi)^{d/2} |\boldsymbol{\Sigma}_{\mathbf{x}}|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_{\mathbf{x}})^T \boldsymbol{\Sigma}_{\mathbf{x}}^{-1}(\mathbf{x} - \boldsymbol{\mu}_{\mathbf{x}})\right) \quad (4.128)$$

heißt multivariat normalverteilt. Eine mehrdimensionale Normalverteilung wird durch den Mittelwertvektor $\boldsymbol{\mu}_{\mathbf{x}}$ und die **Kovarianzmatrix** $\boldsymbol{\Sigma}_{\mathbf{x}}$ vollständig charakterisiert. Die Punkte gleicher Wahrscheinlichkeitsdichte beschreiben Ellipsoide in d Dimensionen. \diamond

Die Kovarianzmatrix $\boldsymbol{\Sigma}_{\mathbf{x}}$ ist stets symmetrisch und positiv semidefinit. Die Elemente auf der Hauptdiagonalen sind die Varianzen $\sigma_{x_i}^2$ der einzelnen Komponenten x_i :

$$\boldsymbol{\Sigma}_{\mathbf{x}} = \begin{pmatrix} C_{x_1 x_1} & \cdots & C_{x_1 x_d} \\ \vdots & \ddots & \vdots \\ C_{x_d x_1} & \cdots & C_{x_d x_d} \end{pmatrix} = \begin{pmatrix} \sigma_{x_1}^2 & \cdots & \rho_{x_1 x_d} \sigma_{x_1} \sigma_{x_d} \\ \vdots & \ddots & \vdots \\ \rho_{x_d x_1} \sigma_{x_d} \sigma_{x_1} & \cdots & \sigma_{x_d}^2 \end{pmatrix}, \quad (4.129)$$

wobei $\rho_{x_i x_j}$ den Korrelationskoeffizienten (4.41) zwischen x_i und x_j bezeichnet. Die Determinante $|\boldsymbol{\Sigma}_{\mathbf{x}}|$ ist proportional zur Größe der Ellipsoide und somit ein Maß für die Streuung von \mathbf{x} . Ferner beschreiben die Eigenvektoren der Kovarianzmatrix $\boldsymbol{\Sigma}_{\mathbf{x}}$ die Richtung der Hauptachsen der Ellipsoide;

⁴Abweichend von der z-Transformation für zeitdiskrete Signale [42] wird in der mathematischen Statistik die Transformation (4.126) zur Standardisierung von Zufallsvariablen ebenfalls als z-Transformation bezeichnet.

die zugehörigen Eigenwerte entsprechen den Varianzen in Hauptachsenrichtung.

4.3.2 Zentraler Grenzwertsatz

In vielen Anwendungen resultiert der zufällige Fehler e aus einer additiven Überlagerung zahlreicher unabhängiger, zufälliger Ereignisse e_n mit unbekannten Wahrscheinlichkeitsdichten:

$$e = \sum_{n=1}^N e_n . \quad (4.130)$$

Die Wahrscheinlichkeitsdichte der Summe von unabhängigen Zufallsvariablen x_n kann nach Abschn. 4.1.6 allgemein über die Faltung der einzelnen Wahrscheinlichkeitsdichten berechnet werden:

$$f_e(e) = f_{e_1}(e) * f_{e_2}(e) * \cdots * f_{e_N}(e) \quad (4.131)$$

$$\Phi_e(f) = \prod_{n=1}^N \Phi_{e_n}(f) . \quad (4.132)$$

Zur Berechnung der Wahrscheinlichkeitsdichte von e ist der sogenannte **zentrale Grenzwertsatz** von großer Bedeutung [37].

Satz 4.3: Zentraler Grenzwertsatz

4.3

Haben die Zufallsvariablen x_n Verteilungen mit beschränktem zweiten und dritten Moment und sind die Zufallsvariablen x_n voneinander unabhängig, dann nähert sich die Dichte $f_x(x)$ der Summe

$$x = \sum_{n=1}^N x_n \quad (4.133)$$

mit wachsendem Umfang N asymptotisch einer Normalverteilung an:

$$f_x(x) = \frac{1}{\sqrt{2\pi}\sigma_x} \exp\left(-\frac{(x - \mu_x)^2}{2\sigma_x^2}\right) . \quad (4.134)$$

Die Parameter der resultierenden Normalverteilung lassen sich dabei gemäß

$$\mu_x = \sum_{n=1}^N E\{x_n\} , \quad (4.135)$$

$$\sigma_x^2 = \sum_{n=1}^N \sigma_{x_n}^2 \quad (4.136)$$

berechnen. \diamond

Aus diesem Satz können wichtige Folgerungen getroffen werden, die die Bedeutung der Normalverteilung in der statistischen Qualitätskontrolle unterstreichen.

1. Fasst man den Wert eines Stichprobenelementes als Zufallsvariable x_n auf, so folgt, dass der Stichprobenmittelwert \hat{x} näherungsweise normalverteilt ist.
2. Entsteht ein zufälliger Messfehler durch die Überlagerung mehrerer unabhängiger Zufallereignisse, so kann für den Fehler eine Normalverteilung angenommen werden.

Beispiel 4.8 (Zentraler Grenzwertsatz): Gegeben seien $N = 4$ Zufallsvariable x_n mit Wahrscheinlichkeitsdichten nach Abb. 4.11. Berechnet man die Wahrscheinlichkeitsdichte der Summe (4.133) durch Faltung der Einzeldichten, so erhält man das in Abb. 4.12 dargestellte Resultat. Man erkennt, wie gering die Abweichung zur Normalverteilung ist. Bei einer Überlagerung beliebig verteilter Zufallsvariabler kann in den meisten Fällen näherungsweise mit einer resultierenden Normalverteilung gerechnet werden. ■

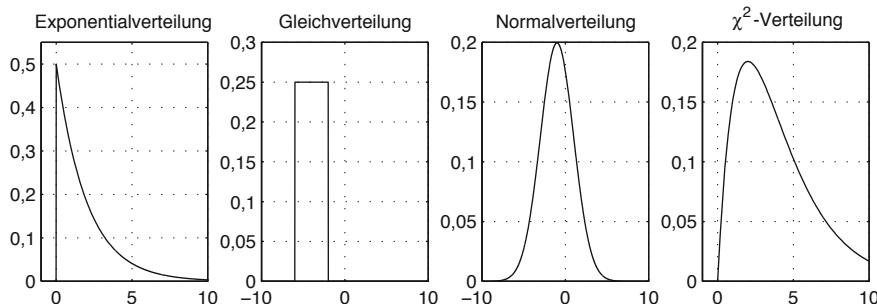


Abbildung 4.11. Wahrscheinlichkeitsdichten einzelner Zufallsvariablen.

4.3.3 χ^2 -Verteilung

Im letzten Unterabschnitt wurden Zufallsvariable additiv überlagert, wie es nach Def. 4.18 bei der Berechnung von Stichprobenmittelwerten erforder-

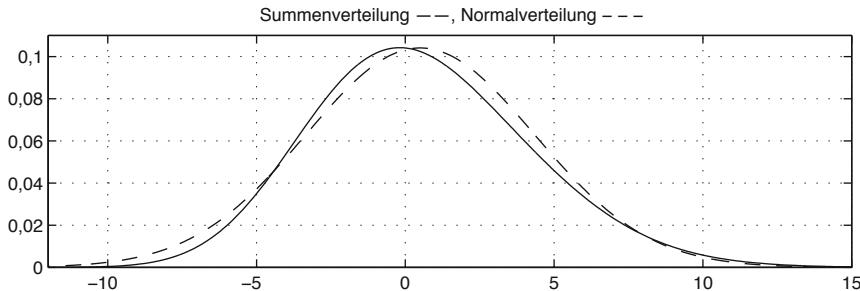


Abbildung 4.12. Wahrscheinlichkeitsdichte der Summe der Dichten aus Abb. 4.11 im Vergleich mit einer Normalverteilung.

lich ist. Für die resultierende Zufallsgröße ergab sich nach dem zentralen Grenzwertsatz näherungsweise eine Normalverteilung.

Nun sollen die Zufallsvariablen vor ihrer Addition quadriert werden. Die Wahrscheinlichkeitsdichte solcher Quadratsummen ist in der Statistik deshalb von Interesse, weil sie die Verteilung der Stichprobenvarianz beschreibt. Bei der Berechnung der Stichprobenvarianz s_x^2 werden nach (4.83) die Stichprobenwerte $x_i - \hat{x}$ quadriert und addiert. Führt man die Zufallsvariablen

$$z_i = x_i - \hat{x} \quad (4.137)$$

ein, so erhält man mit

$$y_n = z_1^2 + z_2^2 + \cdots + z_n^2 \quad (4.138)$$

eine neue Zufallsvariable y_n , die bis auf einen konstanten Faktor $1/(n-1)$ gleich der Stichprobenvarianz s_x^2 ist.

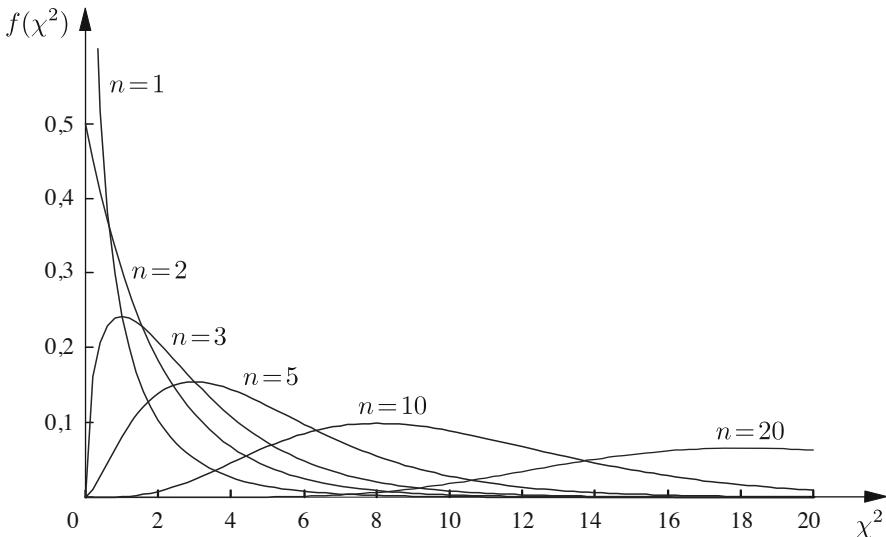
Satz 4.4: χ^2 -Verteilung

Sind n unabhängige Zufallsvariable z_i mit einer Standardnormalverteilung $\mathcal{N}(0, 1)$ gegeben, so hat die Quadratsumme

$$y_n = z_1^2 + z_2^2 + \cdots + z_n^2 \quad (4.139)$$

die folgende Wahrscheinlichkeitsdichte:

$$f_{y_n}(y = \chi^2) = \begin{cases} \frac{1}{\Gamma(\frac{n}{2}) 2^{\frac{n}{2}}} y^{\frac{n}{2}-1} e^{-\frac{y}{2}} & \text{für } \chi^2 \geq 0, \\ 0 & \text{für } \chi^2 < 0. \end{cases} \quad (4.140)$$

Abbildung 4.13. Wahrscheinlichkeitsdichte der χ^2 -Verteilung.

Diese Wahrscheinlichkeitsdichte wird **χ^2 -Verteilung** (sprich: „Chi-Quadrat-Verteilung“) genannt (Abb. 4.13). Dabei ist n ihr einziger Parameter, der die Zahl der Freiheitsgrade der Verteilung beschreibt. Die Gammafunktion $\Gamma(x)$, $x \in \mathbb{R}$, ist dabei eine Verallgemeinerung der Fakultätsfunktion und lässt sich wie letztere ebenfalls rekursiv berechnen:

$$\Gamma(x+1) = x \cdot \Gamma(x) \quad (4.141)$$

mit $\Gamma(\frac{1}{2}) = \sqrt{\pi}$ und $\Gamma(1) = 1$. \diamond

Beweis 4.2 (χ^2 -Verteilung): Der Beweis erfolgt durch vollständige Induktion. Zunächst wird die Dichte für eine einzige Zufallsvariable $y_1 = z_1^2$ berechnet. Nach Satz 4.2 benötigt man dafür die Lösung der Umkehrfunktion

$$z_1 = -z_2 = \sqrt{y}. \quad (4.142)$$

Mit der Annahme einer Standardnormalverteilung der Stichprobe

$$f_{z_1}(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) \quad (4.143)$$

folgt für die Transformation

$$f_{y_1}(y) = f_{z_1}(z_1) \left| \frac{dy(z)}{dz} \right|_{z=z_1}^{-1} + f_{z_1}(z_2) \left| \frac{dy(z)}{dz} \right|_{z=z_2}^{-1} \quad (4.144)$$

$$= 2 \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{y}{2}} \cdot |2\sqrt{y}|^{-1} = \frac{1}{\sqrt{2\pi y}} e^{-\frac{y}{2}}, \quad y \geq 0. \quad (4.145)$$

Mit $\Gamma(\frac{1}{2}) = \sqrt{\pi}$ erhält man

$$f_{y_1}(y) = \begin{cases} \frac{1}{\Gamma(\frac{1}{2}) 2^{\frac{1}{2}}} y^{-\frac{1}{2}} e^{-\frac{y}{2}} & \text{für } y \geq 0, \\ 0 & \text{sonst.} \end{cases} \quad (4.146)$$

Durch inverse Fourier-Transformation⁵ folgt die charakteristische Funktion:

$$\Phi_{y_1}(f) = \frac{1}{\Gamma(\frac{1}{2}) 2^{\frac{1}{2}}} \int_0^\infty y^{-\frac{1}{2}} e^{-(1-j4\pi f)\frac{y}{2}} dy = (1 - j4\pi f)^{-\frac{1}{2}}. \quad (4.147)$$

Für die Quadratsumme von n unabhängigen Zufallsvariablen erhält man analog die charakteristische Funktion der angenommenen Verteilung (4.140) zu

$$\Phi_{y_n}(f) = \frac{1}{\Gamma(\frac{n}{2}) 2^{\frac{n}{2}}} \int_0^\infty y^{\frac{n}{2}-1} e^{-\frac{y}{2}+j2\pi f y} dy = (1 - j4\pi f)^{-\frac{n}{2}}. \quad (4.148)$$

Der Schluss von n auf $n + 1$ basiert darauf, dass die charakteristische Funktion einer Summe von unabhängigen Zufallsvariablen dem Produkt der einzelnen charakteristischen Funktionen entspricht. Mit $y_{n+1} = y_n + y_1$ erhält man

$$\Phi_{y_{n+1}}(f) = \Phi_{y_n}(f) \cdot \Phi_{y_1}(f) = (1 - j4\pi f)^{-\frac{n+1}{2}}. \quad (4.149)$$

Dies ist aber genau die charakteristische Funktion einer χ^2 -Verteilung von $n + 1$ unabhängigen Zufallsvariablen. Man sagt, die χ^2 -Verteilung hat $n + 1$ Freiheitsgrade. ♦

⁵Zur Lösung des Fourier-Integrals ist der folgende Zusammenhang hilfreich [47]:

$$\int_0^\infty y^k e^{-ay} dy = \frac{\Gamma(k+1)}{a^{k+1}}.$$

Der Mittelwert und das zweite Moment der χ^2 -Verteilung berechnen sich über die charakteristische Funktion (vgl. Abb. 4.13):

$$E\{y_n\} = \frac{1}{j2\pi} \left. \frac{d\Phi_{y_n}(f)}{df} \right|_{f=0} = n, \quad (4.150)$$

$$E\{y_n^2\} = \frac{1}{(j2\pi)^2} \left. \frac{d^2\Phi_{y_n}(f)}{df^2} \right|_{f=0} = n^2 + 2n. \quad (4.151)$$

Daraus kann die Varianz berechnet werden:

$$\sigma_{y_n}^2 = E\{y_n^2\} - (E\{y_n\})^2 = 2n. \quad (4.152)$$

Durch Variablentransformation ergibt sich auch für allgemein normalverteilte Zufallsvariable $x_i \sim \mathcal{N}(\mu_x, \sigma_x^2)$ eine χ^2 -verteilte Größe:

$$\chi_n^2 = \frac{(x_1 - \mu_x)^2 + (x_2 - \mu_x)^2 + \cdots + (x_n - \mu_x)^2}{\sigma_x^2}. \quad (4.153)$$

Nachdem die Stichprobenvarianz (4.83)

$$s_x^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \hat{x})^2 \quad (4.154)$$

χ^2 -verteilt ist, stellt sich nun die Frage, wie viele Freiheitsgrade die Verteilung bei einem Stichprobenumfang von n Werten aufweist. Da der Mittelwert

$$\hat{x} = \frac{1}{n} (x_1 + x_2 + \cdots + x_n) \quad (4.155)$$

ebenfalls aus der Stichprobe geschätzt wird, hängt x_n von den übrigen x_i und dem Stichprobenmittelwert \hat{x} ab:

$$-(x_n - \hat{x}) = \sum_{i=1}^{n-1} (x_i - \hat{x}). \quad (4.156)$$

Durch Normierung der Quadratsumme aus der Stichprobenvarianz ergibt sich die χ^2 -verteilte Größe

$$\chi_n^2 = \frac{1}{\sigma_x^2} \sum_{i=1}^n (x_i - \hat{x})^2 = \frac{1}{\sigma_x^2} \left[\sum_{i=1}^{n-1} (x_i - \hat{x})^2 + (x_n - \hat{x})^2 \right] \quad (4.157)$$

$$= \frac{1}{\sigma_x^2} \left[\sum_{i=1}^{n-1} (x_i - \hat{x})^2 + \left(\sum_{i=1}^{n-1} (x_i - \hat{x}) \right)^2 \right] \quad (4.158)$$

$$= \frac{1}{\sigma_x^2} \left[\sum_{i=1}^{n-1} (\mathbf{x}_i - \hat{\mathbf{x}})^2 + \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} (\mathbf{x}_i - \hat{\mathbf{x}})(\mathbf{x}_j - \hat{\mathbf{x}}) \right] \quad (4.159)$$

$$= \frac{2}{\sigma_x^2} \sum_{i=1}^{n-1} (\mathbf{x}_i - \hat{\mathbf{x}})^2 + \underbrace{\frac{1}{\sigma_x^2} \sum_{i=1}^{n-1} \sum_{j=1, j \neq i}^{n-1} (\mathbf{x}_i - \hat{\mathbf{x}})(\mathbf{x}_j - \hat{\mathbf{x}})}_{\approx 0} . \quad (4.160)$$

Die normierte Stichprobenvarianz

$$\chi_n^2 = \frac{n-1}{\sigma_x^2} s_x^2 = \frac{2}{\sigma_x^2} \sum_{i=1}^{n-1} (\mathbf{x}_i - \hat{\mathbf{x}})^2 \quad (4.161)$$

bei n unabhängigen Messwerten ist also χ^2 -verteilt mit $n-1$ Freiheitsgraden.

4.3.4 Student'sche t-Verteilung

Die **Student'sche t-Verteilung** ist die Grundlage wichtiger statistischer Tests, die in Abschn. 4.4 behandelt werden. Sie wurde 1908 von W. S. Gosset unter dem Pseudonym „Student“ veröffentlicht.

Satz 4.5: Student'sche t-Verteilung

4.5

Es seien zwei unabhängige Zufallsvariable x und y gegeben. Dabei besitze x eine Standardnormalverteilung $\mathcal{N}(0, 1)$ und y eine χ^2 -Verteilung mit n Freiheitsgraden. Dann hat die Zufallsvariable

$$t = \frac{x}{\sqrt{y/n}} \quad (4.162)$$

die Wahrscheinlichkeitsdichte [48]

$$f_t(t) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi} \Gamma\left(\frac{n}{2}\right)} \cdot \frac{1}{\left(1 + \frac{t^2}{n}\right)^{(n+1)/2}} . \quad (4.163)$$

Die Zufallsvariable t wird t -verteilt mit n Freiheitsgraden genannt. \diamond

Mit wachsendem n strebt die t -Verteilung gegen die Standardnormalverteilung $\mathcal{N}(0, 1)$. Praktisch wird für $n \geq 30$ die t -Verteilung meist durch die Standardnormalverteilung approximiert (Abb. 4.14).

Die Bedeutung der t -Verteilung kommt aus der Stichprobenuntersuchung. Der Stichprobenmittelwert \hat{x} ist eine normalverteilte Zufallsvariable,

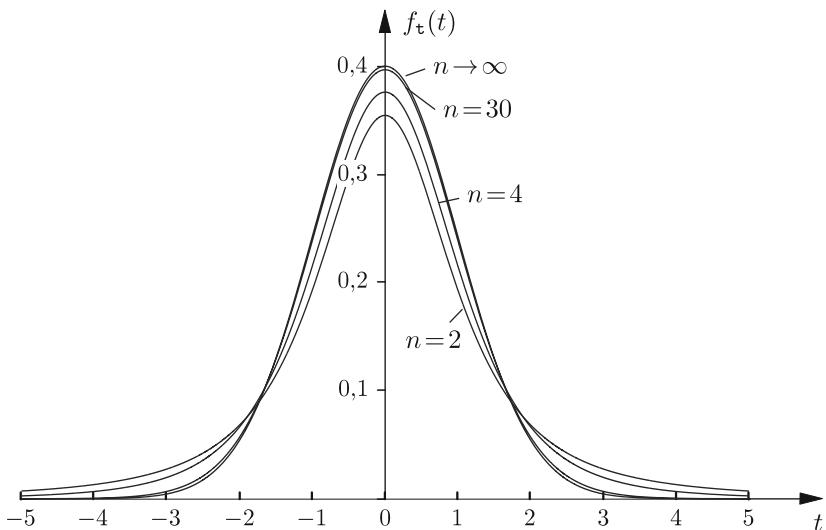


Abbildung 4.14. Wahrscheinlichkeitsdichte der t-Verteilung.

die Stichprobenvarianz s_x^2 ist χ^2 -verteilt. Dann ist das Verhältnis

$$t = \frac{\hat{x}}{\sqrt{s_x^2/n}} \quad (4.164)$$

des Stichprobenmittelwertes zu seiner Standardabweichung gerade t-verteilt.

4.4 Statistische Testverfahren

Die mathematische Statistik stellt Mittel zur Verfügung, um aus einer Stichprobe Aussagen über die zugrunde liegende Verteilung abzuleiten. Dazu gehören neben den in Abschn. 4.2 behandelten Schätzverfahren auch die **statistischen Testverfahren**, bei denen es um das Treffen von „Ja/Nein“-Entscheidungen geht. Hier sind speziell zwei Fragestellungen von großer praktischer Bedeutung. Einerseits stellt sich beim Stichprobenmittelwert die Frage, ob dieser Schätzwert repräsentativ für eine bestimmte angenommene Verteilung ist; die Antwort liefert der **Signifikanztest** für den Stichprobenmittelwert (Abschn. 4.4.3). Andererseits interessiert man sich bei einer Stichprobe, ob diese einem bestimmten Verteilungsmodell folgt; dazu dient der **χ^2 -Anpassungstest** (Abschn. 4.4.4). Es liegt im Wesen der Statistik, dass Testverfahren keine absolut sicheren Aussagen liefern können; Testentschei-

dungen können vielmehr nur mit einer gewissen statistischen Sicherheit getroffen werden (Abschn. 4.4.1 und 4.4.2).

⌚ 4.4.1 Konfidenzintervall und statistische Sicherheit

In der Messtechnik ist es oftmals notwendig, Auskunft über die Zuverlässigkeit einer Schätzung zu geben. Beispielsweise ist die Schätzung des Mittelwertes μ_x durch den Stichprobenmittelwert \hat{x} bei einer kleinen Stichprobe weniger vertrauenswürdig als bei einer großen. Ein Messwert ist daher nur dann aussagekräftig, wenn die mit dessen Schätzung verbundene Messunsicherheit bekannt ist. In vielen Fällen wird dazu das **Konfidenzintervall** (auch Vertrauensintervall) angegeben. Ein bestimmtes Intervall $[\mu_x - x_\alpha, \mu_x + x_\alpha]$ enthält den zu schätzenden Parameter mit der **Irrtumswahrscheinlichkeit**

$$\alpha = P(|x - \mu_x| > x_\alpha). \quad (4.165)$$

Das zweiseitige Konfidenzintervall schließt also mit einer **statistischen Sicherheit** von

$$1 - \alpha = P(|x - \mu_x| \leq x_\alpha) \quad (4.166)$$

den wahren Parameter ein (Abb. 4.15). Das sogenannte einseitige Problem führt dagegen zu der Aussage, dass ein Parameter mit der Irrtumswahrscheinlichkeit α beispielsweise nicht größer ist als eine bestimmte Grenze.

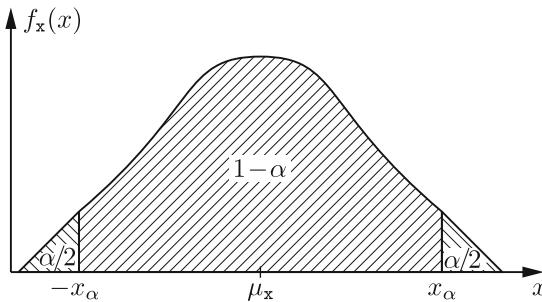


Abbildung 4.15. Zweiseitiges Konfidenzintervall.

⌚ 4.4.1.1 Konfidenzintervall bei bekannter Standardabweichung

In der Praxis geht man meist von einer Normalverteilung aus. Das Konfidenzintervall wird dann in Vielfachen c der Standardabweichung σ_x ausgedrückt:

$$\mu_x - c \sigma_x \leq x \leq \mu_x + c \sigma_x. \quad (4.167)$$

Die statistische Sicherheit erhält man dann durch Integration der Dichte $f_x(x)$ der Normalverteilung (Abb. 4.16) zu

$$P(c) = 1 - \alpha = P\left(\frac{|x - \mu_x|}{\sigma_x} \leq c\right) = \int_{\mu_x - c\sigma_x}^{\mu_x + c\sigma_x} f_x(x) dx. \quad (4.168)$$

Leider lässt sich das Integral (4.168) analytisch nicht lösen, weshalb es meistens durch die Gauß'sche Fehlerfunktion $\text{erf}(c)$ ausgedrückt wird, die sich wiederum numerisch berechnen oder in Tabellen nachschlagen lässt (vgl. Tab. A.1).

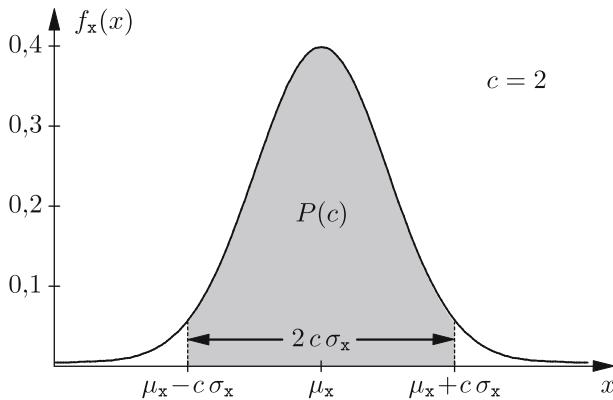


Abbildung 4.16.
Konfidenzintervall bei
Normalverteilung.

4.25

Definition 4.25: Gauß'sche Fehlerfunktion

Die **Gauß'sche Fehlerfunktion** (engl. *error function*) ist durch das folgende Integral definiert:

$$\text{erf}(c) = \frac{2}{\sqrt{\pi}} \int_0^c e^{-x^2} dx. \quad (4.169)$$

Die Gauß'sche Fehlerfunktion ist eine ungerade Funktion. \diamond

Um das Integral (4.168) in Abhängigkeit von der Fehlerfunktion auszudrücken, ist es zweckmäßig, die Verteilung durch die Transformation

$$z = \frac{x - \mu_x}{\sigma_x} \quad (4.170)$$

in eine Standardnormalverteilung zu überführen. Die statistische Sicherheit $P(c)$ hängt dann nur noch vom Parameter c ab und kann durch die

Gauß'sche Fehlerfunktion wie folgt angegeben werden:

$$P(c) = \frac{1}{\sqrt{2\pi}} \int_{-c}^c \exp\left(-\frac{z^2}{2}\right) dz = \sqrt{\frac{2}{\pi}} \int_0^c \exp\left(-\frac{z^2}{2}\right) dz = \operatorname{erf}\left(\frac{c}{\sqrt{2}}\right).$$

Man erhält die folgenden statistischen Sicherheiten in Abhängigkeit vom Parameter c (vgl. Abb. 4.17 und Tab. A.1):

$$c = 1 : \quad \text{Konfidenzintervall: } \mu_x \pm \sigma_x, \quad P(c) = 68,27\%;$$

$$c = 2 : \quad \text{Konfidenzintervall: } \mu_x \pm 2\sigma_x, \quad P(c) = 95,45\%;$$

$$c = 3 : \quad \text{Konfidenzintervall: } \mu_x \pm 3\sigma_x, \quad P(c) = 99,73\%.$$

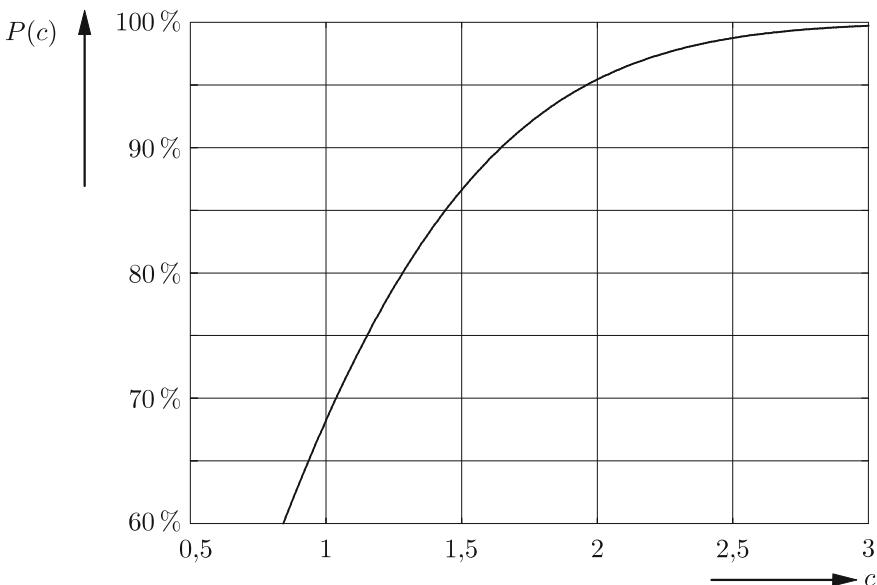


Abbildung 4.17. Statistische Sicherheit bei bekannter Standardabweichung (vgl. Tab. A.1).

Der für die Qualität eines Fertigungsprozesses entscheidende Parameter ist jedoch die Standardabweichung σ_x . Sie beschreibt, wie stark die Istmaße fertigungsbedingt streuen. Für eine statistische Sicherheit von beispielsweise $P(c) = 95,45\%$ ($c = 2$) erhält man für verschiedene Standardabweichungen σ_x unterschiedlich breite Intervalle (Abb. 4.18). Eine schmale Verteilung lässt sich aber leichter in ein Toleranzfeld einer bestimmten Breite einpassen, ohne dass es zu hohen Ausschusswahrscheinlichkeiten kommt (Abschn. 4.5.1).

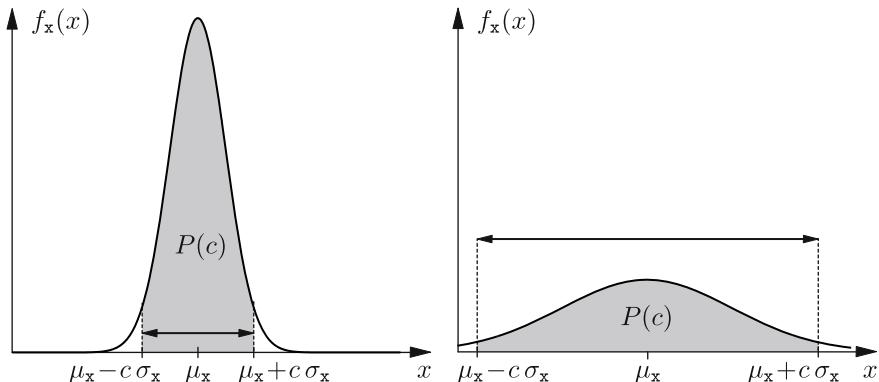


Abbildung 4.18. Breite des Konfidenzintervalls bei schmaler Normalverteilung (kleines σ_x) und breiter Normalverteilung (großes σ_x).

② 4.4.1.2 Konfidenzintervall bei geschätzter Standardabweichung

Das Konfidenzintervall für eine Zufallsgröße x wurde in (4.167) als Vielfaches c der Standardabweichung σ_x angegeben. Möchte man das Konfidenzintervall für den Stichprobenmittelwert \hat{x} einer Messreihe von n unabhängigen Messungen ermitteln, so muss nach (4.82) die Standardabweichung des Stichprobenmittelwertes

$$\sigma_{\hat{x}} = \frac{\sigma_x}{\sqrt{n}} \quad (4.171)$$

herangezogen werden:

$$\mu_x - c \frac{\sigma_x}{\sqrt{n}} \leq \hat{x} \leq \mu_x + c \frac{\sigma_x}{\sqrt{n}}. \quad (4.172)$$

Die Messunsicherheit $u_{\hat{x}}$ ist abhängig von c und wird wie folgt definiert:

$$u_{\hat{x}} = c \cdot \sigma_{\hat{x}} = c \frac{\sigma_x}{\sqrt{n}}. \quad (4.173)$$

In der Praxis ist jedoch die Standardabweichung σ_x nicht bekannt. Daher wird in (4.172) die empirische Standardabweichung s_x der Stichprobe als Schätzwert für die Standardabweichung σ_x eingesetzt:

$$-c \leq \frac{\hat{x} - \mu_x}{s_x/\sqrt{n}} \leq c. \quad (4.174)$$

Die Transformation

$$t = \frac{\hat{x} - \mu_x}{s_x/\sqrt{n}} \quad (4.175)$$

ergibt nach Satz 4.5 eine t-verteilte Zufallsvariable mit $n - 1$ Freiheitsgraden. Zur Berechnung der statistischen Sicherheit aus einer Stichprobe muss daher anstelle der Normalverteilung die Dichte der t-Verteilung (4.163) integriert werden (vgl. Abb. 4.19 und Tab. A.1):

$$P_n(c) = P\left(\frac{|\hat{x} - \mu_x|}{s_x/\sqrt{n}} = |t| \leq c\right) \quad (4.176)$$

$$= \int_{-c}^c \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi} \Gamma(\frac{n}{2})} \cdot \frac{1}{(1 + \frac{t^2}{n})^{(n+1)/2}} dt. \quad (4.177)$$

Damit ist die statistische Sicherheit bei geschätzter Standardabweichung abhängig vom Stichprobenumfang n .

$$P_n(c) = P\{|t| \leq c\}$$

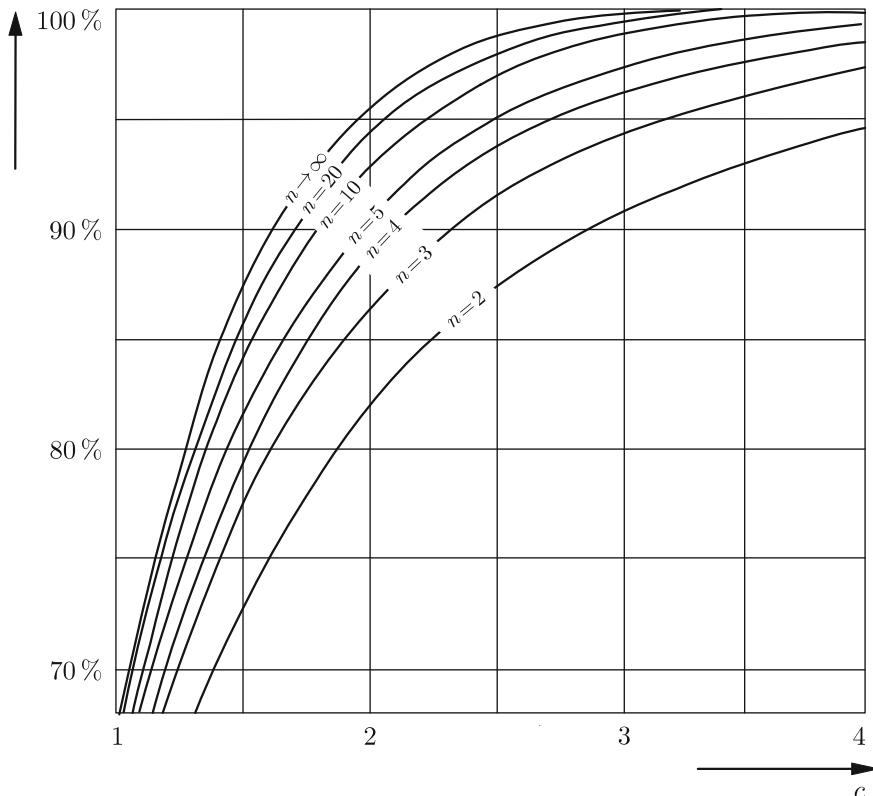


Abbildung 4.19. Statistische Sicherheit bei geschätzter Standardabweichung s_x (vgl. Tab. A.1).

Zur Ermittlung des Konfidenzintervalls geht man in der Praxis folgendermaßen vor:

- Man wählt eine statistische Sicherheit $P\{|t| \leq c\}$ und liest zusammen mit dem Stichprobenumfang n den Wert c aus Abb. 4.19 oder aus Tab. A.1 ab.
- Die Standardabweichung der Stichprobe wird nach Def. 4.22 berechnet:

$$s_x = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \hat{x})^2}. \quad (4.178)$$

- Mit der Messunsicherheit $u_{\hat{x}}$ des Stichprobenmittelwertes

$$u_{\hat{x}} = c \cdot \frac{s_x}{\sqrt{n}} \quad (4.179)$$

ergibt sich das Vertrauensintervall zu:

$$\mu_x - c \frac{s_x}{\sqrt{n}} \leq \hat{x} \leq \mu_x + c \frac{s_x}{\sqrt{n}}. \quad (4.180)$$

Soll die Messunsicherheit u_x einer Einzelmessung unabhängig von der Zahl der Messwerte angegeben werden, wird man die Standardabweichung der Stichprobe s_x aus vielen unabhängigen Messwerten gemäß (4.83) berechnen und den Wert c zu der geforderten statistischen Sicherheit $P_n(c)$ aus Abb. 4.19 ablesen. Die Messunsicherheit der Einzelmessung

$$u_x = c \cdot s_x \quad (4.181)$$

wird größer als die des Stichprobenmittelwertes sein.

Für $n \rightarrow \infty$ geht die Varianz $\sigma_{\hat{x}}^2$ des Stichprobenmittelwertes gegen null, d. h. die statistische Sicherheit $P_n(c)$ konvergiert gegen das Integral $P(c)$ der Dichte der Normalverteilung (vgl. Abb. 4.19).

Strengh von der Messunsicherheit zu unterscheiden ist der Begriff der Fehlergrenze. Die Fehlergrenzen sind in der Messtechnik die vereinbarten oder garantierten, zugelassenen äußersten Abweichungen von einem vorgeschriebenen Wert der Messgröße. Damit die spezifizierten Fehlergrenzen sicher eingehalten werden können, muss die Messunsicherheit erheblich kleiner als der durch eine Fehlergrenze gegebene Bereich sein. Ist ein Messergebnis mittelbar durch die Funktion der Messwerte gegeben, so ist die Messunsicherheit bei ausreichend großer Stichprobe n nach dem sogenannten **Gauß'schen Fehlerfortpflanzungsgesetz** gegeben (Abschn. 4.6).

4.4.2 Hypothesen und statistische Tests

Das Ziel statistischer Tests besteht darin, eine präzise formulierte Behauptung – die sogenannte **Nullhypothese** H_0 – zu überprüfen. Ein Beispiel für H_0 wäre: „Eine gegebene Stichprobe entstamme einer bestimmten Grundgesamtheit.“ Nach Popper lassen sich aber wissenschaftstheoretisch Hypothesen nur widerlegen, nicht jedoch beweisen [39]. Der Nullhypothese H_0 steht daher die komplementäre **Alternativhypothese** H_1 gegenüber, die es (mit einer möglichst großen Wahrscheinlichkeit) zu bestätigen – d. h. zu erkennen – gilt. Sollte dies gelingen, wird die Nullhypothese H_0 verworfen. Andernfalls darf bis auf Weiteres von der Gültigkeit der Nullhypothese ausgegangen werden.

Für statistische Tests benötigt man darüber hinaus den Begriff des **Signifikanzniveaus**. Das Signifikanzniveau α ist identisch mit jener Irrtumswahrscheinlichkeit, die man bereit ist zu akzeptieren, falls das Testverfahren eine Ablehnung einer tatsächlich zutreffenden Nullhypothese H_0 ergibt („**Fehler 1. Art**“). Es entspricht einem **Falschalarm**, also dem Risiko, einen Unterschied zu „erfinden“ oder zu behaupten, der in Wirklichkeit nicht vorhanden ist. Daher wird das Signifikanzniveau in der Regel konservativ (d. h. sehr klein) gewählt, normalerweise in der Größenordnung

$$0,001 \leq \alpha \leq 0,05. \quad (4.182)$$

Da das Signifikanzniveau α vorgegeben werden muss, liegt der Gedanke nahe, dieses so klein wie möglich zu wählen (z. B. $\alpha = 10^{-9}$), damit es praktisch kaum zu der erwähnten Fehlentscheidung kommt. Bei der Durchführung statistischer Tests ist aber zu beachten, dass noch eine zweite Art von Fehlentscheidungen möglich ist (Tab. 4.2).

Wenn man α verkleinert, steigt die Wahrscheinlichkeit eines **Fehlers 2. Art**. Während ein Fehler 1. Art als Falschalarm interpretiert werden kann, ist ein Fehler 2. Art ein unterbliebener Alarm („Schlupf“), d. h. trotz signifikanter Abweichung wird die Hypothese H_0 „bestätigt“. Die Wahrscheinlichkeit für den Fehler 2. Art kann nicht allgemein angegeben werden.

Tabelle 4.2. Mögliche Testentscheidungen beim Prüfen der Nullhypothese H_0 .

| Testentscheidung H_0 wird abgelehnt bestätigt | Tatsächlicher Zustand | |
|--|--------------------------|-------------------------|
| | H_0 trifft zu | H_0 trifft nicht zu |
| abgelehnt | α (Fehler 1. Art) | $1 - \beta$ |
| bestätigt | $1 - \alpha$ | β (Fehler 2. Art) |

4.4.3 Signifikanztest für den Stichprobenmittelwert

Gegenstand des Signifikanztests für den Stichprobenmittelwert ist die Beantwortung der Frage, ob eine Stichprobe zu einer vorgegebenen Grundgesamtheit gehört. Dazu geht man meist von einer Normalverteilung $\mathcal{N}(\mu_0, \sigma_x^2)$ aus und prüft, ob der Stichprobenmittelwert \hat{x} nahe genug am wahren Mittelwert μ_0 der Verteilung liegt. Ist dies nicht der Fall, so ist die Abweichung nicht zufällig, sondern signifikant: die Stichprobe wird als nicht repräsentativ abgelehnt. Da man bei diesem Test nicht eine Wahrscheinlichkeitsdichte $f_x(x)$, sondern einen Parameter \hat{x} einer vorgegebenen Normalverteilung prüft, handelt es sich um ein *parametrisches* Prüfverfahren oder **Parametertest**.

Die Vorgehensweise des Signifikanztests für den Stichprobenmittelwert ist die folgende:

1. Voraussetzungen prüfen:
 - Unabhängigkeit der Messwerte;
 - Normalverteilung der Grundgesamtheit mit Erwartungswert μ_0 .
2. Ermittlung des Stichprobenmittelwertes \hat{x} und – bei unbekannter Varianz σ_x^2 der Grundgesamtheit – der Standardabweichung der Stichprobe s_x .
3. Aufstellen der Nullhypothese:

$$H_0 : \hat{x} = \mu_0 , \quad H_1 : \hat{x} \neq \mu_0 . \quad (4.183)$$

4. Festlegen der Prüfgröße:

- Bei bekannter Varianz σ_x^2 kann mit einer Normalverteilung gerechnet werden:

$$z = \frac{|\hat{x} - \mu_0|}{\sigma_{\hat{x}}} = \frac{|\hat{x} - \mu_0|}{\sigma_x} \sqrt{n} = c . \quad (4.184)$$

- Wird die Varianz durch die Stichprobenvarianz geschätzt, so ist die Prüfgröße t-verteilt:

$$t = \frac{|\hat{x} - \mu_0|}{s_x} \sqrt{n} = c . \quad (4.185)$$

Die Anzahl der Freiheitsgrade beträgt dabei $f = n - 1$.

5. Festlegen des Signifikanzniveaus α . Daraus ergibt sich die maximale statistische Sicherheit $1 - \alpha$, die mit der Nullhypothese verträglich ist.

6. Bestimmen der Wahrscheinlichkeit der Prüfgröße:

Aus einer Tabelle (Abb. 4.17, Abb. 4.19 oder Tab. A.1) wird der Wert für $P(c)$ bzw. $P_n(c)$ entnommen.

7. Testentscheidung:

- Annahme der Nullhypothese, falls

$$P(c) \leq 1 - \alpha. \quad (4.186)$$

- Ablehnung der Nullhypothese, falls

$$P(c) > 1 - \alpha. \quad (4.187)$$

Beispiel 4.9 (Signifikanztest für den Stichprobenmittelwert): Ein Werkstück

habe das Sollmaß $\mu_0 = 12,0$ mm. Es wird eine Stichprobe bestehend aus $n = 90$ Werkstücken vermessen. Der Stichprobenmittelwert wird als $\hat{x} = 12,075$ mm, die Standardabweichung als $s_x = 0,229$ mm ermittelt. Wegen des großen Stichprobenumfangs ($n \geq 30$) geht die t-Verteilung näherungsweise in eine Normalverteilung über. Die Standardabweichung des Stichprobenmittelwertes ist damit näherungsweise $\sigma_{\hat{x}} \approx s_x / \sqrt{n} = 0,0241$ mm. Als Signifikanzniveau wird die $\pm 3\sigma_x$ -Spanne der Normalverteilung zugrunde gelegt:

$$\alpha = 0,0027 \quad \Rightarrow \quad 1 - \alpha = 0,9973. \quad (4.188)$$

Unter der Voraussetzung, dass die Messwerte normalverteilt sind, erhält man für die Wahrscheinlichkeit der Prüfgröße einen hohen Wert von

$$P(c) = P\left(\frac{|\hat{x} - \mu_0|}{\sigma_{\hat{x}}}\right) = P(3,112) = 0,9981 > 1 - \alpha. \quad (4.189)$$

Die Abweichung des Stichprobenmittelwertes \hat{x} vom wahren Mittelwert μ_0 bezogen auf die Standardabweichung $\sigma_{\hat{x}}$ ist zu groß und somit signifikant. Die Stichprobe muss deshalb als nicht repräsentativ abgelehnt werden. ■

4.4.4 χ^2 -Anpassungstest

Im vorigen Abschnitt wurde geprüft, ob die Werte eines oder mehrerer Parameter einer Stichprobe, z. B. Mittelwerte oder Varianzen, signifikant oder rein zufällig voneinander abweichen. Man nennt solche Prüfverfahren *Parametertests*. Eine andere Art von Verfahren, welche die Hypothese H_0 prüfen, ob eine Stichprobe einer Grundgesamtheit mit einer vorgegebenen, beliebigen Wahrscheinlichkeitsdichte $f_x(x)$ entstammt, nennt man **Anpassungstests**. Der wichtigste Test aus dieser Gruppe ist der χ^2 -Test.

Man teilt dabei den gesamten Wertebereich der Zufallsgröße x in k disjunkte Intervalle $\Delta_1, \dots, \Delta_k$, die Klassen, ein (ähnlich wie beim Histogramm, vgl. Abschn. 4.2.1). Die Stichprobe habe den Umfang n . Durch Integration der vorgegebenen Wahrscheinlichkeitsdichte $f_x(x)$ im Intervall Δ_i erhält man die theoretische Wahrscheinlichkeit p_i dafür, dass x in Δ_i fällt:

$$p_i = \int_{\Delta_i} f_x(x) dx \quad \text{mit} \quad \sum_{i=1}^k p_i = 1. \quad (4.190)$$

Die Wahrscheinlichkeit f_{n_i} dafür, dass bei einer Stichprobe mit Umfang n gerade n_i Elemente anstelle der theoretischen $n \cdot p_i$ in die Klasse Δ_i fallen, wird durch die **Binomialverteilung** beschrieben:

$$f_{n_i} = \binom{n}{n_i} p_i^{n_i} (1 - p_i)^{n - n_i}. \quad (4.191)$$

Nach dem **Moivre-Laplace-Theorem** – einem Spezialfall des zentralen Grenzwertsatzes – geht die Binomialverteilung für $n \rightarrow \infty$ in eine Normalverteilung über [37]. Daher kann die letzte Gleichung bei fester Wahrscheinlichkeit p_i und hinreichend großem Stichprobenumfang n näherungsweise durch die Normalverteilung

$$f_{n_i} \approx \frac{1}{\sqrt{2\pi n p_i (1 - p_i)}} \exp\left(-\frac{(n_i - n p_i)^2}{2n p_i (1 - p_i)}\right) \quad (4.192)$$

mit dem Mittelwert der i -ten Klasse Δ_i

$$E\{n_i\} \approx n p_i \quad (4.193)$$

dargestellt werden. Für $p_i \ll 1$, d. h. bei einer ausreichend großen Zahl von Klassen, ist die Varianz für die Elementenzahl n_i

$$\sigma_{n_i}^2 \approx n p_i. \quad (4.194)$$

Die Summe der auf die jeweiligen Varianzen $\sigma_{n_i}^2$ normierten quadratischen Abweichungen der tatsächlichen Elementenzahl n_i zum Erwartungswert $n \cdot p_i$ ist ein Maß für die Abweichung beider Verteilungen voneinander und genügt nach Satz 4.4 näherungsweise einer χ^2 -Verteilung mit $(k - 1)$ Freiheitsgraden:

$$\chi^2 \approx \sum_{i=1}^k \frac{(n_i - n p_i)^2}{n p_i}. \quad (4.195)$$

Die Klasseneinteilung ist weitgehend willkürlich. Einerseits wünscht man sich viele Klassen, um die Wahrscheinlichkeitsdichte $f_x(x)$ möglichst gut zu approximieren. Andererseits sollten die n_i genügend groß sein, damit die Testgröße als χ^2 -verteilt betrachtet werden kann. Als Faustregel für die Praxis sollten die n_i mindestens 1 bei Randklassen, ansonsten mindestens 5 betragen.

Beim χ^2 -Anpassungstest wird wie folgt verfahren:

1. Voraussetzungen prüfen:
 - Unabhängigkeit der Messwerte,
 - Möglichst großer Stichprobenumfang.
2. Erstellen eines Histogramms:
 - Festlegen der k Klassen Δ_i .
 - Ermitteln der absoluten Häufigkeiten n_i innerhalb der Klassen.
 - Sind die Bedingungen $n_i \geq 5$ oder $n_{i,\text{Rand}} \geq 1$ nicht erfüllt, dann Nachbarklassen zu einer gemeinsamen Klasse zusammenfassen.
3. Aufstellen der Nullhypothese:

$$H_0 : f_x(x) = f_0(x), \quad H_1 : f_x(x) \neq f_0(x). \quad (4.196)$$

4. Festlegen des Signifikanzniveaus α :

Daraus bestimmt man die maximale statistische Sicherheit

$$P = 1 - \alpha. \quad (4.197)$$

Häufig wird beim χ^2 -Anpassungstest eine gegenüber den Parametertests auf $\alpha = 0,05$ vergrößerte Irrtumswahrscheinlichkeit gewählt. Da bei Anpassungstests keine Voraussetzungen über die Wahrscheinlichkeitsdichte gemacht werden, sind diese weniger wirksam als entsprechende Parametertests.

5. Festlegen der Prüfgröße:

$$\chi^2 \approx \sum_{i=1}^k \frac{(n_i - n p_i)^2}{n p_i}. \quad (4.198)$$

6. Bestimmung der Freiheitsgrade:

$$m = k - 1 - \text{Anzahl der geschätzten Parameter}. \quad (4.199)$$

Meist sind die Parameter der zu prüfenden Verteilung nur durch Stichproben bekannt. Werden bei einer Normalverteilung die Parameter $(\mu_x; \sigma_x^2)$ durch $(\hat{x}; s_x^2)$ geschätzt, so verringert sich der Freiheitsgrad zusätzlich um zwei: $m = k - 2 - 1$.

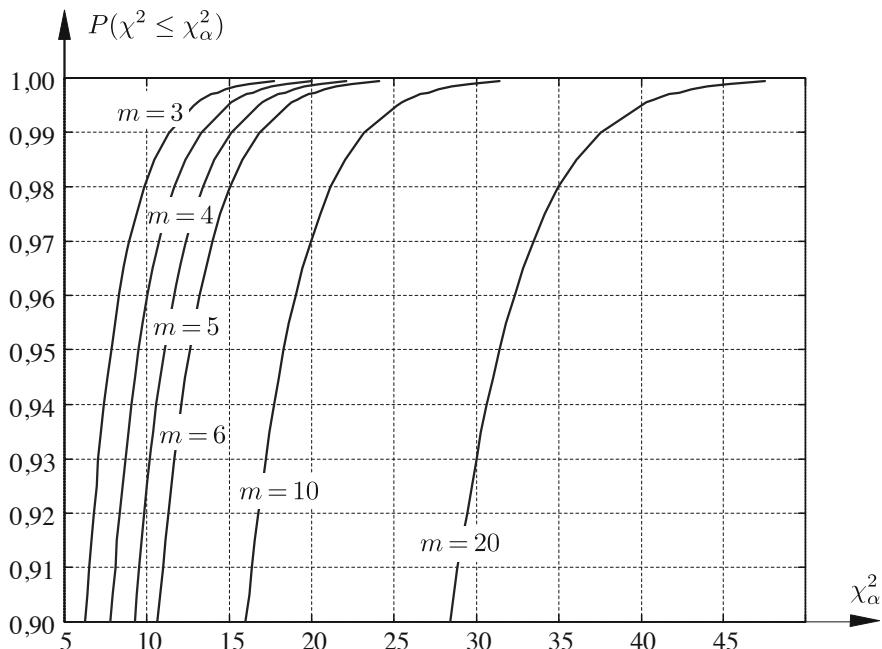


Abbildung 4.20. Wahrscheinlichkeit von $\chi^2 \leq \chi^2_\alpha$ bei $m = k - 1$ Freiheitsgraden.

7. Bestimmen der Wahrscheinlichkeit der Prüfgröße:

Aus Abb. 4.20 für die statistische Sicherheit bei χ^2 -Verteilung entnimmt man den Wert für χ^2_α für das vorgegebene Signifikanzniveau

$$P(\chi^2 \leq \chi^2_\alpha) = 1 - \alpha. \quad (4.200)$$

8. Testentscheidung:

– Annahme der Nullhypothese:

$$\chi^2 \leq \chi^2_\alpha. \quad (4.201)$$

– Ablehnung der Nullhypothese:

$$\chi^2 > \chi^2_\alpha. \quad (4.202)$$

Beispiel 4.10 (χ^2 -Test auf Gleichverteilung): Für einen Würfel mit $k = 6$ Augen soll die Hypothese überprüft werden, ob für $f_x(x)$ eine Gleichverteilung der Augenzahlen vorliegt:

$$H_0 : f_x(x) \hat{=} \text{Gleichverteilung}. \quad (4.203)$$

Tabelle 4.3. Würfelexperiment.

| Augenzahl | 1 | 2 | 3 | 4 | 5 | 6 | Summe |
|---------------------------------|-----|------|------|-----|------|------|-------|
| Anzahl n_i | 14 | 27 | 15 | 24 | 13 | 27 | 120 |
| $n_i - n p_i$ | -6 | 7 | -5 | 4 | -7 | 7 | 0 |
| $\frac{(n_i - n p_i)^2}{n p_i}$ | 1,8 | 2,45 | 1,25 | 0,8 | 2,45 | 2,45 | 11,2 |

Es werden $n = 120$ Testwürfe durchgeführt (Tab. 4.3). Die theoretische Elementezahl $n p_i = 20$ ist bei der angenommenen Gleichverteilung in allen Klassen gleich. Das Signifikanzniveau sei

$$\alpha = 0,05 \quad \Rightarrow \quad P(\chi^2 \leq \chi_{\alpha}^2) \leq 1 - \alpha = 0,95. \quad (4.204)$$

Für die Wahrscheinlichkeit $P(\chi^2 \leq \chi_{\alpha}^2) = 0,95$ bei $m = k - 1 = 5$ Freiheitsgraden entnimmt man Abb. 4.20 den Wert

$$\chi_{\alpha}^2 = 11,0. \quad (4.205)$$

Die Summe der Abweichungsquadrate (Prüfgröße) $\chi^2 = 11,2$ ist nach Tab. 4.3 größer als der Grenzwert:

$$\chi^2 > \chi_{\alpha}^2. \quad (4.206)$$

Die Abweichungen sind signifikant und die Hypothese H_0 wird deshalb fälschlicherweise abgelehnt (Fehler 1. Art). Durch Erhöhung der Zahl von Testwürfen kann aber die Hypothese bestätigt werden, dass eine Gleichverteilung vorliegt. ■

4.5 Qualitätssicherung

4.5

4.5.1 Beurteilung von Fertigungsprozessen

Zur Bewertung der Qualität von Fertigungsprozessen wird geprüft, inwie weit $99,73\% (\mu_x \pm 3\sigma_x)$ der Istmaße x eines Werkstücks innerhalb der spezifizierten **Toleranzgrenzen** $[x_{\min}; x_{\max}]$ liegen, wobei x_{\min} und x_{\max} das kleinste bzw. größte zulässige Grenzmaß bezeichnen. Für die Istmaße wird eine Normalverteilung angenommen, deren Parameter hinreichend genau bekannt sein müssen, z. B. durch eine ausreichend große Stichprobe. Im Fol-

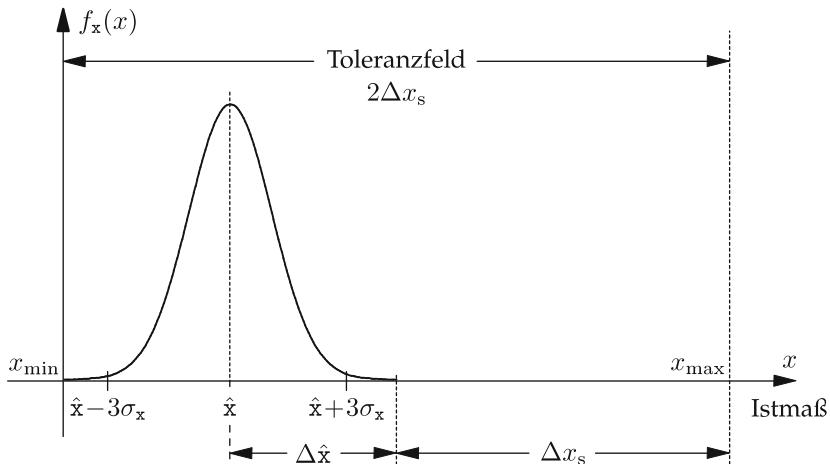


Abbildung 4.21. Spezifikation von Fertigungstoleranzen.

genden bezeichnen

$$2\Delta x_s = x_{\max} - x_{\min} \quad (4.207)$$

die Breite des **Toleranzfeldes** und

$$\Delta \hat{x} = \left| \frac{1}{2}(x_{\max} + x_{\min}) - \hat{x} \right| \quad (4.208)$$

die Abweichung des Stichprobenmittelwertes \hat{x} von der Toleranzfeldmitte (Abb. 4.21).

Als **Prozessfähigkeitsindex**

$$c_p = \frac{x_{\max} - x_{\min}}{6\sigma_x} = \frac{\Delta x_s}{3\sigma_x} \quad (4.209)$$

wird das Verhältnis der Breite des Toleranzfeldes $2\Delta x_s$ zum Vertrauensbereich $6\sigma_x$ definiert. Je größer der Index c_p ist, desto geringer ist die Streuung der Istmaße x im Verhältnis zur spezifizierten Toleranz.

Zur Beurteilung der Eignung eines Fertigungsprozess interessiert auch die Lage des Stichprobenmittelwertes \hat{x} innerhalb des Toleranzfeldes. Eine Abweichung $\Delta \hat{x}$ engt den Abstand des $3\sigma_x$ -Bereichs zur Toleranzgrenze einseitig ein. Zur Abschätzung wird der **Prozessbrauchbarkeitsindex** definiert:

$$c_{pk} = \frac{\Delta x_s - \Delta \hat{x}}{3\sigma_x} = c_p \left(1 - \frac{\Delta \hat{x}}{\Delta x_s} \right). \quad (4.210)$$

Für beide Indizes muss

$$c_p \geq 1, \quad c_{pk} \geq 1 \quad (4.211)$$

gelten, damit der Fertigungsprozess einen geringen Ausschuss ($< 0,27\%$) aufweist. Dann liegt das Istmaß x nur bei einem geringen Prozentsatz der gefertigten Erzeugnisse außerhalb des zulässigen Bereichs. Bei Betrachtung der unterschiedlich breiten Normalverteilungen aus Abb. 4.18 wird sofort klar, dass die Güte einer Fertigung unmittelbar mit der Breite der zugrunde liegenden Normalverteilung zusammenhängt. So ist eine schmale Normalverteilung wesentlich robuster gegenüber Schwankungen des Mittelwertes \hat{x} , ohne dass dies gleich zu erhöhten Ausschussraten führen muss.

Die **Ausschussrate** p berechnet sich bei Normalverteilung über die Gauß'sche Fehlerfunktion $\text{erf}(x)$ zu

$$p \approx \frac{1}{2} \left(1 - \text{erf}\left(3 c_{pk} / \sqrt{2} \right) \right). \quad (4.212)$$

Sie wird häufig in „dpm“ (*defects per million*) angegeben. Bei einer qualitativ hochwertigen Fertigung wird ein Prozessbrauchbarkeitsindex c_{pk} von

$$c_{pk} > 1,67 \dots 2 \quad (4.213)$$

gefordert, mit Ausschussraten von

$$p < 0,3 \dots 0,001 \text{ dpm}. \quad (4.214)$$

Beispiel 4.11 (Länge eines Werkstücks): Bei einem feinmechanischen Werkstück sei das Längenmaß auf

$$x = 0,609 \text{ mm} \quad (4.215)$$

spezifiziert. Zulässige Fertigungstoleranzen werden wie folgt angegeben:

$$x_{\min} = 0,591 \text{ mm} \leq x \leq x_{\max} = 0,627 \text{ mm}. \quad (4.216)$$

Aus einer Stichprobenmessung werden der Mittelwert \hat{x} und die Standardabweichung s_x ermittelt:

$$\hat{x} = 0,600 \text{ mm}, \quad s_x = 0,003 \text{ mm} \approx \sigma_x. \quad (4.217)$$

Mit den Werten

$$\Delta x_s = \frac{1}{2}(0,627 \text{ mm} - 0,591 \text{ mm}) = 0,018 \text{ mm}, \quad (4.218)$$

$$\Delta \hat{x} = \frac{1}{2}(0,591 \text{ mm} + 0,627 \text{ mm}) - 0,600 \text{ mm} = 0,009 \text{ mm} \quad (4.219)$$

erhält man den Prozessfähigkeitsindex

$$c_p = \frac{\Delta x_s}{3\sigma_x} = \frac{0,018 \text{ mm}}{0,009 \text{ mm}} = 2 \quad (4.220)$$

und den Prozessbrauchbarkeitsindex

$$c_{pk} = c_p \left(1 - \frac{\Delta \hat{x}}{\Delta x_s} \right) = 2 \left(1 - \frac{0,009 \text{ mm}}{0,018 \text{ mm}} \right) = 1. \quad (4.221)$$

Die Verteilung liegt unsymmetrisch im Toleranzfeld. Die Ausschussrate beträgt nach (4.212)

$$p \approx \frac{1}{2} \left(1 - \operatorname{erf} \left(3 c_{pk} / \sqrt{2} \right) \right) = 0,00135 = 1350 \text{ dpm}, \quad (4.222)$$

wobei zur Berechnung nur die Spitze der Wahrscheinlichkeitsdichte berücksichtigt wurde, die über die untere Toleranzgrenze hinausragt. ■

4.5.2 Bestimmung der Ausfallrate

Die Hersteller elektronischer Geräte beziehen große Mengen elektronischer Bauelemente mit vertraglich festgelegten **Ausfallraten**. Zur Überprüfung der Qualität der Bauelemente will der Gerätehersteller die Ausfallraten messtechnisch bestimmen. Wegen des großen Aufwands kann die Überprüfung nur stichprobenweise erfolgen. Es seien

- n die Zahl der Bauelemente in der Stichprobe,
- p die Ausfallwahrscheinlichkeit des Bauelementes und
- k die Zahl der in der Stichprobe registrierten Ausfälle.

Dann ist die Wahrscheinlichkeit, dass in der Stichprobe zwischen k_1 und k_2 der n Bauelemente ausgefallen sind, durch die Binomialverteilung

$$P_n(k_1 \leq i \leq k_2) = \sum_{i=k_1}^{k_2} \binom{n}{i} p^i (1-p)^{n-i} \quad (4.223)$$

gegeben [37]. Für den Grenzfall $n \rightarrow \infty$ und $p \rightarrow 0$ gilt bei konstantem Produkt np das **Poisson'sche Theorem**, nach dem die Binomialverteilung in

die **Poisson-Verteilung** übergeht:

$$P_n(k_1 \leq i \leq k_2) = \sum_{i=k_1}^{k_2} \frac{n^i}{i!} p^i e^{-np}. \quad (4.224)$$

Die Wahrscheinlichkeit, dass in der Stichprobe weniger als k Bauelemente ausgefallen sind ($k_1 = 0, k_2 = k$), ist dann

$$P_n(i \leq k) = e^{-np} \sum_{i=0}^k \frac{(np)^i}{i!}. \quad (4.225)$$

Beispiel 4.12 (Ausfallwahrscheinlichkeit): Die Stichprobengröße sei $n = 3000$, die Ausfallwahrscheinlichkeit $p = 10^{-3}$. Dann ergibt sich mit

$$P_n(i \leq 5) = e^{-3 \cdot 10^3 \cdot 10^{-3}} \sum_{i=0}^5 \frac{(3 \cdot 10^3 \cdot 10^{-3})^i}{i!} = 0,916 \quad (4.226)$$

die Wahrscheinlichkeit, dass in der Stichprobe höchstens 5 Bauelemente defekt sind. ■

Im Prüffeld wird häufig mit der Ausfallrate λ gearbeitet, die im Folgenden als konstant angenommen wird und dann dem Kehrwert der mittleren Lebensdauer (*mean time to failure*, MTTF) entspricht. Bei der Prüfung von Bauelementen definiert man

- nt : Zahl der „Bauelementestunden“, d. h. das Produkt aus der Zahl n der Bauelemente in der Stichprobe und der Prüfzeit t ,
- λ : Ausfallwahrscheinlichkeit p der Bauelemente bezogen auf die Prüfzeit t .

Mit $\lambda \cdot nt = np$ kann (4.225) geschrieben werden als

$$P_n(i \leq k) = e^{-\lambda nt} \sum_{i=0}^k \frac{(\lambda \cdot nt)^i}{i!}. \quad (4.227)$$

In der Stichprobenprüfung interessiert man sich für die Ausfallrate λ der Bauelemente. Um diese zu bestimmen, wird die Zahl der Ausfälle k nach der Zahl der durchlaufenen Bauelementestunden nt gemessen. Dabei setzt man die Wahrscheinlichkeit, dass für die zu berechnende Ausfallrate λ höchstens k Ausfälle in der Stichprobe auftreten, mit einer kleinen Wahrscheinlichkeit an:

$$P_n(i \leq k) = 0,1. \quad (4.228)$$

Diese entspricht dem Signifikanzniveau α . Dann ist das Konfidenzniveau

$$P_n(i > k) = 1 - P_n(i \leq k) = 0,9 \quad (4.229)$$

entsprechend hoch, dass für den Wert λ in der Stichprobe sogar mehr als die gemessenen k Ausfälle auftreten und damit registriert werden können. Die Bestimmung von λ liegt damit auf der sicheren Seite. Die Gleichung

$$P_n(i \leq k) = 0,1 = e^{-\lambda \cdot nt} \sum_{i=0}^k \frac{(\lambda \cdot nt)^i}{i!} \quad (4.230)$$

ordnet der im Test registrierten Zahl von Ausfällen k genau einen Wert $\lambda \cdot nt$ zu. Man kann deshalb (4.230) auch als

$$\lambda \cdot nt = f(k) \quad (4.231)$$

interpretieren. Durch numerische Berechnung erhält man bei einer konstanten Ausfallrate λ für verschiedene Konfidenzniveaus $1 - \alpha$ die in der Tab. 4.4 aufgeführten Werte $f(k)$.

Tabelle 4.4. Korrekturfaktoren $f(k)$ zur Berechnung der Ausfallrate.

| Konfidenzniveau $1 - \alpha$ | Zahl der Ausfälle k | | | | | |
|---------------------------------|-----------------------|------|------|-------|-------|-------|
| | 0 | 1 | 2 | 3 | 4 | 5 |
| 0,70 | 1,39 | 2,69 | 3,92 | 5,06 | 6,27 | 7,42 |
| 0,80 | 1,61 | 2,99 | 4,28 | 5,52 | 6,72 | 7,91 |
| 0,90 | 2,30 | 3,89 | 5,32 | 6,68 | 7,99 | 9,27 |
| 0,95 | 3,00 | 4,74 | 6,30 | 7,75 | 9,15 | 10,60 |
| 0,99 | 4,60 | 6,64 | 8,41 | 10,04 | 11,60 | 13,11 |

Die Beziehung (4.231) kann auf doppelt-logarithmischem Papier als Geradenschar

$$\log \lambda = \log f(k) - \log(nt) \quad (4.232)$$

mit dem Parameter k dargestellt werden (Abb. 4.22). Zu Beginn des Tests ist $k = 0$: die Ausfallrate λ nimmt mit wachsender Bauelementestunden-Zahl nt entlang der Kurve $f(k = 0)$ ab. Wenn ein Ausfall registriert wird ($k = 1$), springt man auf die Kurve $f(k = 1)$ zu einer höheren Ausfallrate λ usw. Nach ausreichend langer Testzeit erreicht man einen asymptotischen Wert für λ . Aufgrund von sogenannten Frühausfällen ist die Ausfallrate bei kurzen Testzeiten meist höher als der asymptotische Wert.

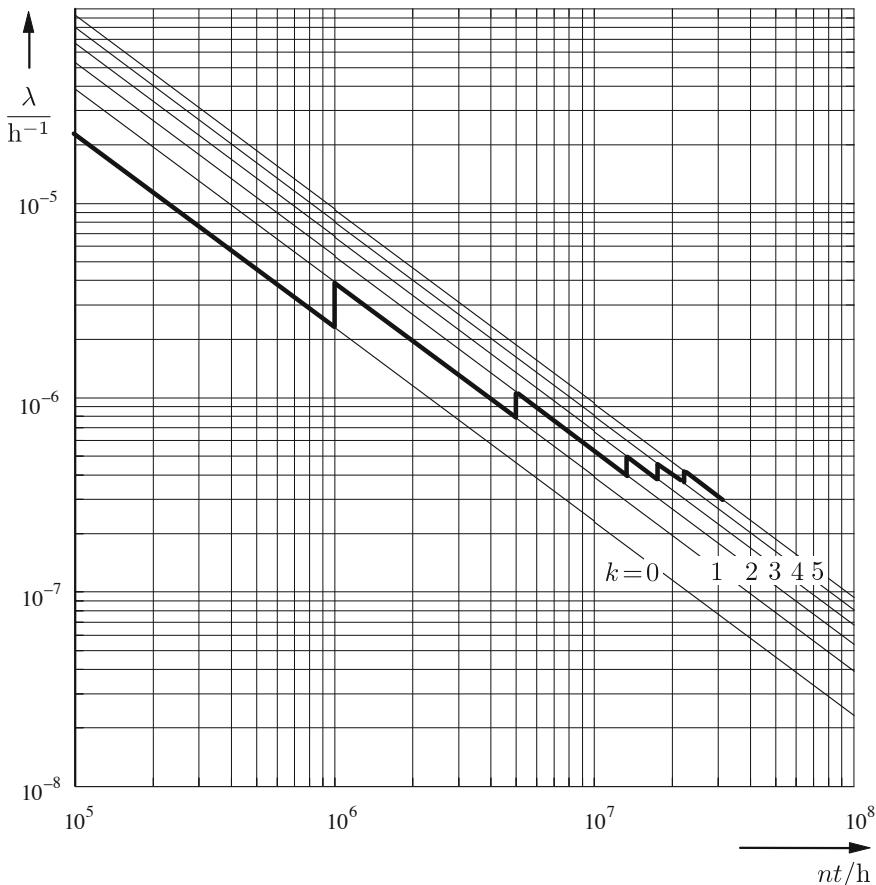


Abbildung 4.22. Bestimmung der Ausfallrate von elektronischen Bauelementen.

Zur Reduktion der Testzeiten werden die Bauelemente zumeist unter verschärften Testbedingungen geprüft. Man betreibt die Bauelemente z. B. bei höheren Temperaturen oder unterwirft sie Temperaturzyklen. Die erhöhte Beanspruchung lässt sich für spezielle Ausfallmechanismen in einen zeitlichen **Raffungsfaktor** r umrechnen, um den die Testzeit gekürzt werden kann.

Beispiel 4.13 (Ausfallrate): Eine Firma verarbeitet $3 \cdot 10^6$ Bauelemente eines Typs pro Jahr. Davon werden $n = 3000$ über 30 Tage, d. h. $t = 720 \text{ h}$, getestet. Der Test findet bei erhöhten Umgebungstemperaturen statt, woraus sich ein Raffungsfaktor von $r \approx 10$ gegenüber den normalen Einsatzbedingungen ergibt. Tabelle 4.5 zeigt die im Test aufgetretenen Ausfälle.

Tabelle 4.5. Gemessene Ausfallzeiten von Bauelementen.

| k | 1 | 2 | 3 | 4 | 5 |
|-----------------------|--------|----------------|------------------|------------------|-------------------|
| t/h | 33 | 167 | 433 | 567 | 720 |
| $r \cdot nt/\text{h}$ | 10^6 | $5 \cdot 10^6$ | $1,3 \cdot 10^7$ | $1,7 \cdot 10^7$ | $2,16 \cdot 10^7$ |

Bei $t = 720 \text{ h}$ wird der Test abgebrochen. Bei der Berechnung der Ausfallrate

$$\lambda = \frac{f(k)}{r \cdot nt} = \frac{9,27}{2,16 \cdot 10^7 \text{ h}} = 4,3 \cdot 10^{-7} \text{ h}^{-1} \quad (4.233)$$

wird nur die letzte Spalte von Tab. 4.5 berücksichtigt, um den Einfluss von Frühhausfällen zu unterdrücken. ■

Die Zahl n der in der Stichprobe getesteten Bauteile sollte möglichst groß sein. Dies ist von entscheidender Bedeutung für die Bewertung der Ausfallwahrscheinlichkeit, was am folgenden Beispiel verdeutlicht werden soll.

Beispiel 4.14 (Größe der Stichprobe): Für drei Lieferanten werden die folgenden Ausfälle gemessen:

Lieferant A: $k = 0$ von $n = 500$, d. h. 0 dpm

Lieferant B: $k = 1$ von $n = 2000$, d. h. 500 dpm

Lieferant C: $k = 6$ von $n = 10000$, d. h. 600 dpm

Man könnte nun versucht sein, den Lieferanten A aufgrund dieser Messergebnisse als den besten einzustufen, und Lieferanten C als den schlechtesten. Dies ist aber deshalb nicht richtig, da die Zahl n der getesteten Einheiten nicht korrekt berücksichtigt wurde. Die Rechnung muss lauten:

Lieferant A: $k < 1$ Ausfälle von $n = 500$, d. h. < 2000 dpm

Lieferant B: $k < 2$ Ausfälle von $n = 2000$, d. h. < 1000 dpm

Lieferant C: $k < 7$ Ausfälle von $n = 10000$, d. h. < 700 dpm

Damit ergibt sich aus der statistisch relevanten Abschätzung der Messresultate eine Umkehrung der Bewertungsskala. Der Lieferant C ist am besten! Bei Lieferant A müssten 1430 Bauelemente ohne Fehler getestet werden, um die gleiche Bewertung wie Lieferant C zu erlangen. Lieferant A hätte dann $k < 1$ Ausfälle von 1430, d. h. $10^6 / 1430 = 700$ dpm. Je niedriger die nachzuweisenden Ausfallraten, desto größer muss die Stichprobe sein, die getestet wird. Hohe Produktqualität (d. h. niedrige Ausfallraten) lässt sich deshalb nur bei großen Fertigungsstückzahlen

wirtschaftlich realisieren, bei denen man sich große Stichproben leisten kann. ■

4.5.3 Statistische Prozessüberwachung

In einem Fertigungsprozess unterliegen bestimmte Merkmale der gefertigten Güter Schwankungen – teils zufälliger, teils systematischer Art. Zufällige, mittelwertfreie Störungen wird man im Allgemeinen nicht verhindern können. Im Rahmen der Qualitätssicherung müssen aber systematische Fehler erkannt werden (vgl. Abschn. 4.5.1). Ausgehend von den Messwerten der relevanten Merkmale gilt es, die darin enthaltenen zufälligen Fertigungsfehler zu unterdrücken, um den systematischen Fehler zu erhalten. Eines der einfachsten Verfahren dafür ist die Mittelwertbildung (vgl. Stichprobenmittelwert)

$$\hat{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad (4.234)$$

bei der nach Abschn. 4.2.5 zufällige, mittelwertfreie Fehler herausfallen. Eine systematische Abweichung bleibt dagegen erhalten.

Beispiel 4.15 (Systematischer Fehler): Die Länge eines mechanischen Bauteils hat ein Sollmaß von 100 mm. Nun werden sechs Bauteile aus der laufenden Produktion herausgenommen. Sie besitzen folgende Längen:

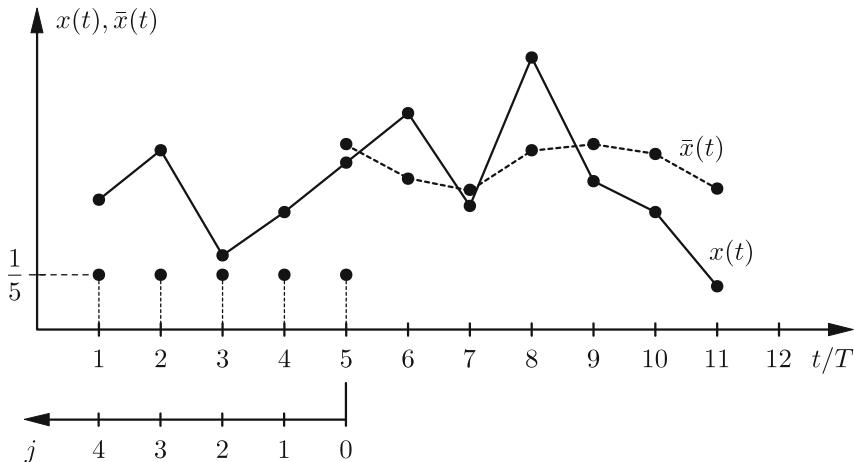
| | | | | | |
|----------|----------|---------|----------|---------|----------|
| 100,1 mm | 100,5 mm | 99,8 mm | 100,0 mm | 99,9 mm | 100,3 mm |
|----------|----------|---------|----------|---------|----------|

Hierbei ergibt sich ein Stichprobenmittelwert von $\hat{x} = 100,1$ mm. Unter der Annahme, dass die statistische Sicherheit bei diesem Probenumfang gewährleistet ist, kann man die Aussage treffen, dass bei der Fertigung ein systematischer Fehler von 0,1 mm auftritt. ■

Da systematische Fehler im Allgemeinen auch zeitabhängig sind, wird man bei instationären Prozessen mit der einfachen Mittelwertbildung nicht weiterkommen. Deshalb wird der **gleitende Mittelwert** (engl. *moving average*, MA) eingeführt. Er wird definiert als zeitabhängiger Mittelwert über die letzten m Messungen, die im Abstand T aufgenommen wurden:

$$\bar{x}(t) = \frac{1}{m} \sum_{j=0}^{m-1} x(t - jT). \quad (4.235)$$

Zum Zeitpunkt t wird also der älteste Wert $x(t - mT)$ „vergessen“ und der neueste Wert $x(t)$ hinzugenommen. Anschaulich ausgedrückt, wird über

Abbildung 4.23. Gleitender Mittelwert (MA-Filter) und Filterkoeffizienten für $m = 5$.

die Zeitreihe ein „Fenster“ der Breite mT geschoben, innerhalb dessen der Mittelwert gebildet wird (Abb. 4.23). Die einzelnen Werte $x(t)$ der Zeitreihe werden innerhalb des Fensters gleich gewichtet.

Nun schreibt man (4.235) auf symmetrische Summationsgrenzen um,

$$\bar{x}(t) = \frac{1}{2M+1} \sum_{j=-M}^{M} x(t-jT), \quad (4.236)$$

wobei $m = 2M + 1$ gesetzt wurde, damit jeweils eine ungerade Anzahl an Werten $x(t)$ gemittelt wird, was aber keine Einschränkung bedeutet. Der symmetrische gleitende Mittelwert $\bar{x}(t)$ kann nicht schritthalbend in Echtzeit berechnet werden, da der gegenwärtige Ausgangswert $\bar{x}(t)$ vom künftigen Eingangswert $x(t+MT)$ abhängig ist und das Filter somit aksual ist. Insofern muss bei der kausalen MA-Filterung eine Verzögerung (Totzeit) des Ausgangssignals der Länge MT in Kauf genommen werden.

Beispiel 4.16 (MA-Filterung): Eine Zeitreihe stelle die Differenz zwischen den Messwerten und dem Sollwert dar. Dieses Fehlersignal $x(t)$ bestehe aus einem systematischen Anteil $s(t)$ und einem zufälligen, mittelwertfreien Anteil $e(t)$:

$$x(t) = s(t) + e(t). \quad (4.237)$$

Wird M genügend groß gewählt, so wird der zufällige Anteil $\epsilon(t)$ infolge der gleitenden Mittelung unterdrückt:

$$\bar{x}(t) = \frac{1}{2M+1} \sum_{j=-M}^M (s(t-jT) + \epsilon(t-jT)) \quad (4.238)$$

$$= \frac{1}{2M+1} \left[\sum_{j=-M}^M s(t-jT) + \underbrace{\sum_{j=-M}^M \epsilon(t-jT)}_{\approx 0} \right] \quad (4.239)$$

$$= \frac{1}{2M+1} \sum_{j=-M}^M s(t-jT). \quad (4.240)$$

Somit bleibt nur der geglättete systematische Anteil übrig. ■

Bislang wurde über das Fehlersignal $x(t)$ nur ausgesagt, dass es aus einem systematischen und einem zufälligen, mittelwertfreien Anteil besteht. Im Folgenden interessieren vor allem die Eigenschaften des systematischen Anteils $s(t)$. Dieser muss nicht zwangsläufig konstant sein, sondern kann auch einen zeitabhängigen Anteil besitzen, der z. B. durch ein Wegdriften der überwachten Messgröße vom Sollwert verursacht wird. Dann kann der Verlauf von $s(t)$ z. B. durch das folgende **Signalmodell** angenähert werden:

$$s(t) = a_0 + a_1 t + \cdots + a_k t^k. \quad (4.241)$$

Der Modellansatz und die Filterung der zufälligen, mittelwertfreien Störungen $\epsilon(t)$ bieten die Möglichkeit, den systematischen Fehler $s(t)$ der Messwerte analytisch durch Bestimmung der Modellkoeffizienten a_i zu beschreiben. Dies erlaubt eine Prädiktion des zeitlichen Verlaufs von $s(t)$, wodurch sich ein unerwünschtes Wegdriften der Messgröße rechtzeitig erkennen lässt. So kennzeichnet in (4.241) der Parameter a_1 die Drift des Prozesses. Allgemein wird bei der statistischen Prozessüberwachung kontrolliert, ob die Koeffizienten a_i innerhalb eines vorgegebenen Toleranzintervalls liegen.

Der in Abschnitt 2.1.2 eingeführte Least-Squares-Schätzer findet in der statistischen Prozessüberwachung breiten Einsatz, da er zufällige Fehler inhärent unterdrückt und als Ergebnis die gesuchten Modellkoeffizienten a_i liefert. Dabei ist man nicht, wie in der Regressionsrechnung, auf Polynomsätze für den systematischen Fehler $s(t)$ beschränkt. Im Gegensatz zum Approximationsansatz für Kennlinien (2.23) wird bei der statistischen Prozessüberwachung die **zeitliche Abfolge der Messwerte** $y(t)$ modelliert. Das

Signalmodell lautet in zeitkontinuierlicher Form

$$y(t) = a_0 \varphi_0(t) + a_1 \varphi_1(t) + \cdots + a_k \varphi_k(t) + e(t). \quad (4.242)$$

Mit $t = nT$ ergibt sich für $m+1$ vergangene zeitdiskrete Messwerte

$$\hat{\mathbf{y}}_n = \begin{bmatrix} \hat{y}(n) \\ \hat{y}(n-1) \\ \vdots \\ \hat{y}(n-m) \end{bmatrix} = \begin{bmatrix} \varphi_0(nT) & \cdots & \varphi_k(nT) \\ \varphi_0((n-1)T) & \cdots & \varphi_k((n-1)T) \\ \vdots & \ddots & \vdots \\ \varphi_0((n-m)T) & \cdots & \varphi_k((n-m)T) \end{bmatrix} \begin{bmatrix} a_0(n) \\ a_1(n) \\ \vdots \\ a_k(n) \end{bmatrix} \quad (4.243)$$

$$= \Phi_n \mathbf{a}_n. \quad (4.244)$$

Zu jedem Messzeitpunkt nT kann damit der Parametervektor \mathbf{a}_n als Pseudoinverse

$$\mathbf{a}_n = (\Phi_n^T \Phi_n)^{-1} \Phi_n^T \mathbf{y}_n \quad (4.245)$$

bestimmt werden. Auf diese Weise können Veränderungen am Prozess frühzeitig entdeckt werden (Abb. 4.24).

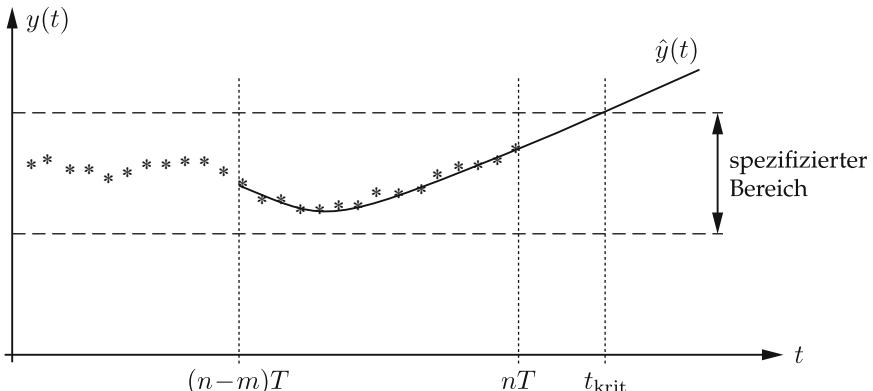


Abbildung 4.24. Least-Squares-Schätzung des zukünftigen Messgrößenverlaufs $y(t)$.

Beispiel 4.17 (Signalmodell): Gegeben sei ein Sinusgenerator, dessen Ausgangsspannung

$$u(t) = a_3 \cdot \sin(2\pi f_g t) \quad (4.246)$$

überwacht werden soll. Aus Vorüberlegungen sei bekannt, dass der Ausgangsverstärker durch Temperatureinflüsse eine lineare Drift $a_1 t$ und

einen Offset a_0 zeigt. Zusätzlich ist dem Ausgangssignal eine harmonische Netzstörung ($f_n = 50 \text{ Hz}$) überlagert. Die Phasenlage der harmonischen Schwingungen sei bekannt. Das messbare Signal einschließlich zufälliger Fehler $\epsilon(t)$ ist

$$y(t) = a_0 + a_1 t + a_2 \sin(2\pi f_n t) + a_3 \sin(2\pi f_g t) + \epsilon(t). \quad (4.247)$$

Die Parameter a_0 , a_1 und a_2 sind die systematischen Störeinflüsse, die es zu überwachen gilt. Für den zugehörigen LS-Schätzer ergibt sich zum Zeitpunkt $t = nT$:

$$\hat{\mathbf{y}}_n = \begin{bmatrix} 1 & nT & \sin(2\pi f_n n T) & \sin(2\pi f_g n T) \\ 1 & (n-1)T & \sin(2\pi f_n (n-1)T) & \sin(2\pi f_g (n-1)T) \\ \vdots & \vdots & \vdots & \vdots \\ 1 & (n-m)T & \sin(2\pi f_n (n-m)T) & \sin(2\pi f_g (n-m)T) \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix}. \quad (4.248)$$

Durch die Bildung der Pseudoinversen im Zeitpunkt $t = nT$

$$\mathbf{a}_n = (\Phi_n^T \Phi_n)^{-1} \Phi_n^T \mathbf{y}_n \quad (4.249)$$

wird der Parametervektor \mathbf{a}_n geschätzt, mit dem die systematischen Störeinflüsse überwacht werden können. Abbildung 4.25 zeigt oben ein stark gestörtes Messsignal mit der Generatorfrequenz $f_g = 70 \text{ Hz}$ und das mittels des LS-Schätzers rekonstruierte Signal (mittlere Grafik). Zum Zeitpunkt $t = 0,045 \text{ s}$ wurde eine Drift aufgeschaltet. Am Verlauf des geschätzten Parameters a_1 (untere Grafik) erkennt man die gute Eignung des LS-Schätzers zur statistischen Prozessüberwachung. ■

4.6 Fehlerfortpflanzung

Vielfach ist das gesuchte Messergebnis y nicht unmittelbar gleich dem Messwert, sondern berechnet sich vielmehr aus mehreren Messwerten x_i nach der Funktion

$$y = f(x_1, x_2, \dots, x_n) = f(\mathbf{x}) \quad \text{mit} \quad \mathbf{x} = (x_1, x_2, \dots, x_n)^T. \quad (4.250)$$

Ein einfaches Beispiel für eine derartige Funktion ist der Stichprobenmittelwert $y = \hat{x} = \frac{1}{n} \sum_i x_i$. In einem weiteren Beispiel kann der Wirkungsgrad eines Dampferzeugers nicht direkt gemessen werden, sondern muss aus den Einzelmesswerten

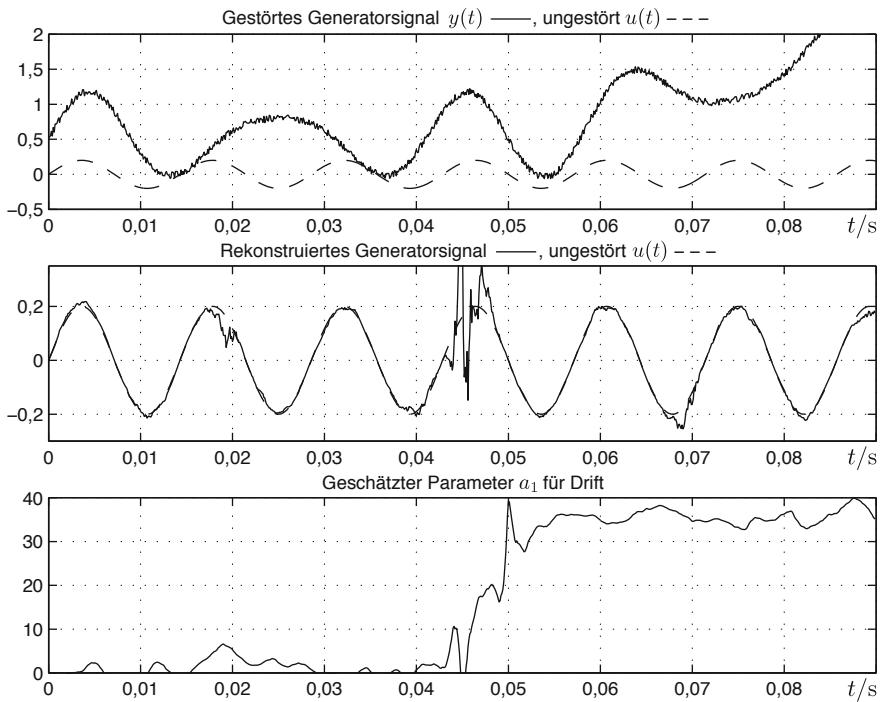


Abbildung 4.25. Statistische Prozessüberwachung eines Signalgenerators mit sprunghafter Änderung des Parameters a_1 bei $t = 0,045$ s.

- Heizwert des Brennstoffes,
- Zuführungsrate des Brennstoffes,
- Dampftemperatur und
- Dampfmenge pro Zeit

bestimmt werden. Da aber die Einzelmesswerte x_i im Allgemeinen fehlerbehaftet sind, also um den Wert

$$\Delta x_i = x_i - x_{i0} \quad (4.251)$$

vom richtigen Wert x_{i0} abweichen, wird auch das daraus berechnete Messergebnis y um einen gewissen Betrag Δy vom richtigen Messergebnis abweichen. Das Fehlerfortpflanzungsgesetz liefert eine Abschätzung dieses Fehlers Δy aus den Einzelmessfehlern Δx_i .

Der Fehler des Messergebnisses y kann mit Hilfe der Taylor-Entwicklung abgeschätzt werden. Bei hinreichend kleinen Messfehlern $|\Delta x_i|$ können die

Glieder höherer Ordnung vernachlässigt werden und man erhält:

$$\Delta y = y - y_0 = \sum_i \left. \frac{\partial f(\mathbf{x})}{\partial x_i} \right|_{\mathbf{x}_0} \Delta x_i . \quad (4.252)$$

Dabei beschreibt $\partial f(\mathbf{x})/\partial x_i$ die Empfindlichkeit der Funktion $f(\mathbf{x})$ für den Messwert x_i .

Manchmal sind nicht die Messfehler Δx_i bekannt, sondern nur ihre Grenzen⁶ $\Delta x_{g,i}$. Dann lässt sich die Fehlergrenze Δy_g des Messergebnisses analog zu (4.252) berechnen:

$$\Delta y_g = \sum_i \left. \left| \frac{\partial f(\mathbf{x})}{\partial x_i} \right| \right|_{\mathbf{x}_0} \Delta x_{g,i} . \quad (4.253)$$

Abhängig von der Art der mathematischen Verknüpfung $f(\mathbf{x})$ der Einzelmesswerte x_i lassen sich zwei wichtige Sonderfälle von (4.252) unterscheiden:

- Ist y eine Linearkombination (Addition/Subtraktion) der Messwerte x_i ,

$$y = f(\mathbf{x}) = a_1 x_1 + a_2 x_2 + \cdots + a_n x_n , \quad (4.254)$$

so ergibt sich der Fehler des Messergebnisses nach (4.252) zu:

$$\Delta y = a_1 \Delta x_1 + a_2 \Delta x_2 + \cdots + a_n \Delta x_n . \quad (4.255)$$

Der absolute Gesamtfehler Δy ist also die Summe aller mit den Koeffizienten a_i gewichteten Einzelfehler.

- Im Falle einer multiplikativen Verknüpfung der Messwerte x_i gemäß

$$y = f(\mathbf{x}) = a_1 x_1^{\alpha_1} \cdot a_2 x_2^{\alpha_2} \cdot \cdots \cdot a_n x_n^{\alpha_n} \quad (4.256)$$

(Produkt-/Quotientenbildung) rechnet man bevorzugt mit dem relativen Fehler. Mit den partiellen Ableitungen

$$\frac{\partial f(\mathbf{x})}{\partial x_i} = a_1 x_1^{\alpha_1} \cdot \cdots \cdot \alpha_i a_i x_i^{\alpha_i-1} \cdot \cdots \cdot a_n x_n^{\alpha_n} = y \cdot \frac{\alpha_i}{x_i} , \quad (4.257)$$

erhält man aus (4.252) den folgenden Ausdruck für den absoluten Gesamtfehler Δy :

$$\Delta y = y \cdot \sum_{i=1}^n \alpha_i \frac{\Delta x_i}{x_i} . \quad (4.258)$$

⁶Fehlergrenzen sind vereinbarte oder garantierter Höchstwerte für betragsmäßige Abweichungen des angezeigten Wertes eines Messsystems vom richtigen Wert.

Der relative Gesamtfehler $\Delta y/y$ resultiert demnach aus der Summe der relativen, mit den jeweiligen α_i gewichteten Einzelfehler $\Delta x_i/x_i$.

Zufällige Messfehler werden meist durch ihre Standardabweichung σ_{x_i} oder durch ihre Varianz $\sigma_{x_i}^2$ beschrieben. Bei betragsmäßig kleinen, zufälligen Messfehlern Δx_i gilt für den Erwartungswert des Messergebnisses $\mu_y \approx y_0$, weshalb sich dessen Varianz σ_y^2 wie folgt approximieren lässt:

$$\sigma_y^2 = E \left\{ (y - \mu_y)^2 \right\} \quad (4.259)$$

$$\approx E \left\{ \left(\sum_i \frac{\partial f(\mathbf{x})}{\partial x_i} \Big|_{\mathbf{x}_0} (x_i - x_{i0}) \right) \left(\sum_j \frac{\partial f(\mathbf{x})}{\partial x_j} \Big|_{\mathbf{x}_0} (x_j - x_{j0}) \right) \right\} \quad (4.260)$$

$$= \sum_i \sum_j \frac{\partial f(\mathbf{x})}{\partial x_i} \Big|_{\mathbf{x}_0} \cdot \frac{\partial f(\mathbf{x})}{\partial x_j} \Big|_{\mathbf{x}_0} E \{ (x_i - x_{i0})(x_j - x_{j0}) \} \quad (4.261)$$

$$\approx \sum_i \sum_j \frac{\partial f}{\partial x_i}(\mathbf{x}_0) \frac{\partial f}{\partial x_j}(\mathbf{x}_0) C_{x_i x_j}, \quad (4.262)$$

wobei in der letzten Zeile für die partiellen Ableitungen die verkürzte Schreibweise gemäß (2.117) verwendet wurde.

Für statistisch unabhängige Messwerte x_i wird die Kovarianz (4.38) zu $C_{x_i x_j} = \sigma_{x_i}^2 \delta_{ij}$. Damit folgt das **Gauß'sche Fehlerfortpflanzungsgesetz**:

$$\sigma_y^2 \approx \sum_i \left[\frac{\partial f}{\partial x_i}(\mathbf{x}_0) \right]^2 \sigma_{x_i}^2. \quad (4.263)$$

Die Varianz des Messergebnisses y resultiert somit aus einer gewichteten Addition der Varianzen der einzelnen Messwerte x_i .

Sind die Messwerte statistisch abhängig, so verwendet man den Korrelationskoeffizienten $\rho_{x_i x_j}$ zur Berechnung der Fehlerfortpflanzung:

$$\sigma_y^2 \approx \sum_i \left[\frac{\partial f}{\partial x_i}(\mathbf{x}_0) \right]^2 \sigma_{x_i}^2 + \sum_i \sum_{j \neq i} \frac{\partial f}{\partial x_i}(\mathbf{x}_0) \frac{\partial f}{\partial x_j}(\mathbf{x}_0) \sigma_{x_i} \sigma_{x_j} \cdot \rho_{x_i x_j}. \quad (4.264)$$

Wird das Messergebnis y als Produkt oder als Quotient mehrerer Messwerte x_i gebildet, so folgt aus (4.258) für den relativen Fehler:

$$\frac{\Delta y}{y_0} = \sum_{i=1}^n \alpha_i \cdot \frac{\Delta x_i}{x_{i0}}. \quad (4.265)$$

Die relative Varianz ist damit

$$\frac{\sigma_y^2}{y_0^2} = E \left\{ \frac{\Delta y^2}{y_0^2} \right\} = E \left\{ \left[\sum_{i=1}^n \alpha_i \cdot \frac{\Delta x_i}{x_{i0}} \right]^2 \right\} \quad (4.266)$$

$$= \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j E \left\{ \frac{\Delta x_i}{x_{i0}} \frac{\Delta x_j}{x_{j0}} \right\} = \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j \frac{C_{x_i x_j}}{x_{i0} x_{j0}}. \quad (4.267)$$

Für statistisch unabhängige Messwerte x_i beträgt die Kovarianz nach (4.38):

$$C_{x_i x_j} = \sigma_{x_i}^2 \delta_{ij}. \quad (4.268)$$

Nach dem Gauß'schen Fehlerfortpflanzungsgesetz erhält man damit die relative Varianz zu

$$\frac{\sigma_y^2}{y_0^2} = \sum_{i=1}^n \alpha_i^2 \frac{\sigma_{x_i}^2}{x_{i0}^2}. \quad (4.269)$$

Beispiel 4.18 (Fehler bei der Massebestimmung): Die Masse einer in einem zylindrischen Tank gelagerten Flüssigkeit soll bestimmt werden. Der Durchmesser des Tanks sei d , der gemessene Flüssigkeitsstand der Flüssigkeit h , die Dichte ρ . Die gelagerte Masse bestimmt sich aus der Beziehung

$$m = \pi \left(\frac{d}{2} \right)^2 \cdot \rho \cdot h. \quad (4.270)$$

Mit $\alpha_d = 2$, $\alpha_\rho = 1$ und $\alpha_h = 1$ errechnet sich die relative Varianz der Masse m bei statistisch unabhängigen Messwerten aus (4.269) zu

$$\left(\frac{\sigma_m}{m} \right)^2 = 4 \left(\frac{\sigma_d}{d} \right)^2 + \left(\frac{\sigma_h}{h} \right)^2 + \left(\frac{\sigma_\rho}{\rho} \right)^2. \quad (4.271)$$

Ist der Durchmesser auf $\sigma_d/d = 1\%$ bestimmt, die Höhe auf $\sigma_h/h = 0,5\%$ und die Dichte auf $\sigma_\rho/\rho = 0,9\%$, so erhält man damit

$$\frac{\sigma_m}{m} = \sqrt{4 + 0,25 + 0,81} \% = 2,2 \% \quad (4.272)$$

für die relative Streuung der Masse. ■



Kapitel 5

Dynamisches Verhalten von Messsystemen

5

5

5 Dynamisches Verhalten von Messsystemen

| | | |
|---------|---|-----|
| 5.1 | Beschreibung von Messsystemen | 181 |
| 5.1.1 | Systemeigenschaften | 181 |
| 5.1.2 | Lineare, zeitinvariante Systeme (LTI-Systeme) | 182 |
| 5.1.3 | Stabilität | 184 |
| 5.2 | Empirische Analyse von Messsystemen | 186 |
| 5.2.1 | Kennwerte der Sprungantwort | 186 |
| 5.2.2 | Nichtlineares dynamisches Verhalten | 187 |
| 5.2.3 | Bestimmung des Frequenzganges | 188 |
| 5.3 | Verbesserung des dynamischen Systemverhaltens | 192 |
| 5.3.1 | Optimierung der Systemparameter | 193 |
| 5.3.2 | Änderung der Systemstruktur | 194 |
| 5.4 | Parameteroptimierung | 195 |
| 5.4.1 | Kriterium „verschwindende Momente der Impulsantwort“ | 198 |
| 5.4.2 | Kriterium „konstanter Amplitudengang für kleine Frequenzen“ | 202 |
| 5.4.3 | Kriterium „konstanter Realteil des Frequenzganges“ . . | 208 |
| 5.4.4 | ITAE-Kriterium | 214 |
| 5.4.4.1 | System 1. Ordnung | 214 |
| 5.4.4.2 | System 2. Ordnung | 215 |
| 5.4.5 | Kriterium „quadratisches Fehlerintegral“ | 219 |
| 5.5 | Strukturänderung zur Optimierung des Zeitverhaltens | 226 |
| 5.5.1 | Kompensation des Zeitverhaltens | 226 |
| 5.5.2 | Zeitverhalten bei Gegenkopplung | 231 |
| 5.5.2.1 | P-Regler | 232 |
| 5.5.2.2 | PI-Regler | 236 |

5 Dynamisches Verhalten von Messsystemen

5.1 Beschreibung von Messsystemen

5.1

5.1.1 Systemeigenschaften

Unter einem **System** versteht man eine Menge zusammenhängender Komponenten, die der Erfüllung eines bestimmten Zweckes dienen. In der Messtechnik ist dieser Zweck meist die Umformung einer zeitabhängigen Messgröße $u(t)$ zu einer Ausgangsgröße $y(t)$. Im Folgenden wird ein System allgemein als eine Vorschrift zur Transformation eines Eingangssignals $u(t)$ in ein Ausgangssignal $y(t)$ aufgefasst. Mathematisch kann diese Vorschrift mit Hilfe des Operators $\mathcal{S}\{\cdot\}$ beschrieben werden:

$$y(t) = \mathcal{S}\{u(t)\}. \quad (5.1)$$

Abhängig von der mathematischen Struktur dieser Abbildung können verschiedene Klassen von Systemen unterschieden werden (Abb. 5.1). Der Einheitlichkeit halber werden im Folgenden die Signale $u(t)$ und $y(t)$ als eindimensional angenommen; es ist jedoch problemlos möglich, diese Betrachtung auf mehrdimensionale Größen zu erweitern.

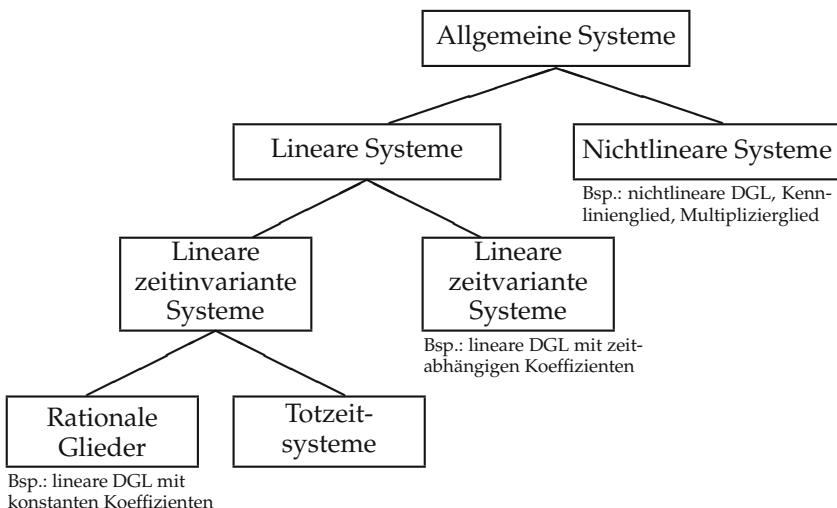


Abbildung 5.1. Klassifikation von Systemen.

Definition 5.1: Linearität

Bei **linearen** Systemen gelten das Superpositionsprinzip (Additivität) und das Verstärkungsprinzip (Homogenität):

- **Superpositionsprinzip:** Treten am Eingang verschiedene Eingangssignale $u_1(t), u_2(t), \dots$ auf, so addieren sich ihre Auswirkungen auf das Ausgangssignal $y(t)$ unabhängig voneinander zur Gesamtwirkung.
- **Verstärkungsprinzip:** Ver- n -facht man das Eingangssignal $u(t)$, so ver- n -facht sich auch das Ausgangssignal $y(t)$. ◇

Besteht demnach ein Eingangssignal aus der gewichteten Überlagerung mehrerer Signale, so bewirkt ein lineares System ein Ausgangssignal, welches der gewichteten Summe der jeweiligen Ausgangssignale entspricht:

$$\mathcal{S}\{c_1 u_1(t) + c_2 u_2(t)\} = c_1 \mathcal{S}\{u_1(t)\} + c_2 \mathcal{S}\{u_2(t)\}, \quad c_i \in \mathbb{R}. \quad (5.2)$$

Bei realen Systemen kann Linearität nur innerhalb gewisser Grenzen erreicht werden.

Definition 5.2: Zeitinvarianz

Ein System $\mathcal{S}\{ \cdot \}$ heißt **zeitinvariant**, wenn es auf ein zeitlich verschobenes Eingangssignal $u(t - t_0)$ mit dem entsprechend zeitlich verschobenen Ausgangssignal $y(t - t_0)$ antwortet:

$$y(t) = \mathcal{S}\{u(t)\} \quad \Rightarrow \quad \mathcal{S}\{u(t - t_0)\} = y(t - t_0). \quad (5.3)$$

Das Systemverhalten ist somit unabhängig von der Zeit t . Systeme, die (5.3) nicht genügen, heißen **zeitvariant**. ◇

Definition 5.3: Kausalität

Ein System $\mathcal{S}\{ \cdot \}$ heißt **kausal**, wenn das Ausgangssignal $y(t)$ zu einem bestimmten Zeitpunkt t_0 nur vom Verlauf des Eingangssignals $u(t)$ zu Zeitpunkten $t \leq t_0$ abhängt. Die Antwort auf eine Erregung kann also nicht vor der Erregung selbst vorliegen. ◇

5.1.2 Lineare, zeitinvariante Systeme (LTI-Systeme)

Lineare, zeitinvariante (engl. *linear time-invariant*, LTI) Systeme lassen eine einfache mathematische Behandlung zahlreicher wichtiger Systeme zu, weshalb sie in der Praxis eine große Bedeutung haben.

Im Zeitbereich werden zeitkontinuierliche LTI-Systeme ohne Totzeit durch **lineare Differentialgleichungen** mit konstanten Koeffizienten be-

schrieben:

$$b_n y^{(n)}(t) + \dots + b_1 \dot{y}(t) + y(t) = K [a_m u^{(m)}(t) + \dots + a_1 \dot{u}(t) + u(t)] \quad (5.4)$$

mit $n \geq m$ für physikalisch realisierbare Systeme, wobei

$$y^{(n)}(t) = \frac{d^n y(t)}{dt^n} \quad (5.5)$$

die n -te zeitliche Ableitung von $y(t)$ bezeichnet. Transformiert man (5.4) in den Laplace-Bereich und setzt dabei alle Anfangswerte zu null, so erhält man:

$$Y(s) [b_n s^n + \dots + b_1 s + 1] = K U(s) [a_m s^m + \dots + a_1 s + 1] \quad (5.6)$$

mit $s \in \mathbb{C}$. Daraus folgt die **Laplace-Übertragungsfunktion**, mit der sich das dynamische Verhalten eines zeitkontinuierlichen LTI-Systems ohne Totzeit im Frequenzbereich beschreiben lässt [42]:

$$G(s) = \frac{Y(s)}{U(s)} = G(0) \frac{1 + \sum_{j=1}^m a_j s^j}{1 + \sum_{i=1}^n b_i s^i}, \quad m \leq n, \quad (5.7)$$

mit $G(0) = K$ und den **Systemparametern**

$$\mathbf{a} = [a_1, a_2, \dots, a_m]^T \quad \text{und} \quad \mathbf{b} = [b_1, b_2, \dots, b_n]^T. \quad (5.8)$$

Für $s = j\omega = j2\pi f$ erhält man aus der Übertragungsfunktion (5.7) den **Frequenzgang** $G(f)$ [42].

Beispiel 5.1 (System 2. Ordnung): Bei einem System mit Verzögerung 2. Ordnung (PT₂-Glied) laute die Übertragungsfunktion

$$G(s) = \frac{1}{T^2 s^2 + 2\delta T s + 1}, \quad (5.9)$$

wobei der Parameter δ die **Dämpfung** bezeichnet. Zur Beschreibung des dynamischen Systemverhaltens kann die **Sprungantwort** herangezogen werden, also die Antwort auf ein sprungförmiges Eingangssignal. Das System (5.9) besitzt die Sprungantwort

$$h(t) = 1 + \frac{1}{T^2(s_{\infty,1} - s_{\infty,2})} \left[\frac{1}{s_{\infty,1}} e^{s_{\infty,1} t} - \frac{1}{s_{\infty,2}} e^{s_{\infty,2} t} \right]. \quad (5.10)$$

Die Dämpfung δ beeinflusst maßgeblich das Einschwingverhalten des Systems. Für Dämpfungen $\delta \geq 1$ ergeben sich zwei reelle Polstellen $s_{\infty,1}$

und $s_{\infty,2}$ – derartige Systeme sind **nicht schwingungsfähig** und können als Reihenschaltung zweier PT₁-Glieder dargestellt werden. Für $\delta < 1$ erhält man dagegen zwei konjugiert komplexe Polstellen, was zu einer schwingenden Einstellung der Sprungantwort führt. Ist für eine gegebene Anwendung kein Überschwingen erlaubt, so wird man den **aperiodischen Grenzfall** mit $\delta = 1$ einstellen, denn für große Dämpfungen $\delta \gg 1$ wird der stationäre Endwert hingegen nur sehr langsam erreicht. ■

Anstelle der mathematischen Modellbildung kann das Systemverhalten auch experimentell ermittelt werden. Dafür werden geeignete **Testsignale** an den Eingang des Systems gegeben. Der Verlauf des Ausgangssignals ist typisch für das jeweils betrachtete System und liefert Informationen über dessen dynamische Eigenschaften. Als Testsignale kommen insbesondere Sprungsignale und harmonische Funktionen in Frage (Abschn. 5.2).

5.1.3 Stabilität

Allen in diesem Kapitel vorgestellten Kriterien zur Verbesserung des dynamischen Verhaltens ist gemeinsam, dass die Stabilität gesondert zu prüfen ist. Die Anwendung eines Optimierungskriteriums ergibt nicht automatisch ein stabiles System.

5.1

Satz 5.1: Stabilität von LTI-Systemen

Ein lineares, zeitinvariantes System ist stabil, wenn alle Polstellen der Übertragungsfunktion $G(s)$ in der linken s -Halbebene liegen, also wenn die Bedingung

$$\Re\{s_{\infty_i}\} < 0 \quad \forall i \tag{5.11}$$

erfüllt ist [42]. ◇

Für Nennerpolynome vom Grad 3 oder höher ist die Berechnung der Polstellen zur Überprüfung auf Stabilität meist nur noch numerisch möglich. Die zahlreichen in der Literatur beschriebenen Stabilitätskriterien vermeiden aber die explizite Berechnung der Polstellen [11, 18, 28]. Das Hurwitz-Kriterium ist eine elegante Art, die Stabilität eines Systems anhand der Systemparameter zu prüfen [11].

Satz 5.2: Stabilitätskriterium nach Hurwitz

Gegeben sei ein Nennerpolynom mit reellen Koeffizienten:

$$N(s) = \lambda_0 s^n + \lambda_1 s^{n-1} + \cdots + \lambda_{n-1} s + \lambda_n, \quad \lambda_0 > 0, \quad \lambda_i \in \mathbb{R}. \quad (5.12)$$

Zur Stabilitätsprüfung bildet man die **Hurwitz-Determinante**

$$H = \begin{vmatrix} \lambda_1 & \lambda_3 & \lambda_5 & \lambda_7 & \cdots \\ \lambda_0 & \lambda_2 & \lambda_4 & \lambda_6 & \cdots \\ 0 & \lambda_1 & \lambda_3 & \lambda_5 & \cdots \\ 0 & \lambda_0 & \lambda_2 & \lambda_4 & \cdots \\ 0 & 0 & \lambda_1 & \lambda_3 & \cdots \\ 0 & 0 & \lambda_0 & \lambda_2 & \cdots \\ \cdot & \cdot & \cdot & \cdot & \cdots \end{vmatrix} \quad (5.13)$$

und von dieser die nordwestlichen Unterdeterminanten

$$H_1 = \lambda_1, \quad H_2 = \begin{vmatrix} \lambda_1 & \lambda_3 \\ \lambda_0 & \lambda_2 \end{vmatrix}, \quad \dots, \quad H_n. \quad (5.14)$$

Sind alle Determinanten

$$H_1, \dots, H_n > 0, \quad (5.15)$$

so liegen alle Nullstellen von $N(s)$ in der linken s -Halbebene. Andernfalls liegt mindestens eine Nullstelle auf der j -Achse oder rechts davon. \diamond

Lässt sich also das Hurwitz-Kriterium für das Nennerpolynom einer Übertragungsfunktion $G(s)$ erfüllen, so ist das System stabil. Man beachte dabei die unterschiedliche Indizierung der Koeffizienten b_i und λ_i im Polynom $N(s)$.

Beispiel 5.2 (Stabilität bei Systemen 3. Ordnung): Gegeben sei eine Übertragungsfunktion $G(s) = Z(s)/N(s)$ mit einem Nennerpolynom dritten Grades:

$$N(s) = \lambda_0 s^3 + \lambda_1 s^2 + \lambda_2 s + \lambda_3 = 0, \quad \lambda_0 > 0. \quad (5.16)$$

Die Hurwitz-Determinante berechnet sich damit zu

$$H = \begin{vmatrix} \lambda_1 & \lambda_3 & 0 \\ \lambda_0 & \lambda_2 & 0 \\ 0 & \lambda_1 & \lambda_3 \end{vmatrix}. \quad (5.17)$$

Das System ist stabil, wenn alle nordwestlichen Unterdeterminanten positiv sind:

$$H_1 = \lambda_1 > 0, \quad (5.18)$$

$$H_2 = \begin{vmatrix} \lambda_1 & \lambda_3 \\ \lambda_0 & \lambda_2 \end{vmatrix} = \lambda_1 \lambda_2 - \lambda_0 \lambda_3 > 0, \quad (5.19)$$

$$H_3 = H = \lambda_3 H_2 > 0. \quad (5.20)$$

Damit sind sämtliche Bedingungen für die Parameter festgelegt. Das System 3. Ordnung ist somit stabil, wenn

$$\lambda_0, \lambda_1, \lambda_3 > 0 \quad \text{und} \quad \lambda_1 \lambda_2 > \lambda_0 \lambda_3 \quad (5.21)$$

gilt. ■

5.2

5.2 Empirische Analyse von Messsystemen

Im idealen Messgerät soll das Ausgangssignal fehlerlos den zeitlichen Verlauf der Messgröße wiedergeben. Zur Beschreibung des dynamischen Verhaltens von Messsystemen wird an den Eingang ein Sprungsignal aufgegeben und die Abweichung der Ausgangsgröße von der idealen **Sprungantwort** als Fehler festgestellt. Abbildung 5.2 zeigt die Anordnung für einen Druckmesser sowie die Auswertung der Sprungantwort. Beim Aufbau der Anordnung muss darauf geachtet werden, dass der Messgrößensprung innerhalb einer Zeitspanne erfolgt, die etwa eine Größenordnung unter der Einstellzeit des Druckmessers liegt. Parallel dazu wird ein schneller Druckmesser mit elektrischem Ausgang – ein sogenanntes Präzisionsinstrument G_n (vgl. Abb. 1.4) – benötigt. Weiter muss die Einstellzeit des Datenerfassungsgerätes erheblich unter der Einstellzeit des Messgerätes liegen.

5.2.1 Kennwerte der Sprungantwort

Beim Auswerten der Sprungantwort kann zwischen schwingender und aperiodischer Einstellung unterschieden werden (Abb. 5.2).

- Bei **schwingender Einstellung** werden folgende Kennwerte abgelesen:

- T_1 : Zeit, die vergeht, bis der stationäre Wert y_∞ zum ersten mal erreicht wird;
- τ : Zeit, die vergeht, bis ein Toleranzband von 1 % der Sprunghöhe y_∞ nicht mehr verlassen wird;

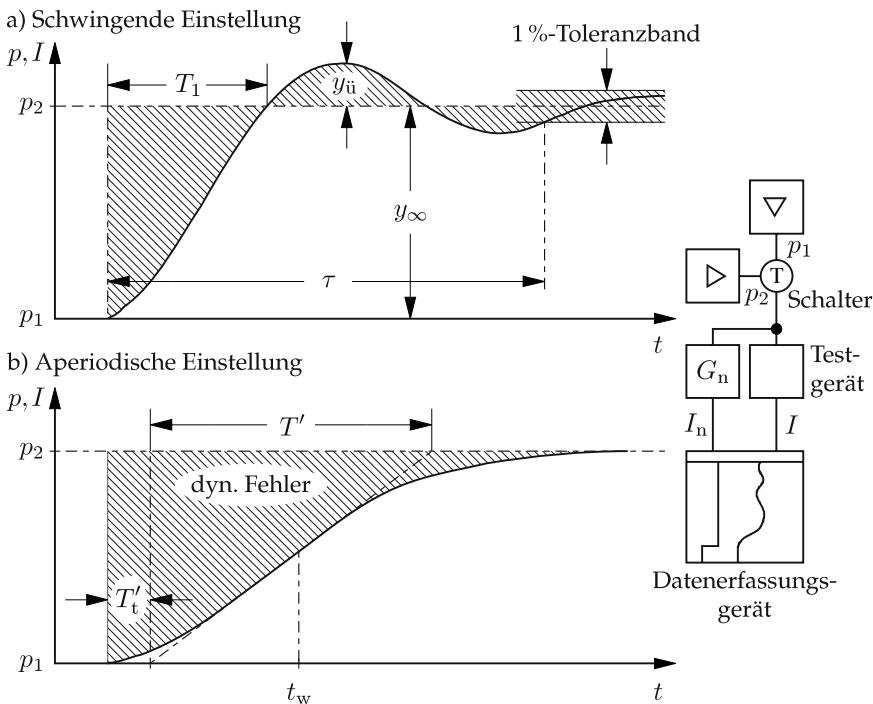


Abbildung 5.2. Messeinrichtung und Kennwerte der Sprungantwort.

$y_{\text{ü}}$: maximale Überschwingweite, angegeben in Prozent der Sprunghöhe y_∞ .

- Bei **aperiodischer Einstellung** lässt sich das Verhalten eines Messsystems durch eine Reihenschaltung eines Totzeitgliedes und eines PT_1 -Gliedes annähern. Durch Anlegen einer Tangente im Wendepunkt t_w der Sprungantwort (**Wendetangentenverfahren**) werden folgende Kennwerte abgelesen:

T'_t : Ersatztotzeit (Verzugszeit zwischen dem Sprungeinsatz und dem Schnittpunkt der Wendetangente mit der Nulllinie);

T' : Ersatzzeitkonstante (Ausgleichszeit zwischen den Schnittpunkten der Wendetangente mit der Nulllinie und dem stationären Wert).

5.2.2 Nichtlineares dynamisches Verhalten

Viele Messeinrichtungen, die statisch eine lineare Kennlinie haben, erweisen sich bei dynamischen Messungen als nichtlinear. Eine Messeinrichtung zeigt ein lineares Verhalten, wenn dem Eingangssignal mit der k -fachen Amplitude zu jedem Zeitpunkt auch das k -fache Ausgangssignal entspricht. In Abb.

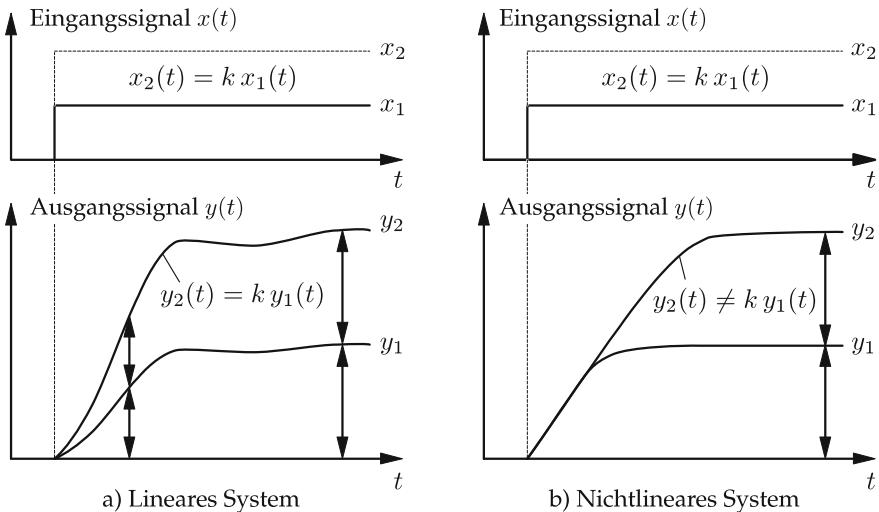


Abbildung 5.3. Lineares und nichtlineares dynamisches Verhalten.

5.3 ist neben linearem Verhalten (links) auch nichtlineares dynamisches Verhalten demonstriert (rechts). Dem doppelt so hohen Sprung entspricht zwar auch beim betrachteten nichtlinearen System *im eingeschwungenen Zustand* ein doppelt so großer Ausschlag. Der stationäre Wert wird aber für beide Sprunghöhen mit der gleichen Maximalgeschwindigkeit (d. h. Steigung der Sprungantwort) erreicht. Ein solches nichtlineares dynamisches Verhalten findet man oft bei Messeinrichtungen nach dem **Kompensationsverfahren**, wenn ein Stellglied in seinen möglichen Ausgangswerten begrenzt ist. Doch auch hier erweist sich die Sprungfunktion als ein geeignetes Signal, um das Zeitverhalten zu charakterisieren. Die Versuche sind dann allerdings mit verschiedenen Sprunghöhen und somit für unterschiedliche Werte der Messgröße durchzuführen.

5.2.3 Bestimmung des Frequenzganges

Eine andere, deutlich aufwendigere Methode zur Bestimmung des dynamischen Fehlers ist die **Messung des Frequenzganges** mit Hilfe harmonischer Anregungen der Messgröße. Der Zeitaufwand ist beträchtlich, da viele Frequenzen der Messgröße angefahren und jeweils der stationäre Zustand des Messsystems abgewartet werden muss. Außerdem lassen sich von vielen technischen Messgrößen kaum harmonische Signale erzeugen, weil lineare Stellglieder nicht existieren.

Abbildung 5.4 a) zeigt die Anordnung für einen Druckmesser. Ein Sinusgeber mit mechanischem oder elektrischem Ausgangssignal steuert einen

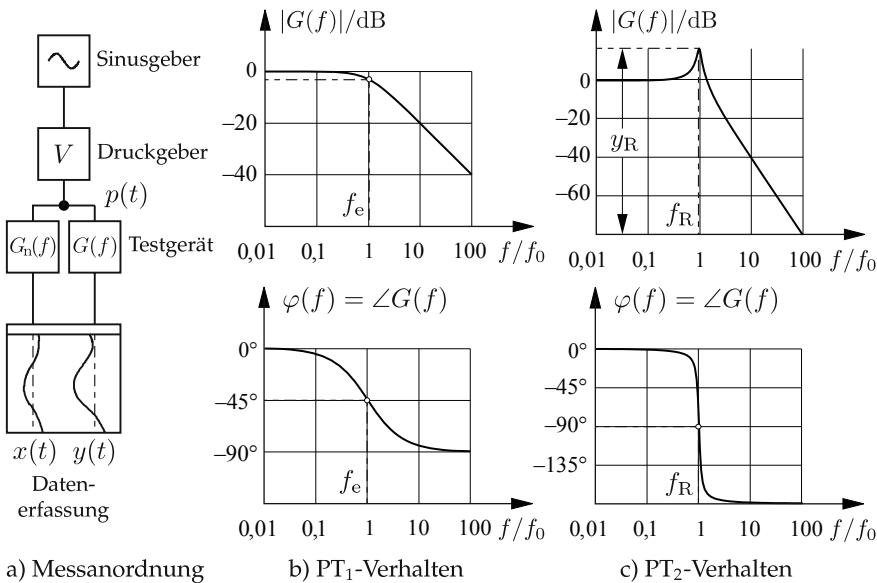


Abbildung 5.4. Aufnahme des Frequenzganges eines Druckmessers.

Druckgeber, der mit linearer Kennlinie und kurzer Ansprechzeit ein harmonisches Drucksignal

$$p(t) = |p_0| \sin(2\pi ft) \quad (5.22)$$

liefert. Dieses Signal wirkt auf das zu untersuchende Messgerät G und parallel dazu auf einen erheblich schnelleren Druckaufnehmer G_n , für welchen **ideales Übertragungsverhalten**

$$G_n(f) = 1 \quad (5.23)$$

angenommen wird (vgl. Abschn. 5.3). Somit entspricht das Signal

$$x(t) = |x_0| \sin(2\pi ft) = |p_0| \sin(2\pi ft) \quad (5.24)$$

am Ausgang des Referenzmessgerätes G_n gleichzeitig dem Signal am Eingang des untersuchten Messgerätes G . Bei Annahme eines linearen, stabilen Systems weist der Ausgang des Messgerätes G im eingeschwungenen Zustand ebenfalls einen sinusförmigen Verlauf

$$y(t) = |y_0(f)| \sin(2\pi ft + \varphi_y(f)) \quad (5.25)$$

auf, der sich jedoch im Allgemeinen in Amplitude und Phasenlage vom Eingangssignal $x(t)$ unterscheidet. Beide Größen hängen von der Frequenz f ab.

Schließlich werden für jede angefahrene Frequenz f beide Signale $x(t)$ und $y(t)$ im eingeschwungenen Zustand aufgezeichnet und, durch Auswertung ihrer Amplituden und Phasenlagen, der gesuchte Frequenzgang des Messsystems $G(f)$ wie folgt ermittelt:

$$G(f) = \frac{|y_0(f)|}{|x_0|} \cdot e^{j\varphi_y(f)} = \frac{|y_0(f)|}{|p_0|} \cdot e^{j\varphi_y(f)}. \quad (5.26)$$

Der resultierende Frequenzgang wird als sogenanntes **Bode-Diagramm** dargestellt [42]. Dabei wird sowohl das Amplitudenverhältnis $|G(f)|$ als auch die Phase $\varphi(f) = \angle G(f)$ über der Frequenz aufgetragen. Als Frequenzachse wird der Logarithmus des Frequenzverhältnisses f/f_0 mit der Bezugsfrequenz f_0 verwendet. Abbildung 5.4 zeigt zwei Beispiele für Systeme mit unterschiedlichem dynamischen Verhalten.

- **Amplitudengang:** Der Betrag des Frequenzganges $|G(f)|$ des Messsystems heißt Amplitudengang und wird in Dezibel aufgetragen, wobei

$$|G(f)|_{\text{dB}} = 20 \lg |G(f)| \quad (5.27)$$

gilt. Oftmals wird der Betrag zuvor noch auf den Wert $|G(0)|$ bezogen, um eine einheitliche Darstellung des Amplitudenganges für Systeme mit unterschiedlichen statischen Verstärkungen $|G(0)|$ zu erhalten.

- **Phasengang:** Der Phasengang, also das Argument $\varphi(f) = \angle G(f)$ des Frequenzganges, wird linear über der logarithmischen Frequenzachse aufgetragen (Abb. 5.4).

Viele Messsysteme weisen Verzögerungsverhalten auf [11]. In den Abbildungen 5.4 b) und c) sind zwei häufig vorkommende Sonderfälle solcher Systeme zu sehen. Aus dem Bode-Diagramm lassen sich wichtige Kennwerte über das dynamische Verhalten des jeweiligen Messsystems ablesen.

- **PT₁-Verhalten:** Bei monoton fallendem Amplitudengang werden die Eckfrequenz f_e und die zugehörige Phase $\varphi_e = \varphi(f_e)$ abgelesen. Die Eckfrequenz ist erreicht, wenn die Amplitude auf $\sqrt{2}/2$ des Maximalwertes $|G(0)|$ oder um 3 dB abgefallen ist. Bei Systemen mit PT₁-Verhalten ergibt sich dabei eine Phase von $\varphi_e = -45^\circ$. Für Frequenzen $f \gg f_e$ kon-

vergibt dann der Amplitudengang gegen eine Asymptote der Steigung -20 dB/Dekade^1 .

- **PT₂-Verhalten:** Systeme mit PT₂-Verhalten sind unter gewissen Umständen schwingungsfähig (vgl. Bsp. 5.1). In diesem Fall kann es zu einer Resonanzüberhöhung y_R kommen, die als größter Wert des Amplitudenganges abgelesen werden kann. Bei hinreichend ausgeprägtem Resonanzmaximum beträgt die Phase für die zugehörige Resonanzfrequenz $\varphi_R \approx -90^\circ$. Nicht schwingungsfähige Systeme lassen sich dagegen durch eine Reihenschaltung zweier PT₁-Glieder darstellen und haben ihr Amplitudenmaximum an der Stelle $f = 0$. In beiden Fällen fällt der Amplitudengang für Frequenzen $f \gg f_0$ mit einer Steigung von -40 dB/Dekade .

Die in den Bildern dargestellten Frequenzgänge weisen beide Tiefpassverhalten auf. Es gibt jedoch auch Messgeräte für dynamische Messungen, die Hoch- oder Bandpassverhalten aufweisen. Dort werden die entsprechenden Werte für die untere Grenzfrequenz mit angegeben. Zu beachten ist hierbei, dass eine Normierung auf $|G(0)|$ natürlich keinen Sinn ergibt, wenn die Übertragungsfunktion für $f = 0$ den Wert null annimmt.

Die aus dem Frequenzgang ermittelten Kennwerte geben die Abweichungen vom idealen Zeitverhalten wieder. Als **ideal** wird ein Messsystem verstanden, das für beliebige Frequenzen den folgenden konstanten Frequenzgang besitzt:

$$G_{\text{ideal}}(f) = 1 \quad \Leftrightarrow \quad |G(f)| = 1 \quad \text{und} \quad \angle G(f) = 0 \quad \forall f. \quad (5.28)$$

Bei nichtlinearem Zeitverhalten müssen zusätzlich verschiedene Amplituden der Eingangsgröße untersucht werden. Die Angabe eines Frequenzganges ist dann allerdings nicht mehr allgemein möglich. Anstelle des Frequenzganges kann aber die zusätzlich noch von der Amplitude der Eingangsschwingung $|x_0|$ abhängige **Beschreibungsfunktion** $N(|x_0|, f)$ angegeben werden [10].

Neben der iterativen Anregung eines Messsystems mit harmonischen Schwingungen gibt es noch eine Reihe anderer Verfahren zur Identifikation des dynamischen Systemverhaltens [21, 29]. Ein weiteres Verfahren zur Identifikation von Übertragungsfunktionen ist in Abschn. 6.5 dargestellt.

¹Eine Dekade bezeichnet ein Frequenzintervall, dessen Grenzen sich um den Faktor 10 unterscheiden.

5.3 Verbesserung des dynamischen Systemverhaltens

Anhand des Beispiels 5.1 wurde deutlich, dass die Systemparameter a und b der Übertragungsfunktion $G(s)$ nach (5.7) das Einschwingverhalten eines Messsystems beeinflussen. Daraus lässt sich unmittelbar eine Vorgehensweise zur Verbesserung des dynamischen Systemverhaltens ableiten: die Parameter sind so einzustellen, dass das Zeitverhalten des Messsystems die Forderungen an Schnelligkeit und Überschwingen erfüllt.

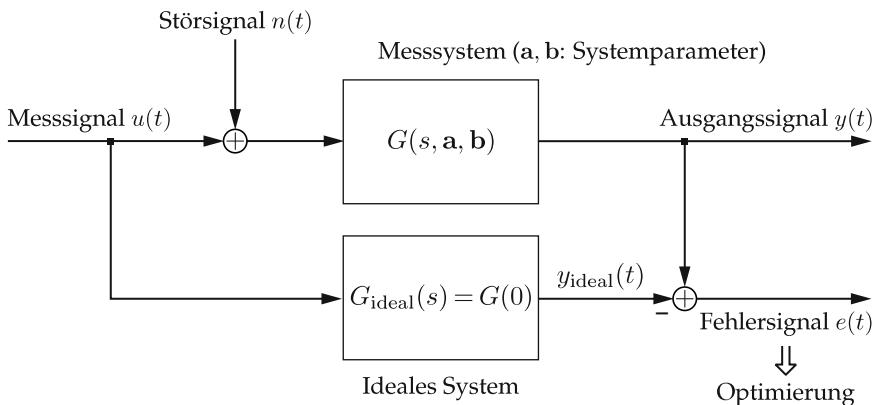


Abbildung 5.5. Verbesserung des dynamischen Verhaltens von Messgeräten.

Die Aufgabe, das Zeitverhalten eines Messsystems zu verbessern, wird in Abb. 5.5 dargestellt. Das Messsignal $u(t)$, überlagert von einem Störsignal $n(t)$, wirkt auf das Messsystem mit der Übertragungsfunktion $G(s, a, b)$. Das Ausgangssignal $y(t)$ wird mit dem idealen, unverfälschten Signal $y_{\text{ideal}}(t)$ verglichen. Das Fehlersignal (engl. *error*)

$$e(t) = y(t) - y_{\text{ideal}}(t) \quad (5.29)$$

wird nun verschiedenen Kriterien unterworfen, um geeignete Parameter a und b für ein gewünschtes Zeitverhalten des Systems zu erhalten. In diesem Kapitel wird der Einfluss der Störgröße zunächst vernachlässigt:

$$n(t) = 0. \quad (5.30)$$

Die Größe $G(0)$ in der Übertragungsfunktion (5.7) ist von besonderer Bedeutung, wie im Folgenden gezeigt wird. Mit dem Endwertsatz der Laplace-Transformation

$$\lim_{t \rightarrow \infty} y(t) = \lim_{s \rightarrow 0} s \cdot Y(s) \quad (5.31)$$

folgt für die Sprungantwort eines Messsystems

$$\lim_{t \rightarrow \infty} y(t) = \lim_{s \rightarrow 0} s \cdot G(s) \frac{1}{s} = G(0). \quad (5.32)$$

Der stationäre Messwert, der nach Beendigung aller Einschwingvorgänge (d. h. für $t \rightarrow \infty$) angenommen wird, entspricht also gerade dem Wert der Übertragungsfunktion $G(s)$ an der Stelle $s = 0$. Dieser Wert $G(0)$ beschreibt die Empfindlichkeit S_i der idealen stationären Kennlinie. Damit ergibt sich folgende Forderung für die ideale Übertragungsfunktion eines Messsystems:

$$G_{\text{ideal}}(s) = G(0). \quad (5.33)$$

Diese Forderung bedeutet, dass das System für alle Frequenzen das gleiche Verhalten zeigen soll wie im stationären Fall. Daraus ergibt sich sofort, dass für Systeme mit $G(0) = 0$, also für Hochpässe und Bandpässe, alle Kriterien versagen, die auf der Bedingung (5.33) basieren. Für solche Systeme sind gesonderte Betrachtungen nötig. In der Messtechnik wird man allerdings zumeist Systeme mit $G(0) \neq 0$ antreffen. Für sie werden Kriterien zur Optimierung des dynamischen Systemverhaltens abgeleitet. Dabei lassen sich zwei prinzipiell verschiedene Vorgehensweisen unterscheiden:

- die **Optimierung der Systemparameter** (Abschn. 5.3.1) und
- die **Änderung der Systemstruktur** (Abschn. 5.3.2).

5.3.1 Optimierung der Systemparameter

Die Kriterien zur Optimierung der Systemparameter gehen alle von der Darstellung im Frequenzbereich gemäß (5.7) aus. Die Aufgabe ist es, durch gezieltes Verändern der Koeffizienten a_i und b_i das dynamische Verhalten des Systems zu optimieren. Die Struktur des Systems wird dabei nicht verändert. Alle im Folgenden aufgelisteten Kriterien werden ausführlich in Abschn. 5.4 behandelt.

- **Momente der Impulsantwort:** Das Kriterium der verschwindenden Momente der Impulsantwort sollte man nur dann anwenden, wenn Geschwindigkeitstreue oder Beschleunigungstreue Systeme verlangt werden. Sind danach noch weitere freie Parameter verfügbar, kann man zusätzlich eines der anderen Kriterien anwenden.
- **Konstanter Amplitudengang:** Dieses Kriterium bringt der Erfahrung nach ein durchaus brauchbares Einschwingverhalten. Meist bleiben die Überschwingweiten der Sprungantwort unter 10 %.

- **Konstanter Realteil:** Bei diesem Kriterium wird ein möglichst konstanter Realteil des Frequenzganges angestrebt. Dahinter steht die allgemeine Aussage, dass Frequenzgänge Überschwingweiten der Sprungantwort von weniger als 20 % bewirken, wenn ihr Realteil im Arbeitsbereich konstant ist und bei höheren Frequenzen abfällt. Diese Überschwingweiten sind zwar meist größer als beim Kriterium vom konstanten Amplitudengang, aber für Messeinrichtungen durchaus noch akzeptabel.
- **ITAE-Kriterium:** Dieses Kriterium liefert, wenn über $n - 1$ Parameter des Nennerpolynoms verfügt werden kann, sehr brauchbare Übergangsfunktionen mit geringen Überschwingweiten. Stehen weniger Parameter für die Dimensionierung zur Verfügung, ist man auf Rechnersimulationen angewiesen. Analytische Lösungen existieren nur für Systeme mit verschwindenden Momenten der Impulsantwort.
- **Quadratisches Fehlerintegral:** Das Integral über den quadratischen Fehler lässt sich analytisch gut in der Frequenzgangdarstellung behandeln. Oft hat die Anwendung länger andauernde, kleine Oszillationen um den stationären Wert zur Folge. Für Systeme höherer Ordnung werden die Beziehungen allerdings unhandlich und erfordern längere numerische Rechnungen.

⌚ 5.3.2 Änderung der Systemstruktur

indexSystemstruktur!Änderung der \sim Anstatt die Koeffizienten a_i und b_i des Systems zu verändern, versucht man hierbei, durch Hinzufügen weiterer Übertragungsglieder das dynamische Verhalten des resultierenden Gesamtsystems zu verbessern. Dies kann durch Reihenschaltung von Kompensationsgliedern oder durch Rückkopplungsschleifen erfolgen. Dabei wird die Struktur des ursprünglichen Systems verändert. Beide Ansätze werden ausführlich in Abschn. 5.5 diskutiert.

- **Kompensation:** Bei der Kompensation werden die Koeffizienten a_i und b_i des Messsystems nicht unmittelbar verändert. Vielmehr wird dort versucht, die das Einschwingverhalten bestimmenden dominanten Pole der Übertragungsfunktion durch entsprechende Nullstellen zu kompensieren. Das Verfahren liefert nur dann befriedigende Ergebnisse, wenn die dominanten Pole bekannt sind, und wenn deren Abhängigkeit von Störgrößen kompensiert werden kann.
- **Gegenkopplung:** In Abschn. 5.5.2 wird gezeigt, dass die proportionale Gegenkopplung von Messsystemen weniger geeignet ist, das Einschwingverhalten zu verbessern. In vielen Fällen wird das Messsystem durch die Gegenkopplung instabil.

5.4 Parameteroptimierung

In diesem Abschnitt wird auf verschiedene Verfahren zur optimalen Einstellung der Systemparameter a und b eingegangen. Die Anwendung der in diesem Abschnitt behandelten Optimierungskriterien soll später jeweils an zwei Beispielen demonstriert werden, die zunächst kurz vorgestellt werden sollen.

Beispiel 5.3 (Federwaage): Als einfaches Messsystem 2. Ordnung wird eine Federwaage nach Abb. 5.6 betrachtet. Mit dem mechanischen Gesetz von Newton, nach welchem die Summe aller Kräfte gleich null ist,

$$\sum_i F_i = 0, \quad (5.34)$$

und den einzelnen Kraftkomponenten,

$m g$: Gewichtskraft,

$m \ddot{x}$: Massenträgheitskraft,

$\beta \dot{x}$: Dämpferkraft und

$c x$: Federkraft,

erhält man die Differentialgleichung des Feder-Masse-Dämpfer-Systems:

$$F = m g = m \ddot{x} + \beta \dot{x} + c x, \quad (5.35)$$

wobei g die Erdbeschleunigung, β die Dämpfungskonstante und c die Federkonstante bezeichnen. Die Übertragungsfunktion der Federwaage resultiert daraus durch Laplace-Transformation:

$$G(s) = \frac{X(s)}{F(s)} = \frac{1}{ms^2 + \beta s + c}. \quad (5.36)$$

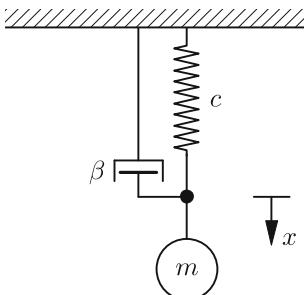


Abbildung 5.6. Federwaage.

Mit den Abkürzungen

$$\omega_0^2 = \frac{c}{m} \quad \text{und} \quad \frac{2\delta}{\omega_0} = \frac{\beta}{c} \quad (5.37)$$

erhält man eine Darstellung, welche die **Dämpfung** δ als Parameter enthält, der Ziel der Optimierung ist:

$$G(s) = \frac{1}{c} \cdot \frac{1}{\frac{s^2}{\omega_0^2} + 2\delta \frac{s}{\omega_0} + 1}. \quad (5.38)$$

Der stationäre Endwert des Systems ist nach dem Endwertsatz der Laplace-Transformation

$$\lim_{s \rightarrow 0} s G(s) \frac{1}{s} = G(0) = \frac{1}{c}. \quad (5.39)$$

Die physikalischen Parameter m , c und β sind immer positiv, so dass nach Satz 5.2 das System immer stabil ist. ■

Beispiel 5.4 (Messsystem 3. Ordnung mit 2 freien Parametern): Am Beispiel eines Messsystems, das als Regelkreis vorliegt, soll eine komplexere Optimierungsaufgabe mit zwei frei wählbaren Parametern vorgestellt werden. Die Struktur des Messsystems ist in Abb. 5.7 dargestellt. Es handelt sich um ein System mit zwei reellen Zeitkonstanten T_1, T_2 und einem vorgeschalteten PI-Regler mit den wählbaren reellen Parametern V und T_k , wobei $T_k \neq T_1, T_2$ gelten soll.

Die Übertragungsfunktion des offenen Kreises lautet

$$G_o(s) = V \frac{1 + T_k s}{s (1 + T_1 s)(1 + T_2 s)}, \quad T_1, T_2, T_k, V > 0. \quad (5.40)$$

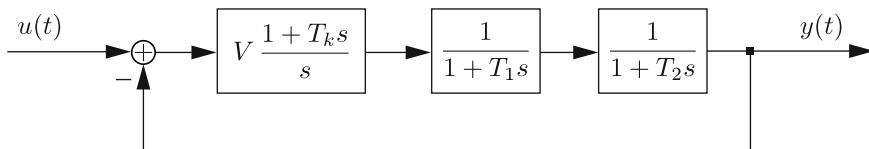


Abbildung 5.7. Strukturbild eines Messsystems mit zwei freien Parametern V und T_k .

Daraus erhält man die Übertragungsfunktion des geschlossenen Kreises zu

$$G(s) = \frac{G_o(s)}{1 + G_o(s)} \quad (5.41)$$

$$= V \frac{1 + T_k s}{T_1 T_2 s^3 + (T_1 + T_2)s^2 + (1 + VT_k)s + V} \quad (5.42)$$

oder, in der Form von (5.7),

$$G(s) = \frac{1 + T_k s}{1 + b_1 s + b_2 s^2 + b_3 s^3} \quad (5.43)$$

mit

$$b_1 = \frac{1 + VT_k}{V}, \quad b_2 = \frac{T_1 + T_2}{V} \quad \text{und} \quad b_3 = \frac{T_1 T_2}{V}. \quad (5.44)$$

Mit dem Endwertsatz der Laplace-Transformation folgt der stationäre Endwert des Systems zu

$$\lim_{s \rightarrow 0} s G(s) \frac{1}{s} = G(0) = 1. \quad (5.45)$$

Die Stabilitätsbedingung für ein System 3. Ordnung lautet nach (5.21)

$$b_1 b_2 > b_3 \Leftrightarrow \frac{1 + VT_k}{V} > \frac{T_1 T_2}{T_1 + T_2} \quad (5.46)$$

$$\Leftrightarrow T_k > T_2 \frac{1}{1 + T_2/T_1} - \frac{1}{V}. \quad (5.47)$$

Wenn $T_k > T_2$ und $V > 0$ gewählt wird, ist die Ungleichung immer erfüllt:

$$T_k > T_2 > T_2 \frac{1}{1 + T_2/T_1} - \frac{1}{V}. \quad (5.48)$$

Für die Beispiele wird mit den Systemparametern $T_1 = 0,2\text{s}$ und $T_2 = 0,4\text{s}$ gearbeitet. Abbildung 5.8 zeigt die Sprungantworten des Beispiel- systems für die Parameter $V = 1$ und $V = 10$ sowie $T_k = 0,2\text{s}$ und $T_k = 4\text{s}$. Die unterschiedlichen Einschwingvorgänge zeigen die Notwendigkeit geeigneter Kriterien für die Bestimmung der Parameter. ■

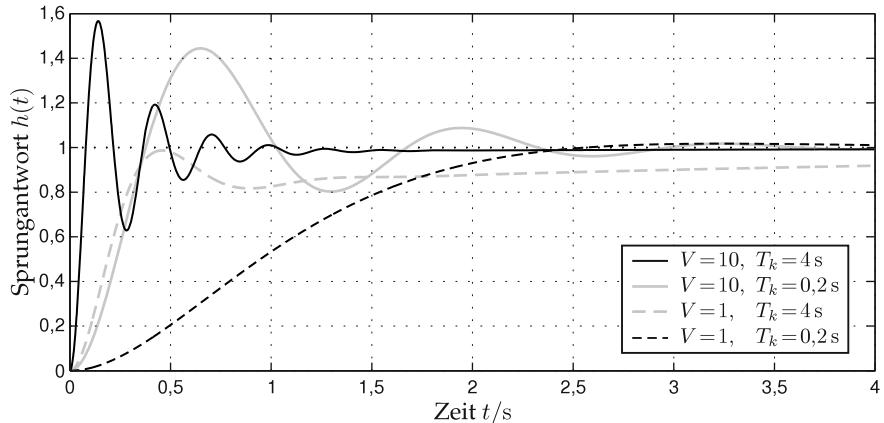


Abbildung 5.8. Sprungantworten des Systems 3. Ordnung bei verschiedenen Parametern.

⦿ 5.4.1 Kriterium „verschwindende Momente der Impulsantwort“

Das Ausgangssignal $y(t)$ des Messsystems erhält man durch Faltung des Eingangssignals $u(t)$ mit der Impulsantwort $g(t)$ des Messsystems:

$$y(t) = g(t) * u(t) = \int_{-\infty}^{\infty} g(\tau) u(t - \tau) d\tau. \quad (5.49)$$

Das Eingangssignal $u(t - \tau)$ wird um t herum in eine Taylor-Reihe

$$u(t - \tau) \approx \sum_{i=0}^n (-1)^i u_i(t) \cdot \tau^i \quad (5.50)$$

mit

$$u_i(t) = \frac{1}{i!} \cdot \left. \frac{d^i u(t - \tau)}{dt^i} \right|_{\tau=0} = \frac{1}{i!} \cdot \frac{d^i u(t)}{dt^i} \quad (5.51)$$

entwickelt. Damit lässt sich das Faltungsintegral (5.49) als Summe schreiben:

$$y(t) \approx \sum_{i=0}^n (-1)^i u_i(t) \cdot \int_{-\infty}^{\infty} \tau^i g(\tau) d\tau. \quad (5.52)$$

Analog zur Wahrscheinlichkeitstheorie (Kap. 4) lassen sich auch von einer Impulsantwort $g(t)$ sogenannte Momente definieren. Das i -te **Moment der**

Impulsantwort berechnet sich wie folgt:

$$g_i = \int_{-\infty}^{\infty} \tau^i g(\tau) d\tau. \quad (5.53)$$

Mit der Bedingung $g(\tau) = 0$ für $\tau < 0$ bei kausalen Systemen können die Momente der Impulsantwort g_i mit der Übertragungsfunktion $G(s)$ verknüpft werden. Dazu wird die Laplace-Transformierte

$$G(s) = \int_0^{\infty} g(\tau) e^{-s\tau} d\tau \quad (5.54)$$

mehrfach nach s abgeleitet:

$$\frac{d^i G(s)}{ds^i} \Big|_{s=0} = (-1)^i \int_0^{\infty} \tau^i g(\tau) d\tau = (-1)^i g_i. \quad (5.55)$$

Die Systemfunktion $G(s)$ ist somit das Analogon zur charakteristischen Funktion (4.58) der Wahrscheinlichkeitstheorie, aus der sich ebenfalls die Momente einer Wahrscheinlichkeitsdichte berechnen lassen. Mit diesen Definitionen lässt sich das Ausgangssignal in (5.52) wie folgt darstellen:

$$y(t) \approx \sum_{i=0}^n (-1)^i u_i(t) \cdot g_i = \sum_{i=0}^n u_i(t) \cdot \frac{d^i G(s)}{ds^i} \Big|_{s=0} \quad (5.56)$$

bzw. mit (5.51)

$$y(t) \approx u(t) G(0) + \dot{u}(t) \frac{dG(s)}{ds} \Big|_{s=0} + \frac{1}{2} \ddot{u}(t) \frac{d^2 G(s)}{ds^2} \Big|_{s=0} + \dots \quad (5.57)$$

Idealerweise sollte die Ausgangsgröße der Messgröße verzögerungsfrei entsprechend

$$y_i(t) = u(t) \cdot G(0) \quad (5.58)$$

folgen, so dass alle Ableitungen

$$\frac{d^i G(s)}{ds^i} \Big|_{s=0} = 0, \quad i \in \{1, 2, \dots, n\}, \quad (5.59)$$

verschwinden sollten. Statt nun die Ableitungen explizit zu berechnen, wird die Differenz ($G(s) - G(0)$) zwischen der Übertragungsfunktion des Mess-

systems $G(s)$ und der idealen Übertragungsfunktion $G(0)$ gebildet:

$$G(s) - G(0) = G(0) \frac{(a_1 - b_1)s + (a_2 - b_2)s^2 + \dots}{1 + b_1 s + b_2 s^2 + \dots}. \quad (5.60)$$

Denkt man sich $(G(s) - G(0))$ in eine Potenzreihe nach s entwickelt, so sind die i -ten Ableitungen in $s = 0$

$$(G(s) - G(0))^{(i)} \Big|_{s=0} = G^{(i)}(s) \Big|_{s=0} = 0 \quad (5.61)$$

gerade null, wenn die entsprechenden Koeffizienten $a_i = b_i$ gleich sind. Für

$$a_1 = b_1 \quad (5.62)$$

verschwindet das erste Moment der Impulsantwort,

$$(G(s) - G(0))' \Big|_{s=0} = G'(s=0) = 0, \quad (5.63)$$

das Messsystem folgt dem Eingangssignal $u(t)$ **geschwindigkeitstreu**. Wird zusätzlich für

$$a_2 = b_2 \quad (5.64)$$

auch das zweite Moment der Impulsantwort gleich null,

$$(G(s) - G(0))'' \Big|_{s=0} = G''(s=0) = 0, \quad (5.65)$$

so folgt das Messsystem dem Eingangssignal $u(t)$ **beschleunigungstreu**. Die Ableitungen $\dot{u}(t)$ und $\ddot{u}(t)$ gehen dann nicht mehr in die Ausgangsgröße $y(t)$ ein. Das Kriterium ergibt schwach gedämpfte Einschwingvorgänge.

Schließlich sei noch auf zwei Nachteile des Kriteriums „verschwindende Momente der Impulsantwort“ hingewiesen.

- Bei Geschwindigkeitstreue ist zwar nach (5.55) das erste Moment der Impulsantwort gleich null,

$$\frac{dG(s)}{ds} \Big|_{s=0} = - \int_0^\infty \tau g(\tau) d\tau = 0, \quad (5.66)$$

bei der Integration können sich aber positive und negative Flächen gegenseitig aufheben (Abb. 5.9).

- Bei vielen Frequenzgängen ist außerdem das Zählerpolynom vom Grad null, so dass das Verfahren überhaupt nicht angewendet werden kann.

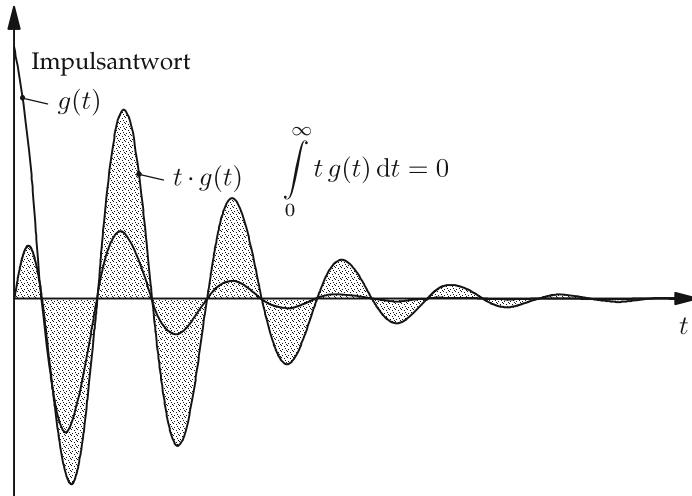


Abbildung 5.9. Impulsantwort und Moment 1. Ordnung.

Beispiel 5.5 (Federwaage): Soll die Federwaage aus Bsp. 5.3 ein geschwindigkeitstreues Messsystem werden, so muss nach (5.62)

$$a_1 = b_1 \quad (5.67)$$

bzw. in (5.38)

$$b_1 = 2\delta = 0 \quad (5.68)$$

gelten. Das System wird nach Satz 5.1 instabil, d. h. ein geschwindigkeitstreues System ist nicht zu erreichen. ■

Beispiel 5.6 (Messsystem 3. Ordnung mit 2 freien Parametern): Ist für das Messsystem 3. Ordnung aus Bsp. 5.4 Geschwindigkeitstreue erwünscht, so muss nach (5.62) $a_1 = b_1$ und demnach

$$T_k = \frac{1 + T_k V}{V} \quad (5.69)$$

gelten, vgl. (5.43). Diese Bedingung ist nur für

$$V \rightarrow \infty \quad (5.70)$$

erfüllbar und lässt sich in der Praxis nur näherungsweise realisieren. Das System bleibt aber für alle V stabil. ■

5.4.2 Kriterium „konstanter Amplitudengang für kleine Frequenzen“

Das Verfahren „konstanter Amplitudengang für kleine Frequenzen“ verlangt lediglich, dass der Betrag des Frequenzganges für kleine Frequenzen konstant bleibt. Über den Phasengang wird hierbei nicht verfügt. Diese Forderung ist erheblich schwächer als die in Abschn. 5.4.1.

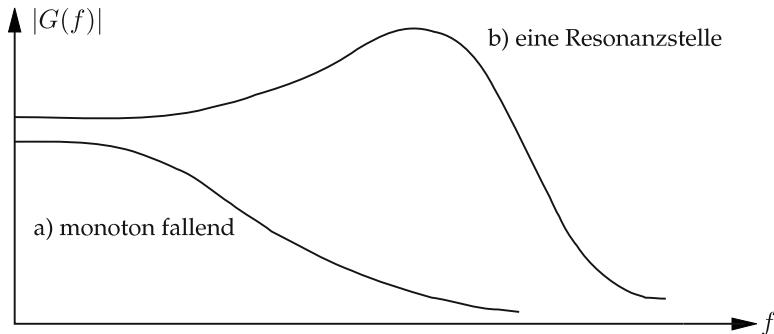


Abbildung 5.10. Konstanter Amplitudengang für kleine Frequenzen.

Abbildung 5.10 zeigt den qualitativen Verlauf zweier Amplitudengänge $|G(f)|$, welche diese Forderung erfüllen, wobei

$$|G(f)|^2 = G(f) G^*(f) \quad (5.71)$$

eine in f gerade Funktion $|G(f)|^2 = F(f^2)$ ist. Um das Kriterium „konstanter Amplitudengang für kleine Frequenzen“ zu erfüllen, setzt man die Ableitungen zu null:

$$\frac{\partial^{(i)} F(f^2)}{\partial (f^2)^{(i)}} = 0, \quad i \in \mathbb{N}. \quad (5.72)$$

Statt die höheren Ableitungen von $F(f^2)$ explizit zu berechnen, stellt man die Funktion $F(f^2)$ am einfachsten wie folgt dar:

$$F(f^2) = G^2(0) \frac{\sum_i A_i (2\pi f)^{2i}}{\sum_k B_k (2\pi f)^{2k}} \quad \text{mit} \quad A_0 = B_0 = 1. \quad (5.73)$$

Durch Subtraktion von $G^2(0)$ erhält man:

$$F(f^2) - G^2(0) = \frac{\sum_i (A_i - B_i) (2\pi f)^{2i}}{\sum_k B_k (2\pi f)^{2k}} G^2(0). \quad (5.74)$$

Wenn für alle $i < j$ die Beziehung $A_i = B_i$ gilt, dann verschwinden alle i -ten Ableitungen von $F(f^2)$ an der Stelle $f^2 = 0$, $i < j$.

Die Koeffizienten A_i und B_i errechnen sich aus dem Frequenzgang:

$$G(f) = G(0) \frac{\sum_{\nu} a_{\nu} (j2\pi f)^{\nu}}{\sum_{\mu} b_{\mu} (j2\pi f)^{\mu}} \quad (5.75)$$

$$= G(0) \frac{\sum_{\nu} a_{2\nu} (-1)^{\nu} (2\pi f)^{2\nu} + j2\pi f \sum_{\nu} a_{2\nu+1} (-1)^{\nu} (2\pi f)^{2\nu}}{\sum_{\mu} b_{2\mu} (-1)^{\mu} (2\pi f)^{2\mu} + j2\pi f \sum_{\mu} b_{2\mu+1} (-1)^{\mu} (2\pi f)^{2\mu}} \quad (5.76)$$

mit $a_0 = b_0 = 1$. Um $F(f^2)$ zu erhalten, muss der Betrag des Zählers und des Nenners berechnet werden. Für die Koeffizienten des Zählers ergibt sich:

$$A_i = (-1)^i \sum_{\nu=0}^i (a_{2\nu} a_{2(i-\nu)} - a_{2\nu+1} a_{2(i-\nu)-1}). \quad (5.77)$$

Im 2. Term der Summendarstellung wird der Index von $a_{2(i-\nu)-1}$ für $i = 0$ negativ. Es gilt $a_{-1} = 0$. Die ersten Koeffizienten A_i lauten ausgeschrieben:

$$A_0 = a_0^2 = 1, \quad (5.78)$$

$$A_1 = -a_0 a_2 - a_2 a_0 + a_1 a_1, \quad (5.79)$$

$$A_2 = a_4 a_0 + a_2 a_2 + a_0 a_4 - a_1 a_3 - a_3 a_1, \quad (5.80)$$

$$A_3 = -a_6 a_0 - a_4 a_2 - a_2 a_4 - a_0 a_6 + a_5 a_1 + a_3 a_3 + a_1 a_5. \quad (5.81)$$

Entsprechend lassen sich die B_i berechnen. Damit lässt sich das Kriterium für einen konstanten Amplitudengang angeben. Ein Frequenzgang der Gestalt

$$G(f) = G(0) \frac{\sum_{\nu} a_{\nu} (j2\pi f)^{\nu}}{\sum_{\mu} b_{\mu} (j2\pi f)^{\mu}} \quad \text{mit} \quad a_0 = b_0 = 1 \quad (5.82)$$

hat für kleine Frequenzen einen konstanten Amplitudengang, wenn nacheinander folgende Bedingungen

$$a_1^2 - 2a_2a_0 = b_1^2 - 2b_2b_0, \quad (5.83)$$

$$a_2^2 + 2a_0a_4 - 2a_1a_3 = b_2^2 + 2b_0b_4 - 2b_1b_3, \quad (5.84)$$

$$a_3^2 + 2a_1a_5 - 2a_0a_6 - 2a_2a_4 = b_3^2 + 2b_1b_5 - 2b_0b_6 - 2b_2b_4, \quad (5.85)$$

oder allgemein

$$\begin{aligned} \sum_{\nu=0}^i (a_{2\nu} a_{2(i-\nu)} - a_{2\nu+1} a_{2(i-\nu)-1}) \\ = \sum_{\nu=0}^i (b_{2\nu} b_{2(i-\nu)} - b_{2\nu+1} b_{2(i-\nu)-1}) \end{aligned} \quad (5.86)$$

erfüllt werden. Sind einige Gleichungen erfüllt, so gibt ein „>“- oder ein „<“-Zeichen bei der ersten nicht erfüllten Gleichung einen Hinweis darüber, ob der Amplitudengang nach dem konstanten Teil für höhere Frequenzen steigt oder fällt.

Das Kriterium geht vom Frequenzgang mit den Koeffizienten a_ν, b_ν der Potenzen von f aus. Diese Koeffizienten sind bei analytischer Behandlung immer bekannt. Das Kriterium ist deshalb einfach anzuwenden. Eine wiederholte Suche nach den Polen ist nicht notwendig, sofern man sicher ist, dass alle Pole auf der linken Halbebene liegen. Ideal wird das Kriterium von Systemen erfüllt, deren Pole und Nullstellen symmetrisch zur j-Achse liegen:

$$G(s) = \prod_i \frac{(s - \alpha_i)(s - \alpha_i^*)}{(s + \alpha_i)(s + \alpha_i^*)}, \quad \Re\{\alpha_i\} > 0 \text{ für Stabilität.} \quad (5.87)$$

Für solche Systeme ist $|G(f)| \equiv 1$. In der Systemtheorie werden solche Systeme als **Allpässe** bezeichnet [42]. Messsysteme sind allerdings von Natur aus keine Allpass-, sondern **Minimalphasensysteme**. Wendet man das Kriterium auf ein Messsystem mit dem Nennergrad n an, so werden nicht alle n Gleichungen erfüllt werden.

Beispiel 5.7 (Butterworth-Filter): Als Beispiel eines Systems mit konstantem Amplitudengang bei kleinen Frequenzen soll das Butterworth-Tiefpassfilter untersucht werden, das durch den folgenden Amplitudengang ge-

geben ist:

$$|G(f)|^2 = \frac{1}{1 + \left(\frac{f}{f_0}\right)^{2n}}. \quad (5.88)$$

Die Pole der Übertragungsfunktion $G(s)$ ergeben sich mit $s = j2\pi f$ aus der Gleichung

$$1 + \left(\frac{s}{j2\pi f_0}\right)^{2n} = 0 \quad (5.89)$$

zu

$$s_{\infty\nu} = 2\pi f_0 \cdot e^{\varphi_\nu} \quad \text{mit} \quad \varphi_\nu = \frac{(2\nu+1)\pi}{2n}, \quad \nu \in \{1, \dots, 2n\}. \quad (5.90)$$

Die Pole liegen auf einem Kreis mit dem Radius $2\pi f_0$ und dem Ursprung als Mittelpunkt. Für ein stabiles System kommen nur die Pole in der linken Halbebene in Frage. Für $n = 2$ erhält man mit $\varphi_1 = \frac{3}{4}\pi$ und $\varphi_2 = \frac{5}{4}\pi$ die Übertragungsfunktion

$$G(s) = \frac{(2\pi f_0)^2}{(s - 2\pi f_0 e^{j\frac{3}{4}\pi})(s - 2\pi f_0 e^{j\frac{5}{4}\pi})} \quad (5.91)$$

und durch inverse Laplace-Transformation die Impulsantwort

$$g(t) = \sqrt{2} \cdot 2\pi f_0 e^{-\sqrt{2}\pi f_0 t} \sin \sqrt{2}\pi f_0 t. \quad (5.92)$$

Die Impulsantwort $g(t)$ und der quadrierte Amplitudengang $|G(f)|^2$ sind in Abb. 5.11 dargestellt. ■

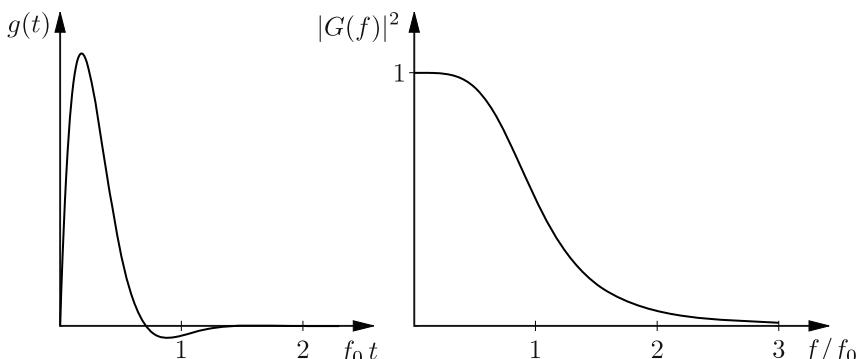


Abbildung 5.11. Impulsantwort $g(t)$ und Amplitudengang $|G(f)|^2$ eines Butterworth-Filters der Ordnung $n = 2$.

Beispiel 5.8 (Federwaage): Nach (5.84) gilt

$$a_1^2 - 2 a_2 a_0 = b_1^2 - 2 b_2 b_0 , \quad (5.93)$$

was bei der Übertragungsfunktion (5.38) der Federwaage aus Bsp. 5.3 zu

$$0 = 4\delta^2 - 2 \quad (5.94)$$

und damit zur Dämpfung

$$\delta = \frac{\sqrt{2}}{2} = 0,707 \quad (5.95)$$

führt. In der Messtechnik heißt diese Dämpfung **Oszillographendämpfung**. Die Sprungantwort $h(t)$ hat eine leichte Überschwingung von 5 % bezogen auf den stationären Endwert (Abb. 5.12). ■

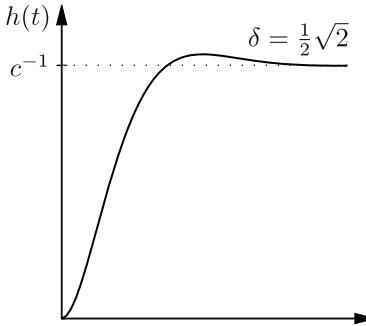


Abbildung 5.12. Sprungantwort $h(t)$ eines Systems 2. Ordnung mit Oszillographendämpfung.

Beispiel 5.9 (Messsystem 3. Ordnung mit 2 freien Parametern): Für das System 3. Ordnung aus Bsp. 5.4 sind zwei Bedingungen für einen konstanten Amplitudengang zu erfüllen. Die erste folgt aus (5.84),

$$a_1^2 - 2 a_2 a_0 = b_1^2 - 2 b_2 b_0 , \quad (5.96)$$

bei Berücksichtigung der Koeffizienten aus (5.43) zu

$$T_k^2 = \frac{(1 + T_k V)^2}{V^2} - 2 \frac{T_1 + T_2}{V} , \quad (5.97)$$

$$T_k V = V(T_1 + T_2) - \frac{1}{2} . \quad (5.98)$$

Die zweite Bedingung aus (5.84),

$$a_2^2 + 2a_0a_4 - 2a_1a_3 = b_2^2 + 2b_0b_4 - 2b_1b_3, \quad (5.99)$$

ergibt

$$0 = \frac{(T_1 + T_2)^2}{V^2} - 2 \frac{T_1 T_2}{V} \frac{1 + T_k V}{V}, \quad (5.100)$$

was mit (5.97) in

$$1 - \frac{1}{2} + V(T_1 + T_2) = \frac{(T_1 + T_2)^2}{2T_1 T_2} \quad (5.101)$$

resultiert. Daraus berechnen sich die gesuchten Parameter zu

$$V = \frac{T_1^2 + T_2^2 + T_1 T_2}{2T_1 T_2(T_1 + T_2)} \quad (5.102)$$

und

$$T_k = \frac{(T_1^2 + T_2^2)(T_1 + T_2)}{T_1^2 + T_2^2 + T_1 T_2}. \quad (5.103)$$

Abbildung 5.13 zeigt die resultierende Sprungantwort und den Ampli-

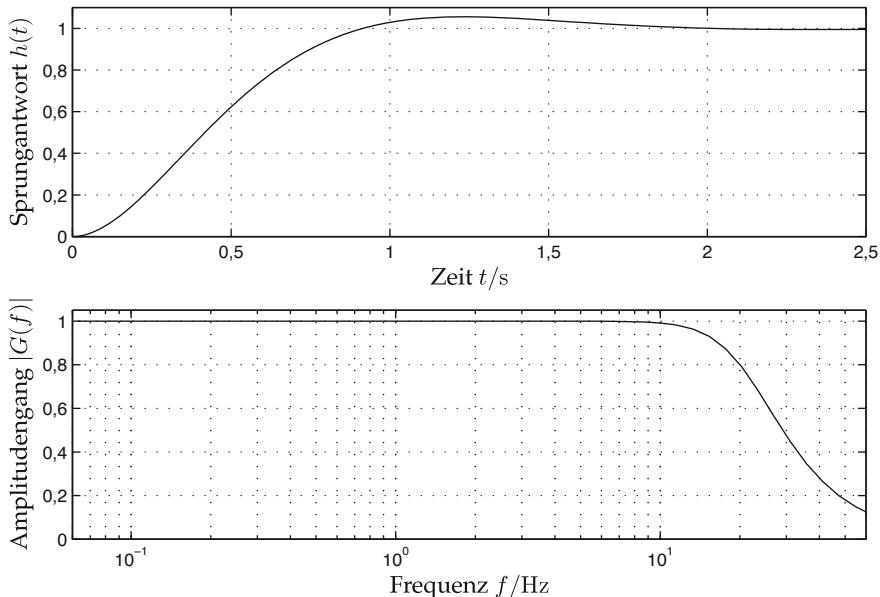


Abbildung 5.13. Sprungantwort und Amplitudengang eines Systems 3. Ordnung nach Anwendung des Kriteriums „konstanter Amplitudengang für kleine Frequenzen“.

tudengang. Die Sprungantwort zeigt wie auch bei der Federwaage (Abb. 5.12) ein nur leichtes Überschwingen von 5 %. Man erkennt sehr gut den konstanten Bereich des Amplitudenganges für kleine Frequenzen. ■

5.4.3 Kriterium „konstanter Realteil des Frequenzganges“

Die Impulsantworten $g(t)$ kausaler Systeme können in einen geraden Teil $g_g(t)$ und einen ungeraden Teil $g_u(t)$ zerlegt werden (vgl. Abb. 5.14):

$$g(t) = g_g(t) + g_u(t). \quad (5.104)$$

Die Fourier-Transformation von $g(t)$ ergibt:

$$G(f) = \int_{-\infty}^{\infty} g(t) e^{-j2\pi ft} dt = \int_{-\infty}^{\infty} g(t) (\cos(2\pi ft) - j \sin(2\pi ft)) dt \quad (5.105)$$

$$\begin{aligned} &= \underbrace{\int_{-\infty}^{\infty} g_u(t) \cos(2\pi ft) dt}_{=0} - j \underbrace{\int_{-\infty}^{\infty} g_g(t) \sin(2\pi ft) dt}_{=0} \\ &\quad + \int_{-\infty}^{\infty} g_g(t) \cos(2\pi ft) dt - j \int_{-\infty}^{\infty} g_u(t) \sin(2\pi ft) dt \end{aligned} \quad (5.106)$$

$$= \Re\{G(f)\} + j \Im\{G(f)\} \quad (5.107)$$

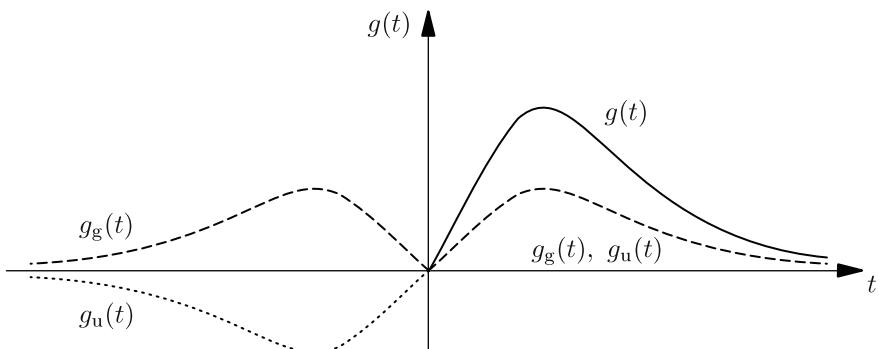


Abbildung 5.14. Zerlegung einer Impulsantwort $g(t)$ in einen geraden Teil $g_g(t)$ und einen ungeraden Teil $g_u(t)$.

mit

$$\Re\{G(f)\} = \int_{-\infty}^{\infty} g_g(t) \cos(2\pi ft) dt, \quad (5.108)$$

$$\Im\{G(f)\} = - \int_{-\infty}^{\infty} g_u(t) \sin(2\pi ft) dt. \quad (5.109)$$

Der Realteil $\Re\{G(f)\}$ lässt sich durch eine Integration von $t = 0$ bis ∞ berechnen, da die Impulsantwort $g_g(t)$ gerade ist:

$$\Re\{G(f)\} = 2 \int_0^{\infty} g_g(t) \cos(2\pi ft) dt. \quad (5.110)$$

Da die Fourier-Transformierte einer geraden Funktion nur Cosinusterme enthält, lässt sich die inverse Fourier-Transformierte $g_g(t)$ wie folgt angeben:

$$g_g(t) = \int_{-\infty}^{\infty} \Re\{G(f)\} \cos(2\pi ft) df = 2 \int_0^{\infty} \Re\{G(f)\} \cos(2\pi ft) df. \quad (5.111)$$

Für $t \geq 0$ gilt:

$$g(t) = 2 g_g(t) = 4 \int_0^{\infty} \Re\{G(f)\} \cos(2\pi ft) df. \quad (5.112)$$

Die Impulsantwort $g(t)$ kann für $t \geq 0$ allein aus dem Realteil $\Re\{G(f)\}$ des Frequenzganges $G(f)$ bestimmt werden. Für ein ideales Einschwingverhalten kann

$$\Re\{G_i(f)\} = G(0) = \text{const.} \quad (5.113)$$

gefordert werden, womit die Impulsantwort

$$g_i(t) = 4 G(0) \cdot \delta(t), \quad t \geq 0, \quad (5.114)$$

und die ideale Ausgangsfunktion des Messsystems

$$y_i(t) = 4 G(0) \cdot \delta(t) * u(t) = 4 G(0) u(t), \quad t \geq 0, \quad (5.115)$$

wären. Die Forderung nach einem konstanten Realteil des Frequenzganges ergibt entsprechende Bedingungen für die Koeffizienten a_ν und b_ν der Über-

tragungsfunktion $G(f)$:

$$G(f) = G(0) \frac{\sum_{\nu=0}^{\infty} a_{\nu} (\mathrm{j}2\pi f)^{\nu}}{\sum_{\mu=0}^{\infty} b_{\mu} (\mathrm{j}2\pi f)^{\mu}} \quad (5.116)$$

$$= G(0) \frac{\sum_{\nu=0}^{\infty} a_{2\nu} (-1)^{\nu} (2\pi f)^{2\nu} + \mathrm{j}2\pi f \sum_{\nu=0}^{\infty} a_{2\nu+1} (-1)^{\nu} (2\pi f)^{2\nu}}{\sum_{\mu=0}^{\infty} b_{2\mu} (-1)^{\mu} (2\pi f)^{2\mu} + \mathrm{j}2\pi f \sum_{\mu=0}^{\infty} b_{2\mu+1} (-1)^{\mu} (2\pi f)^{2\mu}}. \quad (5.117)$$

Die Erweiterung mit dem konjugiert komplexen Nenner $N^*(f)$ ergibt einen realen Nenner $N(f)N^*(f)$. Der Realteil der Übertragungsfunktion ist damit

$$\Re\{G(f)\} = \frac{\Re\{Z N^*\}}{N N^*} \quad \text{mit} \quad (5.118)$$

$$\Re\{Z N^*\} = \sum_{\nu=0}^{\infty} (a_{2\nu} (-1)^{\nu} (2\pi f)^{2\nu}) \sum_{\mu=0}^{\infty} (b_{2\mu} (-1)^{\mu} (2\pi f)^{2\mu}) \quad (5.119)$$

$$\begin{aligned} &+ (2\pi f)^2 \sum_{\nu=0}^{\infty} (a_{2\nu+1} (-1)^{\nu} (2\pi f)^{2\nu}) \sum_{\mu=0}^{\infty} (b_{2\mu+1} (-1)^{\mu} (2\pi f)^{2\mu}) \\ &= \sum_{\nu=0}^{\infty} \sum_{\mu=0}^{\infty} a_{2\nu} b_{2\mu} (-1)^{\nu+\mu} (2\pi f)^{2(\nu+\mu)} \\ &+ (2\pi f)^2 \sum_{\nu=0}^{\infty} \sum_{\mu=0}^{\infty} a_{2\nu+1} b_{2\mu+1} (-1)^{\nu+\mu} (2\pi f)^{2(\nu+\mu)}. \end{aligned} \quad (5.120)$$

Mit der Substitution der Indizes mit

$$k = \nu + \mu \quad \text{und} \quad \ell = \nu + \mu, \quad \nu \leq k, \quad \nu \leq \ell, \quad (5.121)$$

erhält man

$$\begin{aligned} \Re\{Z N^*\} &= \sum_{k=0}^{\infty} \sum_{\nu=0}^k a_{2\nu} b_{2(k-\nu)} (-1)^k (2\pi f)^{2k} \\ &+ (2\pi f)^2 \sum_{\ell=0}^{\infty} \sum_{\nu=0}^{\ell} a_{2\nu+1} b_{2(\ell-\nu)+1} (-1)^{\ell} (2\pi f)^{2\ell}. \end{aligned} \quad (5.122)$$

In den zweiten Summenterm wird der Faktor $(2\pi f)^2$ hereingezogen, wodurch sich der Exponent von $(2\pi f)$ auf $k = \ell + 1$ erhöht. Damit wird

$$\begin{aligned} \Re\{Z N^*\} &= \sum_{k=0}^{\infty} \sum_{\nu=0}^k (a_{2\nu} b_{2(k-\nu)} (-1)^k (2\pi f)^{2k} \\ &\quad - a_{2\nu+1} b_{2(k-\nu)-1} (-1)^k (2\pi f)^{2k}) . \end{aligned} \quad (5.123)$$

Dabei wird der Koeffizient b_{-1} für $\nu = k$ als $b_{-1} = 0$ definiert. Entsprechend erhält man für den Nenner

$$\begin{aligned} N N^* &= \left(\sum_{\nu=0}^{\infty} b_{2\nu} (-1)^\nu (2\pi f)^{2\nu} \right) \left(\sum_{i=0}^{\infty} b_{2i} (-1)^i (2\pi f)^{2i} \right) \\ &\quad + (2\pi f)^2 \left(\sum_{\nu=0}^{\infty} b_{2\nu+1} (-1)^\nu (2\pi f)^{2\nu} \right) \left(\sum_{i=0}^{\infty} b_{2i+1} (-1)^i (2\pi f)^{2i} \right) \\ &= \sum_{k=0}^{\infty} \sum_{\nu=0}^k (b_{2\nu} b_{2(k-\nu)} (-1)^k (2\pi f)^{2k} \\ &\quad - b_{2\nu+1} b_{2(k-\nu)-1} (-1)^k (2\pi f)^{2k}) . \end{aligned} \quad (5.124)$$

Der Realteil der Übertragungsfunktion lässt sich mit

$$A_k = (-1)^k \cdot \sum_{\nu=0}^k (a_{2\nu} b_{2(k-\nu)} - a_{2\nu+1} b_{2(k-\nu)-1}) , \quad (5.126)$$

$$B_k = (-1)^k \cdot \sum_{\nu=0}^k (b_{2\nu} b_{2(k-\nu)} - b_{2\nu+1} b_{2(k-\nu)-1}) \quad (5.127)$$

darstellen als

$$\Re\{G(f)\} = G(0) \frac{\sum_{k=0}^{\infty} A_k (2\pi f)^{2k}}{\sum_{k=0}^{\infty} B_k (2\pi f)^{2k}} . \quad (5.128)$$

Die Bedingung für einen konstanten Realteil $\Re\{G(f)\}$ lautet dann $A_k = B_k$:

$$\begin{aligned} \sum_{\nu=0}^k (a_{2\nu} b_{2(k-\nu)} - a_{2\nu+1} b_{2(k-\nu)-1}) \\ = \sum_{\nu=0}^k (b_{2\nu} b_{2(k-\nu)} - b_{2\nu+1} b_{2(k-\nu)-1}) . \end{aligned} \quad (5.129)$$

Für die ersten Koeffizienten ist dies

$$\begin{aligned} b_0 &= a_0 = 1, \\ -a_0 b_2 + a_1 b_1 - a_2 b_0 &= b_1^2 - 2 b_0 b_2, \\ a_0 b_4 - a_1 b_3 + a_2 b_2 - a_3 b_1 + a_4 b_0 &= b_2^2 + 2 b_0 b_4 - 2 b_1 b_3, \\ -a_0 b_6 + a_1 b_5 - a_2 b_4 + a_3 b_3 - a_4 b_2 + a_5 b_1 - a_6 b_0 &= b_3^2 - 2 b_0 b_6 + 2 b_1 b_5 - 2 b_2 b_4. \end{aligned} \quad (5.130)$$

Die Bedingungen für einen konstanten Realteil $\Re\{G(f)\}$ sind praktisch nur für kleine Frequenzen f erfüllbar. Der Realteil des Frequenzganges fällt hinter dem konstanten Teil ab, wenn in der ersten unerfüllten Gleichung statt des Gleichheitszeichens ein „<“-Zeichen gilt. Umgekehrt erfährt der Realteil eine Erhöhung, wenn in dieser Gleichung ein „>“-Zeichen gilt.

Beispiel 5.10 (Federwaage): Für die Federwaage aus Bsp. 5.3 gilt nach (5.130)

$$-a_0 b_2 + a_1 b_1 - a_2 b_0 = b_1^2 - 2 b_0 b_2, \quad (5.131)$$

woraus man mit (5.38) und $a_0 = 1$ die Dämpfung

$$\delta = \frac{1}{2} \quad (5.132)$$

erhält. Diese Einstellung bringt größere Überschwinger als im Fall des konstanten Amplitudenganges für kleine Frequenzen, wie an der Sprungantwort des Systems Abb. 5.15 zu erkennen ist. ■

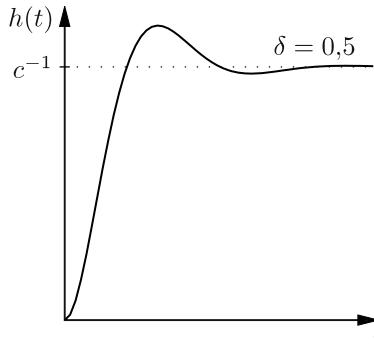


Abbildung 5.15. Sprungantwort $h(t)$ des Systems 2. Ordnung mit der Dämpfung $\delta = 0,5$.

Beispiel 5.11 (Messsystem 3. Ordnung mit 2 freien Parametern): Für das Messsystem 3. Ordnung aus Bsp. 5.4 sind zwei Bedingungen für einen konstanten Realteil des Frequenzganges zu erfüllen. Für die erste Bedin-

gung (5.130) erhält man mit (5.43):

$$-a_0 b_2 + a_1 b_1 - a_2 b_0 = b_1^2 - 2 b_0 b_2, \quad (5.133)$$

$$T_k V = V (T_1 + T_2) - 1. \quad (5.134)$$

Die zweite Bedingung lautet:

$$a_0 b_4 - a_1 b_3 + a_2 b_2 - a_3 b_1 + a_4 b_0 = b_2^2 + 2 b_0 b_4 - 2 b_1 b_3, \quad (5.135)$$

$$-T_k \frac{T_1 T_2}{V} = \frac{(T_1 + T_2)^2}{V^2} - 2 \frac{T_1 T_2}{V} \frac{1 + T_k V}{V}, \quad (5.136)$$

was mit (5.134) die folgende Gleichung ergibt:

$$V (T_1 + T_2) - 1 = \frac{(T_1 + T_2)^2 - 2 T_1 T_2}{T_1 T_2}. \quad (5.137)$$

Daraus berechnen sich die gesuchten Parameter zu

$$V = \frac{T_1^2 + T_2^2 + T_1 T_2}{T_1 T_2 (T_1 + T_2)} \quad \text{und} \quad T_k = \frac{(T_1^2 + T_2^2)(T_1 + T_2)}{T_1^2 + T_2^2 + T_1 T_2}. \quad (5.138)$$

Abbildung 5.16 zeigt die resultierende Sprungantwort und den Realteil des Frequenzganges. Die Sprungantwort zeigt – wie auch die der Feder-

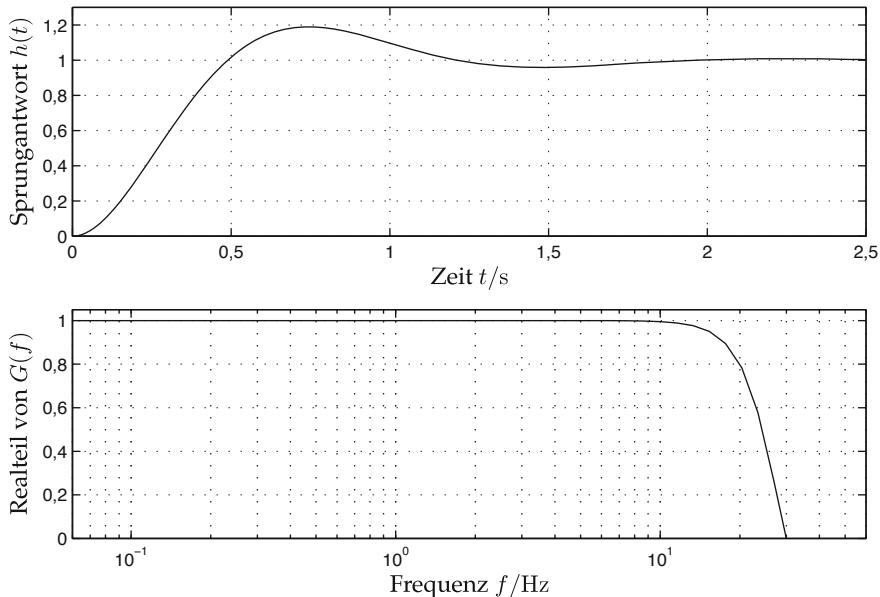


Abbildung 5.16. Sprungantwort und Realteil des Frequenzganges eines Systems 3. Ordnung nach Anwendung des Kriteriums „konstanter Realteil des Frequenzganges“.

waage – größeres Überschwingen. Man erkennt sehr gut den konstanten Bereich des Realteils des Frequenzganges für kleine Frequenzen. ■

⌚ 5.4.4 ITAE-Kriterium

Das dieses Kriterium namensgebende Akronym ITAE steht für *integral of time-multiplied absolute value of error*. Als Testsignal $u(t)$ dient eine Sprungfunktion, weshalb als ideales Ausgangssignal $y_{\text{ideal}}(t)$ ebenfalls eine Sprungfunktion angesetzt wird. Das zu minimierende Gütemaß Q wird als Integral über den Betrag der mit der Zeit t multiplizierten Abweichung $e(t) = y(t) - y_{\text{ideal}}(t)$ des Ausgangssignals vom idealen Ausgangssignal definiert:

$$Q = \int_0^{\infty} t \cdot |e(t)| dt = \int_0^{\infty} t \cdot |y(t) - y_{\text{ideal}}(t)| dt \quad (5.139)$$

$$= \int_0^{\infty} t \cdot |g(t) * \sigma(t) - G(0) \sigma(t)| dt \rightarrow \min . \quad (5.140)$$

Länger andauernde Fehler wirken sich stark aus, da sie mit der Zeit t gewichtet werden. Fehler zu Beginn werden dagegen kaum gewertet. Die Anwendung des Kriteriums führt zu einer leicht schwingenden Einstellung. Die Integration über den Betrag des Fehlersignals ist für Systeme höherer Ordnung nur noch numerisch möglich.

⌚ 5.4.4.1 System 1. Ordnung

Die Übertragungsfunktion eines Messsystems 1. Ordnung sei

$$G_1(s) = G_1(0) \frac{\omega_0}{s + \omega_0}, \quad (5.141)$$

woraus für die Abweichung von der idealen Übertragungsfunktion $G_1(0)$ der Ausdruck

$$\Delta G_1(s) = G_1(s) - G_1(0) = -G_1(0) \frac{s}{s + \omega_0} \quad (5.142)$$

resultiert. Mit dem Eingangssignal

$$\sigma(t) \quad \circ\bullet \quad s^{-1} \quad (5.143)$$

erhält man die Fehlerfunktion $E(s)$ im Laplace-Bereich

$$E(s) = \Delta G_1(s) \cdot \frac{1}{s} = -G_1(0) \cdot \frac{1}{s + \omega_0} \quad (5.144)$$

und durch inverse Laplace-Transformation das zugehörige Zeitsignal

$$e(t) = -G_1(0) \cdot e^{-\omega_0 t}. \quad (5.145)$$

Das zu minimierende ITAE-Gütemaß für das System 1. Ordnung ist damit:

$$Q_1 = \int_0^{\infty} G_1(0) \cdot t \cdot e^{-\omega_0 t} dt = \frac{G_1(0)}{\omega_0^2}. \quad (5.146)$$

Durch eine kleine Zeitkonstante $1/\omega_0$ wird das Gütemaß Q_1 minimiert.

▷ 5.4.4.2 System 2. Ordnung

Bei einem System 2. Ordnung wird der Rechenaufwand bereits wesentlich höher:

$$G_2(s) = G_2(0) \frac{\omega_0^2}{s^2 + b_1 \omega_0 s + \omega_0^2}. \quad (5.147)$$

Die Abweichung von der idealen Übertragungsfunktion $G_2(0)$ ist

$$\Delta G_2(s) = G_2(s) - G_2(0) = -G_2(0) \frac{s^2 + b_1 \omega_0 s}{s^2 + b_1 \omega_0 s + \omega_0^2}, \quad (5.148)$$

woraus sich der folgende Fehler der Sprungantwort im Laplace-Bereich ergibt:

$$E(s) = \Delta G_2(s) \frac{1}{s} = -G_2(0) \frac{s + b_1 \omega_0}{s^2 + b_1 \omega_0 s + \omega_0^2}. \quad (5.149)$$

Mit den Definitionen

$$\alpha = \frac{b_1}{2} \omega_0, \quad \omega = \omega_0 \sqrt{1 - \frac{b_1^2}{4}}, \quad (5.150)$$

$$\delta = \frac{\alpha}{\omega_0} = \frac{b_1}{2} = \cos \varphi, \quad \frac{\omega}{\omega_0} = \sqrt{1 - \frac{b_1^2}{4}} = \sin \varphi, \quad (5.151)$$

$$\frac{\alpha}{\omega} = \cot \varphi \quad (5.152)$$

erhält man durch inverse Laplace-Transformation von (5.149) die Fehlerfunktion im Zeitbereich:

$$e(t) = -\frac{G_2(0)}{\omega} e^{-\alpha t} (\omega \cos(\omega t) + \alpha \sin(\omega t)) \quad (5.153)$$

$$= -\frac{G_2(0)}{\sin \varphi} e^{-\alpha t} \sin(\omega t + \varphi). \quad (5.154)$$

Nun wird die Variablentransformation

$$x = \omega t + \varphi, \quad dx = \omega dt, \quad t = \frac{x - \varphi}{\omega}$$

eingeführt, mit der das ITAE-Gütemaß (5.139) nach der Substitution die folgende Form annimmt:

$$Q_2 = \frac{G_2(0)}{\omega^2 \sin \varphi} e^{\varphi \cot \varphi} \underbrace{\int_{\varphi}^{\infty} e^{-x \cot \varphi} (x - \varphi) |\sin x| dx}_{I(x)}. \quad (5.155)$$

Das Integral $I(x)$ wird schrittweise berechnet, wobei sich an den Integrationszwischengrenzen jeweils das Vorzeichen des Integranden ändert. Auf diese Weise wird die Betragsbildung berücksichtigt:

$$Q_2 = \frac{G_2(0)}{\omega^2 \sin \varphi} e^{\varphi \cot \varphi} \left(-I(\varphi) + 2 \sum_{i=1}^{\infty} (-1)^{i-1} I(i\pi) \right). \quad (5.156)$$

Die Integrale sind

$$I(i\pi) = \frac{\cos(i\pi) \sin^2 \varphi (\varphi - i\pi - \sin(2\varphi))}{e^{i\pi \cot \varphi}}, \quad (5.157)$$

$$I(\varphi) = \frac{\sin^3 \varphi}{e^{\varphi \cot \varphi}} (4 \sin^2 \varphi - 3). \quad (5.158)$$

Damit wird das Gütemaß

$$Q_2 = \frac{G_2(0) \sin^2 \varphi}{\omega_0^2} \left[3 - 4 \sin^2 \varphi + \frac{2}{\sin \varphi} \sum_{i=1}^{\infty} e^{-\cot \varphi (i\pi - \varphi)} (i\pi - \varphi + \sin 2\varphi) \right]; \quad (5.159)$$

das Minimum folgt durch Differentiation für $\varphi = 0,72$ und $b_1 = 2 \cos \varphi = 1,5$. Daraus ergibt sich eine Dämpfung von $\delta = \cos \varphi \approx 1/\sqrt{2}$.

Bei Systemen höherer Ordnung lässt sich das ITAE-Kriterium nicht mehr analytisch berechnen. Für solche Systeme werden daher im Folgenden die durch numerische Optimierung ermittelten Koeffizienten des Nennerpolynoms $N(s)$ angegeben [4, 13]. Dabei wird zwischen Systemen mit konstantem Zähler, geschwindigkeitstreuen Systemen und beschleunigungstreuen Systemen unterschieden.

Zur Anpassung des Nennerpolynoms an die optimalen Polynome wird

$$\omega_0^n = \frac{b_0}{b_n} \quad (5.160)$$

eingeführt und weiter mit

$$q_i = \frac{b_i}{\omega_0^{n-i} b_n}, \quad i = 1, \dots, n, \quad (5.161)$$

gerechnet. Dann ist das Nennerpolynom $N(s) = \sum q_i (s/\omega_0)^i$ für die Anwendung der folgenden Gleichungen geeignet.

- Für Systeme der Ordnung n

$$G(s) = G(0) \frac{1}{\sum_{\nu=0}^n b_\nu s^\nu} = G(0) \frac{1}{N(s)} \quad (5.162)$$

mit konstantem Zählerpolynom wird das Nennerpolynom $N(s)$:

$$\begin{aligned} & s + \omega_0 \\ & s^2 + 1,505 \omega_0 s + \omega_0^2 \\ & s^3 + 1,784 \omega_0 s^2 + 2,172 \omega_0^2 s + \omega_0^3 \\ & s^4 + 1,953 \omega_0 s^3 + 3,347 \omega_0^2 s^2 + 2,648 \omega_0^3 s + \omega_0^4 \\ & s^5 + 2,068 \omega_0 s^4 + 4,498 \omega_0^2 s^3 + 4,674 \omega_0^3 s^2 + 3,257 \omega_0^4 s + \omega_0^5 \\ & s^6 + 2,152 \omega_0 s^5 + 5,629 \omega_0^2 s^4 + 6,934 \omega_0^3 s^3 + 6,793 \omega_0^4 s^2 + 3,740 \omega_0^5 s + \omega_0^6, \end{aligned} \quad (5.163)$$

- für geschwindigkeitstreue Systeme $G(s) = G(0) \frac{b_0 + b_1 s}{\sum_{\nu=0}^n b_\nu s^\nu}$ wird $N(s)$:

$$\begin{aligned} & s^2 + 3,2 \omega_0 s + \omega_0^2 \\ & s^3 + 1,75 \omega_0 s^2 + 3,25 \omega_0^2 s + \omega_0^3 \\ & s^4 + 2,41 \omega_0 s^3 + 4,93 \omega_0^2 s^2 + 5,14 \omega_0^3 s + \omega_0^4 \\ & s^5 + 2,19 \omega_0 s^4 + 6,50 \omega_0^2 s^3 + 6,30 \omega_0^3 s^2 + 5,24 \omega_0^4 s + \omega_0^5 \\ & s^6 + 6,12 \omega_0 s^5 + 6,71 \omega_0^2 s^4 + 8,58 \omega_0^3 s^3 + 7,07 \omega_0^4 s^2 + 6,76 \omega_0^5 s + \omega_0^6, \end{aligned} \quad (5.164)$$

- für geschwindigkeits- und beschleunigungstreue Systeme

$$G(s) = G(0) \frac{b_0 + b_1 s + b_2 s^2}{\sum_{\nu=0}^n b_\nu s^\nu} \quad (5.165)$$

erhält man für $N(s)$:

$$\begin{aligned}
 & s^3 + 2,97 \omega_0 s^2 + 4,94 \omega_0^2 s + \omega_0^3 \\
 & s^4 + 3,71 \omega_0 s^3 + 7,88 \omega_0^2 s^2 + 5,93 \omega_0^3 s + \omega_0^4 \\
 & s^5 + 3,81 \omega_0 s^4 + 9,94 \omega_0^2 s^3 + 13,44 \omega_0^3 s^2 + 7,36 \omega_0^4 s + \omega_0^5 \\
 & s^6 + 3,93 \omega_0 s^5 + 11,68 \omega_0^2 s^4 + 18,56 \omega_0^3 s^3 + 19,3 \omega_0^4 s^2 + 8,06 \omega_0^5 s + \omega_0^6
 \end{aligned} \quad (5.166)$$

Abbildung 5.17 zeigt die Sprungantworten mehrerer ITAE-optimaler Systeme mit konstantem Zählerpolynom nach (5.163). Der stationäre Endwert h_∞ wird nach kurzer Zeit und bei nicht zu starken Überschwingungen erreicht.

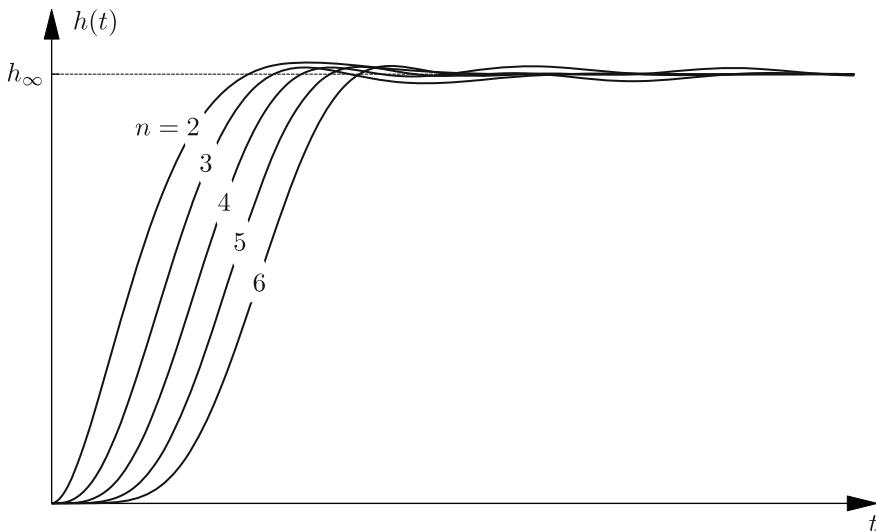


Abbildung 5.17. ITAE-optimale Sprungantworten von Systemen der Ordnung n .

Beispiel 5.12 (Federwaage): Nach (5.163) ist für ein System 2. Ordnung wie die Federwaage aus Bsp. 5.3 der Koeffizient $b_1 = 1,505$, d. h. die Dämpfung ist

$$\delta = 0,7525, \quad (5.167)$$

was nahezu der Oszillographendämpfung entspricht. ■

Beispiel 5.13 (Messsystem 3. Ordnung mit 2 freien Parametern): Das Beispiel-
system 3. Ordnung mit der Übertragungsfunktion (5.43) aus Bsp. 5.4 ist
nach Bsp. 5.6 nicht geschwindigkeitstreu. Daher lassen sich die Nenner-
polynome von (5.164) nicht anwenden. Da die optimalen Nennerpoly-
nome nur für die angegebenen Klassen von Übertragungsfunktionen be-
kannt sind, muss in allen anderen Fällen eine numerische Optimierung
vorgenommen werden. ■

5.4.5 Kriterium „quadratisches Fehlerintegral“

Als Gütemaß für die Systemparameterwahl wird beim Kriterium „quadra-
tisches Fehlerintegral“ das Integral über den quadrierten Fehler $e^2(t)$ bei
sprungförmiger Erregung herangezogen:

$$Q = \int_0^\infty e^2(t) dt = \int_0^\infty (y(t) - y_{\text{ideal}}(t))^2 dt \quad (5.168)$$

$$= \int_0^\infty (g(t) * \sigma(t) - G(0) \cdot \sigma(t))^2 dt \quad \rightarrow \quad \min. \quad (5.169)$$

Das einfache Fehlerintegral $\int_0^\infty e(t) dt$ wäre nicht sinnvoll, da sich im Inte-
gral positive und negative Fehler gegenseitig aufheben würden. Aus (5.168)
lässt sich ablesen, dass große Fehler sehr stark bewertet werden, kleine Feh-
ler dagegen nicht so stark zu Buche schlagen.

Zur analytischen Berechnung von Q wird vorausgesetzt, dass $e(t)$ abso-
lut integrierbar ist und damit dessen Fourier-Transformierte existiert. Für
die Praxis bedeutet dies, dass $\lim_{t \rightarrow \infty} e(t) = 0$ gelten muss, was wegen der
Beschränktheit der Fehlersignale $|e(t)| < M$ eine realistische Annahme ist.
Damit lässt sich Q aus (5.168) mit Hilfe des Parseval'schen Satzes [42] mit
 $E^*(s) = E(-s)$ für $s = j2\pi f$ berechnen:

$$Q = \int_0^\infty e^2(t) dt = \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} E(s) E(-s) ds. \quad (5.170)$$

Das Berechnungsverfahren nutzt die Symmetrieeigenschaft des Integran-
den $E(s) E(-s)$ aus. Die Laplace-Transformierte $E(s)$ habe die Form

$$E(s) = \frac{\sum_{\nu=0}^{n-1} a_\nu s^\nu}{\sum_{\mu=0}^n b_\mu s^\mu} = \frac{A(s)}{B(s)}. \quad (5.171)$$

Der Integrand in (5.170) ist dann

$$E(s) E(-s) = \frac{A(s) A(-s)}{B(s) B(-s)}. \quad (5.172)$$

Er lässt sich in einen kausalen und einen antikausalen Teil aufspalten:

$$\frac{A(s) A(-s)}{B(s) B(-s)} = \underbrace{\frac{C(s)}{B(s)}}_{\text{kausal}} + \underbrace{\frac{D(s)}{B(-s)}}_{\text{antikausal}}. \quad (5.173)$$

Multipliziert man (5.173) mit $B(s) B(-s)$ und bedenkt ferner, dass $A(s) A(-s)$ eine gerade Funktion ist, so folgt aus $A(s) A(-s) = A(-s) A(s)$:

$$C(s) B(-s) + D(s) B(s) = C(-s) B(s) + D(-s) B(-s) \quad (5.174)$$

bzw.

$$D(s) = C(-s). \quad (5.175)$$

Für ein Messsystem n -ter Ordnung lautet das Gütekriterium dann

$$Q_n = \frac{1}{2\pi j} \underbrace{\int_{-\infty}^{j\infty} \frac{C(s)}{B(s)} ds}_{\text{kausal}} + \frac{1}{2\pi j} \underbrace{\int_{-\infty}^{j\infty} \frac{C(-s)}{B(-s)} ds}_{\text{antikausal}} = \frac{2}{2\pi j} \int_{-\infty}^{j\infty} \frac{C(s)}{B(s)} ds. \quad (5.176)$$

Die Zusammenfassung zum Zweifachen des kausalen Teils ist möglich, da die Pole des Integranden symmetrisch zur imaginären Achse liegen. Dies lässt sich leicht zeigen, indem man im zweiten Integral $-s = v$, $-ds = dv$ substituiert. Schließlich kann Q_n mittels der inversen Laplace-Transformation ausgedrückt werden. Mit der Umkehrformel der Laplace-Transformation

$$\mathcal{L}^{-1}\{F(s)\} = \frac{1}{2\pi j} \int_{-\infty}^{j\infty} F(s) e^{st} ds = \begin{cases} f(t) & \text{für } t > 0 \\ 0 & \text{für } t < 0 \\ \frac{1}{2} f(+0) & \text{für } t = 0 \end{cases} \quad (5.177)$$

für $t = 0$ und der Definition $F(s) = C(s)/B(s)$ lässt sich Q_n schreiben als

$$Q_n = 2 \cdot \frac{1}{2} f(+0) = f(+0). \quad (5.178)$$

Mit dem Anfangswertsatz der Laplace-Transformation erhält man

$$f(+0) = \lim_{t \rightarrow +0} f(t) = \lim_{s \rightarrow \infty} s \cdot \frac{C(s)}{B(s)}. \quad (5.179)$$

$B(s)$ ist ein Polynom n -ter Ordnung, $C(s)$ hat die Ordnung $(n-1)$. Zur Bestimmung von Q_n genügt daher allein die Bestimmung zweier Koeffizienten

ten:

$$Q_n = \int_0^\infty e^2(t) dt = \frac{c_{n-1}}{b_n}. \quad (5.180)$$

Beispiel 5.14 (System 1. Ordnung): Für eine rationale Übertragungsfunktion

1. Ordnung sei das Fehlersignal im Laplace-Bereich

$$E(s) = \frac{a_0}{b_0 + b_1 s} = \frac{A(s)}{B(s)} \quad (5.181)$$

und somit

$$E(s) E(-s) = \frac{a_0^2}{(b_0 + b_1 s)(b_0 - b_1 s)} \stackrel{!}{=} \frac{C(s)}{B(s)} + \frac{C(-s)}{B(-s)}. \quad (5.182)$$

$C(s)$ erhält man z. B. durch eine Partialbruchzerlegung, wobei die Pole

$$s_1 = -\frac{b_0}{b_1} \quad \text{und} \quad s_2 = +\frac{b_0}{b_1} \quad (5.183)$$

sind. Damit wird

$$C(s) = \text{Res}\left\{E(s) E(-s), s_1 = -\frac{b_0}{b_1}\right\} = \frac{a_0^2}{b_0 + b_1 \frac{b_0}{b_1}} = \frac{a_0^2}{2b_0}, \quad (5.184)$$

$$C(-s) = \text{Res}\left\{E(s) E(-s), s_2 = +\frac{b_0}{b_1}\right\} = \frac{a_0^2}{b_0 + b_1 \frac{b_0}{b_1}} = \frac{a_0^2}{2b_0}, \quad (5.185)$$

womit das zu minimierende Gütemaß zu

$$Q_1 = \lim_{s \rightarrow \infty} \left(s \frac{C(s)}{B(s)} \right) = \lim_{s \rightarrow \infty} \left(\frac{a_0^2 s}{2b_0 (b_0 + b_1 s)} \right) = \frac{a_0^2}{2b_0 b_1} \quad (5.186)$$

resultiert. ■

Im Folgenden soll noch eine allgemeine Berechnungsvorschrift für den Parameter c_{n-1} angegeben werden. Dazu muss (5.173) unter Verwendung von (5.175) gelöst werden:

$$C(s) B(-s) + C(-s) B(s) = A(s) A(-s) \quad (5.187)$$

bzw.

$$\begin{aligned} \sum_{i=0}^{n-1} c_i s^i \sum_{k=0}^n b_k (-s)^k + \sum_{i=0}^{n-1} c_i (-s)^i \sum_{k=0}^n b_k s^k \\ = \sum_{i=0}^{n-1} a_i s^i \sum_{k=0}^n a_k (-s)^k. \quad (5.188) \end{aligned}$$

Durch Auflösen gelangt man nach einem Koeffizientenvergleich zu folgendem linearen Gleichungssystem:

$$\begin{bmatrix} b_0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ b_2 & b_1 & b_0 & 0 & 0 & \cdots & 0 \\ b_4 & b_3 & b_2 & b_1 & b_0 & \cdots & 0 \\ \vdots & & \ddots & & & \vdots & \\ 0 & 0 & \cdots & b_n & b_{n-1} \end{bmatrix} \cdot \begin{bmatrix} c_0 \\ -c_1 \\ c_2 \\ \vdots \\ (-1)^{n-1} c_{n-1} \end{bmatrix} = \mathbf{B}_n \cdot \mathbf{c}_n$$

$$= \frac{1}{2} \begin{bmatrix} a_0^2 \\ 2a_0 a_2 - a_1^2 \\ -2a_1 a_3 + a_2^2 \\ \vdots \\ (-1)^n 2a_{n-2} a_n + (-1)^{n-1} a_{n-1}^2 \end{bmatrix} = \mathbf{a}_n. \quad (5.189)$$

Die Berechnung von c_{n-1} erfolgt durch Anwendung der Cramer'schen Regel. Das Gleichungssystem sei nach (5.189) $\mathbf{B}_n \cdot \mathbf{c}_n = \mathbf{a}_n$. Die Matrix \mathbf{B}_n^* ergibt sich aus der Matrix \mathbf{B}_n , indem die letzte Spalte durch den Vektor \mathbf{c}_n ersetzt wird. Der gesuchte Koeffizient ist dann der Quotient der Determinanten

$$(-1)^{n-1} c_{n-1} = \frac{1}{2} \frac{|\mathbf{B}_n^*|}{|\mathbf{B}_n|}. \quad (5.190)$$

Damit erhält man die allgemeine Formel für das quadratische Fehlerintegral:

$$Q_n = \int_0^\infty e^2(t) dt = \frac{c_{n-1}}{b_n} = (-1)^{n-1} \frac{|\mathbf{B}_n^*|}{2b_n |\mathbf{B}_n|}. \quad (5.191)$$

Im Folgenden sind einige Fehlerintegrale allgemein berechnet:

$$Q_2 = \frac{a_1^2 b_0 + a_0^2 b_2}{2 b_0 b_1 b_2},$$

$$Q_3 = \frac{a_2^2 b_0 b_1 + (a_1^2 - 2 a_0 a_2) b_0 b_3 + a_0^2 b_2 b_3}{2 b_0 b_3 (b_1 b_2 - b_0 b_3)}, \quad (5.192)$$

$$Q_4 = \frac{a_3^2 (b_0 b_1 b_2 - b_0^2 b_3) + (a_2^2 - 2 a_1 a_3) b_0 b_1 b_4 + (a_1^2 - 2 a_0 a_2) b_0 b_3 b_4 + a_0^2 (b_2 b_3 b_4 - b_1 b_4^2)}{2 b_0 b_4 (-b_0 b_3^2 - b_1^2 b_4 + b_1 b_2 b_3)}.$$

Die Beziehungen für Systeme höherer Ordnung n findet man in [32]. Mit wachsender Ordnung n werden die Gleichungen rasch unhandlich.

Zur Minimierung des Kriteriums wird die Gütefunktion Q_n nach sämtlichen Parametern p_i abgeleitet, die den Einschwingvorgang beeinflussen:

$$\frac{\partial Q_n}{\partial p_i} \stackrel{!}{=} 0, \quad i = 1, \dots, m. \quad (5.193)$$

Auf diese Weise erhält man die Parameter der Übertragungsfunktion, die zu einem optimalen Einschwingverhalten des Messsystems führen.

Beispiel 5.15 (Federwaage): Für die Federwaage aus Bsp. 5.3 wird die Laplace-Transformierte $E(s)$ des Fehlers der Sprungantwort

$$E(s) = (G(s) - G(0)) \frac{1}{s} = \frac{-(2\delta + \frac{s}{\omega_0})}{c\omega_0 \left(\left(\frac{s}{\omega_0}\right)^2 + 2\delta\frac{s}{\omega_0} + 1 \right)} \quad (5.194)$$

betrachtet. Nach (5.192) lautet das Gütemaß für ein System 2. Ordnung

$$Q_2 = \frac{a_1^2 b_0 + a_0^2 b_2}{2 b_0 b_1 b_2} = \left(\frac{1}{4\delta} + \delta \right) \cdot \frac{1}{\omega_0}, \quad (5.195)$$

woraus man durch Differentiation

$$\frac{dQ_2}{d\delta} \stackrel{!}{=} 0 \quad \Rightarrow \quad \delta = \frac{1}{2} \quad (5.196)$$

den gesuchten Parameter δ erhält. ■

Beispiel 5.16 (Messsystem 3. Ordnung mit 2 freien Parametern): Aus (5.43) bildet man die Laplace-Transformierte $E(s)$ des Fehlers der Sprungantwort:

$$E(s) = \frac{1}{s} (G(s) - G(0)) \quad (5.197)$$

$$= \frac{-1 - (T_1 + T_2)s - T_1 T_2 s^2}{V + (1 + T_k V)s + (T_1 + T_2)s^2 + T_1 T_2 s^3}. \quad (5.198)$$

Nach (5.192) erhält man für ein System 3. Ordnung das Gütemaß

$$Q_3 = \frac{a_2^2 b_0 b_1 + (a_1^2 - 2 a_0 a_2) b_0 b_3 + a_0^2 b_2 b_3}{2 b_0 b_3 (b_1 b_2 - b_0 b_3)}. \quad (5.199)$$

Durch Einsetzen der Koeffizienten wird daraus:

$$Q_3 = \frac{(T_1 T_2)^2 V (1 + T_k V) + ((T_1 + T_2)^2 - 2 T_1 T_2) V T_1 T_2 + (T_1 + T_2) T_1 T_2}{V T_1 T_2 ((1 + T_k V) (T_1 + T_2) - V T_1 T_2)}.$$

Durch Differentiation nach den beiden Parametern

$$\frac{\partial Q_3}{\partial V} \stackrel{!}{=} 0 \quad \text{und} \quad \frac{\partial Q_3}{\partial T_k} \stackrel{!}{=} 0 \quad (5.200)$$

erhält man zwei optimale Parametersätze:

$$T_{k1} = 0,25, \quad V_1 = 10,4 \quad \text{sowie} \quad T_{k2} = 0,4, \quad V_2 = 4,06. \quad (5.201)$$

Beide Lösungen entsprechen einer Polkompensation. Die bessere Wahl für ein schnelles Einschwingverhalten ist die Kompensation des Pols mit der größeren Zeitkonstante. Das Überschwingen ist ähnlich wie beim Kriterium „konstanter Realteil des Frequenzganges“, wie in Abb. 5.18 erkannt werden kann. ■

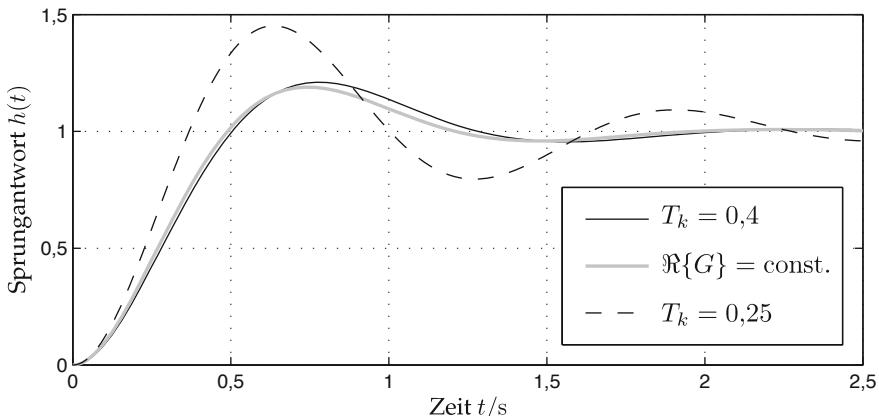


Abbildung 5.18. Sprungantwort bei Verwendung der Kriterien „quadratisches Fehlerintegral“ und „konstanter Realteil des Frequenzganges“.

⑧ Gewichtung des quadratischen Fehlers mit der Zeit

Anders als z. B. das ITAE-Kriterium bewertet das Kriterium „quadratisches Fehlerintegral“ alle Fehler gleich stark, gleichgültig zu welchem Zeitpunkt sie auftreten. In den meisten praktischen Fällen wird man aber Fehler für

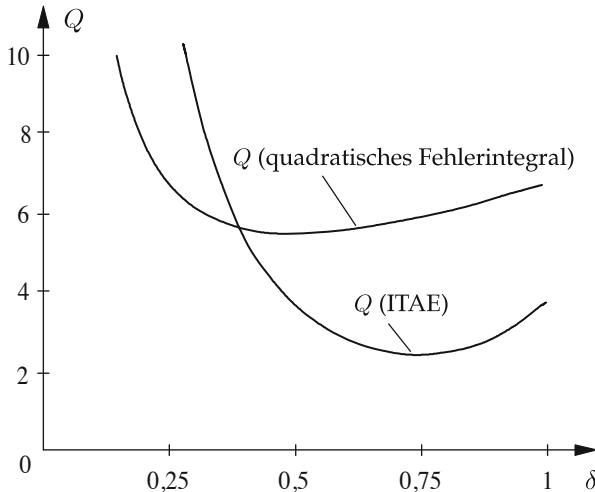


Abbildung 5.19. Gütemaß Q in Abhängigkeit von der Dämpfung δ bei einem System 2. Ordnung.

kleine Zeiten t (d. h. während des Einschwingvorganges) eher tolerieren als Fehler im eingeschwungenen Zustand. Durch die zeitliche Gewichtung des Fehlerbetrags führt das ITAE-Kriterium zu einem besser gedämpftem Einschwingverhalten als das Kriterium „quadratisches Fehlerintegral“, wie man in Abb. 5.19 exemplarisch am Verlauf des Gütemaßes Q in Abhängigkeit von der Dämpfung δ für ein System 2. Ordnung erkennen kann.

Bei einfacheren Systemen liegt es daher nahe, das Kriterium „quadratisches Fehlerintegral“ um eine Gewichtung des quadratischen Fehlers mit Potenzen der Zeit t zu ergänzen:

$$Q_n = \int_0^\infty e^2(t) t^n dt. \quad (5.202)$$

Derartige Kriterien lassen sich zumindest im Prinzip analytisch behandeln. Hierbei gehen Fehler bei großen Zeiten t stärker in das zu minimierende Gütemaß Q_n ein. Nach der Parseval'schen Beziehung ist

$$Q_n = \frac{1}{2\pi j} \int_{-\infty}^{\infty} (-1)^n \frac{d^n E(s)}{ds^n} \cdot E(-s) ds. \quad (5.203)$$

Mit Hilfe einer Partialbruchzerlegung des Integranden können die Pole in den kausalen und den antikausalen Teil aufgeteilt werden. Der kausale Teil wird über die linke Halbebene integriert, der antikausale über die rechte

Halbebene. Der Integrand ist aber nicht mehr unbedingt eine symmetrische Funktion in s . Bei umfangreicheren Systemen können die Pole jedoch nicht mehr analytisch bestimmt werden, wodurch das Verfahren unpraktisch wird.

5.5

5.5 Strukturänderung zur Optimierung des Zeitverhaltens

Alle bisher geschilderten Verfahren hatten eine Optimierung der Parameter der Übertragungsfunktion zum Ziel. Anstatt die Koeffizienten a_i und b_i des Systems zu verändern, kann das dynamische Systemverhalten durch Hinzufügen weiterer Übertragungsglieder verbessert werden. Dies kann durch Reihenschaltung von Kompensationsgliedern (Abschn. 5.5.1) oder durch Rückkopplungsschleifen (Abschn. 5.5.2) erfolgen.

5.5.1 Kompensation des Zeitverhaltens

In diesem Abschnitt wird ein Verfahren diskutiert, welches die *Struktur* des Frequenzganges durch Kompensation ändert, um ein besseres Zeitverhalten zu erreichen. Dabei wird ein schnelles Einschwingen des Messsystems angestrebt.

Das Zeitverhalten eines linearen Systems wird bekanntlich durch die Lage der Pole der Übertragungsfunktion bestimmt. Die Übertragungsfunktion eines Systems ohne Totzeit ist allgemein

$$G(s) = \frac{Z(s)}{N(s)} = c \cdot \frac{\prod_{j=0}^m (s - s_{0j})}{\prod_{i=0}^n (s - s_{\infty i})}, \quad m \leq n. \quad (5.204)$$

Bei der Lösung der Umkehrformel der Laplace-Transformation mit Hilfe des Residuensatzes wird das deutlich: einfache Pole $s_{\infty i}$ führen im Zeitbereich zu Funktionen der Form $e^{s_{\infty i} t}$. Lediglich die Faktoren vor den Exponentialfunktionen werden durch den Zähler der Übertragungsfunktion gegeben. Mit Hilfe der Partialbruchzerlegung erhält man bei einfachen Polen die Übertragungsfunktion in Summendarstellung

$$G(s) = \sum_{i=0}^n \frac{A_i}{s - s_{\infty i}}, \quad (5.205)$$

die man einfach in den Zeitbereich zurücktransformieren kann:

$$g(t) = \sum_{i=0}^n A_i e^{s_{\infty i} t} \cdot \sigma(t). \quad (5.206)$$

Sind die Polstellen $s_{\infty i}$ komplex, so treten auch konjugiert komplexe Pole bei $s_{\infty i}^*$ auf, da $g(t)$ eine reelle Funktion ist. Bei stabilen Systemen liegen alle Pole $s_{\infty i}$ und $s_{\infty i}^*$ in der linken s -Halbebene [42].

Die beiden Pole, die der imaginären Achse am nächsten liegen, welche also den betragsmäßig kleinsten Realteil haben, werden im Folgenden mit $s_{\infty 0} = -\delta_0 + j2\pi f_0$ und $s_{\infty 0}^* = -\delta_0 - j2\pi f_0$ bezeichnet. Diese sogenannten **dominanten Pole** bestimmen weitgehend das Zeitverhalten der Impulsantwort für große Zeiten:

$$\lim_{t \rightarrow \infty} g(t) \approx e^{-\delta_0 t} (A_0 e^{j2\pi f_0 t} + A_0^* e^{-j2\pi f_0 t}). \quad (5.207)$$

Es liegt nun nahe zu versuchen, diese beiden dominanten Pole durch ein nachgeschaltetes Kompensationsglied $G_K(s)$ zu eliminieren (Abb. 5.20). Dafür wird der Übertragungsfunktion

$$G(s) = \frac{Z(s)}{N(s)} = \frac{Z(s)}{M(s) \cdot (s^2 - (s_{\infty 0} + s_{\infty 0}^*) s + s_{\infty 0} s_{\infty 0}^*)} \quad (5.208)$$

in Reihe ein Kompensationsglied mit der Übertragungsfunktion

$$G_K(s) = \frac{(s^2 - (s_{0K} + s_{0K}^*) s + s_{0K} s_{0K}^*)}{(s^2 - (s_{\infty K} + s_{\infty K}^*) s + s_{\infty K} s_{\infty K}^*)} \quad (5.209)$$

hinzugefügt. Zur Kompensation der dominanten Pole, damit ein insgesamt schnelleres Einschwingverhalten des Gesamtsystems resultiert, muss gelten:

$$s_{0K} = s_{\infty 0}, \quad (5.210)$$

$$\Re\{s_{\infty K}\} < \Re\{s_{\infty 0}\}. \quad (5.211)$$

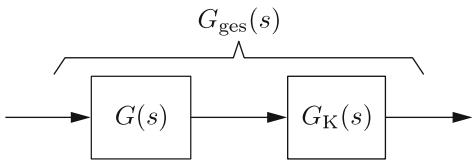


Abbildung 5.20. Strukturbild zur Kompensation des Zeitverhaltens.

Die Übertragungsfunktion der gesamten Anordnung ist damit

$$G_{\text{ges}}(s) = G(s) \cdot G_K(s) \quad (5.212)$$

$$= \frac{Z(s)}{M(s) \cdot (s^2 - (s_{\infty K} + s_{\infty K}^*) s + s_{\infty K} s_{\infty K}^*)}. \quad (5.213)$$

Bei idealer Kompensation sind die dominanten Pole $(s_{\infty 0}, s_{\infty 0}^*)$ nicht mehr in $G_{\text{ges}}(s)$ enthalten. In der praktischen Anwendung ist eine ideale Kompensation aus zwei Gründen nicht möglich:

- Die Systemparameter $(s_{\infty 0}, s_{\infty 0}^*)$ ändern sich im Betrieb – beispielsweise durch Temperatureinflüsse oder Alterung:

$$(s_{\infty 0}, s_{\infty 0}^*) \rightarrow (s_{\infty}, s_{\infty}^*). \quad (5.214)$$

- Die Systemparameter $(s_{\infty 0}, s_{\infty 0}^*)$ sind nicht exakt bekannt und können z. B. durch eine Systemidentifikation nur ungenau bestimmt werden:

$$s_{\infty} = s_{\infty 0} + \Delta s_{\infty 0}. \quad (5.215)$$

Die dominanten Pole können daher nicht vollständig kompensiert werden. Die Übertragungsfunktion des unvollständig kompensierten Systems ist

$$\tilde{G}_{\text{ges}}(s) = G_{\text{ges}}(s) \cdot \frac{s^2 - (s_{\infty 0} + s_{\infty 0}^*) s + s_{\infty 0} s_{\infty 0}^*}{s^2 - (s_{\infty} + s_{\infty}^*) s + s_{\infty} s_{\infty}^*}. \quad (5.216)$$

Die inverse Laplace-Transformierte von $\tilde{G}_{\text{ges}}(s)$ ist die Impulsantwort des kompensierten Systems. Sie lässt sich nicht allgemein angeben. Wegen der unvollständigen Kompensation enthält das Nennerpolynom von $\tilde{G}_{\text{ges}}(s)$ wieder den Term

$$s^2 - (s_{\infty} + s_{\infty}^*) s + s_{\infty} s_{\infty}^* = s^2 + 2\delta s + \delta^2 + (2\pi f)^2. \quad (5.217)$$

Dieser bewirkt einen zusätzlichen Einschwingvorgang in der Impulsantwort, der mit $e^{-\delta t}$ langsam abklingt.

Beispiel 5.17 (Ideale Kompensation bei der Temperaturmessung): Ein Berührungsthermometer steckt in einer Flüssigkeit der Temperatur T . Seine Wärmekapazität sei C , der Wärmeübergang sei durch α gegeben (Abb. 5.21). Der Energieerhaltungssatz besagt, dass die auf das Thermometer übergehende Wärme gleich der Erhöhung der inneren Energie des Thermometers ist,

$$\Phi = \alpha(T - T_m) = C \frac{dT_m}{dt} = \frac{dE}{dT}, \quad (5.218)$$

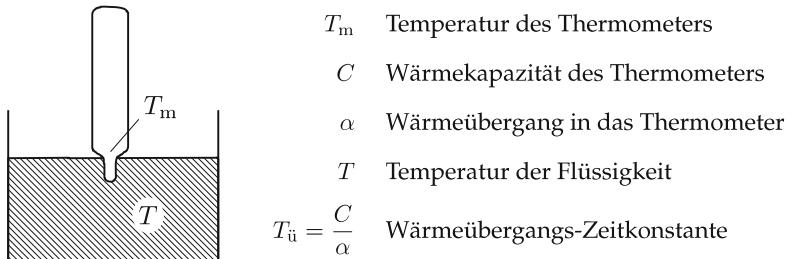


Abbildung 5.21. Temperaturmessung.

oder in den Laplace-Bereich transformiert:

$$T_m(s) = \frac{\alpha}{Cs + \alpha} T(s) = \frac{1}{T_{\ddot{u}} s + 1} T(s). \quad (5.219)$$

Im stationären Zustand ist wegen

$$\lim_{s \rightarrow 0} G(s) = \lim_{s \rightarrow 0} \frac{1}{T_{\ddot{u}} s + 1} = 1 \quad (5.220)$$

die Temperatur des Thermometers T_m gleich der Flüssigkeitstemperatur T :

$$T_m = T \quad \text{für} \quad t \rightarrow \infty. \quad (5.221)$$

Mit der Kompensationsübertragungsfunktion

$$G_K(s) = \frac{T_{\ddot{u}} s + 1}{T_K s + 1}, \quad T_K \ll T_{\ddot{u}}, \quad (5.222)$$

erhält man eine ideal kompensierte Gesamtübertragungsfunktion

$$G_{\text{ges}}(s) = G(s) G_K(s) = \frac{1}{T_K s + 1} \quad (5.223)$$

mit einer kleineren Zeitkonstante T_K und damit eine Impulsantwort

$$g_{\text{ges}}(t) = \frac{1}{T_K} e^{-t/T_K} \quad (5.224)$$

mit einem schnelleren Einschwingverhalten. ■

Beispiel 5.18 (Kompensation bei veränderlichen Parametern): Der Wärmeübergang α aus Bsp. 5.17 ist nun aber nicht konstant, sondern hängt von der Oberfläche und Form des Thermometers, aber auch von Größen ab, die sich im Betrieb ändern, wie die Anströmgeschwindigkeit, die Dichte

und Art des Messstoffes.

Betrachtet man die Abhängigkeit von der Strömungsgeschwindigkeit v des Messstoffes, so erhält man nach einem empirischen Gesetz

$$\frac{\alpha}{\alpha_0} = \left(\frac{v}{v_0} \right)^m, \quad m \approx 0,8. \quad (5.225)$$

Für kleine Änderungen gilt die Näherung

$$\frac{\Delta\alpha}{\alpha_0} \approx m \frac{\Delta v}{v}. \quad (5.226)$$

Die Strömungsgeschwindigkeit ändere sich im Betrieb um $\Delta v/v = 0,5$, wodurch $\Delta\alpha/\alpha_0 = 0,4$ wird. Dadurch ändert sich die Systemzeitkonstante:

$$T_{\ddot{u}} = \frac{C}{\alpha} = \frac{C}{\alpha_0 + \Delta\alpha} = \frac{C}{\alpha_0} \cdot \frac{1}{1 + \frac{\Delta\alpha}{\alpha_0}} = T_{\ddot{u}0} \frac{1}{1 + \frac{\Delta\alpha}{\alpha_0}}. \quad (5.227)$$

Die unvollständig kompensierte Übertragungsfunktion ist

$$\tilde{G}_{\text{ges}}(s) = \frac{1}{T_K s + 1} \cdot \frac{T_{\ddot{u}0} s + 1}{T_{\dot{u}} s + 1} \quad (5.228)$$

$$\begin{aligned} &= \frac{(T_{\ddot{u}0} - T_K)(1 + \frac{\Delta\alpha}{\alpha_0})}{T_{\ddot{u}0} - T_K(1 + \frac{\Delta\alpha}{\alpha_0})} \cdot \frac{1}{T_K s + 1} \\ &\quad - \frac{T_{\ddot{u}0} \frac{\Delta\alpha}{\alpha_0}}{T_{\ddot{u}0} - T_K(1 + \frac{\Delta\alpha}{\alpha_0})} \cdot \frac{1}{\left[T_{\ddot{u}0} / (1 + \frac{\Delta\alpha}{\alpha_0}) \right] s + 1} \end{aligned} \quad (5.229)$$

und die zugehörige Impulsantwort

$$\begin{aligned} \tilde{g}_{\text{ges}}(t) &= \frac{1}{T_K} \cdot \frac{(T_{\ddot{u}0} - T_K)(1 + \frac{\Delta\alpha}{\alpha_0})}{T_{\ddot{u}0} - T_K(1 + \frac{\Delta\alpha}{\alpha_0})} \exp\left(-\frac{t}{T_K}\right) \\ &\quad - \frac{(1 + \frac{\Delta\alpha}{\alpha_0}) \frac{\Delta\alpha}{\alpha_0}}{T_{\ddot{u}0} - T_K(1 + \frac{\Delta\alpha}{\alpha_0})} \exp\left(-t \frac{1 + \frac{\Delta\alpha}{\alpha_0}}{T_{\ddot{u}0}}\right). \end{aligned} \quad (5.230)$$

Mit wachsender Parameterabweichung $\Delta\alpha$ wird der Einschwingvorgang durch die größere Zeitkonstante $T_{\dot{u}}$ bestimmt. In Abb. 5.22 sind die verschiedenen Einschwingvorgänge dargestellt. Der Fehler wird im Wesentlichen durch den langsam abklingenden zweiten Term bestimmt. ■

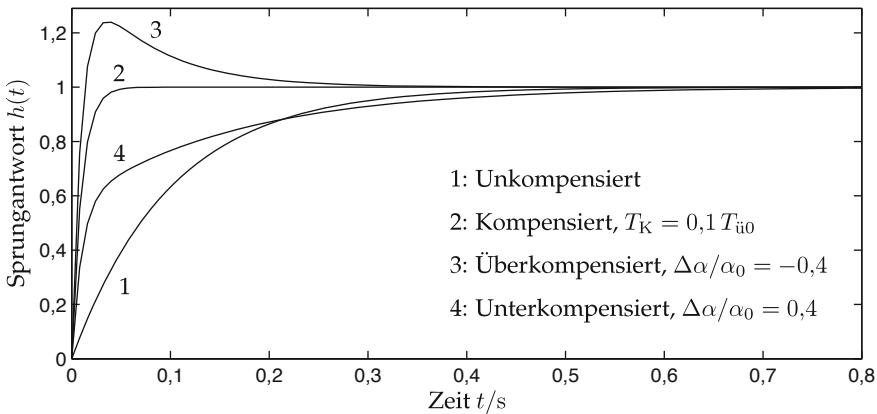


Abbildung 5.22. Sprungantwort eines Berührungsthermometers mit einem Kompensationsnetzwerk in Reihe.

5.5.2 Zeitverhalten bei Gegenkopplung

Die Gegenkopplung soll auf die Möglichkeit zur Veränderung des Zeitverhaltens untersucht werden. Dazu wird das System mit der Übertragungsfunktion

$$G_R(s) = V \cdot G_K(s) \quad (5.231)$$

in Reihe zum Messsystem $G(s)$ geschaltet. Der Systemausgang $y(t)$ wird auf den Eingang zurückgekoppelt (Abb. 5.23). Diese Anordnung ist auch als Regelkreis bekannt. In der Regelungstechnik bezeichnet man daher das System $G_{\text{R}}(s)$ auch als Regler. Die Übertragungsfunktion des geschlossenen Kreises ist

$$G_{\text{ges}}(s) = \frac{G_{\text{R}}(s) G(s)}{1 + G_{\text{R}}(s) G(s)} = \frac{V G_K(s) G(s)}{1 + V G_K(s) G(s)}. \quad (5.232)$$

Erhöht man formal die Verstärkung V , so geht die Übertragungsfunktion $G_{\text{ges}}(s)$ unabhängig von $G(s)$ und $G_K(s)$ in die ideale Übertragungsfunktion

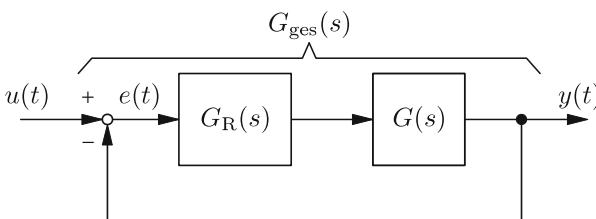


Abbildung 5.23. Strukturbild zur Gegenkopplung.

über:

$$\lim_{V \rightarrow \infty} G_{\text{ges}}(s) = G_{\text{ideal}}(s) = 1. \quad (5.233)$$

Damit *scheint* eine hohe Regelkreisverstärkung V das geeignete Mittel zu sein, um das Zeitverhalten von Messsystemen zu verbessern. Leider wird für große Verstärkungen V der Regelkreis meist instabil, selbst wenn die einzelnen Teilsysteme stabil sind. Daher ist die Gegenkopplung nur bedingt geeignet, das dynamische Verhalten von Messsystemen zu verbessern.

④ 5.5.2.1 P-Regler

Um die Problematik zu verdeutlichen, wird im Folgenden ein **P-Regler** mit der konstanten Übertragungsfunktion

$$G_R(s) = V \quad (5.234)$$

betrachtet und gleichzeitig eine Übertragungsfunktion $G(s)$ des Messsystems ohne Nullstellen angenommen:

$$G(s) = \frac{(-1)^n \prod_{i=1}^n s_{\infty i}}{\prod_{i=1}^n (s - s_{\infty i})}, \quad (5.235)$$

wobei $G(0) = 1$ ist. Das Nennerpolynom von $G(s)$ wird ausmultipliziert:

$$N(s) = s^n - s^{n-1} \sum_{i=1}^n s_{\infty i} + \dots + (-1)^{n-1} s \cdot \prod_{i=1}^n s_{\infty i} \sum_{j=1}^n \frac{1}{s_{\infty j}} + (-1)^n \prod_{i=1}^n s_{\infty i}. \quad (5.236)$$

Die komplexen Pole sind jeweils paarweise konjugiert komplex zueinander:

$$s_{\infty i} = -\delta_i + j2\pi f_i, \quad s_{\infty i}^* = -\delta_i - j2\pi f_i. \quad (5.237)$$

Die Summe und das Produkt zweier konjugiert komplexer Pole führen auf

$$s_{\infty} + s_{\infty}^* = -\delta + j2\pi f - \delta - j2\pi f = -2\delta = 2\Re\{s_{\infty}\} \quad (5.238)$$

bzw.

$$s_{\infty} s_{\infty}^* = \delta^2 + (2\pi f)^2 = r^2, \quad (5.239)$$

wobei r den Abstand der Pole vom Ursprung der komplexen Ebene bezeichnet. Damit lassen sich die Summen- und Produktterme des ausmultiplizier-

ten Nennerpolynoms $N(s)$ ausdrücken:

$$\sum_{i=1}^n s_{\infty i} = \sum_{i=1}^n \Re\{s_{\infty i}\}, \quad (5.240)$$

$$\prod_{i=1}^n s_{\infty i} = \prod_{i=1}^n r_i. \quad (5.241)$$

Die Übertragungsfunktion des geschlossenen Kreises ist

$$G_{\text{ges}}(s) = \frac{(-1)^n V \cdot \prod_{i=1}^n s_{\infty i}}{\prod_{i=1}^n (s - s_{\infty i}) + (-1)^n V \cdot \prod_{i=1}^n s_{\infty i}}. \quad (5.242)$$

Durch Ausmultiplizieren ergibt sich das Nennerpolynom des geschlossenen Kreises zu:

$$\begin{aligned} N_{\text{ges}}(s) &= s^n - s^{n-1} \sum_{i=1}^n s_{g\infty i} + \dots \\ &\quad + (-1)^{n-1} s \cdot \prod_{i=1}^n s_{g\infty i} \sum_{j=1}^n \frac{1}{s_{g\infty j}} + (-1)^n \prod_{i=1}^n s_{g\infty i}. \end{aligned} \quad (5.243)$$

Durch Vergleich der beiden Nennerpolynome findet man die Beziehung zwischen der Lage der ursprünglichen Pole $s_{\infty i}$ und derer im gegengekoppelten Kreis $s_{g\infty i}$.

– Koeffizientenvergleich für s^{n-1} :

$$\sum_{i=1}^n s_{\infty i} = \sum_{i=1}^n s_{g\infty i}. \quad (5.244)$$

Mit (5.240) ergibt sich eine unveränderte Summe der Realteile aller Pole:

$$\sum_{i=1}^n \Re\{s_{\infty i}\} = \sum_{i=1}^n \Re\{s_{g\infty i}\}. \quad (5.245)$$

– Koeffizientenvergleich für s :

$$\prod_{i=1}^n s_{\infty i} \cdot \sum_{j=1}^n \frac{1}{s_{\infty j}} = \prod_{i=1}^n s_{g\infty i} \cdot \sum_{j=1}^n \frac{1}{s_{g\infty j}}. \quad (5.246)$$

Aus dem Produkt und der Summe der paarweise konjugiert komplexen Pole wird mit (5.238)

$$\frac{1}{s_\infty} + \frac{1}{s_\infty^*} = \frac{s_\infty^*}{r^2} + \frac{s_\infty}{r^2} = \frac{2\Re\{s_\infty\}}{r^2}, \quad (5.247)$$

so dass man schreiben kann:

$$\prod_{i=1}^n r_i \sum_{j=1}^n \frac{\Re\{s_{\infty j}\}}{r_j^2} = \prod_{i=1}^n r_{gi} \sum_{j=1}^n \frac{\Re\{s_{g\infty j}\}}{r_{gj}^2}. \quad (5.248)$$

- Koeffizientenvergleich für s^0 :

$$\prod_{i=1}^n s_{g\infty i} = (V+1) \prod_{i=1}^n s_{\infty i}. \quad (5.249)$$

Mit (5.241) ergibt sich ein Zusammenhang zwischen den Abständen der Pole zum Ursprung:

$$\prod_{i=1}^n r_{gi} = (V+1) \prod_{i=1}^n r_i. \quad (5.250)$$

Führt man einen mittleren Realteil der Pole ein,

$$\overline{\left(\frac{\Re\{s_\infty\}}{r^2} \right)} = \frac{1}{n} \sum_{j=1}^n \frac{\Re\{s_{\infty j}\}}{r_j^2}, \quad (5.251)$$

so lassen sich die Pole mit und ohne Gegenkopplung vergleichen. Mit (5.248) und (5.250) erhält man

$$\overline{\left(\frac{\Re\{s_{g\infty}\}}{r_g^2} \right)} = \frac{1}{V+1} \overline{\left(\frac{\Re\{s_\infty\}}{r^2} \right)}. \quad (5.252)$$

Aufgrund von (5.245) kann man grob abschätzen, dass sich die Realteile der Pole im Mittel wenig verändern. Die Abstände r_{gj}^2 der Pole vom Ursprung vergrößern sich dagegen mit wachsender Verstärkung V . Daraus kann man schließen, dass im Wesentlichen die Imaginärteile mit der Verstärkung anwachsen, was im Mittel zu wenig gedämpften Eigenschwingungen führt. Der mittlere Realteil wird im Fall der Gegenkopplung evtl. dadurch erhalten bleiben, dass einige Pole sogar auf die rechte Halbebene rücken. Damit wird das System instabil.

Beispiel 5.19 (Zweistufiger Verstärker): Ein zweistufiger Verstärker mit der Übertragungsfunktion

$$G(s) = \frac{V\alpha^2}{(s + \alpha)^2}, \quad \alpha \in \mathbb{R}, \quad (5.253)$$

hat die Sprungantwort

$$H(s) = \frac{G(s)}{s} = \frac{V}{s} - \frac{V\alpha}{(s + \alpha)^2} - \frac{V}{s + \alpha} \quad (5.254)$$



$$h(t) = V (1 - e^{-\alpha t} (\alpha t + 1)), \quad (5.255)$$

die für $\alpha > 0$ aperiodisches Einschwingverhalten aufweist. Schließt man nun die Gegenkopplungsschleife, so ist die Übertragungsfunktion des geschlossenen Kreises:

$$G_{\text{ges}}(s) = \frac{V\alpha^2}{(s + \alpha)^2 + V\alpha^2}. \quad (5.256)$$

Die Laplace-Transformierte der Sprungantwort des geschlossenen Kreises ist damit

$$\begin{aligned} H_{\text{ges}}(s) &= \frac{G_{\text{ges}}(s)}{s} = \frac{V\alpha^2}{s((s + \alpha)^2 + V\alpha^2)} \\ &= \frac{V}{1 + V} \left(\frac{1}{s} - \frac{s + \alpha}{(s + \alpha)^2 + V\alpha^2} - \frac{1}{\sqrt{V}} \cdot \frac{\sqrt{V}\alpha}{(s + \alpha)^2 + V\alpha^2} \right). \end{aligned} \quad (5.257)$$

Daraus berechnet sich die Sprungantwort im Zeitbereich zu

$$h_{\text{ges}}(t) = \frac{V}{1 + V} \left[1 - e^{-\alpha t} \left(\cos \sqrt{V}\alpha t + \frac{1}{\sqrt{V}} \cdot \sin \sqrt{V}\alpha t \right) \right]. \quad (5.258)$$

Der stationäre Endwert der Sprungantwort ist

$$\lim_{t \rightarrow \infty} h_{\text{ges}}(t) = \frac{V}{1 + V} < 1. \quad (5.259)$$

Stationäre Genauigkeit wird nur für $V \rightarrow \infty$ erreicht. Setzt man im Fall der Gegenkopplung $V = 100$, so wird aus einer aperiodischen Einstellung bei etwa gleichen Dämpfungseigenschaften α in der Gegenkopplungsschaltung eine schwingende Einstellung. Abbildung 5.24 zeigt die auf den Endwert h_∞ normierten Sprungantworten $h(t)$ und $h_{\text{ges}}(t)$ für die Parametereinstellungen $V = 100$ und $\alpha = 10$. ■

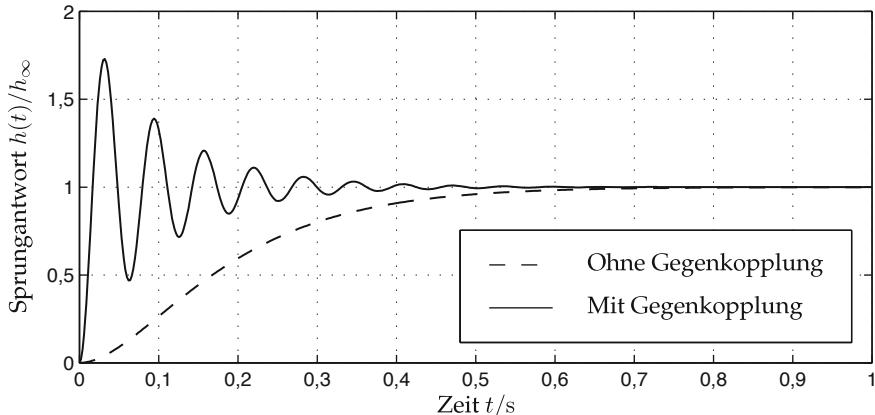


Abbildung 5.24. Sprungantwort eines zweistufigen Verstärkers mit und ohne Gegenkopplung.

5.5.2.2 PI-Regler

Die Berechnungen und das Beispiel haben gezeigt, dass allein die Erhöhung der Verstärkung im geschlossenen Regelkreis kein geeignetes Mittel ist, das dynamische Verhalten von Messsystemen zu verbessern. Kombiniert man allerdings die Rückkopplung mit dem Kompensationsverfahren nach Abschn. 5.5.1 zur Kompensation der größten Zeitkonstante in der Übertragungsfunktion $G(s)$ des Messsystems, so gelangt man zum **PI-Regler**:

$$G_R(s) = V \frac{1 + T_K s}{s}. \quad (5.260)$$

Mit den Grenzwertsätzen der Laplace-Transformation folgt für die Sprungantwort:

$$h_\infty = \lim_{t \rightarrow \infty} h(t) = \lim_{s \rightarrow 0} s \cdot V \frac{1 + T_K s}{s} \cdot \frac{1}{s} = \infty, \quad (5.261)$$

$$h(0) = \lim_{t \rightarrow 0} h(t) = \lim_{s \rightarrow \infty} s \cdot V \frac{1 + T_K s}{s} \cdot \frac{1}{s} = V T_K. \quad (5.262)$$

Damit ist die stationäre Verstärkung unendlich groß, während die Verstärkung für hohe Frequenzen begrenzt ist. Damit treten weniger Stabilitätsprobleme auf als bei konstanter Verstärkung im P-Regler. Zusätzlich ist der Regelkreis stationär genau. Es sind aber auch andere Übertragungsfunktionen für den Regler $G_R(s)$ denkbar. Die Untersuchung und Dimensionierung dieser Systeme ist allerdings Aufgabe der Regelungstechnik. Für eine genauere Beschreibung sei daher auf die Literatur [6, 11, 57] verwiesen. Hier sollen lediglich an einem Beispiel die Vorteile eines PI-Reglers gegenüber dem P-Regler demonstriert werden.

Beispiel 5.20 (Gegenkopplung mit P- und PI-Regler): Gegeben sei ein Messsystem, das durch drei in Reihe geschaltete PT₁-Glieder

$$G(s) = \frac{1}{1+T_1s} \frac{1}{1+T_2s} \frac{1}{1+T_3s} \quad (5.263)$$

mit den Zeitkonstanten $T_1 = 0,53\text{ s}$, $T_2 = 0,055\text{ s}$ und $T_3 = 0,005\text{ s}$ charakterisiert ist.

P-Regler: Setzt man einen P-Regler mit

$$G_R(s) = V \quad (5.264)$$

an, so erhält man die Übertragungsfunktion des geschlossenen Kreises

$$G_{\text{ges},P}(s) = \frac{V}{V + (1+T_1s)(1+T_2s)(1+T_3s)}. \quad (5.265)$$

Für die geforderte Stabilität müssen die Pole von $G_{\text{ges},P}(s)$ in der linken s -Halbebene liegen. Dazu wird der Nenner ausmultipliziert, zu null gesetzt und man gelangt auf diese Weise zur charakteristischen Gleichung:

$$T_1 T_2 T_3 s^3 + (T_1 T_2 + T_1 T_3 + T_2 T_3)s^2 + (T_1 + T_2 + T_3)s + V + 1 = 0. \quad (5.266)$$

Zur Prüfung auf Stabilität bietet sich das Hurwitz-Kriterium (Satz 5.2) an. Für eine allgemeine Gleichung 3. Grades

$$a_0 s^3 + a_1 s^2 + a_2 s + a_3 = 0 \quad (5.267)$$

ergibt sich die Hurwitz-Determinante zu

$$\begin{vmatrix} a_1 & a_3 & 0 \\ a_0 & a_2 & 0 \\ 0 & a_1 & a_3 \end{vmatrix}. \quad (5.268)$$

Das System ist stabil, wenn alle nordwestlichen Unterdeterminanten positiv sind:

$$H_1 = a_1 \stackrel{!}{>} 0, \quad (5.269)$$

$$H_2 = \begin{vmatrix} a_1 & a_3 \\ a_0 & a_2 \end{vmatrix} = a_1 a_2 - a_0 a_3 \stackrel{!}{=} 0, \quad (5.270)$$

$$H_3 = a_3 H_2 \stackrel{!}{>} 0. \quad (5.271)$$

Da alle Zeitkonstanten positiv sind, gilt dies auch für a_1 und a_3 . Prüft man die Bedingung $H_2 > 0$, so führt dies nach einigen Umformungen

auf die Stabilitätsbedingung

$$V < (T_1 + T_2 + T_3) \left(\frac{1}{T_1} + \frac{1}{T_2} + \frac{1}{T_3} \right) - 1. \quad (5.272)$$

Setzt man die Zahlenwerte ein, so ist der geschlossene Kreis stabil für

$$V < 128,9. \quad (5.273)$$

Für die praktische Anwendung wird man V um einiges kleiner wählen, will man starke Einschwingvorgänge vermeiden. Dies stellt einen Kompromiss mit der stationären Genauigkeit dar, die nur für große V erreicht werden kann.

PI-Regler: Setzt man einen PI-Regler ein, so kann die größte Zeitkonstante des Messsystems $T_1 = 0,53\text{s}$ kompensiert werden. Die Übertragungsfunktion des PI-Reglers lautet dann

$$G_R(s) = V \frac{1 + T_1 s}{s}. \quad (5.274)$$

Damit wird die Übertragungsfunktion des geschlossenen Kreises zu

$$G_{\text{ges,PI}}(s) = \frac{V}{V + s(1 + T_2 s)(1 + T_3 s)}. \quad (5.275)$$

Die Stabilitätsuntersuchung erfolgt ebenfalls mit den Hurwitz-Kriterium. Man erhält als Stabilitätsbedingung

$$V < \frac{1}{T_2} + \frac{1}{T_3} = 218,2 = V_G, \quad (5.276)$$

wobei V_G die Verstärkung an der Stabilitätsgrenze bezeichnet. Man erkennt bereits, dass V prinzipiell im Vergleich zum P-Regler viel größer gewählt werden kann. In der Praxis sind allerdings nur Werte

$$V = 0,05 V_G \quad \text{bis} \quad V = 0,1 V_G \quad (5.277)$$

für ein brauchbares Einschwingverhalten sinnvoll. Abbildung 5.25 zeigt die Sprungantworten für $V = 60$ bei Einsatz eines P-Reglers und für $V = 20$ bei einem PI-Regler. Zum Vergleich sind auch die Sprungantworten des kompensierten und des unkompensierten Systems ohne Rückkopplung nach Abschn. 5.5.1 dargestellt.

Man erkennt, dass die Kompensation sehr gute Resultate liefert. Sie ist allerdings empfindlich gegenüber Schwankungen von T_1 . Der PI-Regler zeigt ein schnelleres Anstiegsverhalten bei nur geringem Überschwingen und ist stationär genau. Der P-Regler zeigt starke Schwin-

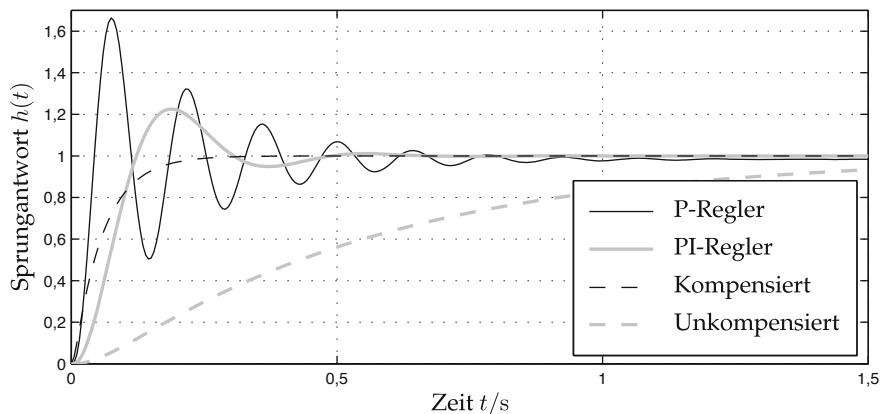
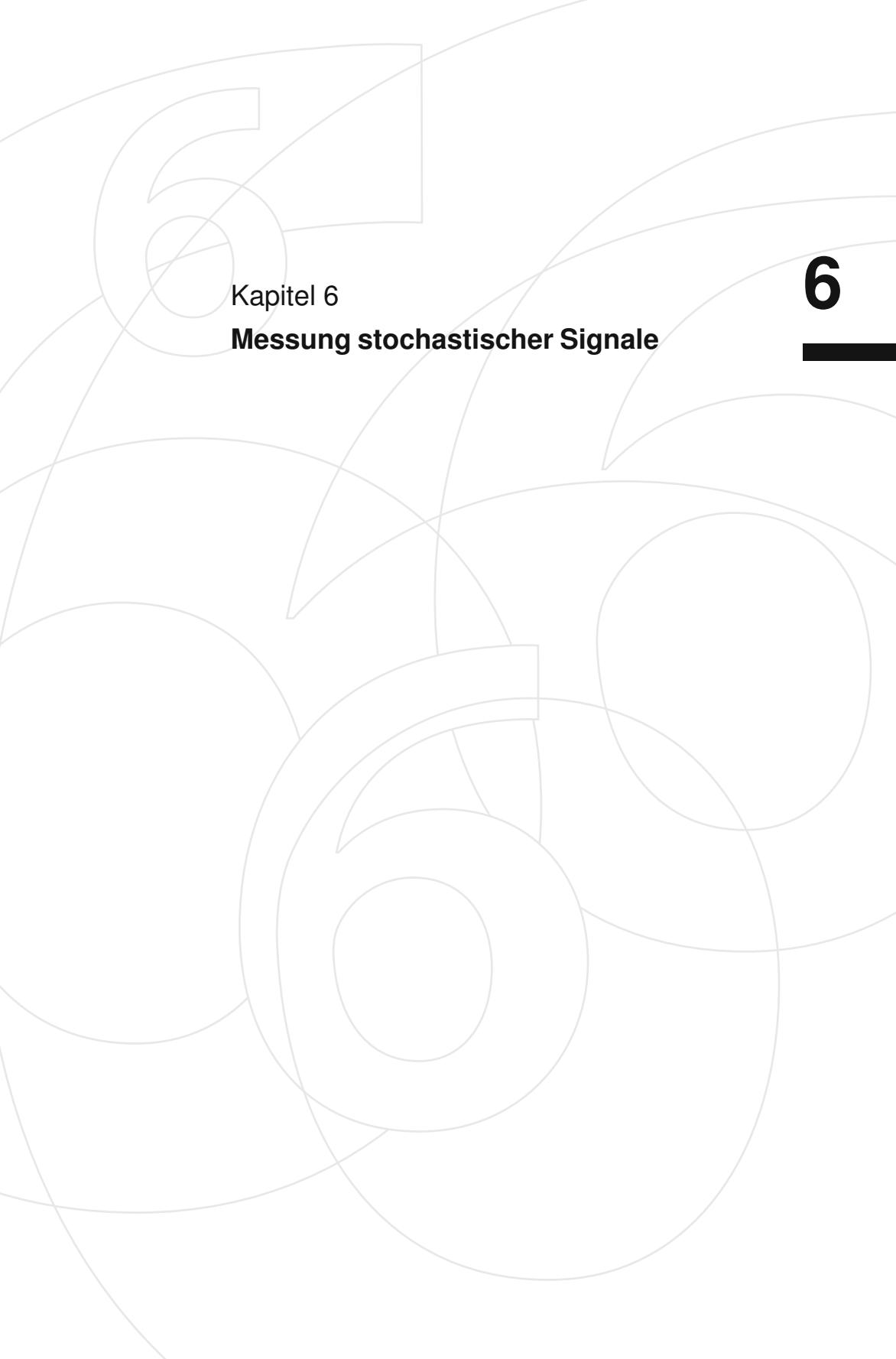


Abbildung 5.25. Einschwingverhalten bei verschiedenen Maßnahmen zur Optimierung des Zeitverhaltens.

gungsneigung. Trotz der großen Verstärkung ist dieser Regler stationär ungenau. ■



Kapitel 6

Messung stochastischer Signale

6

6 Messung stochastischer Signale

| | | |
|---------|---|-----|
| 6.1 | Stochastische Prozesse | 245 |
| 6.1.1 | Einführung | 245 |
| 6.1.2 | Wahrscheinlichkeitsverteilung und Wahrscheinlichkeitsdichte | 247 |
| 6.1.3 | Schar- und Zeitmittelwerte | 250 |
| 6.1.4 | Momente der Statistik 1. Ordnung | 250 |
| 6.1.5 | Momente der Statistik 2. Ordnung | 252 |
| 6.1.6 | Stationäre Prozesse | 254 |
| 6.1.7 | Ergodische Prozesse | 256 |
| 6.2 | Korrelationsfunktionen | 261 |
| 6.2.1 | Signalklassen | 261 |
| 6.2.1.1 | Ergodische Prozesse | 263 |
| 6.2.1.2 | Funktionenräume | 263 |
| 6.2.2 | Korrelation von Leistungssignalen | 265 |
| 6.2.3 | Korrelation von Energiesignalen | 267 |
| 6.2.4 | Eigenschaften von Auto- und Kreuzkorrelationsfunktion . | 269 |
| 6.3 | Korrelationsmesstechnik | 271 |
| 6.3.1 | Messung von Korrelationsfunktionen | 271 |
| 6.3.2 | Ähnlichkeit von Signalen, Laufzeitmessung | 272 |
| 6.3.3 | Closed-loop-Korrelation | 278 |
| 6.3.4 | Polaritätskorrelation | 280 |
| 6.3.5 | Ähnlichkeit von Spektren, Dopplerfrequenzmessung . | 281 |
| 6.3.6 | Selbstähnlichkeit | 284 |
| 6.4 | Spektrale Darstellung stochastischer Signale | 285 |
| 6.4.1 | Leistungsdichtespektrum | 285 |
| 6.4.2 | Rauschen | 288 |
| 6.4.3 | Überlagerung zufälliger Störsignale | 293 |
| 6.4.4 | Übertragung stochastischer Signale durch LTI-Systeme | 295 |
| 6.5 | Systemidentifikation | 298 |
| 6.5.1 | Schätzung des Leistungsdichtespektrums | 299 |
| 6.5.2 | Systemidentifikation bei geschätzter Leistungsdichte . | 301 |
| 6.5.2.1 | Quotientenbildung gemittelter Periodogramme | 301 |
| 6.5.2.2 | Quotientenbildung gemittelter Kreuzleistungsdichten | 302 |
| 6.5.3 | Dynamische Systemidentifikation | 303 |
| 6.6 | Wiener-Filter | 304 |

| | | |
|-------|---|-----|
| 6.6.1 | Signalmodell | 305 |
| 6.6.2 | Herleitung des Wiener-Filters | 307 |
| 6.6.3 | Wiener-Filter bei linearer Verzerrung und additivem Rauschen | 308 |

6

6 Messung stochastischer Signale

Bisher wurden zufällige Messfehler behandelt, die *einzelne* Messwerte stationärer Messgrößen verfälschen. Will man hingegen *Funktionsverläufe* $u(t)$ messen, so erwartet man auch hier für jeden Zeitpunkt t einen zufälligen Messfehler. Dieser wird gewöhnlich als Rauschen interpretiert. Für zufällige Signale – wie Rauschen oder chaotische Schwingungen – ist die Angabe eines Funktionsverlaufes unmöglich. Herkömmliche Analysemethoden gehen davon aus, dass die Signale durch mathematische Funktionen beschreibbar sind (z. B. durch eine Interpolationsfunktion). Damit kann man mit den Werten der Signalverläufe rechnen, Signale addieren und subtrahieren, verzögern oder Ableitungen und Integrale bilden. Speziell für Fourier-, Laplace- oder Faltungsintegrale ist nicht sichergestellt, ob diese für zufällige Signale überhaupt existieren.

Um dennoch wie gewohnt mit Eingangs- und Ausgangsgrößen von Messsystemen rechnen zu können, muss auch für Zufallssignale eine geeignete Beschreibung gefunden werden. Hier hilft das Konzept der **stochastischen Prozesse**, die gewöhnlich durch Erwartungswerte, Korrelationsfunktionen und Leistungsdichtespektren beschrieben werden.

6.1 Stochastische Prozesse

6.1

6.1.1 Einführung

Bei der Beschreibung von Systemen, in denen zufällige Signale (oder Signalanteile) auftreten, will man sich nicht auf wenige, explizit bekannte Störungen beschränken. Vielmehr sind allgemeine Lösungen gefordert, die nicht nur für einzelne Signale gelten, sondern für eine große Anzahl *möglicher* Signale mit gewissen gemeinsamen Eigenschaften. Ein mathematisches Modell für eine derartige Schar von Signalen ist der **stochastische Prozess** – auch **Zufallsprozess** genannt. Genaue Ausführungen hierzu findet man in [19,37].

Definition 6.1: Stochastischer Prozess

6.1

Ein stochastischer Prozess $y(t, \xi)$ besteht aus einem Zufallsexperiment und einer Zuordnung von deterministischen Funktionen $y(t, \xi_\nu)$ – den sogenannten **Musterfunktionen** (oder Realisierungen) des stochastischen Prozesses – zu den Elementarereignissen ξ_ν des Zufallsexperimentes (vgl. Abb. 6.1). Für jeden festen Zeitpunkt $t = t_i$ geht der stochastische Prozess $y(t, \xi)$

in eine Zufallsvariable $y(t_i, \xi)$ über, die sich mit Hilfe einer Wahrscheinlichkeitsverteilung beschreiben lässt. \diamond

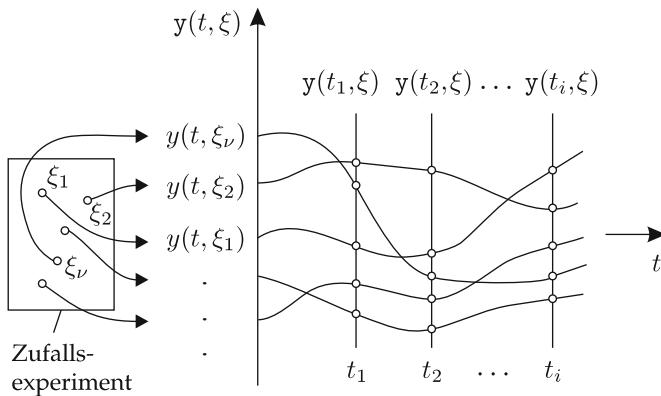


Abbildung 6.1. Illustration eines stochastischen Prozesses.

Ein stochastischer Prozess $y(t, \xi)$ hat zwei Parameter. Abhängig davon, ob diese jeweils als fest oder variabel angenommen werden, ergeben sich vier verschiedene Interpretationen (vgl. Tab. 6.1):

1. $\xi = \xi_\nu$ ist fest, t ist variabel:

$y(t, \xi_\nu)$ ist eine einzelne Realisierung des stochastischen Prozesses, also eine deterministische Musterfunktion (vgl. Bsp. 6.2).

2. ξ ist variabel, $t = t_i$ ist fest:

$y(t_i, \xi)$ ist eine Zufallsvariable, die jedem Ereignis ξ_ν den entsprechenden Funktionswert der zugehörigen Musterfunktion $y(t, \xi_\nu)$ zuordnet.

3. $\xi = \xi_\nu$ ist fest, $t = t_i$ ist fest:

$y(t_i, \xi_\nu)$ ist ein Zahlenwert.

4. ξ ist variabel, t ist variabel:

$y(t, \xi)$ ist ein stochastischer Prozess, d. h. eine Schar von Musterfunktionen.

Tabelle 6.1. Interpretationen eines stochastischen Prozesses.

| Elementar- ereignis ξ | Zeit t | |
|------------------------------|-------------------------------|--------------------------------|
| | fest | variabel |
| fest | Zahlenwert $y(t_i, \xi_\nu)$ | Musterfunktion $y(t, \xi_\nu)$ |
| | Zufallsvariable $y(t_i, \xi)$ | Zufallsprozess $y(t, \xi)$ |

Da ein Zufallsprozess $y(t, \xi)$ als eine Familie von Zufallsvariablen aufgefasst werden kann, können alle Gesetze für Zufallsvariable aus der Wahrscheinlichkeitstheorie (Abschn. 4.1) darauf angewendet werden. Im Folgenden wird bei Zufallsprozessen $y(t, \xi)$ die Abhängigkeit vom Elementarergebnis ξ unterdrückt und somit die kürzere Schreibweise $y(t)$ verwendet.

Beispiel 6.1 (Rauschspannung): Bedingt durch thermisches Rauschen ist die Spannung $u(t)$ an einem Widerstand R verrauscht und kann selbst bei Kenntnis des durchfließenden Stromes $i(t)$ nicht genau vorhergesagt werden. Bei der Spannung $u(t)$ handelt es sich daher um einen stochastischen Prozess, dessen Musterfunktionen verschiedene Verläufe aufweisen. Die Schar aller möglichen Funktionen ist in diesem Fall unendlich groß. ■

Beispiel 6.2 (Harmonische Schwingung mit zufälliger Phase): Ein Zufallsprozess $y(t)$ erzeugt Musterfunktionen

$$y_i(t) = \sin(\omega t + \varphi_i), \quad (6.1)$$

bei denen die Kreisfrequenz ω fest ist, die Phase φ_i aber rein zufällig ist. Für feste Zeiten t_0 ist $y(t_0)$ eine Zufallsvariable, die alle Amplitudenwerte der Sinusschwingung annehmen kann. Jede einzelne Musterfunktion $y_i(t)$ ist eine deterministische Funktion. ■

6.1.2 Wahrscheinlichkeitsverteilung und Wahrscheinlichkeitsdichte

Bei stochastischen Prozessen $y(t)$ sind die Wahrscheinlichkeitsverteilungen $F_{y(t)}(y)$ und die Wahrscheinlichkeitsdichten $f_{y(t)}(y)$ abhängig von der Zeit t . Die Definition der Wahrscheinlichkeitsverteilung folgt unmittelbar aus der klassischen Wahrscheinlichkeitsrechnung. Daraus wird die Wahrscheinlichkeitsdichte abgeleitet.

Definition 6.2: Wahrscheinlichkeitsverteilung

Die Wahrscheinlichkeitsverteilung

$$F_{y(t)}(y) = P(y(t) \leq y) \quad (6.2)$$

eines stochastischen Prozesses $y(t)$ gibt die Wahrscheinlichkeit P an, mit welcher der Funktionswert von $y(t)$ zum Zeitpunkt t kleiner oder höchstens gleich y ist. \diamond

Hieraus folgt sofort die Wahrscheinlichkeitsdichte als partielle Ableitung der Wahrscheinlichkeitsverteilung.

6.3 Definition 6.3: Wahrscheinlichkeitsdichte

Die Wahrscheinlichkeitsdichte eines stochastischen Prozesses $y(t)$ ist

$$f_{y(t)}(y) = \frac{\partial F_{y(t)}(y)}{\partial y} \quad (6.3)$$

mit

$$\int_{-\infty}^{\infty} f_{y(t)}(y) dy = 1; \quad (6.4)$$

siehe (4.7). \diamond

Sowohl die Wahrscheinlichkeitsverteilung als auch die Wahrscheinlichkeitsdichte hängen im Allgemeinen von der Zeit t ab. Die partielle Ableitung in (6.3) ist als verallgemeinerte Ableitung anzusehen. An Stellen, an denen in der Wahrscheinlichkeitsverteilung Sprünge auftreten, enthält die Wahrscheinlichkeitsdichte δ -Distributionen („Dirac-Impulse“) [42].

Diese Definitionen lassen sich auf mehrdimensionale stochastische Prozesse ausdehnen. Wir wollen uns hier auf zweidimensionale Prozesse beschränken. Betrachtet man zwei verschiedene Zufallsprozesse $x(t)$ und $y(t)$ zu festen Zeitpunkten t_1 und t_2 , so lassen sich die Verbundverteilung und die Verbunddichte analog zu Abschn. 4.1 definieren, indem man darüber hinaus die Zeitabhängigkeit berücksichtigt.

6.4

Definition 6.4: Verbundwahrscheinlichkeitsverteilung zweier Zufallsprozesse

Die Verbundwahrscheinlichkeitsverteilung (oder gemeinsame Wahrscheinlichkeitsverteilung)

$$F_{x(t_1)y(t_2)}(x, y) = P(x(t_1) \leq x \cap y(t_2) \leq y) \quad (6.5)$$

zweier stochastischer Prozesse gibt die Wahrscheinlichkeit P an, mit welcher der Funktionswert von $x(t_1)$ zum Zeitpunkt t_1 kleiner oder höchstens gleich x ist und der Funktionswert von $y(t_2)$ zum Zeitpunkt t_2 kleiner oder höchstens gleich y ist. \diamond

Definition 6.5: Verbundwahrscheinlichkeitsdichte zweier Zufallsprozesse

Die Verbundwahrscheinlichkeitsdichte (oder gemeinsame Wahrscheinlichkeitsdichte) zweier stochastischer Prozesse $x(t)$ und $y(t)$ ist durch

$$f_{x(t_1)y(t_2)}(x, y) = \frac{\partial^2 F_{x(t_1)y(t_2)}(x, y)}{\partial x \partial y} \quad (6.6)$$

definiert. \diamond

Durch die Betrachtung eines einzigen stochastischen Prozesses $x(t)$ zu zwei Zeitpunkten t_1 und t_2 lässt sich ebenfalls eine Verbundwahrscheinlichkeitsverteilung angeben.

Definition 6.6: Verbundwahrscheinlichkeitsverteilung eines Zufallsprozesses

Die Verbundwahrscheinlichkeitsverteilung (oder gemeinsame Wahrscheinlichkeitsverteilung) eines stochastischen Prozesses $x(t)$ beschreibt die Wahrscheinlichkeit P , mit welcher der Funktionswert von $x(t_1)$ zum Zeitpunkt t_1 kleiner oder höchstens gleich x_1 ist und der Funktionswert von $x(t_2)$ zum Zeitpunkt t_2 kleiner oder höchstens gleich x_2 ist. \diamond

$$F_{x(t_1)x(t_2)}(x_1, x_2) = P(x(t_1) \leq x_1 \cap x(t_2) \leq x_2) \quad (6.7)$$

eines stochastischen Prozesses $x(t)$ beschreibt die Wahrscheinlichkeit P , mit welcher der Funktionswert von $x(t_1)$ zum Zeitpunkt t_1 kleiner oder höchstens gleich x_1 ist und der Funktionswert von $x(t_2)$ zum Zeitpunkt t_2 kleiner oder höchstens gleich x_2 ist. \diamond

Definition 6.7: Verbundwahrscheinlichkeitsdichte eines Zufallsprozesses

Die Verbundwahrscheinlichkeitsdichte (oder gemeinsame Wahrscheinlichkeitsdichte) eines stochastischen Prozesses $x(t)$ wird entsprechend

$$f_{x(t_1)x(t_2)}(x_1, x_2) = \frac{\partial^2 F_{x(t_1)x(t_2)}(x_1, x_2)}{\partial x_1 \partial x_2} \quad (6.8)$$

definiert. \diamond

Der in Abschn. 4.1 eingeführte Begriff der statistischen Unabhängigkeit zweier Zufallsvariabler kann ebenfalls auf stochastische Prozesse angewandt werden.

Definition 6.8: Statistische Unabhängigkeit

Zwei stochastische Prozesse $x(t)$ und $y(t)$ nennt man statistisch unabhängig, wenn für alle Zeiten t_1, t_2 gilt:

$$F_{x(t_1)y(t_2)}(x, y) = F_{x(t_1)}(x) \cdot F_{y(t_2)}(y) \quad (6.9)$$

bzw.

$$f_{x(t_1)y(t_2)}(x, y) = f_{x(t_1)}(x) \cdot f_{y(t_2)}(y). \quad (6.10)$$

◊

Statistische Unabhängigkeit ist eine Eigenschaft, die experimentell höchstens näherungsweise nachgewiesen werden kann. Bei der Formulierung eines Modells für ein Messsystem kann sie meist nur als Voraussetzung angenommen werden. Diese Voraussetzung ist in der Regel berechtigt, wenn die stochastischen Signale unterschiedliche Ursachen haben. Der große Vorteil der statistischen Unabhängigkeit liegt in der wesentlich vereinfachten modellgestützten Analyse. Man wird daher oft statistische Unabhängigkeit annehmen, selbst wenn die Quellen der stochastischen Signale nicht völlig unabhängig voneinander sind, vorhandene Abhängigkeiten aber nicht interessieren.

6.1.3 Schar- und Zeitmittelwerte

Stochastische Prozesse $y(t, \xi)$ sind Funktionen zweier Parameter t und ξ . Bei der Beschreibung zufälliger Signale dominiert meist der Zeitparameter t . In der Literatur wird daher sehr oft nur dieser Parameter explizit angegeben. Bei der Bildung von Mittelwerten gibt es jedoch, entsprechend den beiden Parametern, zwei mögliche Vorgehensweisen (Abb. 6.2):

1. *Mittelwertbildung über alle Musterfunktionen $y(t, \xi_1), y(t, \xi_2), \dots$ bei festem t :*
Diesen Mittelwert nennt man **Erwartungswert** oder **Scharmittelwert**. Er gibt an, welchen Wert der Zufallsprozess zu einem Zeitpunkt t im Mittel annimmt und hängt im Allgemeinen von der Zeit t ab.
2. *Mittelwertbildung über den Parameter t einer einzigen Musterfunktion $y(t, \xi_\nu)$:*
Diesen Mittelwert nennt man **Zeitmittelwert**. Im Allgemeinen sind die Zeitmittelwerte einzelner Musterfunktionen eines Zufallsprozesses verschieden und damit vom Parameter ξ_ν abhängig.

Im Allgemeinen liefern beide Vorgehensweisen verschiedene Ergebnisse.

6.1.4 Momente der Statistik 1. Ordnung

Erwartungswerte stochastischer Prozesse sind immer Erwartungswerte über alle Musterfunktionen $y(t, \xi_1), y(t, \xi_2), \dots$ zu einem festen Zeitpunkt $t = t_i$. Entsprechend der Momentendefinition für Zufallsvariable sind die Momente stochastischer Prozesse um den Parameter t erweitert.

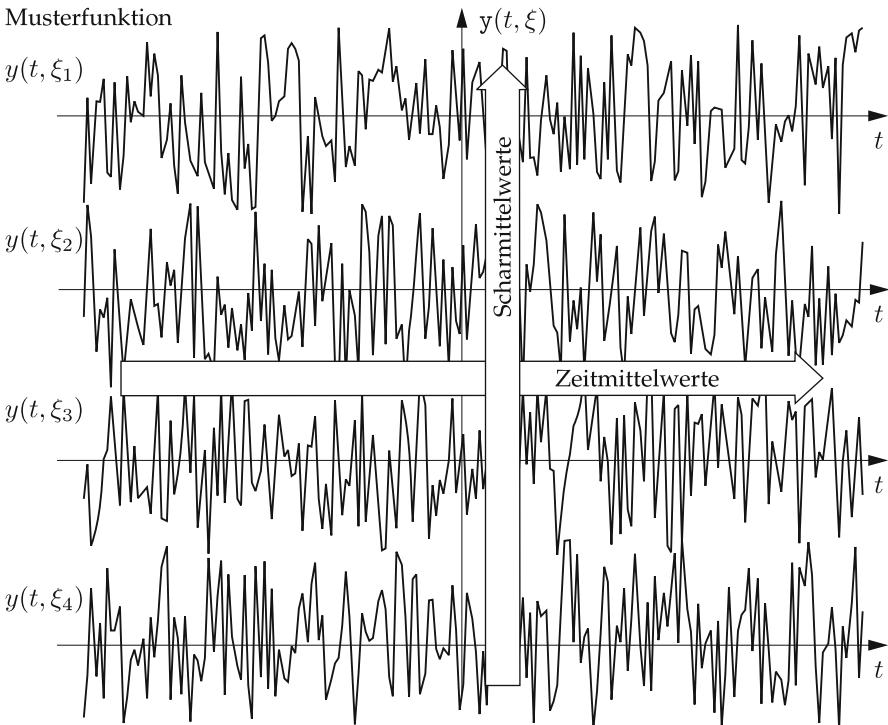


Abbildung 6.2. Mittelwertbildung als Zeitmittelwert oder Scharmittelwert.

Definition 6.9: Moment eines stochastischen Prozesses

6.9

Das m -te Moment eines stochastischen Prozesses $y(t)$ ist definiert zu

$$\mu_{y,m}(t) = E\{y^m(t)\} = \int_{-\infty}^{\infty} y^m f_{y(t)}(y) dy . \quad (6.11)$$

Das erste Moment $\mu_y(t) = \mu_{y,1}(t)$ ist der zeitabhängige Mittelwert oder auch Scharmittelwert. \diamond

Definition 6.10: Zentrales Moment eines stochastischen Prozesses

6.10

Das m -te zentrale Moment eines stochastischen Prozesses $y(t)$ ist definiert zu

$$E\{(y(t) - \mu_y(t))^m\} = \int_{-\infty}^{\infty} (y - \mu_y(t))^m f_{y(t)}(y) dy \quad (6.12)$$

mit $\mu_y(t) = E\{y(t)\}$. ◊

Das zweite zentrale Moment ist die zeitabhängige Varianz $\sigma_y^2(t)$.

Es ist einleuchtend, dass in den meisten Fällen die Erwartungswertbildung über alle Musterfunktionen praktisch nicht durchführbar ist. Selbst mit einer begrenzten Zahl an Musterfunktionen zu arbeiten, ist in der Praxis mit einem deutlich höheren Aufwand verbunden, benötigt man doch den parallelen Einsatz mehrerer identischer Systeme, um verschiedene Musterfunktionen desselben stochastischen Prozesses zu erzeugen. Unter bestimmten Voraussetzungen, bei den sogenannten ergodischen Prozessen (Abschn. 6.1.7), können jedoch Scharmittelwerte durch Zeitmittelwerte ersetzt werden. Dies erleichtert die messtechnische Erfassung der Momente ungemein.

❷ 6.1.5 Momente der Statistik 2. Ordnung

Zu den Momenten der Statistik *zweiter* Ordnung gelangt man, indem man Zufallsprozesse zu *zwei* Zeitpunkten t_1 und t_2 betrachtet (vgl. Def. 4.9). Bei Beschränkung auf einen einzigen stochastischen Prozess $y(t)$ erhält man als wichtigste Momente der Statistik 2. Ordnung die **Autokorrelationsfunktion** (zweites Moment der Statistik 2. Ordnung) und die **Autokovarianzfunktion** (zweites zentrales Moment der Statistik 2. Ordnung).

6.11

Definition 6.11: Autokorrelationsfunktion

Die Autokorrelationsfunktion (AKF) eines stochastischen Prozesses $y(t)$ ist als

$$r_{yy}(t_1, t_2) = E\{y(t_1) y^*(t_2)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y_1 y_2^* f_{y(t_1)y(t_2)}(y_1, y_2) dy_1 dy_2 \quad (6.13)$$

definiert. ◊

6.12

Definition 6.12: Autokovarianzfunktion

Die Autokovarianzfunktion (AKV) eines Zufallsprozesses $y(t)$ ist definiert als

$$\begin{aligned} C_{yy}(t_1, t_2) &= E\{(y(t_1) - \mu_y(t_1))(y(t_2) - \mu_y(t_2))^*\} \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (y_1 - \mu_y(t_1))(y_2 - \mu_y(t_2))^* f_{y(t_1)y(t_2)}(y_1, y_2) dy_1 dy_2 \end{aligned} \quad (6.14)$$

mit $\mu_y(t) = E\{y(t)\}$. ◊

Die Autokorrelationsfunktion unterscheidet sich von der Autokovarianzfunktion nur um das Produkt der Mittelwerte:

$$C_{yy}(t_1, t_2) = r_{yy}(t_1, t_2) - \mu_y(t_1) \mu_y^*(t_2). \quad (6.15)$$

Betrachtet man zwei verschiedene stochastische Prozesse zu zwei Zeitpunkten t_1 und t_2 , so erhält man die **Kreuzkorrelationsfunktion** und die **Kreuzkovarianzfunktion** als die entsprechenden gemeinsamen Momente der Statistik 2. Ordnung.

Definition 6.13: Kreuzkorrelationsfunktion

6.13

Die Kreuzkorrelationsfunktion (KKF) zweier stochastischer Prozesse $x(t)$ und $y(t)$ ist durch

$$r_{xy}(t_1, t_2) = E\{x(t_1) y^*(t_2)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x y^* f_{x(t_1)y(t_2)}(x, y) dx dy \quad (6.16)$$

gegeben. ◊

Definition 6.14: Kreuzkovarianzfunktion

6.14

Die Kreuzkovarianzfunktion (KKV) zweier Zufallsprozesse $x(t)$ und $y(t)$ ist definiert als

$$C_{xy}(t_1, t_2) = E\{(x(t_1) - \mu_x(t_1))(y(t_2) - \mu_y(t_2))^*\} \quad (6.17)$$

$$= \iint_{-\infty}^{\infty} (x - \mu_x(t_1))(y - \mu_y(t_2))^* f_{x(t_1)y(t_2)}(x, y) dx dy \quad (6.18)$$

mit $\mu_x(t_1) = E\{x(t_1)\}$ und $\mu_y(t_2) = E\{y(t_2)\}$. ◊

Durch die Betrachtung von Prozessen zu verschiedenen Zeitpunkten ist es möglich, Aussagen über die „Zufälligkeit“ des Prozesses zu machen (Erhaltungstendenz). Auch die Ähnlichkeit zweier zufälliger Signale zu unterschiedlichen Zeiten kann mit Hilfe der Momente der Statistik 2. Ordnung untersucht werden. Deshalb haben Korrelationsfunktionen in der Praxis eine große Bedeutung. Ihre Eigenschaften und Anwendungen werden daher in den Abschnitten 6.2 bzw. 6.3 gesondert besprochen.

6.15

Definition 6.15: Unkorreliertheit

Zwei Zufallsprozesse $x(t)$ und $y(t)$ heißen **unkorreliert**, wenn für alle Zeiten t_1 und t_2 der Zusammenhang

$$E\{x(t_1)y^*(t_2)\} = E\{x(t_1)\} \cdot E\{y^*(t_2)\} \quad (6.19)$$

$$\Leftrightarrow C_{xy}(t_1, t_2) = 0 \quad (6.20)$$

gilt. \diamond

Sind zwei Prozesse nach Def. 6.8 statistisch unabhängig, so folgt daraus *immer* ihre Unkorreliertheit. Die Umkehrung des Satzes ist jedoch allgemein *nicht* zugelassen. Sie gilt nur für den Spezialfall normalverteilter Prozesse! Unkorreliertheit ist somit – verglichen mit statistischer Unabhängigkeit – die schwächere Eigenschaft. Physikalisch drücken beide Eigenschaften fehlende Kopplungen zwischen den Prozessen aus, die sich bei Unkorreliertheit jedoch nur auf den Fall einer linearen statistischen Abhängigkeit beziehen.

6.16

Definition 6.16: Orthogonalität

Zwei Zufallsprozesse $x(t)$ und $y(t)$ heißen zueinander **orthogonal**, wenn die Bedingung

$$r_{xy}(t_1, t_2) = E\{x(t_1)y^*(t_2)\} = 0 \quad (6.21)$$

für alle Zeiten t_1 und t_2 erfüllt ist. \diamond

6.1.6 Stationäre Prozesse

Allgemein sind die Wahrscheinlichkeitsverteilungen und -dichten eines stochastischen Prozesses zeitabhängig. Dies hat zur Folge, dass auch die zugehörigen Momente Funktionen der Zeit t sind. Die Korrelationsfunktion und die Kovarianzfunktion sind sogar zweidimensionale Funktionen der Zeiten t_1 und t_2 . Allerdings treten wesentliche Vereinfachungen ein, wenn sich die statistischen Eigenschaften eines Prozesses bei Verschiebung der Zeitachse nicht ändern. Dies ist bei **stationären Prozessen** der Fall.

6.17

Definition 6.17: Stationarität

Ein stochastischer Prozess heißt (streng) stationär, wenn seine statistischen Eigenschaften invariant gegenüber Verschiebungen der Zeit sind. \diamond

Definition 6.18: Verbundene Stationarität zweier Zufallsprozesse

Zwei stochastische Prozesse heißen verbunden stationär, wenn beide stationär sind und ferner ihre gemeinsamen statistischen Eigenschaften invariant gegenüber Verschiebungen der Zeit sind. \diamond

Aus Def. 6.17 folgt, dass für (streng) stationäre Prozesse die Zeitabhängigkeit der Wahrscheinlichkeitsdichten und die der Momente verschwinden. Für die Wahrscheinlichkeitsverteilung eines streng stationären Zufallsprozesses gilt:

$$F_{y(t)}(y) = F_{y(t+t_0)}(y) = F_y(y) \quad (6.22)$$

und für die Wahrscheinlichkeitsdichten entsprechend:

$$f_{y(t)}(y) = f_{y(t+t_0)}(y) = f_y(y). \quad (6.23)$$

Abhängigkeiten von zwei Zeitpunkten t_1 und t_2 vereinfachen sich bei Stationarität auf die Abhängigkeit von der Differenz $\tau = t_1 - t_2$ dieser Zeitpunkte:

$$f_{y(t_1)y(t_2)}(y_1, y_2) = f_{y(t_1+t_0)y(t_2+t_0)}(y_1, y_2) = f_{y(t+\tau)y(t)}(y_1, y_2). \quad (6.24)$$

Für zwei verbunden stationäre Prozesse ergibt sich zusätzlich:

$$f_{x(t_1)y(t_2)}(x, y) = f_{x(t_1+t_0)y(t_2+t_0)}(x, y) = f_{x(t+\tau)y(t)}(x, y). \quad (6.25)$$

Insbesondere sind bei Stationarität die m -ten Momente der Statistik 1. Ordnung unabhängig von der Zeit t :

$$\mu_{y,m}(t) = E\{y^m(t)\} = \mu_{y,m}. \quad (6.26)$$

Die Momente der Statistik 2. Ordnung sind nur noch abhängig von der Zeitdifferenz τ :

$$r_{yy}(t_1, t_2) = E\{y(t+\tau) y^*(t)\} = r_{yy}(\tau), \quad (6.27)$$

$$r_{xy}(t_1, t_2) = E\{x(t+\tau) y^*(t)\} = r_{xy}(\tau). \quad (6.28)$$

Definition 6.19: Schwache Stationarität

Ein stochastischer Prozess $y(t)$ heißt schwach stationär, wenn dessen Erwar-

tungswert $\mu_y(t)$ und dessen Autokorrelationsfunktion¹ $r_{yy}(t_1, t_2)$ invariant gegenüber Verschiebungen der Zeit sind, also wenn

$$\mu_y(t) = \mu_y = \text{const.} \quad (6.29)$$

und

$$r_{yy}(t_1, t_2) = r_{yy}(t_1 - t_2) = r_{yy}(\tau) \quad (6.30)$$

gilt. \diamond

Aus Def. 6.19 folgt, dass bei schwacher Stationarität die Wahrscheinlichkeitsverteilung und die Wahrscheinlichkeitsdichte selbst nicht notwendigerweise invariant gegenüber Zeitverschiebungen sein müssen. Allerdings sind bei einem schwach stationären Prozess neben dem Erwartungswert μ_y und der Autokorrelationsfunktion $r_{yy}(\tau)$ nach (6.15) auch die Autokovarianzfunktion

$$C_{yy}(\tau) = r_{yy}(\tau) - |\mu_y|^2 \quad (6.31)$$

und die Varianz

$$\sigma_y^2 = C_{yy}(0) = r_{yy}(0) - |\mu_y|^2 \quad (6.32)$$

unabhängig vom absoluten Zeitpunkt.

Anmerkung

Streng stationäre Prozesse sind stets schwach stationär, aber die Umkehrung dieser Regel ist im Allgemeinen nicht zulässig. Eine Ausnahme bilden jedoch die normalverteilten Prozesse. Da diese durch den Erwartungswert und die Autokovarianzfunktion vollständig charakterisiert sind, folgt bei diesen Prozessen aus der schwachen Stationarität automatisch die strenge Stationarität.

6.1.7 Ergodische Prozesse

Die Bestimmung der Momente eines Zufallsprozesses erfolgt allgemein nach Abschn. 6.1.4 für einen festen Zeitpunkt t durch Integration über die Schar aller möglichen Musterfunktionen $y(t, \xi_\nu)$, was im Allgemeinen zu zeitabhängigen Momenten $\mu_{y,m}(t)$ führt. Oftmals sind jedoch nicht sämtliche Musterfunktionen eines Zufallsprozesses bekannt, weshalb eine Mitte-

¹ Manche Autoren fordern alternativ, dass die Autokovarianzfunktion invariant gegenüber Verschiebungen der Zeit sein muss. Wegen (6.15) sind bei konstantem Erwartungswert $\mu_y(t) = \mu_y$ beide Forderungen äquivalent.

lung über die gesamte Schar nicht in Frage kommt. Durch Integration über alle Zeiten t bei einer festen Musterfunktion $y(t, \xi_0)$ ergeben sich hingegen Zeitmittelwerte, die im Allgemeinen für jede Musterfunktion verschieden sind. Es gibt jedoch eine für die Anwendung wichtige Klasse stationärer Zufallsprozesse, bei denen Scharmittelwerte und Zeitmittelwerte identische Ergebnisse liefern. Derartige Prozesse nennt man **ergodisch**.

Definition 6.20: Ergodizität

6.20

Ein Zufallsprozess $y(t)$ heißt (streng) ergodisch, wenn die Zeitmittelwerte einer beliebigen Musterfunktion $y(t, \xi_\nu)$ mit der Wahrscheinlichkeit eins mit den entsprechenden Scharmittelwerten des Prozesses übereinstimmen. ◇

Ergodische Prozesse sind stets stationär, weshalb Stationarität in jedem Fall eine Voraussetzung für Ergodizität ist. Dies geht schon daraus hervor, dass die Momente instationärer Prozesse zeitabhängig sind und somit nicht für alle Zeiten mit den zeitunabhängigen Zeitmittelwerten übereinstimmen können.

Definition 6.21: Schwache Ergodizität

6.21

Ein Zufallsprozess $y(t)$ heißt schwach ergodisch, wenn die anhand einer beliebigen Musterfunktion $y(t, \xi_\nu)$ berechneten Zeitmittelwerte für das erste Moment und für die Autokorrelationsfunktion mit den entsprechenden Scharmittelwerten mit der Wahrscheinlichkeit eins übereinstimmen, also wenn die Bedingungen

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T y(t, \xi_\nu) dt = E\{y(t)\} = \mu_y \quad (6.33)$$

und

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T y(t + \tau, \xi_\nu) y^*(t, \xi_\nu) dt = E\{y(t + \tau) y^*(t)\} = r_{yy}(\tau) \quad (6.34)$$

erfüllt sind. Die linken Seiten der Gleichungen (6.33) und (6.34) stellen eine Integration *einer* beliebigen Musterfunktion $y(t, \xi_\nu)$ über t dar, also eine Zeitmittelung. In den rechten Seiten dieser Gleichungen erfolgt durch die Erwartungswertbildung eine Mittelung über die Schar aller Musterfunktionen für einen festen Zeitpunkt t . ◇

Praktisch heißt das, dass bei einem schwach ergodischen Prozess jede Musterfunktion bezüglich der ersten beiden Momente der Statistik 1. und 2. Ordnung die Schar – d. h. den ganzen Prozess – vertreten kann.

Für die praktische Anwendung bringt die Ergodizität große Vorteile mit sich. Es ist damit möglich, sich auf die Untersuchung einer einzigen Musterfunktion zu beschränken. Ihr zeitlicher Verlauf beinhaltet alle statistischen Eigenschaften des zugrunde liegenden Zufallsprozesses. Allerdings lässt sich der mathematisch strenge Nachweis der Ergodizität höchstens in Sonderfällen erbringen. In den meisten Anwendungen kann dagegen die Ergodizität des Zufallsprozesses nur postuliert werden.

Unter der Annahme einer festen Musterfunktion $y(t, \xi_0)$ gelten für die Berechnung der Momente eines ergodischen Prozesses folgende Regeln:

6.1

Satz 6.1: Moment eines ergodischen Zufallsprozesses

Das m -te Moment eines ergodischen Zufallsprozesses $y(t)$ berechnet sich als Zeitmittelwert

$$\mu_{y,m} = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T y^m(t, \xi_0) dt \quad (6.35)$$

einer beliebigen Musterfunktion $y(t, \xi_0)$. \diamond

6.2

Satz 6.2: Kreuzkorrelationsfunktion zweier ergodischer Zufallsprozesse

Die Kreuzkorrelationsfunktion zweier ergodischer Zufallsprozesse $x(t)$ und $y(t)$ mit den Musterfunktionen $x(t, \xi_0)$ bzw. $y(t, \xi_1)$ berechnet sich entsprechend zu

$$r_{xy}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x(t + \tau, \xi_0) y^*(t, \xi_1) dt. \quad (6.36)$$

Für $x(t, \xi_0) = y(t, \xi_1)$ resultiert aus dieser Gleichung die Autokorrelationsfunktion $r_{xx}(\tau)$. \diamond

6.3

Satz 6.3: Kreuzkovarianzfunktion zweier ergodischer Zufallsprozesse

Die Kreuzkovarianzfunktion zweier ergodischer Zufallsprozesse $x(t)$ und

$y(t)$ mit den Musterfunktionen $x(t, \xi_0)$ bzw. $y(t, \xi_1)$ berechnet sich zu

$$C_{xy}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T (x(t + \tau, \xi_0) - \mu_x)(y(t, \xi_1) - \mu_y)^* dt. \quad (6.37)$$

Für $x(t, \xi_0) = y(t, \xi_1)$ ergibt sich aus dieser Gleichung die Autokovarianzfunktion $C_{xx}(\tau)$. \diamond

Anmerkung

Natürlich kann in praktischen Anwendungen der Grenzübergang $T \rightarrow \infty$ nicht durchgeführt werden. Bei der Bestimmung der Momente ist man daher bei der Auswertung der Integrale auf ein endliches Zeitintervall $[-T, T]$ beschränkt. Entsprechend erhält man lediglich Schätzwerte für die Momente des Prozesses. So wird beispielsweise der Mittelwert durch folgenden Ausdruck geschätzt:

$$\hat{\mu}_y(t) = \overline{y(t)} = \frac{1}{2T} \int_{t-T}^{t+T} y(t) dt. \quad (6.38)$$

Man nennt ihn **gleitenden Mittelwert** (engl. *moving average*). Er ist – im Gegensatz zum wahren Mittelwert μ_y eines schwach ergodischen Zufallsprozesses – im Allgemeinen zeitabhängig.

Beispiel 6.3 (Schwach stationärer, ergodischer Prozess): Gegeben sei der Zufallsprozess $y(t)$ aus Bsp. 6.2 mit den Musterfunktionen

$$y_i(t) = \sin(\omega t + \varphi_i), \quad (6.39)$$

wobei die Kreisfrequenz ω fest, die Phase φ_i jedoch zufällig sei. Die Phasenwinkel seien gleichverteilt. Damit der Prozess schwach stationär ist, müssen dessen Erwartungswert und dessen Autokorrelationsfunktion unabhängig von t sein. Für den Mittelwert $\mu_y(t)$ lässt sich dies leicht nachweisen. Die Autokorrelationsfunktion $r_{yy}(t_1, t_2)$ ist allgemein als Scharmittelwert definiert. Mit dem Zusammenhang $\sin \alpha \cdot \sin \beta = \frac{1}{2}(\cos(\alpha - \beta) - \cos(\alpha + \beta))$ folgt:

$$r_{yy}(t_1, t_2) = E\{\sin(\omega t_1 + \varphi) \sin(\omega t_2 + \varphi)\} \quad (6.40)$$

$$= \frac{1}{2\pi} \int_0^{2\pi} \sin(\omega t_1 + \varphi) \sin(\omega t_2 + \varphi) d\varphi \quad (6.41)$$

$$= \frac{1}{2\pi} \int_0^{2\pi} \frac{1}{2} \left(\cos(\omega t_1 - \omega t_2) - \underbrace{\cos(\omega t_1 + \omega t_2 + 2\varphi)}_{\int d\varphi \rightarrow 0} \right) d\varphi \quad (6.42)$$

$$= \frac{1}{2} \cos(\omega t_1 - \omega t_2) = \frac{1}{2} \cos(\omega\tau) \quad \text{mit} \quad \tau = t_1 - t_2. \quad (6.43)$$

Da die Autokorrelationsfunktion $r_{yy}(\tau)$ nur von der Differenz τ der Beobachtungszeitpunkte abhängt, ist der Prozess schwach stationär. Damit der Prozess darüber hinaus auch schwach ergodisch ist, müssen Scharmittelwerte und Zeitmittelwerte übereinstimmen. Für die Autokorrelationsfunktion eines ergodischen Prozesses erhält man nach Satz 6.2:

$$\overline{y_i(t + \tau) y_i(t)} = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \sin(\omega t + \omega\tau + \varphi) \sin(\omega t + \varphi) dt \quad (6.44)$$

$$= \lim_{T \rightarrow \infty} \frac{1}{4T} \left[\int_{-T}^T \cos(\omega\tau) dt - \int_{-T}^T \cos(2\omega t + \omega\tau + 2\varphi) dt \right] \quad (6.45)$$

$$= \lim_{T \rightarrow \infty} \frac{1}{4T} \left[2T \cos(\omega\tau) - \frac{1}{2\omega} \sin(2\omega t + \omega\tau + 2\varphi) \Big|_{-T}^T \right] \quad (6.46)$$

$$= \frac{1}{2} \cos(\omega\tau) - \underbrace{\lim_{T \rightarrow \infty} \frac{1}{4T\omega} [\sin(2\omega T + \omega\tau + 2\varphi)]}_{=0} \quad (6.47)$$

$$= \frac{1}{2} \cos(\omega\tau) = r_{yy}(\tau). \quad (6.48)$$

Schar- und Zeitmittelwerte stimmen überein, weshalb der Zufallsprozess ebenfalls schwach ergodisch ist. ■

Beispiel 6.4 (Schwach stationärer, nicht ergodischer Prozess): Gegeben sei der Zufallsprozess $y(t)$ mit folgenden Musterfunktionen

$$y_i(t) = a_i \sin(\omega t + \varphi_i), \quad (6.49)$$

wobei die Kreisfrequenz ω fest, die Phase φ und die Amplitude a jedoch zufällig und voneinander statistisch unabhängig seien. Die Phasenwinkel und Amplituden seien gleichverteilt. Aus dem vorherigen Beispiel

ergibt sich die Korrelationsfunktion zu

$$r_{yy}(t_1, t_2) = E\{\mathbf{a} \sin(\omega t_1 + \varphi) \mathbf{a} \sin(\omega t_2 + \varphi)\} \quad (6.50)$$

$$= E\{\mathbf{a}^2\} E\{\sin(\omega t_1 + \varphi) \sin(\omega t_2 + \varphi)\} \quad (6.51)$$

$$= E\{\mathbf{a}^2\} \frac{1}{2} \cos(\omega t_1 - \omega t_2) \quad (6.52)$$

$$= E\{\mathbf{a}^2\} \frac{1}{2} \cos(\omega \tau) \quad \text{mit} \quad \tau = t_1 - t_2. \quad (6.53)$$

Der Prozess ist schwach stationär, da die Autokorrelationsfunktion $r_{yy}(\tau)$ unabhängig von t ist. Führt man die Berechnung der Autokorrelationsfunktion durch Zeitmittelung ausgehend von einer einzigen Musterfunktion durch, so erhält man analog zum obigen Beispiel:

$$\overline{y_i(t + \tau) y_i(t)} = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T a_i^2 \sin(\omega t + \omega \tau + \varphi) \sin(\omega t + \varphi) dt \quad (6.54)$$

$$= a_i^2 \frac{1}{2} \cos(\omega \tau) \neq r_{yy}(\tau). \quad (6.55)$$

Der Prozess ist nicht ergodisch, da sich Zeit- und Scharmittelwerte unterscheiden. ■

6.2 Korrelationsfunktionen

6.2

6.2.1 Signalklassen

Mit Hilfe der Korrelation kann die Erhaltungstendenz von stochastischen Prozessen oder deren Musterfunktionen beschrieben werden. Damit gelingt es, zwei zunächst völlig regellose – insbesondere zufällige – Signale miteinander zu vergleichen. Es soll nun untersucht werden, wie man die Korrelation für die verschiedenen Signalklassen berechnen kann.

Ausgehend von einem allgemeinen stochastischen Prozess, der beliebige Zeitfunktionen als Musterfunktionen beinhalten kann, sollen die Musterfunktionen in drei Klassen gegliedert werden. Dabei soll zunächst angenommen werden, dass die Funktionen im Zeitintervall $(-\infty, \infty)$ definiert sind.

6.22

Definition 6.22: Energiesignal

Ein beschränktes, stückweise stetiges Signal $x(t)$, für das

$$\int_{-\infty}^{\infty} x(t) x^*(t) dt = \int_{-\infty}^{\infty} |x(t)|^2 dt < \infty \quad (6.56)$$

gilt, nennt man **Energiesignal**. ◊

Die Bezeichnung „Energiesignal“ kommt von der physikalischen Interpretation, da das Integral (6.56) als **Energie** interpretiert werden kann. Damit die Konvergenz des Integrals gesichert ist, müssen Energiesignale für betragsmäßig große Zeiten verschwinden:

$$\lim_{t \rightarrow \pm\infty} x(t) = 0. \quad (6.57)$$

Typische Vertreter von Energiesignalen sind die Impulsantworten stabiler LTI-Systeme sowie **Fensterfunktionen** [42].

6.23

Definition 6.23: Leistungssignal

Ein beschränktes, stückweise stetiges Signal $x(t)$, für welches das Integral

$$\int_{-\infty}^{\infty} x(t) x^*(t) dt = \int_{-\infty}^{\infty} |x(t)|^2 dt \quad (6.58)$$

divergiert (unendliche Energie), aber der Grenzwert

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x(t) x^*(t) dt = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T |x(t)|^2 dt < \infty \quad (6.59)$$

existiert, nennt man **Leistungssignal**. ◊

Der Grenzwert lässt sich physikalisch als **mittlere Leistung** interpretieren. Bei Energiesignalen ist die mittlere Leistung dagegen null. Als Beispiel von Leistungssignalen sei die Klasse der amplitudenbeschränkten periodischen Signale genannt.

Definition 6.24: Sonstige Signale

Alle Zeitfunktionen, für welche die Integrale

$$\int_{-\infty}^{\infty} x(t) x^*(t) dt \quad \text{und} \quad \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^{T} x(t) x^*(t) dt \quad (6.60)$$

nicht existieren, welche nicht stückweise stetig oder nicht beschränkt sind, werden als **sonstige Signale** klassifiziert. ◇

In die Klasse der sonstigen Signale fallen zwar viele Funktionen, jedoch haben diese nur eine geringe praktische Relevanz. Im weiteren Verlauf werden lediglich Energie- und Leistungssignale betrachtet, da für diese die Analysemethoden der Systemtheorie anwendbar sind.

④ **6.2.1.1 Ergodische Prozesse**

Nach der Unterteilung in Signalklassen soll kurz untersucht werden, welche Signale Musterfunktionen ergodischer Prozesse sein können. Nach Bsp. 6.3 ist eine harmonische Schwingung mit zufälliger Phase ein ergodischer Prozess. Daraus schließt man, dass Leistungssignale Musterfunktionen ergodischer Prozesse sein können. Die Bildung von Zeitmittelwerten ist also erlaubt.

Energiesignale hingegen können *keine* Musterfunktionen ergodischer Prozesse sein, da sie nicht stationär sind:

$$\lim_{t \rightarrow \pm\infty} x(t) = 0 \quad \Rightarrow \quad \lim_{t \rightarrow \pm\infty} \mu_{x,m}(t) = 0. \quad (6.61)$$

Da nach Voraussetzung Energiesignale für $t \rightarrow \infty$ verschwinden, verschwinden auch deren Momente der Statistik 1. Ordnung für $t \rightarrow \infty$. Damit ist die Zeitunabhängigkeit der Momente für Energiesignale im Allgemeinen verletzt.

④ **6.2.1.2 Funktionenräume**

Für Energie- und Leistungssignale lassen sich jeweils **unitäre Funktionenräume** mit Norm und Innenprodukt definieren, wenn man als Norm $\|x(t)\|$ die Wurzel des jeweils konvergierenden Integralausdrucks setzt [42].

6.25

Definition 6.25: Norm und Innenprodukt von EnergiesignalenDie Norm eines Energiesignals $x(t)$ ist

$$\|x(t)\| = \sqrt{\int_{-\infty}^{\infty} x(t) x^*(t) dt} < \infty. \quad (6.62)$$

Das Innenprodukt zweier Energiesignale $x(t)$ und $y(t)$ ist durch das Integral

$$\langle x(t), y(t) \rangle = \int_{-\infty}^{\infty} x(t) y^*(t) dt \quad (6.63)$$

definiert. ◊

6.26

Definition 6.26: Norm und Innenprodukt von LeistungssignalenDie Norm eines Leistungssignals $x(t)$ ist

$$\|x(t)\| = \sqrt{\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^{T} x(t) x^*(t) dt} < \infty. \quad (6.64)$$

Das Innenprodukt zweier Leistungssignale $x(t)$ und $y(t)$ wird durch

$$\langle x(t), y(t) \rangle = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^{T} x(t) y^*(t) dt \quad (6.65)$$

definiert. ◊In beiden Funktionenräumen gilt die **Schwarz'sche Ungleichung**, die hier ohne Beweis angegeben wird [42].

6.4

Satz 6.4: Schwarz'sche Ungleichung

In unitären Funktionenräumen gilt die Schwarz'sche Ungleichung in der folgenden Form:

$$|\langle x(t), y(t) \rangle|^2 \leq \|x(t)\|^2 \cdot \|y(t)\|^2. \quad (6.66)$$

In (6.66) gilt das Gleichheitszeichen, wenn die Signale $x(t)$ und $y(t)$ linear abhängig sind [42]. ◊

In der praktischen Anwendung ist ein Zeitintervall $(-\infty, \infty)$ zur Messung von Signalen nicht möglich. Man wird es daher *immer* mit zeitbegrenzten Funktionen zu tun haben. Im endlichen Zeitintervall $[t_1, t_2]$ konvergiert das Integral

$$\int_{t_1}^{t_2} x(t) x^*(t) dt < \infty \quad (6.67)$$

für beschränkte, stückweise stetige Funktionen immer. Bei der späteren Signalanalyse geht man allerdings vielfach implizit von einer periodischen Fortsetzung des Signals außerhalb des Zeitintervalls über alle Zeiten aus (vgl. Fourier-Reihe, DFT). Trotz einer messtechnisch bedingten Begrenzung des Signals geht man daher oftmals von Leistungssignalen aus. Die Tabelle 6.2 gibt einen Überblick über die möglichen Verfahren zur Signalanalyse für die verschiedenen Signalklassen.

Tabelle 6.2. Signalklassen und Verfahren (\otimes : zulässig, $[\otimes]_{\text{Dist}}$: eingeschränkt zulässig unter Verwendung der Distributionstheorie, $*)$: Korrelation nicht im statistischen Sinn).

| | Energiesignale | Leistungssignale | Sonstige |
|-------------------------------------|----------------|---------------------------|-----------|
| Stochastik | \otimes | \otimes | \otimes |
| Ergodizität | — | \otimes | — |
| Korrelation mit Zeitfunktion | $\otimes^*)$ | \otimes | — |
| Fourier-Reihe | — | \otimes | — |
| Fourier-Transformation | \otimes | $[\otimes]_{\text{Dist}}$ | — |
| zeitdiskrete Fourier-Transformation | \otimes | $[\otimes]_{\text{Dist}}$ | — |
| diskrete Fourier-Transformation | \otimes | \otimes | — |

6.2.2 Korrelation von Leistungssignalen

Für Leistungssignale ist die Kreuzkorrelationsfunktion über das Innenprodukt definiert:

$$r_{xy}(\tau) = \langle x(t + \tau), y(t) \rangle = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x(t + \tau) y^*(t) dt. \quad (6.68)$$

Diese Definition ist identisch mit der Berechnung der Korrelationsfunktion ergodischer Prozesse nach Satz 6.2. Dies ist auch nicht weiter verwunderlich, da zuvor festgestellt wurde, dass Leistungssignale Musterfunktionen ergodischer Prozesse sein können. Da es sich jedoch bei den Signalen $x(t)$ und $y(t)$ nicht zwangsläufig um Musterfunktionen stochastischer Prozesse

handeln muss, tauchen in (6.68) – anders als in (6.36) – die Elementarereignisse ξ_0 und ξ_1 nicht auf.

Durch Einsetzen des Innenproduktes (6.68) in die Schwarz'sche Ungleichung (6.66)

$$|\langle x(t + \tau), y(t) \rangle|^2 \leq \|x(t + \tau)\|^2 \cdot \|y(t)\|^2 \quad (6.69)$$

erhält man:

$$\begin{aligned} \left| \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x(t + \tau) y^*(t) dt \right|^2 &\leq \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x(t + \tau) x^*(t + \tau) dt \\ &\times \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T y(t) y^*(t) dt, \end{aligned} \quad (6.70)$$

$$|r_{xy}(\tau)|^2 \leq P_x \cdot P_y < \infty. \quad (6.71)$$

Das Betragsquadrat des Innenproduktes der Leistungssignale $x(t)$ und $y(t)$ ist somit kleiner gleich dem Produkt der **Signalleistungen** P_x und P_y . Die Leistung eines Signals folgt aus (6.68) für $y(t) = x(t)$ und $\tau = 0$:

$$r_{xx}(0) = \langle x(t), x(t) \rangle = \|x(t)\|^2 = P_x. \quad (6.72)$$

Beispiel 6.5 (Korrelationsfunktion zweier Leistungssignale): Gegeben seien die beiden Signale

$$x(t) = \sin(2\pi f_0 t) \quad \text{und} \quad y(t) = \cos(2\pi f_0 t). \quad (6.73)$$

Ihre Kreuzkorrelationsfunktion ist nach (6.68)

$$r_{xy}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x(t + \tau) y^*(t) dt \quad (6.74)$$

$$= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \sin(2\pi f_0(t + \tau)) \cos(2\pi f_0 t) dt \quad (6.75)$$

$$\begin{aligned} &= \lim_{T \rightarrow \infty} \left[\cos(2\pi f_0 \tau) \left[\frac{1}{8\pi f_0 T} \sin^2(2\pi f_0 t) \right]_{-T}^T \right. \\ &\quad \left. + \sin(2\pi f_0 \tau) \left[\frac{t}{4T} + \frac{1}{16\pi f_0 T} \sin(4\pi f_0 t) \right]_{-T}^T \right] \end{aligned} \quad (6.76)$$

$$= \frac{1}{2} \sin(2\pi f_0 \tau), \quad (6.77)$$

wobei zur Umformung des Ausdrucks $\sin(2\pi f_0(t + \tau))$ das Additionstheorem $\sin(\alpha + \beta) = \sin \alpha \cos \beta + \sin \beta \cos \alpha$ herangezogen wurde. Die Kreuzkorrelationsfunktion der zwei harmonischen Signale ist somit ebenfalls ein harmonisches Signal der gleichen Frequenz. ■

6.2.3 Korrelation von Energiesignalen

Da Energiesignale nicht zu den ergodischen Prozessen gehören, ist die Berechnung einer Korrelation nach Satz 6.2 nicht möglich. Für die wichtige Anwendung des Signalvergleichs lässt sich aber auch für Energiesignale eine Korrelationsfunktion über das Innenprodukt angeben. Für quadratisch integrierbare Signale $x(t)$ und $y(t)$ lautet die Kreuzkorrelationsfunktion

$$r_{xy}^E(\tau) = \langle x(t + \tau), y(t) \rangle = \int_{-\infty}^{\infty} x(t + \tau) y^*(t) dt < \infty. \quad (6.78)$$

Der Zusatz „E“ in $r_{xy}^E(\tau)$ kennzeichnet die abweichende Berechnungsvorschrift sowie die Tatsache, dass es sich hierbei nicht mehr um eine Korrelationsfunktion im stochastischen Sinn handelt. Vielfach spricht man bei (6.78) oftmals auch von einer **Impulskorrelation**, da Energiesignale als Impulsantworten stabiler LTI-Systeme aufgefasst werden können.

Die Korrelation in (6.78) kann auch als Faltung interpretiert werden:

$$r_{xy}^E(\tau) = x(\tau) * y^*(-\tau). \quad (6.79)$$

Die Autokorrelationsfunktion

$$r_{xx}^E(\tau) = \langle x(t + \tau), x(t) \rangle = \int_{-\infty}^{\infty} x(t + \tau) x^*(t) dt \quad (6.80)$$

eines Energiesignals ist für $\tau = 0$ gerade gleich der **Signalenergie**:

$$r_{xx}^E(0) = \langle x(t), x(t) \rangle = \|x(t)\|^2 = E_x. \quad (6.81)$$

Aufgrund der Schwarz'schen Ungleichung (6.66)

$$|\langle x(t + \tau), y(t) \rangle|^2 \leq \|x(t)\|^2 \cdot \|y(t)\|^2 \quad (6.82)$$

konvergiert das Korrelationsintegral, wenn $x(t)$ und $y(t)$ eine endliche Norm besitzen, d. h. eine endliche Signalenergie haben:

$$\left| \int_{-\infty}^{\infty} x(t + \tau) y^*(t) dt \right|^2 \leq \int_{-\infty}^{\infty} x(t) x^*(t) dt \cdot \int_{-\infty}^{\infty} y(t) y^*(t) dt, \quad (6.83)$$

$$|r_{xy}^E(\tau)|^2 \leq E_x \cdot E_y < \infty. \quad (6.84)$$

Für Energiesignale geht die **Energiedichte im Zeitbereich**

$$s_{xx}(t) = x(t) x^*(t) = |x(t)|^2 \quad (6.85)$$

für große Zeiten gegen null, weil nur so das Integral konvergiert:

$$\lim_{t \rightarrow \pm\infty} s_{xx}(t) = 0. \quad (6.86)$$

Das Gleiche gilt für die **Energiedichte im Frequenzbereich**:

$$\lim_{f \rightarrow \pm\infty} S_{xx}(f) = 0. \quad (6.87)$$

Andernfalls würde nach einer Integration der Energiedichte im Frequenzbereich $S_{xx}(f)$ keine endliche Signalenergie resultieren.

Aufgrund der Parseval'schen Beziehung kann die Signalenergie sowohl aus der Energiedichte im Zeitbereich als auch aus der Energiedichte im Frequenzbereich durch Integration ermittelt werden [42]:

$$E_x = \int_{-\infty}^{\infty} s_{xx}(t) dt = \|x(t)\|^2 = \|X(f)\|^2 = \int_{-\infty}^{\infty} S_{xx}(f) df. \quad (6.88)$$

Das Korrelationsintegral (6.78) zur Berechnung von $r_{xy}^E(\tau)$ konvergiert auch dann, wenn nur eines der beiden Signale $x(t)$ oder $y(t)$ ein Energiesignal und das andere ein Leistungssignal ist. Das liegt daran, dass das Energiesignal den Integranden $x(t + \tau) y^*(t)$ für große Zeiten t zu null werden lässt. Das Energiesignal im Produkt $x(t + \tau) y^*(t)$ wirkt für das zweite Signal mit unendlich hoher Energie (Leistungssignal) als Fensterfunktion.

Beispiel 6.6 (Korrelation von Energie- und Leistungssignal): Gegeben seien die beiden Signale

$$x(t) = r_T(t) = \begin{cases} T^{-1} & \text{für } |t| \leq T/2 \\ 0 & \text{für } |t| > T/2 \end{cases} \quad \text{und} \quad (6.89)$$

$$y(t) = \sin(2\pi f_0 t). \quad (6.90)$$

Die Signalenergien sind nach (6.88)

$$E_x = \int_{-\infty}^{\infty} r_T^2(t) dt = \int_{-T/2}^{T/2} \frac{1}{T^2} dt = \frac{1}{T}, \quad (6.91)$$

$$E_y = \int_{-\infty}^{\infty} \sin^2(2\pi f_0 t) dt = \left[\frac{1}{2}t - \frac{1}{8\pi f_0} \sin(4\pi f_0 t) \right]_{-\infty}^{\infty} = \infty. \quad (6.92)$$

Dabei ist zu beachten, dass die Fensterfunktion $r_T(t)$ auf die Fläche eins normiert ist, nicht auf die Signalenergie eins. Die Kreuzkorrelationsfunktion der beiden Signale

$$r_{xy}^E(\tau) = \int_{-\infty}^{\infty} r_T(t + \tau) \sin(2\pi f_0 t) dt = \frac{1}{T} \int_{\tau - \frac{T}{2}}^{\tau + \frac{T}{2}} \sin(2\pi f_0 t) dt \quad (6.93)$$

$$= \frac{1}{2\pi f_0 T} \left[-\cos(2\pi f_0 t) \right]_{\tau - \frac{T}{2}}^{\tau + \frac{T}{2}} = \sin(2\pi f_0 \tau) \frac{\sin(\pi f_0 T)}{\pi f_0 T} \quad (6.94)$$

ist aufgrund der Fensterwirkung des Energiesignals endlich. ■

6.2.4 Eigenschaften von Auto- und Kreuzkorrelationsfunktion

Tabelle 6.3 fasst die wichtigsten Eigenschaften der Korrelationsfunktionen stationärer Prozesse zusammen. Die Darstellung erfolgt allgemein für komplexwertige Prozesse.

Beweis 6.1 (Eigenschaften von Auto- und Kreuzkorrelationsfunktion): Zunächst sollen die Abschätzungen für die **Maximalwerte** der Korrelationsfunktionen bewiesen werden. Als Ausgangspunkt dient die folgende nichtnegative Größe:

$$E\{|x(t + \tau) - y(t)|^2\} \geq 0. \quad (6.95)$$

Durch Multiplikation und Aufspalten der Erwartungswerte ergibt sich:

$$\begin{aligned} & E\{|x(t + \tau)|^2\} - E\{x(t + \tau)y^*(t)\} \\ & - E\{\underbrace{x^*(t + \tau)y(t)}_{[x(t+\tau) y^*(t)]^*}\} + E\{|y(t)|^2\} \geq 0, \end{aligned} \quad (6.96)$$

Tabelle 6.3. Eigenschaften von Auto- und Kreuzkorrelationsfunktion.

| Eigenschaften der Autokorrelationsfunktion $r_{xx}(\tau)$ | |
|--|---|
| Maximalwert: | $\Re\{r_{xx}(\tau)\} \leq r_{xx}(0) = \sigma_x^2 + \mu_x^2$ |
| Symmetrie: | $r_{xx}(\tau) = r_{xx}^*(-\tau)$ |
| Unkorreliertheit für $ \tau \rightarrow \infty$: | $\lim_{ \tau \rightarrow \infty} r_{xx}(\tau) = \mu_x^2$ |
| Periodische Funktionen (T : Periode): | $r_{xx}(\tau) = r_{xx}(\tau + T)$ |
| Eigenschaften der Kreuzkorrelationsfunktion $r_{xy}(\tau)$ | |
| Maximalwert: | $\Re\{r_{xy}(\tau)\} \leq 1/2(r_{xx}(0) + r_{yy}(0))$ |
| Symmetrie: | $r_{xy}(\tau) = r_{yx}^*(-\tau) \neq r_{xy}^*(-\tau)$ |
| Unkorreliertheit für $ \tau \rightarrow \infty$: | $\lim_{ \tau \rightarrow \infty} r_{xy}(\tau) = \mu_x \cdot \mu_y^*$ |
| Unkorreliertheit von $x(t)$ und $y(t)$: | $r_{xy}(\tau) = \mu_x \cdot \mu_y^* \quad \forall \tau$ |

$$\mathbb{E}\{|x(t+\tau)|^2\} + \mathbb{E}\{|y(t)|^2\} \geq 2\Re\{\mathbb{E}\{x(t+\tau)y^*(t)\}\}, \quad (6.97)$$

$$\Re\{r_{xy}(\tau)\} \leq \frac{1}{2}(r_{xx}(0) + r_{yy}(0)). \quad (6.98)$$

Für die Autokorrelationsfunktion $r_{xx}(\tau)$ erhält man daraus mit $y(t) = x(t)$:

$$\Re\{r_{xx}(\tau)\} \leq r_{xx}(0). \quad (6.99)$$

Bei Leistungssignalen beschreibt die Autokorrelationsfunktion für $\tau = 0$ die Leistung des Signals.

Die **Symmetrie** der Autokorrelationsfunktion $r_{xx}(\tau)$ bezüglich $\tau = 0$ kann durch Substitution von $t' = t + \tau$ gezeigt werden. Dies ist erlaubt, da $r_{xx}(\tau)$ nicht vom absoluten Zeitpunkt t abhängig ist:

$$\mathbb{E}\{x(t+\tau)x^*(t)\} = \mathbb{E}\{x(t')x^*(t'-\tau)\} = \mathbb{E}\{[x(t'-\tau)x^*(t')]^*\}. \quad (6.100)$$

Daraus folgt unmittelbar die Symmetrieeigenschaft

$$r_{xx}(\tau) = r_{xx}^*(-\tau). \quad (6.101)$$

Sind zwei Signale unkorreliert, so zerfällt nach (6.19) das zweite Moment der Statistik 2. Ordnung in das Produkt der zwei ersten Momente der Statistik 1. Ordnung. Für die Korrelationsfunktion zweier unkorrelierter

Signale $x(t)$ und $y(t)$ gilt dann:

$$r_{xy}(\tau) = E\{x(t + \tau)y^*(t)\} = E\{x(t + \tau)\} \cdot E\{y^*(t)\} = \mu_x \mu_y^*. \quad (6.102)$$

Die **Unkorreliertheit** gilt meist nur für $\tau \rightarrow \infty$, da dann weit auseinander liegende Signalwerte keine Verwandtschaft mehr besitzen. Periodische Signale sind auch für $\tau \rightarrow \infty$ nicht unkorreliert! ◆

6.3 Korrelationsmesstechnik

6.3

Im Folgenden werden wichtige Anwendungen der Korrelationsfunktionen in der Messtechnik vorgestellt. Auf die Vielzahl der Anwendungen im Bereich der Signalübertragung und Nachrichtentechnik kann in diesem Rahmen aber nicht eingegangen werden.

6.3.1 Messung von Korrelationsfunktionen

Korrelationsfunktionen sind mathematisch – und nicht durch eine physikalische Messvorschrift – definiert. Durch Messungen lassen sich daher immer nur Schätzwerte bestimmen. Hierbei tritt eine Reihe von Schwierigkeiten auf, die in dem Gegensatz zwischen der mathematischen Definition und den messtechnischen Möglichkeiten begründet liegen.

1. Aus der Schar der Musterfunktionen eines Zufallsprozesses ist in der Regel nur eine einzelne Musterfunktion $y(t)$ für eine Messung verfügbar.
2. Anstelle des Scharmittelwertes muss folglich ein Zeitmittelwert gemessen werden.
3. Bei Messungen in Echtzeit, d. h. bei der Auswertung von nicht zuvor aufgezeichneten Signalen, lassen sich die momentanen Signalwerte nur mit Werten aus der Vergangenheit verknüpfen (kausales System). Es lassen sich also nur negative Werte von τ berücksichtigen.
4. Die Messung ist auf ein endliches Zeitintervall beschränkt.

Der **erste Punkt** dieser Einschränkung bereitet formal dann keine Schwierigkeiten, wenn der zugrundeliegende Prozess ergodisch ist. In diesem Fall kann eine *beliebige* Musterfunktion für die Messung benutzt werden. Darf für den Prozess Ergodizität nicht angenommen werden, so schließt das nicht aus, dass einzelne Musterfunktionen – aber eben nicht alle möglichen Musterfunktionen – repräsentativ für den Prozess sind. In diesem Fall ist eine geeignete Auswahl notwendig.

Der **zweite Punkt** hängt eng mit dem ersten zusammen. Für einzelne Musterfunktionen lassen sich nur Zeitmittelwerte bestimmen, was bei er-

godischen Prozessen keine Einschränkung darstellt. Es bedeutet aber, dass Korrelationsfunktionen nur für stationäre Prozesse gemessen werden können. Nur die Korrelation für Energiesignale erlaubt für diese spezielle Signalklasse die Messung einer Korrelation für nicht stationäre Signale.

Eine Messung in Echtzeit (**dritter Punkt**) verlangt negative Zeitverschiebungen τ . Dies stellt aber insofern keine Einschränkung dar, weil die Symmetrieeigenschaften der Korrelationsfunktionen bekannt sind.

Der **vierte Punkt** stellt das gravierendste Problem bei der Messung statistischer Kenngrößen dar: die Messung muss nach endlicher Zeit abgebrochen werden. Im vorliegenden Fall bedeutet dies, dass nur **Kurzzeitkorrelationsfunktionen** gemessen werden können. Man erhält damit einen Schätzwert $\hat{r}_{xx}(\tau)$ für die Korrelationsfunktion, der selbst wieder eine Zufallsgröße ist:

$$\hat{r}_{xx}(\tau) = \frac{1}{T} \int_0^T x(t + \tau) x^*(t) dt. \quad (6.103)$$

Heutzutage werden Korrelationsfunktionen meist rechnergestützt ermittelt. Dabei stehen nur N zeitdiskrete Abtastwerte x_n des Signals $x(t)$ zur Verfügung. Die Korrelationsfunktion wird dabei durch die folgende Summe approximiert [24]:

$$\hat{r}_{xx}(k) = \frac{1}{N} \sum_{n=0}^{N-1-k} x_{n+k} x_n^* \quad \text{mit} \quad |k| = 0, 1, 2, \dots, M \ll N. \quad (6.104)$$

❸ 6.3.2 Ähnlichkeit von Signalen, Laufzeitmessung

Die Kreuzkorrelation ist ein Maß für die Ähnlichkeit zweier um τ zeitverschobener, reeller Signale $x(t)$ und $y(t)$. Die Distanz der beiden Signale ist die quadratische Norm

$$\|x(t + \tau) - y(t)\|^2 = \langle x(t + \tau) - y(t), x(t + \tau) - y(t) \rangle \quad (6.105)$$

$$\begin{aligned} &= \langle x(t + \tau), x(t + \tau) \rangle + \langle y(t), y(t) \rangle \\ &\quad - \underbrace{\langle x(t + \tau), y(t) \rangle - \langle y(t), x(t + \tau) \rangle}_{2\Re\{\langle x(t+\tau), y(t) \rangle\} = 2\Re\{r_{xy}(\tau)\}} \end{aligned} \quad (6.106)$$

$$= \|x(t)\|^2 + \|y(t)\|^2 - 2\Re\{r_{xy}(\tau)\}. \quad (6.107)$$

Bei minimaler Distanz der Signale ist deren Ähnlichkeit maximal, d. h. die Kreuzkorrelationsfunktion $r_{xy}(\tau)$ hat für die entsprechende Verschiebung τ ihr globales Maximum. Für reelle Signale lässt sich daraus die Kreuzkorre-

lationsfunktion wie folgt abschätzen:

$$\|x(t + \tau) - y(t)\|^2 = \underbrace{\|x(t)\|^2}_{r_{xx}(0)} + \underbrace{\|y(t)\|^2}_{r_{yy}(0)} - 2\Re\{r_{xy}(\tau)\} \geq 0 \quad (6.108)$$

$$\Rightarrow r_{xy}(\tau) \leq \frac{1}{2}(r_{xx}(0) + r_{yy}(0)). \quad (6.109)$$

Für $y(t) = x(t)$ folgt daraus, dass die Autokorrelationsfunktion ihr Maximum bei der Verschiebung $\tau = 0$ hat:

$$r_{xx}(\tau) \leq r_{xx}(0). \quad (6.110)$$

Zur Bewertung der Ähnlichkeit zweier zeitlich gegeneinander verschobener Signale wird häufig der Kreuzkorrelationskoeffizient verwendet, der sich aus der Kreuzkorrelationsfunktion $r_{xy}(\tau)$ durch Normierung errechnet:

$$r_{xy,\text{norm}}(\tau) = \frac{r_{xy}(\tau)}{\sqrt{r_{xx}(0) \cdot r_{yy}(0)}}. \quad (6.111)$$

Beispiel 6.7 (Orten von Leckagen mit Körperschallmikrofonen): Eine Anwendung, die ihre praktische Bewährungsprobe bestanden hat, ist die Leckortung an in der Erde verlegten Wasser- oder Gasleitungen mit Hilfe von Körperschallmikrofonen. Abbildung 6.3 zeigt das Messprinzip. Zwei Körperschallmikrofone sind in bekanntem Abstand l , vorzugsweise an den Absperrschiebern, akustisch an die defekte Rohrleitung gekoppelt. Das ausströmende Medium erzeugt an der Leckstelle Geräusche $u(t)$, die über die Rohrleitung zu den beiden Mikrofonen geleitet werden. Die Schallgeschwindigkeit c im Medium wurde vorab an einer anderen Rohrleitung ähnlicher Bauart aufgenommen und ist deshalb bekannt. Das Geräuschsignal $u(t)$ an der Leckstelle wird auf den Wegen l_x und l_y zu den beiden Mikrofonen verzögert, weshalb man folgende Signale misst:

$$x(t) = u\left(t - \frac{l_x}{c}\right) \quad (6.112)$$

und

$$y(t) = u\left(t - \frac{l_y}{c}\right). \quad (6.113)$$

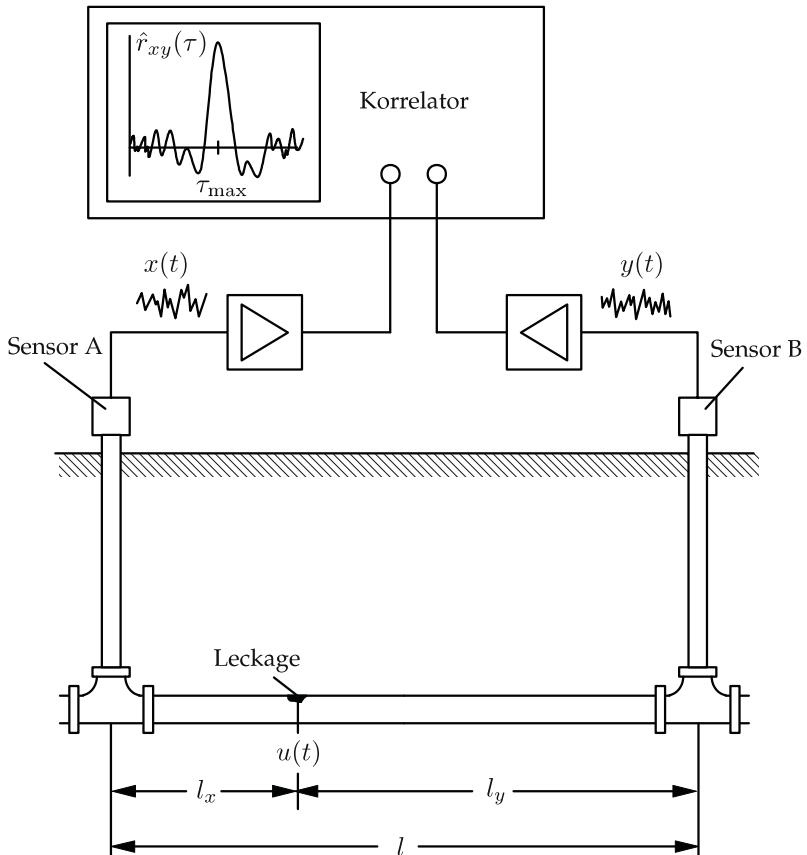


Abbildung 6.3. Anordnung zur Ortung von Leckagen an unterirdischen Rohrleitungen mit Hilfe von Körperschallmikrofonen.

Der Gesamtabstand $l = l_x + l_y$ der Mikrofone sei bekannt. Eine der beiden Längen, z. B. l_y , soll durch die Kreuzkorrelationsfunktion

$$\hat{r}_{xy}(\tau) = \frac{1}{2T} \int_{-T}^T x(t + \tau) y(t) dt \quad (6.114)$$

$$= \frac{1}{2T} \int_{-T}^T u\left(t + \tau - \frac{l_x}{c}\right) u\left(t - \frac{l_y}{c}\right) dt \quad (6.115)$$

bestimmt werden. Bei $\hat{r}_{xy}(\tau)$ handelt es sich um einen Schätzwert, da die Integration auf ein endliches Intervall $[-T, T]$ beschränkt ist. Die Lage

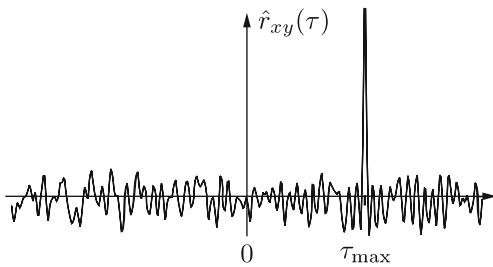


Abbildung 6.4. Typische Korrelationsfunktion bei der Ortung von Leckagen.

des Maximums von $\hat{r}_{xy}(\tau)$ entspricht der Laufzeitdifferenz

$$\tau_{\max} = \frac{l_x}{c} - \frac{l_y}{c} = \frac{l - 2l_y}{c}. \quad (6.116)$$

Mit der Laufzeitdifferenz τ_{\max} erhält man den Leckageort zentimetergenau zu

$$l_y = \frac{1}{2}(l - \tau_{\max} c), \quad (6.117)$$

ohne umfangreiche Ausgrabungen vornehmen zu müssen. Für $\tau_{\max} = 0$ wäre die Laufzeitdifferenz gerade null, das Leck läge dann in der Mitte der Strecke l . Abbildung 6.4 zeigt die Korrelationsfunktion einer typischen Messung. ■

Beispiel 6.8 (Vergleich von Schusswaffenspuren): In der Kriminaltechnik werden optische Aufnahmen von Geschossen und Patronenhülsen herangezogen, um Spurenzusammenhänge zwischen Munitionsteilen aufzudecken. Dabei liefern die Spuren an einem Munitionsteil Hinweise über die verwendete Tatwaffe [40]. Bei Geschossen sind vor allem die riefenartigen Verfeuerungsspuren im Mantelbereich, die durch die schraubenlinienförmigen Züge im Lauf der Waffe verursacht werden, vergleichsrelevant (Abb. 6.5 links). Auf Patronenhülsen sind es hauptsächlich die Spuren auf dem Hülsensoden, die Gleitspuren des Auswurfers und des Ausziehers und nicht zuletzt das Zündhütchen, welches einen Abdruck der Schlagbolzenspitze aufweist (Abb. 6.5 rechts).

Mittels Korrelationsverfahren lässt sich ein quantitativer Vergleich der relevanten Riefenspuren auf Geschossen vornehmen [41]. Dafür muss zunächst von jedem Munitionsteil eine Rundumaufnahme des Geschossmantels gewonnen werden. Abbildung 6.6 zeigt unten die Aufnahmen zweier Geschosse, die aus derselben Waffe stammen. Nach einer Vorverarbeitung der Grauwertbilder $g_i(\mathbf{x})$, wobei $\mathbf{x} = (x, y)^T$ den Orts-

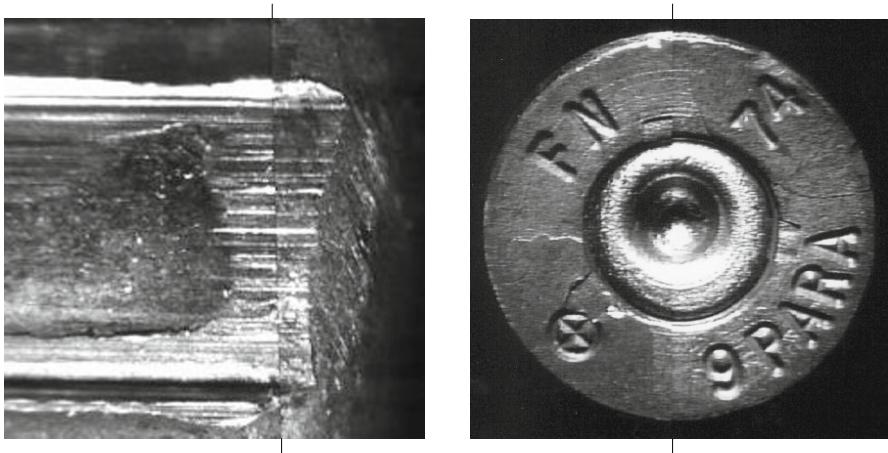


Abbildung 6.5. Spurenübereinstimmungen: (links) Spuren auf zwei Geschossen; (rechts) Spuren auf dem Hülsenboden und dem Zündhütchen zweier Patronenhülsen (Quelle: BKA).

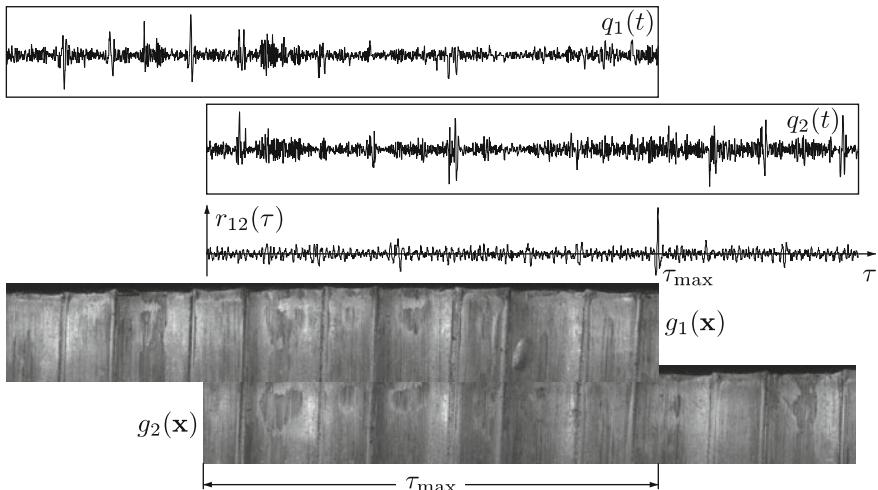


Abbildung 6.6. Bestimmung der Ähnlichkeit von Projektilen mit Korrelationsverfahren.

vektor bezeichnet, erfolgt eine Projektion (d. h. Summation) der Bildintensitäten in Riefenrichtung. Dadurch wird die Nutzinformation konzentriert und gleichzeitig werden Störungen unterdrückt. Aus dem resultierenden eindimensionalen Projektionssignal werden mittels eines nichtlinearen Filters feine Strukturen herausgearbeitet, da diese die für den Vergleich relevanten Informationen darstellen. Abbildung 6.6 zeigt oben typische Verläufe der resultierenden Signale $q_i(t)$, wobei im vorliegenden Beispiel die Variable t eine örtliche Dimension darstellt. Die

Kreuzkorrelationsfunktion $r_{12}(\tau)$ der Signale $q_1(t)$ und $q_2(t)$ weist ein ausgeprägtes Maximum auf, dessen Lage τ_{\max} die Verschiebung beider Signale wiedergibt. Der Kreuzkorrelationskoeffizient $r_{12,\text{norm}}(\tau_{\max})$ ist ein Maß für die Ähnlichkeit beider Geschosse [41]. ■

Beispiel 6.9 (Berührungslose Geschwindigkeitsmessung): Eine weitere Anwendung der Korrelation ist die berührungslose Geschwindigkeitsmessung nach dem Prinzip der **Laufzeitkorrelation**. Abbildung 6.7 zeigt die Anordnung zur Geschwindigkeitsmessung von Schüttgut auf einem Förderband. Dazu werden zwei Lichtquellen und zwei optische Sensoren im festen Abstand d montiert. Sie messen die Reflexionen des Schüttgutes an der jeweiligen Sensorposition. Die gemessenen Signale sind ähnlich. Im Idealfall ist das zweite Signal $y(t)$ lediglich um die Laufzeit τ des Förderbandes gegenüber dem ersten Signal $x(t)$ verschoben:

$$y(t) = x(t - \tau). \quad (6.118)$$

Das Maximum der Kreuzkorrelationsfunktion erhält man genau für diese Zeitverschiebung $\tau_{\max} = \tau$, woraus sich die Fördergeschwindigkeit

$$v = \frac{d}{\tau_{\max}} \quad (6.119)$$

berechnen lässt. Das Verfahren findet Anwendung zur Messung der Transportgeschwindigkeit von Walzgut, Papier oder Textilien sowie – mit geeigneten Sensoren – zur Messung der Strömungsgeschwindigkeit von Gasen und Flüssigkeiten [30]. ■

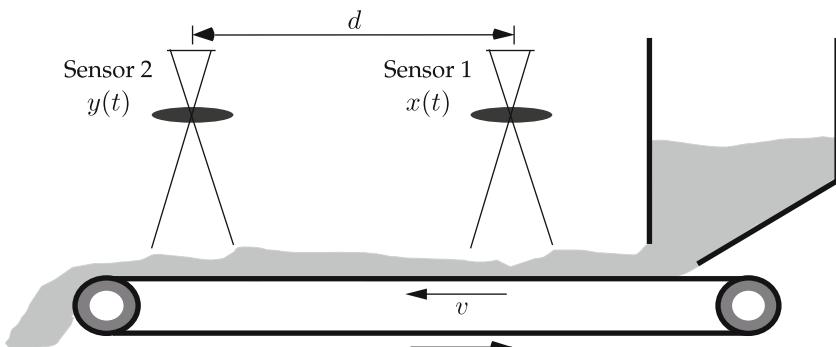


Abbildung 6.7. Berührungslose Geschwindigkeitsmessung eines Förderbandes mit dem Laufzeitkorrelationsverfahren.

6.3.3 Closed-loop-Korrelation

Das vorgestellte Verfahren zur Laufzeitkorrelation hat in einigen Anwendungen mehrere gravierende Nachteile.

1. Das Ergebnis der Korrelation

$$\hat{r}_{xy}(\tau) = \frac{1}{2T} \int_{-T}^T x(t + \tau) y(t) dt \quad (6.120)$$

steht erst am Ende des Integrationsintervalls zur Verfügung. Die Bestimmung von τ_{\max} ist dadurch sehr träge. Für dynamische Messgrößen (z. B. veränderliche Transportgeschwindigkeiten in Bsp. 6.9) ist das Verfahren nur schlecht geeignet.

2. Das Verfahren ist numerisch aufwendig. Zum einen müssen die Produkte der zeitverschobenen Signale $x(t + \tau) \cdot y(t)$ gebildet werden, zum anderen muss nach der Integration noch eine Maximumsuche durchgeführt werden.

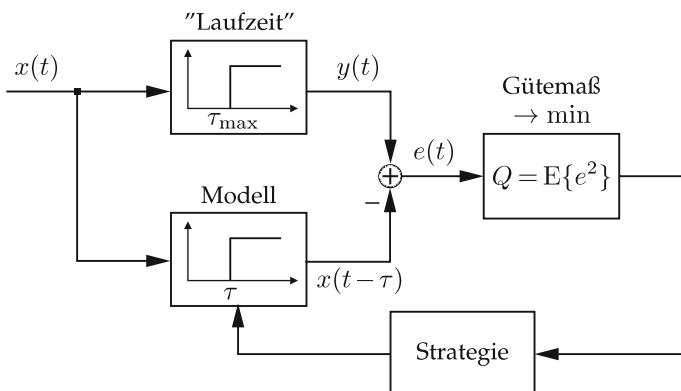


Abbildung 6.8. Schätzung der Verzögerungszeit τ beim Laufzeitkorrelator.

Zur Verbesserung der Dynamik des Verfahrens kann die Laufzeitkorrelation als **Identifikationsproblem** für die unbekannte Laufzeit $\tau_{\max} = \tau$ interpretiert werden [30]. Damit ergibt sich die in Abb. 6.8 dargestellte Blockstruktur. Die wahre Laufzeit τ_{\max} wird mit einer Modelltotzeit τ verglichen. Diese wird entsprechend einer Strategie so lange verstellt, bis ein Gütemaß Q ein Minimum erreicht. Als Gütfunktion wählt man den Erwartungswert des quadratischen Fehlers

$$Q = E\{e^2(t, \tau)\} \rightarrow \min, \quad (6.121)$$

der bezüglich τ minimiert werden soll. Mit dem Fehlersignal

$$e(t) = y(t) - x(t - \tau) \quad (6.122)$$

erhält man durch Differentiation von Q die notwendige Bedingung für ein Minimum:

$$\frac{dQ}{d\tau} = \frac{dE\{e^2(t, \tau)\}}{d\tau} = 2E\left\{ e(t, \tau) \cdot \frac{de(t, \tau)}{d\tau} \right\} = 0 \quad (6.123)$$

oder, nach Einsetzen von $e(t)$ und $\frac{de(t, \tau)}{d\tau}$,

$$E\{[y(t) - x(t - \tau)] \dot{x}(t - \tau)\} = 0. \quad (6.124)$$

Der Erwartungswert (6.124) hat links und rechts vom Abgleichpunkt $\tau_{\max} = \tau$ verschiedene Vorzeichen und verhält sich damit wie die Regelabweichung eines normalen Regelkreises. Damit kann die Maximumsuche in eine einfache Regelung für τ überführt werden (Abb. 6.9). Die Erwartungswertbildung wird von einem Integralregler übernommen, der die Modelltotzeit τ verstellt. Der Regler arbeitet also als **Nachlaufkorrelator** oder Laufzeit-Tracker. Er sucht die Laufzeit $\tau_{\max} = \tau$ und verfolgt sie bei Änderungen. Veränderliche Transportgeschwindigkeiten in Bsp. 6.9 können damit schneller erfasst werden. Wegen der Regelkreisstruktur des Korrelators nennt man das Verfahren auch **Closed-loop-Korrelation**.

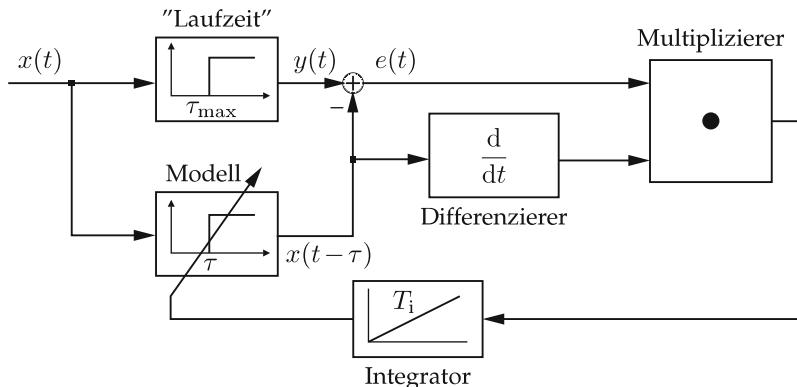


Abbildung 6.9. Schätzung der Verzögerungszeit τ im geschlossenen Regelkreis durch Ableitung der Korrelationsfunktion.

④ Differentiation der Kreuzkorrelationsfunktion

Der Erwartungswert aus (6.124) soll nun nochmals genauer betrachtet werden. Bekanntlich lässt sich das Maximum einer Funktion durch Differen-

tiation bestimmen. Für die Kreuzkorrelationsfunktion, deren Maximum für die Laufzeitmessung gesucht ist, gilt:

$$\frac{dr_{xy}(\tau)}{d\tau}\Big|_{\tau=\tau_{\max}} = \dot{r}_{xy}(\tau)\Big|_{\tau=\tau_{\max}} = 0. \quad (6.125)$$

Für die Ableitung der Korrelationsfunktion kann man analog zu (8.29)

$$\dot{r}_{xy}(\tau) = -r_{xy}(\tau) \quad (6.126)$$

schreiben. Multipliziert man (6.124) aus, so erhält man

$$E\{\dot{x}(t-\tau)y(t)\} - E\{\dot{x}(t-\tau)x(t-\tau)\} = 0, \quad (6.127)$$

$$r_{\dot{x}y}(-\tau) - r_{x\dot{x}}(\tau=0) = 0. \quad (6.128)$$

Da die Autokorrelation allgemein eine gerade Funktion mit dem Maximum bei $\tau = 0$ ist, folgt daraus für die Ableitung der Korrelationsfunktion

$$\dot{r}_{xx}(\tau=0) = 0 = r_{x\dot{x}}(0) = E\{x(t)\dot{x}(t)\} = 0. \quad (6.129)$$

Daraus folgt schließlich

$$r_{\dot{x}y}(-\tau) = r_{y\dot{x}}(\tau) = -\dot{r}_{yx}(\tau) = \dot{r}_{xy}(\tau) = 0. \quad (6.130)$$

Die Bildung des Erwartungswertes nach (6.124) ist somit identisch mit der Differentiation der Kreuzkorrelationsfunktion.

6.3.4 Polaritätskorrelation

Der Closed-loop-Korrelator vermeidet die vollständige numerische Berechnung der Kreuzkorrelationsfunktion und erlaubt die Anpassung an sich ändernde Laufzeiten τ_{\max} . Es müssen jedoch weiterhin die Produkte der zeitverschobenen Signale $x(t+\tau) \cdot y(t)$ gebildet werden, was numerisch immer noch aufwendig ist. Abhilfe schafft der **Polaritätskorrelator**, der mit grob quantisierten 1-Bit-Signalen arbeitet. Dadurch reduziert sich die Multiplikation der zeitverschobenen Signale auf einfache Operationen, die mit Hilfe von logischen Gattern realisiert werden können. Es kann gezeigt werden, dass unter bestimmten Voraussetzungen statistische Eigenschaften von quantisierten Signalen mit denen der kontinuierlichen Signale übereinstimmen (Abschn. 7.2) [54, 55]. Insbesondere bleibt bei der Polaritätskorrelation die Lage des Maximums erhalten, so dass daraus weiterhin die Laufzeit bestimmt werden kann (Abb. 6.10). Die Polaritätskorrelation für abgetastete

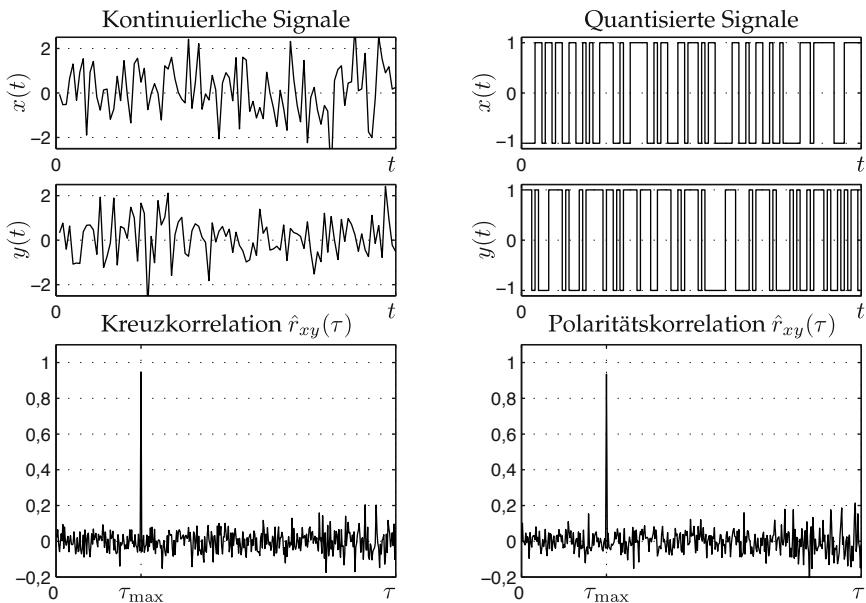


Abbildung 6.10. Vergleich zwischen der Kreuzkorrelation kontinuierlicher Signale und der Polaritätskorrelation von mit einem Bit quantisierten Signalen.

Signale lautet für $\tau = k t_A$

$$\hat{r}_{xy}(k) = \frac{1}{N} \sum_{n=0}^{N-1-k} \text{sign}[x_{n+k}] \text{sign}[y_n], \quad |k| = 0, 1, \dots, M \ll N. \quad (6.131)$$

Abbildung 6.11 zeigt das Prinzipbild eines Laufzeitkorrelators mit Polaritätskorrelation. Die Modelllaufzeit τ wird durch ein einfaches Schieberegister implementiert, welches durch eine variable Taktfrequenz angesteuert wird. Dazu muss der Ausgang des Integrators in eine entsprechende Frequenz umgesetzt werden. Dies kann z. B. mit einem spannungsgesteuerten Oszillator (VCO, *voltage-controlled oscillator*) erfolgen.

6.3.5 Ähnlichkeit von Spektren, Dopplerfrequenzmessung

In Analogie zur Korrelationsfunktion $r_{xy}(\tau)$ für um τ zeitverschobene Signale kann im Frequenzbereich eine Korrelationsfunktion $\varrho_{XY}(\vartheta)$ für um ϑ frequenzverschobene Signale definiert werden:

$$\varrho_{XY}(\vartheta) = \langle X(f + \vartheta), Y(f) \rangle = \int_{-\infty}^{\infty} X(f + \vartheta) Y^*(f) df < \infty. \quad (6.132)$$

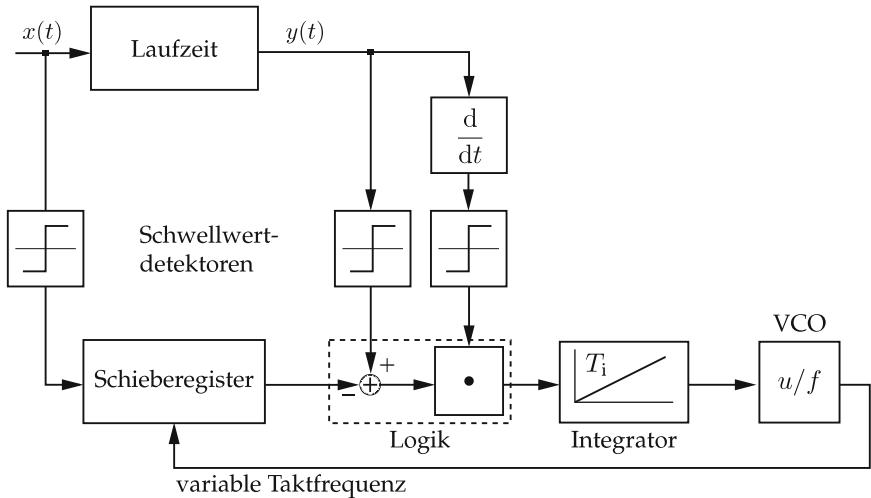


Abbildung 6.11. Laufzeitkorrelator mit Polaritätskorrelation. Die Modelllaufzeit wird durch ein getaktetes Schieberegister realisiert.

Sie ist ein Maß für die Ähnlichkeit der Spektren $X(f + \vartheta)$ und $Y(f)$. Die Distanz der um die Frequenz ϑ verschobenen Spektren ist analog zu (6.105)

$$\begin{aligned} \|X(f + \vartheta) - Y(f)\|^2 &= \langle X(f + \vartheta), X(f + \vartheta) \rangle + \langle Y(f), Y(f) \rangle \\ &\quad - 2\Re\{\langle X(f + \vartheta), Y(f) \rangle\} \end{aligned} \quad (6.133)$$

$$= \|X(f)\|^2 + \|Y(f)\|^2 - 2\Re\{\varrho_{XY}(\vartheta)\}. \quad (6.134)$$

Nach dem Satz von Parseval ist

$$\varrho_{XY}(\vartheta) = \int_{-\infty}^{\infty} x(t) e^{-j2\pi\vartheta t} y^*(t) dt = \langle x(t) e^{-j2\pi\vartheta t}, y(t) \rangle. \quad (6.135)$$

Die Frequenzverschiebung des Spektrums entspricht einer Modulation des Signals im Zeitbereich.

Beispiel 6.10 (Geschwindigkeitsmessung basierend auf dem Doppler-Effekt):

Die Durchflussgeschwindigkeit einer Strömung soll berührungslos mit der Anordnung in Abb. 6.12 gemessen werden. Eine Ultraschallsonde sendet ein Signal $x(t)$ mit einer Frequenz f_0 in das Medium. Die Schallwellen werden vom Medium gestreut und werden von der Sonde als $y(t)$ wieder empfangen. Durch die Relativbewegung ergibt sich nach dem **Doppler-Effekt** eine **Frequenzverschiebung** am Empfänger.

Bewegt sich der Empfänger mit der Relativgeschwindigkeit v auf den

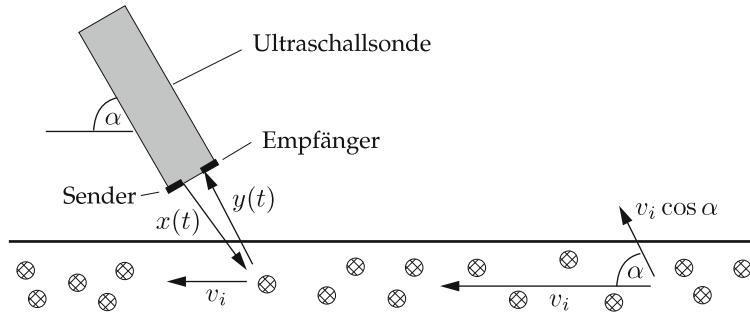


Abbildung 6.12. Geschwindigkeitsmessung nach dem kontinuierlichen Doppler-Verfahren.

ruhenden Sender zu, so gilt für die registrierte Frequenz:

$$f = f_0 \left(1 + \frac{v}{c} \right). \quad (6.136)$$

Dabei bezeichnen f_0 die Sendefrequenz und c die Schallgeschwindigkeit im Medium. Ruht der Empfänger und bewegt sich der Sender auf den Empfänger zu, so ergibt sich für die registrierte Frequenz

$$f = f_0 \left(\frac{1}{1 - \frac{v}{c}} \right). \quad (6.137)$$

Bei der Geschwindigkeitsmessung trifft der Ultraschall auf ein Teilchen i mit der Relativgeschwindigkeit $v_i \cos \alpha$. Das bewegte Teilchen streut den Ultraschall, ist also gleichzeitig Empfänger und Sender. Das Echo trifft auf den ruhenden Empfänger. Die registrierte Frequenz wird damit

$$f = f_0 \frac{1 + \frac{v_i}{c} \cos \alpha}{1 - \frac{v_i}{c} \cos \alpha}. \quad (6.138)$$

Durch Darstellung des Nenners als geometrische Reihe und Approximation für kleine Strömungsgeschwindigkeiten $v_i \ll c$ erhält man die Frequenz des empfangenen Signals:

$$f = f_0 \left(1 + \frac{v_i}{c} \cos \alpha \right) \cdot \sum_{n=0}^{\infty} \left(\frac{v_i}{c} \cos \alpha \right)^n \quad (6.139)$$

$$\approx f_0 \left(1 + \frac{v_i}{c} \cos \alpha \right) \cdot \left(1 + \frac{v_i}{c} \cos \alpha \right). \quad (6.140)$$

Berücksichtigt man nur die linearen Terme von v_i , so lässt sich die Empfangsfrequenz f als Summe der Sendefrequenz f_0 und einer Frequenz-

verschiebung f_i darstellen:

$$f \approx f_0 \left(1 + 2 \frac{v_i}{c} \cos \alpha \right) = f_0 + f_i . \quad (6.141)$$

Das Empfangssignal ist damit proportional zum frequenzverschobenen Sendesignal

$$Y(f) = k \cdot X(f + f_i) . \quad (6.142)$$

Das Maximum der Korrelationfunktion der Signalspektren

$$\varrho_{XY}(\vartheta) = k \cdot \int_{-\infty}^{\infty} X(f + \vartheta) X^*(f + f_i) df \quad (6.143)$$

ergibt sich gerade für $\vartheta = f_i$, woraus sich die Strömungsgeschwindigkeit

$$v_i = \vartheta \frac{c}{2f_0 \cos \alpha} \quad (6.144)$$

berechnen lässt. ■

6.3.6 Selbstähnlichkeit

Die Autokorrelationsfunktion ist hervorragend geeignet, um **periodische Signalanteile** aufzufinden, die im Vergleich zu einem Störsignal sehr klein sind. Exemplarisch zeigt Abb. 6.13 ein Signal $y(t)$, das keine sichtbare periodische Schwingung enthält. Tatsächlich handelt es sich jedoch um eine harmonische Schwingung, die von starkem Rauschen überlagert ist. Es sei

$$y(t) = x(t) + n(t) \quad (6.145)$$

ein stochastischer Prozess, der sich aus einem periodischen Anteil $x(t)$ nach Bsp. 6.3 mit Musterfunktionen $x_i(t) = \sin(\omega t + \varphi_i)$ und einem mittelwertfreien Rauschen $n(t)$ zusammensetzt. Beide Signale stammen aus nicht gekoppelten Quellen und dürfen daher als unkorreliert angenommen werden. Dann gilt:

$$r_{yy}(\tau) = r_{xx}(\tau) + r_{nn}(\tau) = \frac{1}{2} \cos(\omega\tau) + r_{nn}(\tau) . \quad (6.146)$$

Nach Bsp. 6.3 ist $r_{xx}(\tau)$ periodisch, während bei Rauschprozessen $r_{nn}(\tau)$ für hinreichend große τ verschwindet (vgl. Abschn. 6.4.2):

$$r_{nn}(\tau) \approx 0 \quad \text{für} \quad |\tau| \gg 0 . \quad (6.147)$$

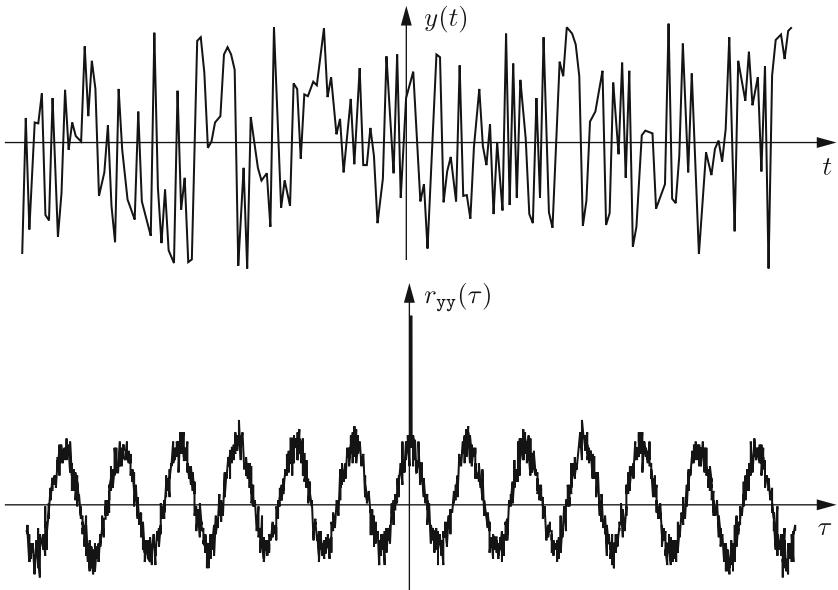


Abbildung 6.13. Harmonische Schwingung überlagert von starkem Rauschen und zugehörige Autokorrelationsfunktion $r_{yy}(\tau)$.

Es gilt daher für große τ :

$$r_{yy}(\tau) \approx r_{xx}(\tau) = \frac{1}{2} \cos(\omega\tau) \quad \text{für} \quad |\tau| \gg 0. \quad (6.148)$$

Das Verfahren wird beispielsweise zum Auffinden von Brummstörungen in elektrischen Signalen, zur Schwingungsanalyse oder auch zur rechtzeitigen Diagnose von Maschinenschäden eingesetzt [1].

6.4 Spektrale Darstellung stochastischer Signale

6.4

6.4.1 Leistungsdichtespektrum

Bei der Analyse von Signalen und Systemen ist die Beschreibung im Frequenzbereich mit Hilfe der Laplace- und Fourier-Transformation von großer Hilfe. Für Musterfunktionen stationärer Prozesse existiert im Allgemeinen das Fourier-Integral jedoch nicht, da diese für $|t| \rightarrow \infty$ nicht abklingen. Abhilfe schafft eine Fourier-Transformation der Korrelationsfunktion. Hierbei werden die Eigenschaften des Prozesses zuerst durch eine Erwartungswertbildung im Zeitbereich erfasst und dann mittels der Fourier-Transformation in den Frequenzbereich überführt. Dadurch lässt sich insbesondere be-

schreiben, wie sich stochastische Prozesse bei ihrer Übertragung durch LTI-Systeme verhalten. Im Folgenden wird von schwach stationären Zufallsprozessen ausgegangen.

6.27

Definition 6.27: Autoleistungsdichespektrum

Das **Autoleistungsdichespektrum** eines schwach stationären Zufallsprozesses $x(t)$ ist die Fourier-Transformierte der Autokorrelationsfunktion:

$$S_{xx}(f) = \int_{-\infty}^{\infty} r_{xx}(\tau) e^{-j2\pi f\tau} d\tau. \quad (6.149)$$

Dieser Zusammenhang ist als **Wiener-Khintchine-Theorem** bekannt. ◇

Das Autoleistungsdichespektrum gibt an, welche Frequenzen im statistischen Mittel wie stark am Aufbau einer Musterfunktion beteiligt sind. Ferner können auch Kreuzkorrelationsfunktionen in den Frequenzbereich überführt werden.

6.28

Definition 6.28: Kreuzleistungsdichespektrum

Das **Kreuzleistungsdichespektrum** (kurz: Kreuzspektrum) zweier stochastischer Prozesse $x(t)$ und $y(t)$ ist durch die Fourier-Transformierte

$$S_{xy}(f) = \int_{-\infty}^{\infty} r_{xy}(\tau) e^{-j2\pi f\tau} d\tau \quad (6.150)$$

ihrer Kreuzkorrelationsfunktion definiert. ◇

Die Umkehrung der Gleichungen (6.149) und (6.150) erhält man durch inverse Fourier-Transformation:

$$r_{xx}(\tau) = \int_{-\infty}^{\infty} S_{xx}(f) e^{j2\pi f\tau} df, \quad (6.151)$$

$$r_{xy}(\tau) = \int_{-\infty}^{\infty} S_{xy}(f) e^{j2\pi f\tau} df. \quad (6.152)$$

Der Begriff „Leistungsdichte“ deutet an, dass $S_{xx}(f)$ die Leistungsverteilung über den Frequenzen f darstellt. Integriert man $S_{xx}(f)$ über einen Frequenzbereich $f_1 \leq f \leq f_2$, so ist das Ergebnis proportional zur **Leistung** des Prozesses in diesem Spektralbereich. Insbesondere gilt der folgende Satz:

Satz 6.5: Leistung

Für die Leistung P_x eines Zufallsprozesses $x(t)$ gilt:

$$P_x = E\{x^2(t)\} = r_{xx}(0) = \int_{-\infty}^{\infty} S_{xx}(f) df. \quad (6.153)$$

Die Leistung P_x eines Leistungssignals $x(t)$ lässt sich analog berechnen. ◇

Anmerkungen

- Genau genommen müsste man in Satz 6.5 von der **mittleren Leistung** des Prozesses bzw. des Signals sprechen, da die Berechnung der Leistung im Frequenzbereich sämtliche Zeiten berücksichtigt. Wie in der Literatur üblich wird hier jedoch der Begriff der Leistung bevorzugt. Streng hier-von zu unterscheiden ist die **Momentanleistung**, die über ein begrenztes Zeitintervall ermittelt wird und nicht aus dem Leistungsdichtespektrum eines Signals abgeleitet werden kann.
- Korrekt erweise müsste man noch ein **Energiedichtespektrum** für die Fourier-Transformierte der Korrelationsfunktion für Energiesignale einführen:

$$S_{xx}^E(f) = \int_{-\infty}^{\infty} r_{xx}^E(\tau) e^{-j2\pi f\tau} d\tau. \quad (6.154)$$

Die Unterscheidung zwischen Energie- und Leistungsdichtespektrum wird im weiteren Verlauf allerdings unterlassen, da im Spektralbe-reich nicht mehr festgestellt werden kann, ob das Spektrum von einem Energie- oder von einem Leistungssignal stammt. Im Folgenden wird da-her meist nur noch von Leistungsdichtespektren gesprochen.

In (6.79) wurde die Korrelation von Energiesignalen durch eine Faltung dar-gestellt. Die Faltung entspricht einer Multiplikation im Frequenzbereich:

$$r_{xx}^E(\tau) = x(\tau) * x^*(-\tau) \quad (6.155)$$

$$\begin{array}{c} \circ \\ \bullet \\ \circ \end{array} \\ S_{xx}^E(f) = X(f) X^*(f) = |X(f)|^2. \quad (6.156)$$

Interpretiert man $x(t)$ als Impulsantwort eines stabilen LTI-Systems, so ent-spricht die zugehörige spektrale Energiedichte $S_{xx}^E(f)$ gerade dem Betrags-quadrat $|X(f)|^2$ der Übertragungsfunktion $X(f)$ des Systems. Die Phasen-information geht dabei verloren.

⌚ Eigenschaften des Leistungsdichtespektrums

Im Folgenden werden die Eigenschaften des Leistungsdichtespektrums $S_{xx}(f)$ zusammengestellt.

1. $S_{xx}(f)$ ist für reelle Prozesse eine gerade Funktion:

$$S_{xx}(f) = S_{xx}(-f). \quad (6.157)$$

2. $S_{xx}(f)$ ist reell für alle Frequenzen f :

$$S_{xx}(f) = S_{xx}^*(f). \quad (6.158)$$

Die Eigenschaften 1 und 2 folgen daraus, dass die Autokorrelationsfunktion $r_{xx}(\tau)$ eines reellen, stationären Prozesses $x(t)$ selbst reell und gerade ist.

3. $S_{xx}(f)$ ist nichtnegativ für alle Frequenzen f :

$$S_{xx}(f) \geq 0. \quad (6.159)$$

Dies lässt sich dadurch erklären, dass der quadratische Erwartungswert $P_x = E\{x^2(t)\}$ immer größer gleich null sein muss. Er wird aber nach (6.153) über das Integral der Leistungsdichte $S_{xx}(f)$ berechnet. Die Bedingung $E\{x^2(t)\} \geq 0$ kann allgemein nur für $S_{xx}(f) \geq 0$ sichergestellt werden.

4. Für die Kreuzleistungsdichtespektren gilt die Beziehung:

$$S_{xy}(f) = S_{yx}(-f). \quad (6.160)$$

Sind die beiden Prozesse $x(t)$ und $y(t)$ reell, so gilt zusätzlich:

$$S_{xy}(f) = S_{xy}^*(-f) = S_{yx}(-f) = S_{yx}^*(f). \quad (6.161)$$

Da die Kreuzkorrelationsfunktion im Allgemeinen nicht symmetrisch ist, ist auch das Kreuzleistungsdichtespektrum reellwertiger Prozesse nicht reellwertig.

⌚ 6.4.2 Rauschen

Viele Störeinflüsse – wie Verstärkerrauschen oder Störungen bei der Signalübertragung – lassen sich durch Zufallsprozesse modellieren, deren Leistungsdichtespektrum über einen weiten Frequenzbereich nahezu konstant ist. Zur Modellierung derartiger Störungen verwendet man meist einen idealisierten Prozess, dessen Leistungsdichte über alle Frequenzen konstant ist. Man nennt einen solchen Prozess **weißes Rauschen**.

Definition 6.29: Weißes Rauschen

Einen Zufallsprozess $n(t)$ nennt man weißes Rauschen, wenn sein Leistungsdichtespektrum für alle Frequenzen f konstant ist:

$$S_{nn}(f) = a^2 = \text{const.} \quad \forall f; \quad (6.162)$$

siehe Abb. 6.14 links.

◊

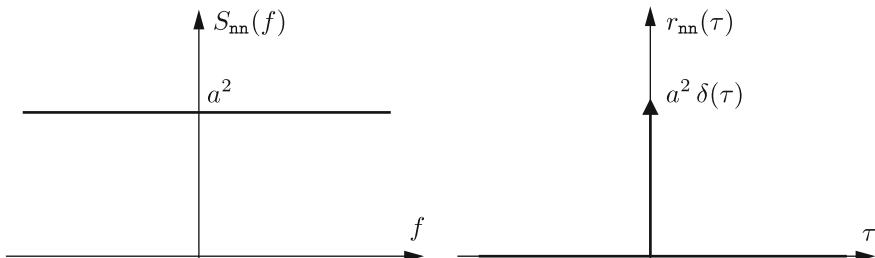


Abbildung 6.14. Leistungsdichtespektrum $S_{nn}(f)$ und Autokorrelationsfunktion $r_{nn}(\tau)$ von weißem Rauschen.

Das Symbol $n(t)$ röhrt von der Bezeichnung für Störgeräusche (engl. *noise*) her und deutet an, dass weißes Rauschen häufig zur Beschreibung rauscharakteriger Störungen verwendet wird.

Die Autokorrelationsfunktion von weißem Rauschen ergibt sich aus $S_{nn}(f)$ durch inverse Fourier-Transformation:

$$r_{nn}(\tau) = \mathcal{F}^{-1}\{a^2\} = a^2 \delta(\tau). \quad (6.163)$$

Die Autokorrelationsfunktion $r_{nn}(\tau)$ verschwindet also für alle Werte $\tau \neq 0$ (Abb. 6.14 rechts). Das bedeutet, dass benachbarte Signalwerte von $n(t)$ grundsätzlich unkorreliert sind, so dass sich die Signalamplituden unendlich schnell ändern müssten. Dies ist physikalisch nicht möglich, was sich auch in der Leistung nach (6.153) zeigt, die für weißes Rauschen unendlich groß wird:

$$P_n = r_{nn}(0) = \int_{-\infty}^{\infty} a^2 df \rightarrow \infty. \quad (6.164)$$

Weißes Rauschen ist daher eine Idealisierung, die sich physikalisch nicht realisieren lässt. Trotzdem wird dieses Modell in der Praxis gerne verwendet, weil die Autokorrelationsfunktion und die Leistungsdichte eine einfache Form besitzen. Oft nimmt man darüber hinaus normalverteilte Signalamplituden an und spricht dann von **weißem Gauß'schen Rauschen**, um

aus der Eigenschaft der Unkorreliertheit benachbarter Signalwerte auf ihre statistische Unabhängigkeit schließen zu können.

Die Verwendung von weißem Rauschen als Modellansatz ist gerechtfertigt, wenn das Störsignal auf Systeme mit Tiefpass- oder Bandpasscharakter trifft. Die hohen Frequenzanteile werden dann unterdrückt und die Leistung am Systemausgang bleibt insgesamt endlich. Rauschprozesse, deren Leistungsdichtespektrum zu hohen Frequenzen hin abfällt, bezeichnet man als **farbiges Rauschen**.

Beispiel 6.11 (Farbiges Rauschen): Die hochfrequenten Anteile eines weißen Rauschprozesses mit dem Effektivwert a werden durch einen Tiefpass erster Ordnung mit der Übertragungsfunktion

$$G(f) = \frac{1}{1 + j\frac{2\pi f}{2\pi f_g}} = \frac{2\pi f_g}{2\pi f_g + j2\pi f} \quad (6.165)$$

gedämpft. Das resultierende farbige Rauschen $n(t)$ hat die Leistungsdichte

$$S_{nn}(f) = \frac{a^2 \cdot (2\pi f_g)^2}{(2\pi f_g)^2 + (2\pi f)^2}, \quad (6.166)$$

wie später in Bsp. 6.13 gezeigt wird. Durch inverse Fourier-Transformation erhält man die zugehörige Autokorrelationsfunktion

$$r_{nn}(\tau) = \mathcal{F}^{-1} \left\{ a^2 \cdot \frac{(2\pi f_g)^2}{(2\pi f_g)^2 + (2\pi f)^2} \right\} = a^2 \pi f_g e^{-2\pi f_g |\tau|}. \quad (6.167)$$

Die Leistung des farbigen Rauschprozesses folgt für $\tau = 0$ und ist endlich:

$$P_n = r_{nn}(0) = a^2 \cdot \pi f_g. \quad (6.168)$$

Abbildung 6.15 zeigt das Leistungsdichtespektrum $S_{nn}(f)$ und die zugehörige Autokorrelationsfunktion $r_{nn}(\tau)$ des gefilterten Rauschprozesses. Man erkennt, dass die AKF im Vergleich zum weißen Rauschen breiter geworden ist, d. h. zeitlich benachbarte Werte sind nun miteinander korreliert. Als Maß für die Breite der Korrelationsfunktion wird oftmals die Korrelationslänge τ_κ herangezogen. Sie beschreibt die Zeitverschiebung τ , bei welcher die AKF auf den Wert $r_{nn}(0)/e$ (ca. 37 % vom Maximalwert) abgefallen ist. ■

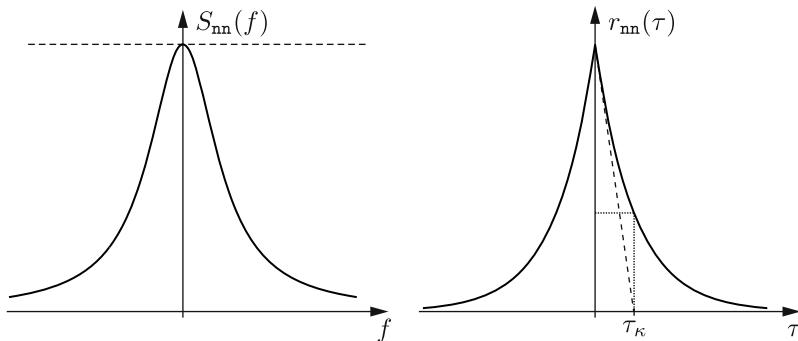


Abbildung 6.15. Leistungsdichtespektrum $S_{nn}(f)$ und Autokorrelationsfunktion $r_{nn}(\tau)$ eines farbigen Rauschprozesses.

▷ Pseudostochastische Binärfolgen

Neben dem Rauschen als unerwünschtem Störsignal ist es manchmal zweckmäßig – beispielsweise zur Systemidentifikation (Abschn. 6.5) – ein Rauschsignal auf ein System als Testsignal zu schalten. Es stellt sich die Frage, wie ein möglichst ideales weißes Rauschen praktisch auf einfache Art und Weise generiert werden kann. Eine einfache Realisierung von annähernd weißem Rauschen sind **pseudostochastische Binärfolgen** (engl. *pseudo-random binary sequences*, PRBS). Dazu wird ein Schieberegister der Länge N herangezogen, dessen Information mit der Taktzeit t_A entweder nach links oder nach rechts geschoben wird. Abbildung 6.16 zeigt entsprechende Schieberegisterschaltungen für die Wortlänge $N = 6$ basierend auf D-Flipflops und einem Äquivalenz-Gatter:

- **Rechts-Schieben:** Am Eingang D_1 des ersten D-Flipflops x_1 liegt der Ausgang der Äquivalenz nach Tab. 6.4 a.
- **Links-Schieben:** Am Eingang D_N des N -ten D-Flipflops x_N liegt der Ausgang der Äquivalenz nach Tab. 6.4 b.

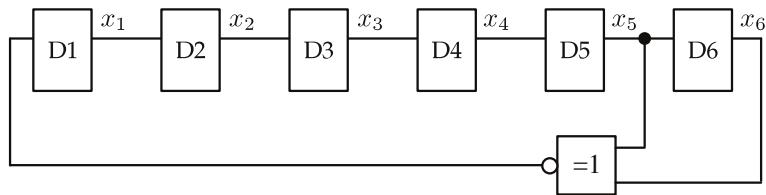
Die Beschaltung für größere Wortlängen N findet man in [52]. Die Periode ist immer $2^N - 1$ und deshalb ungerade, weil durch eine hier nicht dargestellte Zusatzschaltung der Fall ausgeschlossen wird, dass alle Flipflops den Wert eins haben.

Ordnet man den Ausgangswerten des Flipflops x_N das Ausgangssignal $u(t)$ wie folgt zu:

$$x_N = 1 \quad \mapsto \quad u(t) = a, \quad (6.169)$$

$$x_N = 0 \quad \mapsto \quad u(t) = -a, \quad (6.170)$$

a) Rechts-Schieben



b) Links-Schieben

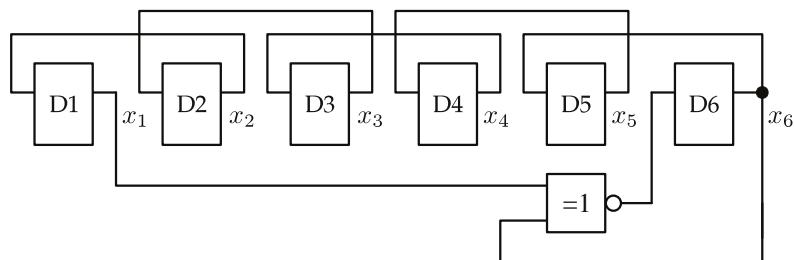


Abbildung 6.16. Rückgekoppelte Schieberegisterschaltungen zur Erzeugung von pseudostochastischen Binärfolgen.

Tabelle 6.4. Beschaltung rückgekoppelter Schieberegister zur Erzeugung von PRBS-Folgen.

| Wortlänge N des Schieberegisters | Äquivalenz-Beschaltung | | Periode |
|------------------------------------|--|--|---------|
| | a. Rechts-Schieben | b. Links-Schieben | |
| 1 | $0 \equiv x_1$ | $0 \equiv x_1$ | 1 |
| 2 | $x_1 \equiv x_2$ | $x_1 \equiv x_2$ | 3 |
| 3 | $x_2 \equiv x_3$ | $x_1 \equiv x_3$ | 7 |
| 4 | $x_3 \equiv x_4$ | $x_1 \equiv x_4$ | 15 |
| 5 | $x_3 \equiv x_5$ | $x_1 \equiv x_4$ | 31 |
| 6 | $x_5 \equiv x_6$ | $x_1 \equiv x_6$ | 63 |
| 7 | $x_4 \equiv x_7$ | $x_1 \equiv x_5$ | 127 |
| 8 | $(x_3 \equiv x_5) \equiv (x_7 \equiv x_8)$ | $(x_4 \equiv x_6) \equiv (x_1 \equiv x_8)$ | 255 |
| 9 | $x_5 \equiv x_9$ | $x_1 \equiv x_6$ | 511 |
| 10 | $x_7 \equiv x_{10}$ | $x_1 \equiv x_8$ | 1023 |
| 11 | $x_9 \equiv x_{11}$ | $x_1 \equiv x_{10}$ | 2047 |

so kommt innerhalb der Periode $(2^N - 1) t_A$ der Fall $u(t) = -a$ um eine Taktzeit t_A häufiger vor als der Fall $u(t) = a$ (Abb. 6.17).

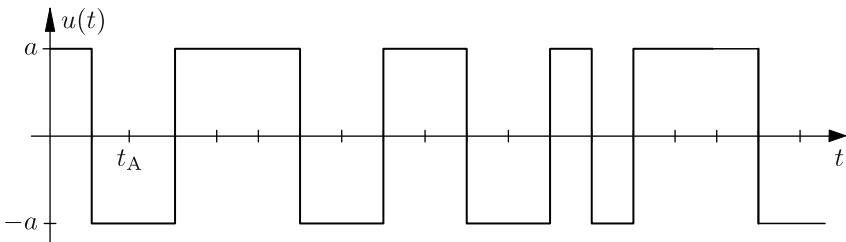


Abbildung 6.17. Beispiel einer pseudostochastischen Binärfolge.

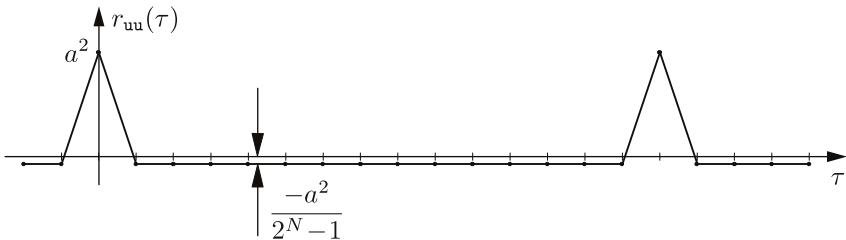


Abbildung 6.18. Autokorrelationsfunktion einer pseudostochastischen Binärfolge.

Da die PRBS-Folge periodisch ist, ist auch ihre Autokorrelationsfunktion periodisch mit der Periode $(2^N - 1) t_A$. Sie ist gegeben durch die Funktion

$$r_{uu}(\tau) = \begin{cases} a^2 & \text{für } \tau = k \cdot (2^N - 1) t_A, \\ \frac{-a^2}{2^N - 1} & \text{sonst,} \end{cases} \quad (6.171)$$

siehe Abb. 6.18. In der Praxis müssen ausreichend lange Perioden verwendet werden, damit die PRBS-Folgen zur Systemidentifikation geeignet sind. Mit der Wahl der Taktzeit t_A lässt sich auch die gewünschte Bandbreite der resultierenden Leistungsdichte einstellen.

6.4.3 Überlagerung zufälliger Störsignale

Die additive Überlagerung von Zufallssignalen ist das einfachste, aber auch das am häufigsten verwendete Modell zur Beschreibung von rauschähnlichen Störungen. Man kann beispielsweise das Ausgangssignal $y(t)$ eines Messsystems darstellen als die Summe aus einem idealen Messsignal $u(t)$ und einem zufälligen Messrauschen $n(t)$ (Abb. 6.19). Das Modell ermöglicht eine sehr einfache Systemanalyse mit Hilfe des Leistungsdichtespektrums und der Korrelationsfunktion.

Für das Summensignal $y(t)$ aus Abb. 6.19 gilt:

$$y(t) = u(t) + n(t). \quad (6.172)$$

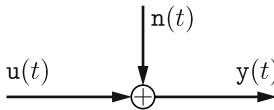


Abbildung 6.19. Modell zur Überlagerung zufälliger Störsignale.

Für die Autokorrelationsfunktion der Summe erhält man mit den Rechenregeln für den Erwartungswertoperator:

$$r_{yy}(\tau) = E\{y(t + \tau) y^*(t)\} \quad (6.173)$$

$$= E\{[u(t + \tau) + n(t + \tau)][u(t) + n(t)]^*\} \quad (6.174)$$

$$\begin{aligned} &= E\{u(t + \tau) u^*(t)\} + E\{u(t + \tau) n^*(t)\} \\ &\quad + E\{n(t + \tau) u^*(t)\} + E\{n(t + \tau) n^*(t)\} \end{aligned} \quad (6.175)$$

$$= r_{uu}(\tau) + r_{un}(\tau) + r_{nu}(\tau) + r_{nn}(\tau). \quad (6.176)$$

Die Autokorrelationsfunktion $r_{yy}(\tau)$ setzt sich also aus den Autokorrelationsfunktionen der beiden Eingangssignale $u(t)$ und $n(t)$ und ihren Kreuzkorrelationen zusammen. Für das Leistungsdichtespektrum erhält man entsprechend durch Fourier-Transformation:

$$S_{yy}(f) = S_{uu}(f) + S_{un}(f) + S_{nu}(f) + S_{nn}(f). \quad (6.177)$$

Bei reellen Signalen erhält man mit der Symmetrieeigenschaft (6.161) der Kreuzleistungsdichte den Zusammenhang

$$S_{un}(f) + S_{nu}(f) = S_{un}(f) + S_{un}^*(f) = 2\Re\{S_{un}(f)\} \quad (6.178)$$

und damit für die Leistungsdichte $S_{yy}(f)$

$$S_{yy}(f) = S_{uu}(f) + S_{nn}(f) + 2\Re\{S_{un}(f)\}. \quad (6.179)$$

Der Ausdruck vereinfacht sich, wenn die Prozesse $u(t)$ und $n(t)$ unkorreliert sind und mindestens einer der beiden Prozesse mittelwertfrei ist (typischerweise $\mu_n = 0$). Dann entfallen nach Abschn. 6.3 die Kreuzterme

$$r_{un}(\tau) = 0 \quad \text{---} \quad S_{un}(f) = 0 \quad (6.180)$$

und man erhält die wichtigen Beziehungen für **unkorrelierte Prozesse**:

$$r_{yy}(\tau) = r_{uu}(\tau) + r_{nn}(\tau), \quad (6.181)$$

$$S_{yy}(f) = S_{uu}(f) + S_{nn}(f). \quad (6.182)$$

Die Annahme der Unkorreliertheit gilt in vielen Fällen, in denen $n(t)$ ein dem Messsignal $u(t)$ überlagertes Störsignal darstellt. Beispielsweise ist das Grundrauschen eines Messverstärkers eine derartige Störung. Da in der Praxis viele additiv überlagerte Störungen ebenfalls mittelwertfrei sind, können die Beziehungen aus (6.182) verwendet werden, in denen einfach die Autokorrelationsfunktionen bzw. die Leistungsdichten der jeweiligen Signale addiert werden.

6.4.4 Übertragung stochastischer Signale durch LTI-Systeme

Die Übertragung deterministischer Signale durch lineare, zeitinvariante Systeme (LTI-Systeme) wird im Zeitbereich durch die Impulsantwort $g(t)$ des Systems vollständig charakterisiert. Durch Fourier-Transformation der Impulsantwort $g(t)$ erhält man die Übertragungsfunktion $G(f)$ des Systems. Damit lässt sich die Systemantwort $y(t)$ bei deterministischen Eingangssignalen $x(t)$ auch im Frequenzbereich angeben:

$$y(t) = g(t) * x(t) \quad \circ\bullet \quad Y(f) = G(f) \cdot X(f). \quad (6.183)$$

Wegen der Divergenz der quadratischen Norm existiert für stationäre Zufallsprozesse keine Laplace-Transformierte, die Fourier-Transformation nur in Ausnahmefällen. Es wird daher eine andere Beschreibung benötigt. Schaltet man Zufallssignale $x(t)$ an den Eingang von LTI-Systemen, so kann die Übertragungsfunktion aus den Leistungsdichten von Eingangs- und Ausgangssignal berechnet werden.

Zunächst soll die Klasse der Energiesignale betrachtet werden, für welche die Fourier-Transformation $X(f) = \mathcal{F}\{x(t)\}$ existiert und die Systemantwort $Y(f)$ nach (6.183) mit der Übertragungsfunktion bestimmt werden kann. Über die Faltungsbeziehung der Korrelation (6.156) erhält man für das Energiedichtespektrum des Ausgangssignals:

$$S_{yy}^E(f) = Y(f) Y^*(f) = G(f) X(f) \cdot G^*(f) X^*(f) \quad (6.184)$$

$$= |G(f)|^2 S_{xx}^E(f). \quad (6.185)$$

Entsprechendes gilt für das Kreuzenergiedichtespektrum von Eingangs- und Ausgangssignal:

$$S_{xy}^E(f) = X(f) Y^*(f) = X(f) \cdot G^*(f) X^*(f) = G^*(f) S_{xx}^E(f). \quad (6.186)$$

Die eben gefundenen Zusammenhänge lassen sich nicht ohne Weiteres auf die Klasse der Leistungssignale übertragen. Für sie existiert in der Regel die Fourier-Transformierte $X(f)$ des Zeitsignals $x(t)$ nicht. Allenfalls für peri-

odische Signale lässt sich eine Fourier-Transformierte unter Verwendung von Dirac-Impulsen im Frequenzbereich angeben. Dann steht man allerdings vor dem Problem, das Produkt zweier Dirac-Impulse zur Berechnung von $X(f)X^*(f)$ angeben zu müssen, welches jedoch nicht definiert ist.

Beispiel 6.12 (Periodisches Signal): Gegeben ist eine harmonische Schwingung mit der Frequenz f_0 . Das Zeitsignal und seine Fourier-Transformierte sind angebar mit

$$x(t) = a \cos(2\pi f_0 t) \quad (6.187)$$

$$\begin{array}{c} \bullet \\ \circ \\ X(f) = \frac{a}{2} [\delta(f - f_0) + \delta(f + f_0)] \end{array} \quad (6.188)$$

Die Leistungsdichte lässt sich aber nicht über das Produkt $S_{xx}(f) = X(f)X^*(f)$ angeben, da das Produkt $\delta(f) \cdot \delta(f)$ nicht definiert ist. ■

Bei Leistungssignalen muss man daher einen anderen Weg einschlagen. Interpretiert man Leistungssignale als Musterfunktionen ergodischer Zufallsprozesse, so ist es einleuchtend, dass man eine Erwartungswertbildung durchführen muss, um ihre Übertragung durch LTI-Systeme zu beschreiben. Dies wird anschaulich klar, wenn man berücksichtigt, dass man die Eigenschaften des gesamten ergodischen Prozesses erfassen möchte und nicht nur einer speziellen Musterfunktion $x_i(t)$, die später nicht mehr exakt identisch auftritt.

6.6

Satz 6.6: Übertragung ergodischer Prozesse durch LTI-Systeme

Für einen ergodischen Zufallsprozess $x(t)$, der ein LTI-System mit der Impulsantwort $g(t)$ und der Übertragungsfunktion $G(f)$ durchläuft, gelten die folgenden Beziehungen:

$$r_{xy}(\tau) = r_{xx}(\tau) * g^*(-\tau) \quad \circ \bullet \quad S_{xy}(f) = S_{xx}(f) G^*(f), \quad (6.189)$$

$$r_{yy}(\tau) = r_{xy}(\tau) * g(\tau) \quad \circ \bullet \quad S_{yy}(f) = S_{xy}(f) G(f), \quad (6.190)$$

$$r_{yy}(\tau) = r_{xx}(\tau) * r_{gg}^E(\tau) \quad \circ \bullet \quad S_{yy}(f) = S_{xx}(f) |G(f)|^2, \quad (6.191)$$

wobei $y(t)$ den Prozess am Ausgang des Systems bezeichnet.

Beweis 6.2 (Übertragung ergodischer Prozesse durch LTI-Systeme): Allgemein ist die Antwort $y(t)$ eines LTI-Systems mit der Impulsantwort $g(t)$ auf einen Eingangsprozess $x(t)$ durch die Faltung gegeben:

$$y(t) = \int_{-\infty}^{\infty} x(t - \alpha) g(\alpha) d\alpha. \quad (6.192)$$

Multipliziert man den konjugiert komplexen Prozess $y^*(t)$ nach (6.192) mit dem verschobenen Prozess $x(t + \tau)$ und bildet den Erwartungswert, so erhält man

$$E\{x(t + \tau) y^*(t)\} = \int_{-\infty}^{\infty} E\{x(t + \tau) x^*(t - \alpha)\} g^*(\alpha) d\alpha. \quad (6.193)$$

Da $E\{x(t + \tau) x^*(t - \alpha)\}$ gleich der AKF $r_{xx}(\tau + \alpha)$ ist, folgt unmittelbar (6.189):

$$r_{xy}(\tau) = \int_{-\infty}^{\infty} r_{xx}(\tau + \alpha) g^*(\alpha) d\alpha = \int_{-\infty}^{\infty} r_{xx}(\tau - \beta) g^*(-\beta) d\beta \quad (6.194)$$



$$S_{xy}(f) = S_{xx}(f) G^*(f). \quad (6.195)$$

Durch Multiplikation von (6.192) mit $y^*(t - \tau)$ und anschließende Erwartungswertbildung erhält man die Beziehung (6.190):

$$E\{y(t) y^*(t - \tau)\} = \int_{-\infty}^{\infty} E\{x(t - \alpha) y^*(t - \tau)\} g(\alpha) d\alpha, \quad (6.196)$$

$$r_{yy}(\tau) = \int_{-\infty}^{\infty} r_{xy}(\tau - \alpha) g(\alpha) d\alpha \quad (6.197)$$



$$S_{yy}(f) = S_{xy}(f) G(f). \quad (6.198)$$

Setzt man schließlich $r_{xy}(\tau)$ aus (6.189) in (6.190) ein, so erhält man (6.191):

$$r_{yy}(\tau) = r_{xx}(\tau) * g(\tau) * g^*(-\tau) \quad (6.199)$$



$$S_{yy}(f) = S_{xx}(f) G(f) G^*(f) = S_{xx}(f) |G(f)|^2. \quad (6.200)$$

Die Faltung der Impulsantworten

$$g(\tau) * g^*(-\tau) = \int_{-\infty}^{\infty} g(t + \tau) g^*(t) dt = r_{gg}^E(\tau) \quad (6.201)$$

ist dabei nach (6.80) eine Autokorrelationsfunktion für Energiesignale oder Impulskorrelation. ♦

Beispiel 6.13 (Erzeugung von farbigem Rauschen): Es soll hier nochmals das Bsp. 6.11 aufgegriffen werden. Gegeben sei ein weißer Rauschprozess $x(t)$ mit konstantem Leistungsdichtespektrum

$$S_{xx}(f) = a^2 \quad \forall f. \quad (6.202)$$

Das Signal werde auf einen Tiefpass erster Ordnung mit der Grenzfrequenz f_g gegeben. Dessen Übertragungsfunktion lautet nach Bsp. 6.11

$$G(f) = \frac{1}{1 + j f/f_g}. \quad (6.203)$$

Das Leistungsdichtespektrum $S_{yy}(f)$ am Ausgang des Filters ist nach (6.191)

$$S_{yy}(f) = |G(f)|^2 S_{xx}(f) = \frac{a^2}{1 + (f/f_g)^2} = \frac{a^2 f_g^2}{f_g^2 + f^2}. \quad (6.204)$$

Das Ergebnis entspricht dem in Bsp. 6.11 vorgestellten Leistungsdichtespektrum eines farbigen Rauschprozesses. Solche Prozesse lassen sich sehr einfach durch Tiefpassfilterung von weißem Rauschen (beispielsweise einer PRBS-Folge) erzeugen. ■

6.5 Systemidentifikation

Mit Hilfe von Leistungsdichtespektren lässt sich die Übertragungsfunktion $G(f)$ eines unbekannten Systems auf einfache Weise identifizieren. Ein breitbandiges Rauschsignal $x(t)$, beispielsweise eine PRBS-Folge, wird an den Eingang des unbekannten Systems gelegt (Abb. 6.20). Der Eingangsprozess $x(t)$ kann näherungsweise als weißes Rauschen beschrieben werden, wenn die Bandbreite der Rauschquelle viel größer ist als die des untersuchten Systems. Das Ausgangssignal $y(t)$ des Systems sei zusätzlich von einem unabhängigen, mittelwertfreien Messrauschen $n(t)$ überlagert. Für das weiße

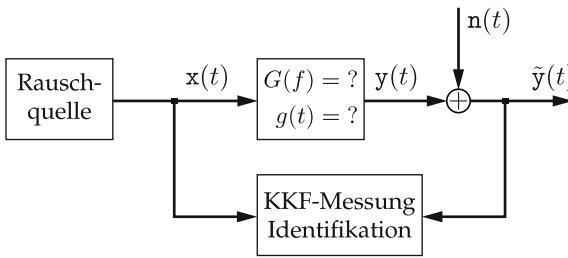


Abbildung 6.20. Identifikation einer Übertragungsfunktion.

Rauschen am Eingang des Systems gilt

$$S_{xx}(f) = a^2 \quad \forall f. \quad (6.205)$$

Ohne überlagertes Messrauschen $n(t) = 0$ erhält man die Übertragungsfunktion $G(f)$ des unbekannten Systems aus der Kreuzleistungsdichte nach (6.189)

$$S_{yx}(f) = G(f) S_{xx}(f) = a^2 G(f) \quad (6.206)$$

$$\begin{aligned} r_{yx}(\tau) &= g(\tau) * a^2 \delta(\tau) = a^2 g(\tau). \end{aligned} \quad (6.207)$$

Für die Kreuzkorrelationsfunktion zwischen dem gestörten Ausgang $\tilde{y}(t)$ und dem Eingang $x(t)$ ergibt sich:

$$r_{\tilde{y}x}(\tau) = E\{\tilde{y}(t + \tau) x(t)\} \quad (6.208)$$

$$= E\{[y(t + \tau) + n(t + \tau)] x(t)\} \quad (6.209)$$

$$= E\{y(t + \tau) x(t)\} + \underbrace{E\{n(t + \tau) x(t)\}}_{= 0} = r_{yx}(\tau) \quad (6.210)$$

mit der unabhängigen, mittelwertfreien Messstörung $n(t)$. Das bedeutet, dass selbst ein überlagertes mittelwertfreies Messrauschen die Identifikation nicht stört, solange $x(t)$ und $n(t)$ unkorreliert sind.

6.5.1 Schätzung des Leistungsdichtespektrums

Im Folgenden werden einige Effekte diskutiert, die bei der Identifikation von realen Systemen auftreten. In Abschn. 6.3.1 wurde bereits auf die Schwierigkeiten hingewiesen, die beim Messen von Korrelationsfunktionen entstehen. Da man auf ein endliches Messintervall beschränkt ist, erhält man auch für die Leistungsdichten nur Schätzwerte $\hat{S}_{xy}(f)$. Dadurch kann es passieren, dass trotz vorhandener Unkorreliertheit zweier Prozesse $x(t)$ und $n(t)$ die zugehörigen Kreuzleistungsdichten nicht exakt zu null werden. Der

Vorteil des endlichen Messintervalls ist aber, dass in jedem Fall die Fourier-Transformierte $Y(f)$ des gemessenen Signals $y(t)$ existiert. Damit kann in der Praxis ein *Schätzwert* für das Leistungsdichtespektrum direkt aus der Fourier-Transformierten berechnet werden:

$$\hat{S}_{yy}(f) = Y(f) Y^*(f), \quad (6.211)$$

$$\hat{S}_{xy}(f) = X(f) Y^*(f). \quad (6.212)$$

Den Schätzwert für das Autoleistungsdichtespektrum $\hat{S}_{yy}(f)$ bezeichnet man auch als **Periodogramm**.

Durch eine Mittelung über mehrere Schätzvorgänge lässt sich die Schätzung der spektralen Leistungsdichte verbessern:

$$\overline{\hat{S}_{yy}(f)} = \frac{1}{N} \sum_{i=1}^N \hat{S}_{yy,i}(f) = \frac{1}{N} \sum_{i=1}^N Y_i(f) Y_i^*(f), \quad (6.213)$$

$$\overline{\hat{S}_{xy}(f)} = \frac{1}{N} \sum_{i=1}^N \hat{S}_{xy,i}(f) = \frac{1}{N} \sum_{i=1}^N X_i(f) Y_i^*(f). \quad (6.214)$$

Beispiel 6.14 (Erkennung periodischer Strukturen in Bildern): Ein wichtiges Einsatzgebiet des Autoleistungsdichtespektrums ist die Erkennung periodischer Signalanteile. Die Riefentextur in Abb. 6.21(a) zeigt einen vergrößerten Ausschnitt einer Wellendichtfläche eines Automobils, die aus einer Überlagerung periodischer Drallriefen und stochastisch platzierte Schleifriefen besteht. Erstere sind unerwünscht, da sie Leckage oder Tro-

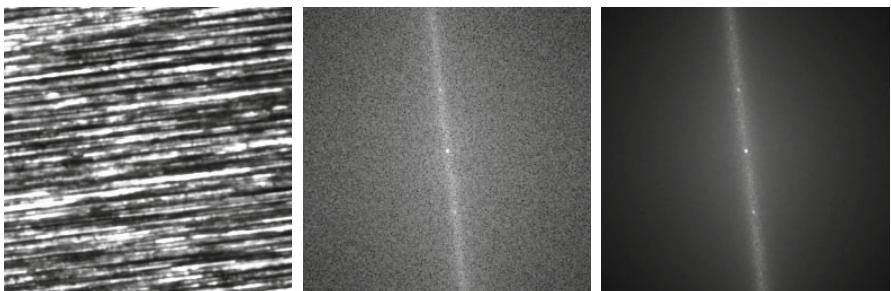


Abbildung 6.21. Schätzung des Leistungsdichtespektrums zur Erkennung periodischer Strukturen: (a) Riefentextur $g(\mathbf{x})$; (b) Periodogramm $\hat{S}_{gg}(\mathbf{f})$ der Riefentextur $g(\mathbf{x})$; (c) Ergebnis $\overline{\hat{S}_{gg}(\mathbf{f})}$ der Mittelung von $N = 120$ Periodogrammen.

ckenlauf verursachen können, weshalb sie zuverlässig erkannt werden müssen. Da die Textur aus geraden, parallelen Riefen besteht, konzentrieren sich im Periodogramm Abb. 6.21 (b) die Leistungsanteile längs einer Geraden durch den Ursprung senkrecht zur Riefenschar [2]. Aufgrund ihrer Periodizität schlagen sich Drallriefen in Form von Dirac-Impulsen nieder.

Am stark verrauschten Eindruck wird deutlich, dass das Periodogramm kein besonders guter Schätzer für die Leistungsdichte ist – tatsächlich liegt seine Standardabweichung in der Größenordnung der zu schätzenden Größe selbst [24]. Bei schwach ausgeprägten Drallstrukturen sind die Dirac-Impulse daher im Periodogramm kaum zu erkennen. Abhilfe schafft eine Mittelung mehrerer Periodogramme, was die Verfügbarkeit weiterer Aufnahmen der Riefentextur desselben Werkstücks unter vergleichbaren Bedingungen voraussetzt. Abbildung 6.21 (c) zeigt das Ergebnis $\hat{S}_{gg}(f)$ der Mittelung von $N = 120$ Periodogrammen. Bei Annahme unkorrelierter Periodogramme weist dieser Schätzer eine um den Faktor $1/N$ geringere Varianz auf als ein einzelnes Periodogramm $\hat{S}_{gg}(f)$. Aufgrund der Reduktion der statistischen Schätzfehler lassen sich darin die periodischen Strukturen – d. h. die Dirac-Impulse – weitaus robuster erkennen als in Abb. 6.21 (b). ■

6.5.2 Systemidentifikation bei geschätzter Leistungsdichte

Für die Identifikation der Übertragungsfunktion $G(f)$ bieten sich nun zwei unterschiedliche Vorgehensweisen an: die Verwendung gemittelter Periodogramme und die Verwendung gemittelter Kreuzleistungsdichten.

6.5.2.1 Quotientenbildung gemittelter Periodogramme

Für das Periodogramm des Ausgangssignals gilt:

$$\hat{S}_{yy}(f) = [G(f) X(f) + N(f)] \cdot [G(f) X(f) + N(f)]^* \quad (6.215)$$

$$= |G(f)|^2 \hat{S}_{xx}(f) + G(f) \hat{S}_{xn}(f) + G^*(f) \hat{S}_{nx}(f) + \hat{S}_{nn}(f) \quad (6.216)$$

$$= |G(f)|^2 \hat{S}_{xx}(f) + 2\Re\{G(f) \hat{S}_{xn}(f)\} + \hat{S}_{nn}(f). \quad (6.217)$$

Durch die Mittelung verschwindet der Kreuzterm näherungsweise, da $x(t)$ und $n(t)$ als unkorreliert angenommen wurden:

$$\overline{\hat{S}_{yy}(f)} = |G(f)|^2 \overline{\hat{S}_{xx}(f)} + 2\Re\{G(f) \underbrace{\overline{\hat{S}_{xn}(f)}}_{\approx 0}\} + \overline{\hat{S}_{nn}(f)}. \quad (6.218)$$

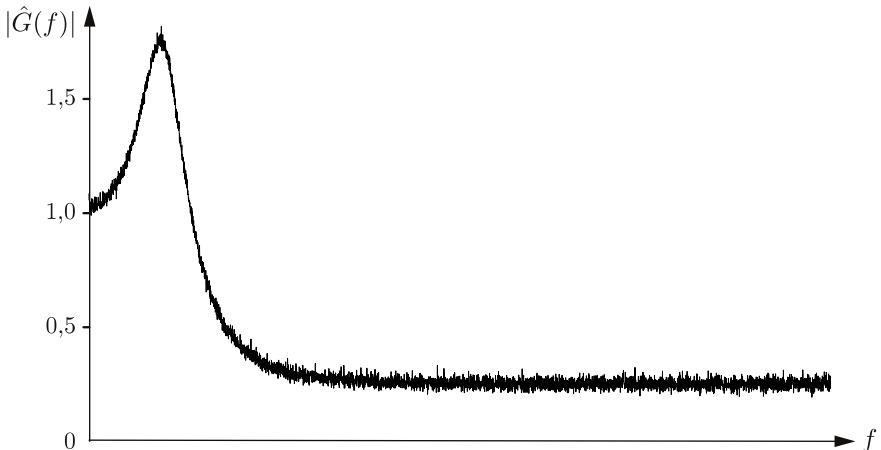


Abbildung 6.22. Identifikation der Übertragungsfunktion eines PT₂-Gliedes durch Mittelung von $N = 100$ Periodogrammen.

Die Übertragungsfunktion kann damit durch den folgenden Quotienten geschätzt werden:

$$|\hat{G}(f)|^2 = \frac{\hat{S}_{yy}(f)}{\hat{S}_{xx}(f)} = |G(f)|^2 + \frac{\hat{S}_{nn}(f)}{\hat{S}_{xx}(f)}, \quad (6.219)$$

wobei durch die im Ausgangssignal überlagerte Störung $n(t)$ ein Fehler auftritt. In Abb. 6.22 ist die geschätzte Übertragungsfunktion $|\hat{G}(f)|$ eines PT₂-Gliedes dargestellt. Für große Frequenzen f geht die Funktion $|\hat{G}(f)|$ nicht gegen null, sondern in den Quotienten

$$\frac{\hat{S}_{nn}(f)}{\hat{S}_{xx}(f)} \quad (6.220)$$

über. Der Betrag der Übertragungsfunktion wird also bei Störungen zu groß geschätzt. Ein weiterer Nachteil dieser Methode ist, dass nur der Amplitudengang $|\hat{G}(f)|$, nicht aber der Phasengang $\angle \hat{G}(f)$ geschätzt werden kann.

⌚ 6.5.2.2 Quotientenbildung gemittelter Kreuzleistungsdichten

Das Kreuzleistungsdichtespektrum

$$\hat{S}_{yx}(f) = [G(f) X(f) + N(f)] X^*(f) = G(f) \hat{S}_{xx}(f) + \hat{S}_{nx}(f) \quad (6.221)$$

ist im Vergleich zum Periodogramm eine komplexe Funktion. Die Mittelung muss daher getrennt nach Betrag und Phase oder getrennt nach Real- und Imaginärteil erfolgen, was im Vergleich zur ersten Methode einen höheren

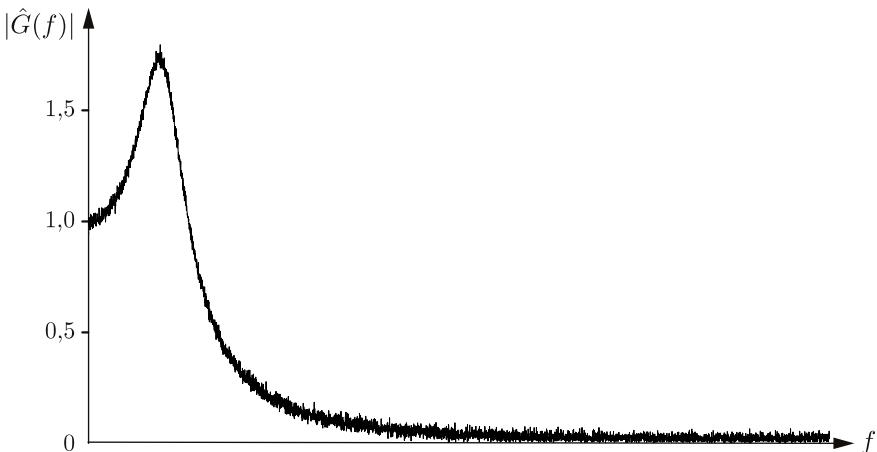


Abbildung 6.23. Identifikation der Übertragungsfunktion eines PT₂-Gliedes durch Mittelung von $N = 100$ Kreuzleistungsdichtespektren.

Rechenaufwand mit sich bringt:

$$\overline{\hat{S}_{yx}(f)} = \overline{G(f) \hat{S}_{xx}(f)} + \underbrace{\overline{\hat{S}_{nx}(f)}}_{\approx 0} . \quad (6.222)$$

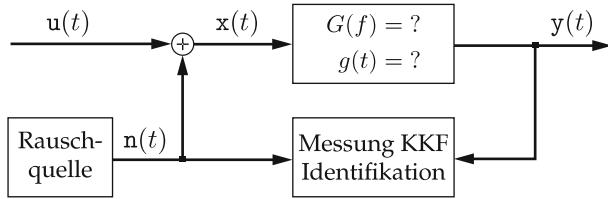
Die Übertragungsfunktion wird schließlich durch den folgenden Quotienten geschätzt:

$$\hat{G}(f) = \frac{\overline{\hat{S}_{yx}(f)}}{\overline{\hat{S}_{xx}(f)}} = \frac{\overline{G(f) \hat{S}_{xx}(f)}}{\overline{\hat{S}_{xx}(f)}} . \quad (6.223)$$

Der Vorteil des höheren Rechenaufwandes liegt zum einen in der verfügbaren Phaseninformation. Aber auch der Einfluss der überlagerten Störung $n(t)$ ist im vorliegenden Fall wesentlich geringer, wie Abb. 6.23 zeigt.

6.5.3 Dynamische Systemidentifikation

In vielen Fällen reicht eine einmalige Identifikation eines Systems nicht aus. Dies ist vor allem dann der Fall, wenn sich die Parameter, die das dynamische Verhalten des Systems bestimmen, zeitlich ändern. Möchte man ein solches System messtechnisch überwachen, so muss man die Identifikation während des laufenden Betriebs durchführen, da man nicht ständig das System für Messzwecke stilllegen kann. Bei der Identifikation während des Echtzeitbetriebs wird dem Eingangssignal $u(t)$ ein unkorreliertes weißes Rauschen $n(t)$ als Testsignal überlagert (Abb. 6.24). Die Amplituden a des Testsignals betragen im Vergleich zum Nutzsignal $u(t)$ nur wenige Prozent.

Abbildung 6.24. Identifikation der Übertragungsfunktion $G(f)$ während des Echtzeitbetriebs.

Das Testsignal $n(t)$ und das Ausgangssignal $y(t)$ werden gemessen und daraus das Periodogramm $\hat{S}_{nn}(f)$ und das Kreuzleistungsdichtespektrum $\hat{S}_{yn}(f)$ geschätzt. Mit den Fourier-Transformierten der Eingangssignale

$$X(f) = U(f) + N(f) \quad (6.224)$$

erhält man die Ausgangsfunktion

$$Y(f) = G(f) U(f) + G(f) N(f) \quad (6.225)$$

und daraus die geschätzte Kreuzleistungsdichte

$$\hat{S}_{yn}(f) = G(f) \hat{S}_{un}(f) + G(f) \hat{S}_{nn}(f). \quad (6.226)$$

Durch Mittelung über mehrere Einzelmessungen ergibt sich die verbesserte Schätzung:

$$\overline{\hat{S}_{yn}(f)} = \underbrace{\mathbb{E}\{G(f) \hat{S}_{un}(f)\}}_{\approx 0} + \overline{G(f) \hat{S}_{nn}(f)}. \quad (6.227)$$

Damit erhält man für die geschätzte Übertragungsfunktion

$$\hat{G}(f) = \frac{\overline{G(f) \hat{S}_{nn}(f)}}{\overline{\hat{S}_{nn}(f)}} = \frac{\overline{\hat{S}_{yn}(f)}}{\overline{\hat{S}_{nn}(f)}} = \frac{\overline{\hat{S}_{yn}(f)}}{a^2}. \quad (6.228)$$

Als Testsignale eignen sich beispielsweise PRBS-Folgen (Abschn. 6.4.2).

6.6 Wiener-Filter

In Abschn. 5.2 wurde diskutiert, wie Messsysteme den zeitlichen Verlauf eines Messsignals $u(t)$ beeinflussen. Ziel der Überlegungen war es, das Messsystem so zu dimensionieren, dass das Fehlersignal $e(t)$ zwischen Systemeingang und Systemausgang möglichst klein wird. Dies führte auf das Problem der **Parameteroptimierung**, die konstruktiv frei wählbaren Para-

meter einer Übertragungsfunktion so zu wählen, dass die Verfälschung des Messsignals möglichst gering wird.

Einen anderen Ansatz verfolgt die **Signalschätzung**. Diese befasst sich mit der Rekonstruktion eines Signals, das durch die Dynamik eines nicht-idealnen (Mess-)Systems sowie durch überlagerte Störungen verfälscht wurde. Unter Verwendung statistischer Eigenschaften des Signals und der verschiedenen Störquellen wird aus dem verfälschten Signal das ursprüngliche Messsignal soweit wie möglich wiederhergestellt. Das **Wiener-Filter** ist ein **Optimalfilter** dafür. Das Problem tritt in vielen technischen Anwendungen auf:

- **Messtechnik:** Der Zeitverlauf des Messsignals $u(t)$ soll erfasst werden. Da jede Messung notwendigerweise den zu messenden Vorgang beeinflusst, liefert die Messeinrichtung nicht genau das Originalsignal $u(t)$, sondern eine verfälschte Version: die angezeigte Größe $y(t)$. Zusätzlich treten Messfehler auf, die durch Rauschsignale beschrieben werden können.
- **Signalübertragung:** Wird ein Signal $u(t)$ zu einem anderen Ort hin übertragen, so wird es auf dem Weg dorthin durch die nichtidealen Eigenschaften des Übertragungsmediums verändert. Zusätzlich kommen noch Störungen auf dem Übertragungsweg hinzu, so dass das empfangene Signal $y(t)$ nicht mehr dem Sendesignal $u(t)$ entspricht. Dies ist ein klassisches Problem in der Nachrichtentechnik und der Automatisierungstechnik.
- **Speicherung:** Um ein Signal $u(t)$ zu speichern, muss es so verändert werden, dass es den Erfordernissen des Datenträgers entspricht (z. B. Abtastung und Quantisierung; vgl. Kap. 7). Das zu einem späteren Zeitpunkt wieder ausgelesene Signal enthält Störungen, die meist auch durch Rauschen modelliert werden können.

6.6.1 Signalmodell

Das in Abb. 6.25 dargestellte Signalmodell veranschaulicht das Problem der **Signalrekonstruktion**. Das zu messende Signal $u(t)$ ist nicht direkt zugänglich. Stattdessen kann nur das Signal $y(t)$ erfasst werden, das die Einflüsse des dynamischen Systems $G(f)$ und eine überlagerte Störung $n(t)$ aufweist. Es wird angenommen, dass die Kreuzleistungsdichte $S_{yu}(f)$ zwischen dem Ausgang $y(t)$ und dem Messsignal $u(t)$ bekannt ist; beispielsweise kann diese durch eine Messung mit bekanntem Eingangssignal ermittelt werden. Das Kreuzleistungsdichtespektrum $S_{yu}(f)$ beschreibt den Einfluss der Verfälschungen, denen das Originalsignal unterliegt. Das Leistungsdichtespek-

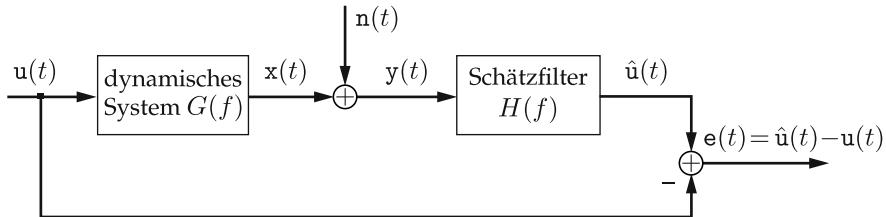


Abbildung 6.25. Signalrekonstruktion mittels eines Schätzfilters.

trum $S_{yy}(f)$ des gemessenen Signals $y(t)$ kann jederzeit messtechnisch bestimmt werden und ist damit auch bekannt.

Das Schätzfilter $H(f)$ in Abb. 6.25 soll nun so entworfen werden, dass sein Ausgangssignal $\hat{u}(t)$ dem Originalsignal $u(t)$ möglichst nahe kommt. Dazu wird die Leistung des Fehlersignals $e(t) = \hat{u}(t) - u(t)$ minimiert:

$$E\{e^2(t)\} = E\{(\hat{u}(t) - u(t))^2\} \rightarrow \min . \quad (6.229)$$

Das resultierende Filter heißt *Optimalfilter* oder, nach seinem Erfinder, *Wiener-Filter* [56].

Anmerkung

Die Bedingung in (6.229) wurde in ähnlicher Form bereits an anderer Stelle angewendet. Das Kriterium „quadratisches Fehlerintegral“ in Abschn. 5.4.5 basiert nach (5.168) auf der Minimierung des folgenden Gütemaßes:

$$Q = \int_{-\infty}^{\infty} e^2(t) dt \rightarrow \min . \quad (6.230)$$

Für Energiesignale entspricht der Ausdruck gerade der quadratischen Norm oder der Autokorrelationsfunktion $r_{ee}^E(\tau = 0)$ an der Stelle $\tau = 0$. Damit lässt sich das Kriterium auch in folgender Form schreiben:

$$Q = \int_{-\infty}^{\infty} e^2(t) dt = \int_{-\infty}^{\infty} S_{ee}(f) df \quad (6.231)$$

$$= \int_{-\infty}^{\infty} E(f) E(-f) df \rightarrow \min . \quad (6.232)$$

Eine weitere verwandte Anwendung ist die Approximation von Kennlinien mittels des Least-Squares-Schätzers (Abschn. 2.1.2).

6.6.2 Herleitung des Wiener-Filters

Obwohl das Eingangssignal $u(t)$ und damit auch der Rekonstruktionsfehler $e(t)$ nicht bekannt sind, lässt sich dennoch die Fehlerleistung statistisch beschreiben. Aus Satz 6.5 ist bekannt, dass die Fehlerleistung $E\{e^2(t)\}$ mit Hilfe des Leistungsdichtespektrums $S_{ee}(f)$ bestimmt werden kann. Für die Leistungsdichte des Rekonstruktionsfehlers erhält man mit (6.179) für die Summe von Zufallsgrößen:

$$S_{ee}(f) = S_{uu}(f) - S_{uu}(f) - S_{uu}(f) + S_{uu}(f). \quad (6.233)$$

Mit den Zusammenhängen (6.189)–(6.191) kann das Leistungsdichtespektrum $S_{ee}(f)$ in Abhängigkeit der Übertragungsfunktion $H(f)$ des gesuchten Schätzfilters dargestellt werden:

$$S_{ee}(f) = S_{yy}(f) |H(f)|^2 - S_{yu}(f) H(f) - S_{uy}(f) H^*(f) + S_{uu}(f). \quad (6.234)$$

Das Schätzfilter $H(f)$, welches zur kleinsten Fehlerleistung führt, erhält man durch Differentiation von (6.234) nach $H(f)$. Dabei ist zu beachten, dass sowohl die Kreuzleistungsdichtespektren als auch $H(f)$ im Allgemeinen komplexwertig sind. Die Ableitung muss daher getrennt nach Betrag und Phase erfolgen. Zur Vereinfachung wird das Argument (f) bis auf Weiteres weggelassen. Mit $H = |H| e^{j\varphi}$ erhält man:

$$S_{ee} = S_{yy} |H|^2 - S_{yu} |H| e^{j\varphi} - S_{uy} |H| e^{-j\varphi} + S_{uu}. \quad (6.235)$$

Die Ableitung von S_{ee} nach dem Betrag $|H|$ der Übertragungsfunktion ergibt

$$\frac{dS_{ee}}{d|H|} = 2 |H| S_{yy} - S_{yu} e^{j\varphi} - S_{uy}^* e^{-j\varphi} \quad (6.236)$$

$$= 2 |H| S_{yy} - 2 \Re\{S_{yu} e^{j\varphi}\} \stackrel{!}{=} 0. \quad (6.237)$$

Den optimalen Amplitudengang $|H|$ des Schätzfilters erhält man durch Nullsetzen der Ableitung:

$$|H| = \frac{\Re\{S_{yu} e^{j\varphi}\}}{S_{yy}}. \quad (6.238)$$

Zur Ermittlung des Phasenganges wird S_{ee} nach der Phase φ abgeleitet:

$$\frac{dS_{ee}}{d\varphi} = -j|H| e^{j\varphi} S_{yu} + j|H| e^{-j\varphi} S_{yu}^* = -j2|H| \Im\{S_{yu} e^{j\varphi}\} \stackrel{!}{=} 0. \quad (6.239)$$

Der optimale Phasengang folgt aus

$$\Im\{S_{yu} e^{j\varphi}\} = \Im\{|S_{yu}| e^{j(\arg\{S_{yu}\} + \varphi)}\} = 0. \quad (6.240)$$

Der Imaginärteil wird gleich null, wenn

$$\varphi = -\arg\{S_{yu}\} \quad (6.241)$$

gilt. Der Phasengang des Wiener-Filters ist also gleich dem negativen Phasengang der Kreuzleistungsdichte S_{yu} . Damit kann auch der optimale Amplitudengang bestimmt werden, der mit (6.241) wie folgt resultiert:

$$|H| = \frac{\Re\{S_{yu} e^{j\varphi}\}}{S_{yy}} = \frac{\Re\{|S_{yu}| e^{j(\arg\{S_{yu}\} + \varphi)}\}}{S_{yy}} = \frac{|S_{yu}|}{S_{yy}}. \quad (6.242)$$

Mit den Gln. (6.241) und (6.242) erhält man die gesuchte Übertragungsfunktion des optimalen Rekonstruktionsfilters zu

$$H(f) = \frac{S_{yu}^*(f)}{S_{yy}(f)} = \frac{S_{uy}(f)}{S_{yy}(f)}, \quad (6.243)$$

wobei die Identität $S_{yu}^*(f) = S_{uy}(f)$ nach (6.161) reelle Prozesse voraussetzt.

6.6.3 Wiener-Filter bei linearer Verzerrung und additivem Rauschen

Bei der Herleitung der allgemeinen Form (6.243) des Wiener-Filters wurden weder Annahmen über die Art der Signalverfälschung noch über die Störung getroffen. Geht man davon aus, dass die Messgröße $u(t)$ gemäß Abb. 6.25 durch ein LTI-System mit der Übertragungsfunktion $G(f)$ verzerrt und zusätzlich von einem additiven Rauschen $n(t)$ gestört wurde, so kommt man auf eine Beschreibung, die für viele reale Messsysteme zutrifft. Das Rauschsignal $n(t)$ fasst dabei alle externen Störeinflüsse zusammen und sei nicht mit der Messgröße $u(t)$ korreliert. Der Ausgang des LTI-Systems sei mit $x(t)$ bezeichnet.

Die Leistungsdichten $S_{uy}(f)$ und $S_{yy}(f)$ in (6.243) können nun durch das Modell genauer beschrieben werden. Mit (6.182) erhält man für das Leistungsdichtespektrum der Summe der unkorrelierten Prozesse $x(t)$ und $n(t)$:

$$S_{yy}(f) = S_{xx}(f) + S_{nn}(f) = S_{uu}(f) |G(f)|^2 + S_{nn}(f). \quad (6.244)$$

Für die Kreuzleistungsdichte folgt

$$S_{uy}(f) = S_{uu}(f) G^*(f) + \underbrace{S_{un}(f)}_{=0} . \quad (6.245)$$

Damit erhält man das Wiener-Filter zur Signalrekonstruktion bei linearer Verzerrung und additivem Rauschen:

$$H(f) = \frac{S_{uy}(f)}{S_{yy}(f)} = \frac{S_{uu}(f) G^*(f)}{S_{uu}(f) |G(f)|^2 + S_{nn}(f)} . \quad (6.246)$$

Die Kenntnis der verzerrenden Übertragungsfunktion $G(f)$ erlaubt es somit, das Kreuzleistungsdichtespektrum $S_{uy}(f)$ auf die Autoleistungsdichtespektren des Messsignals $S_{uu}(f)$ und der Rauschstörung $S_{nn}(f)$ zurückzuführen. Der Frequenzgang $G(f)$ kann dabei explizit bekannt sein oder durch eine Systemidentifikation bestimmt werden.

Durch Umformung von (6.246) lässt sich das Wiener-Filter als Reihenschaltung eines **inversen Filters** $G^{-1}(f)$ und eines adaptiven Störungsumterdrückungsfilters darstellen:

$$H(f) = \frac{1}{G(f)} \cdot \frac{|G(f)|^2}{|G(f)|^2 + \underbrace{\frac{S_{nn}(f)}{S_{uu}(f)}}_{\text{SNR}^{-1}(f)}} . \quad (6.247)$$

Wenn keine Störungen vorhanden sind, geht $H(f)$ in die Übertragungsfunktion des inversen Filters $G^{-1}(f)$ über. Sind dagegen die Störungen stark gegenüber dem Nutzsignal, so „schließt“ das adaptive Filter bei den entsprechenden Frequenzen. Auf diese Weise wird eine Verstärkung der Störanteile an den Nullstellen von $G(f)$ verhindert.

Beispiel 6.15 (Rauschunterdrückung): Einem Messsignal $u(t)$ mit Tiefpasscharakteristik sei eine rauschartige, unkorrelierte Störung $n(t)$ additiv überlagert. Gesucht wird ein Filter, welches das Rauschen optimal unterdrückt, ohne dabei das Nutzsignal zu stark zu verfälschen. Für die Leistungsdichtespektren gelte

$$S_{uu}(f) = \frac{U^2}{1 + (b \cdot 2\pi f)^2} \quad \text{und} \quad S_{nn}(f) = a^2 . \quad (6.248)$$

Das gesuchte Optimalfilter kann mit (6.246) berechnet werden. Da im vorliegenden Fall keine Verzerrung durch ein LTI-System erfolgt, gilt

$$G(f) = 1 . \quad (6.249)$$

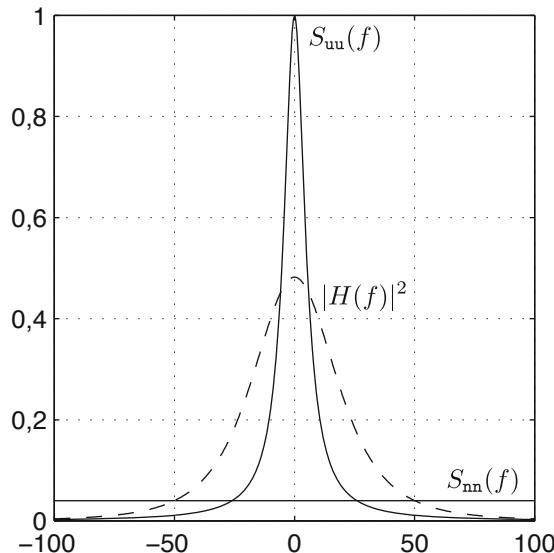


Abbildung 6.26. Leistungsdichtespektren von Nutzsignal $S_{uu}(f)$ und Störung $S_{nn}(f)$ sowie Amplitudengang des Optimalfilters $|H(f)|^2$.

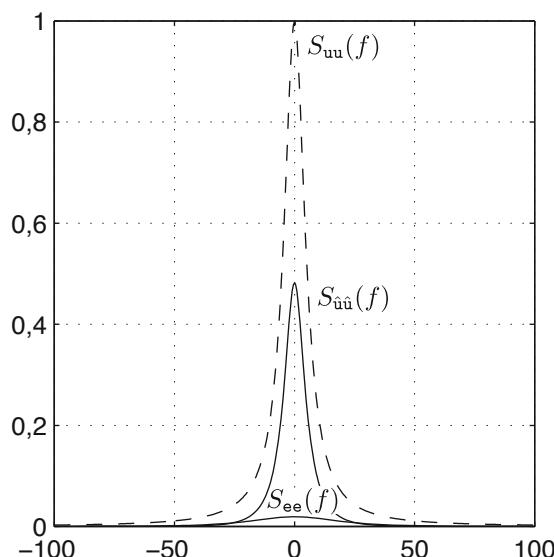


Abbildung 6.27. Leistungsdichtespektren des Nutzsignals $S_{uu}(f)$ sowie der verbleibenden Störung $S_{ee}(f)$ und des Signals $S_{\hat{u}\hat{u}}(f)$ am Ausgang des Optimalfilters.

Damit ergibt sich die Übertragungsfunktion des Wiener-Filters zu

$$H(f) = \frac{S_{uu}(f)}{S_{uu}(f) + S_{nn}(f)} = \frac{U^2}{U^2 + a^2 + (b \cdot a \cdot 2\pi f)^2}. \quad (6.250)$$

Abbildung 6.26 zeigt die Leistungsdichtespektren von Nutzsignal und Störung sowie den Amplitudengang $|H(f)|$ des Optimalfilters. Das Filter lässt die Frequenzen des Nutzsignals weitgehend durch und unterdrückt die Störung außerhalb des Nutzsignalbereiches. Abbildung 6.27 zeigt die Leistungsdichten beider Signalanteile nach der Optimalfilterung im Vergleich zur Leistungsdichte des ungefilterten Nutzsignals. ■

Das Wiener-Filter $H(f)$ ist allgemein aksausal, weshalb sich die Filterung eines Messsignals $y(t)$ nicht in Echtzeit realisieren lässt. Für die Filterung in Echtzeit stehen nur vergangene Werte zur Verfügung, so dass nur der kausale Teil der Impulsantwort berücksichtigt werden kann. Die Impulsantwort folgt aus der Übertragungsfunktion $H(f)$ durch inverse Fourier-Transformation:

$$h(t) = \mathcal{F}^{-1}\{H(f)\}. \quad (6.251)$$

Durch Multiplikation mit der Sigma-Funktion erhält man ein kausales Filter

$$h_{\text{kaus}}(t) = h(t) \cdot \sigma(t), \quad (6.252)$$

welches für die Echtzeitanwendung geeignet ist.

Beispiel 6.16 (Kausales Wiener-Filter): Für das aksuale Störungsunterdrückungsfilter (6.250) soll das entsprechende kausale Wiener-Filter berechnet werden. Dazu formt man die Übertragungsfunktion wie folgt um,

$$H(f) = k \frac{2\lambda}{\lambda^2 + (2\pi f)^2}, \quad (6.253)$$

$$k = \frac{U^2}{2ab\sqrt{U^2 + a^2}}, \quad (6.254)$$

$$\lambda = \sqrt{\frac{U^2}{a^2 b^2} + \frac{1}{b^2}}, \quad (6.255)$$

und erhält durch inverse Fourier-Transformation die Impulsantwort

$$h(t) = k \cdot e^{-\lambda|t|} \quad (6.256)$$

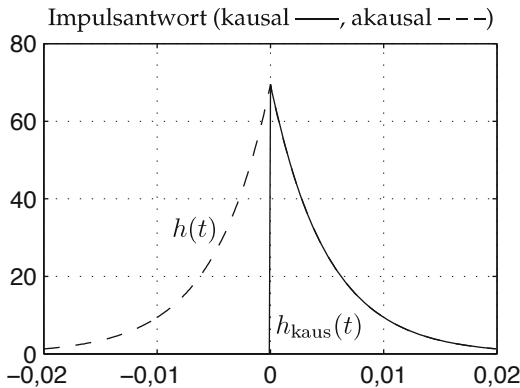


Abbildung 6.28. Impulsantworten $h(t)$ und $h_{\text{kaus}}(t)$ des acausalen bzw. des kausalen Wiener-Filters.

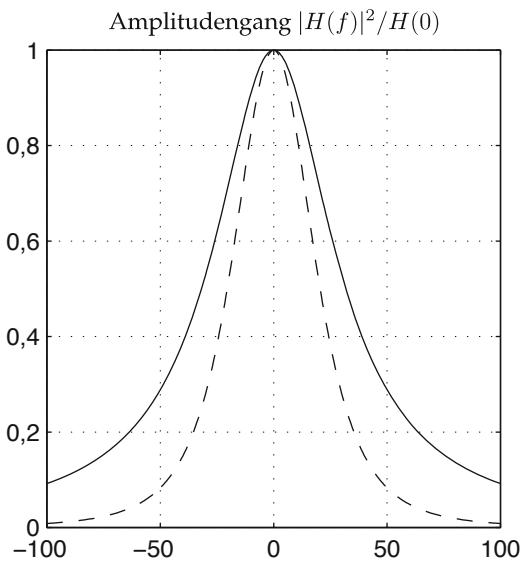


Abbildung 6.29. Normierte Amplitudengänge des kausalen (—) und des acausalen (---) Rauschunterdrückungsfilters.

des acausalen Systems. Setzt man für $t < 0$ die Impulsantwort zu null, so erhält man die kausale Impulsantwort (Abb. 6.28)

$$h_{\text{kaus}}(t) = k \cdot e^{-\lambda t} \cdot \sigma(t) . \quad (6.257)$$

Durch Fourier-Transformation erhält man schließlich den Frequenzgang des kausalen Filters:

$$H_{\text{kaus}}(f) = \mathcal{F}(h_{\text{kaus}}(t)) = k \cdot \frac{1}{\lambda + j2\pi f} . \quad (6.258)$$

Abbildung 6.29 zeigt die normierten Amplitudengänge der beiden Wiener-Filter. Der Durchlassbereich des kausalen Filters ist deutlich breiter

als der des akausalen Filters. Die Rauschunterdrückung gelingt nicht so gut, das Filter kann aber dafür in Echtzeitanwendungen eingesetzt werden. ■

Beispiel 6.17 (Wiener-Filter zur Bildrestauration): Abbildung 6.30 zeigt ein Beispiel zum Einsatz des Wiener-Filters im Kontext der Bildrestauration. Ausgangspunkt ist das gemessene Eingangsbild $y(x)$ in Abb. 6.30 (a), bei welchem bedingt durch eine unscharfe optische Abbildung $G(f)$ Spektralanteile abgeschwächt oder unterdrückt wurden. Dabei bezeichnet $\mathbf{x} = (x, y)^T$ den Ortsvektor und $\mathbf{f} = (f_x, f_y)^T$ den zugehörigen Ortsfrequenzvektor. Der Amplitudengang $|G(f)|$ der unscharfen optischen Abbildung ist rotationssymmetrisch und weist Nullstellen auf konzentrischen Kreisen auf; er ist in den Abbildungen 6.30 (c) und 6.30 (e) dargestellt [2]. Darüber hinaus enthält das Messsignal $y(x)$ ein additives Rauschen $n(x)$.

Neben der Leistungsdichte $S_{nn}(f)$ des Rauschens $n(x)$, die bei einem weißen Prozess konstant ist und im vorliegenden Fall als bekannt angenommen wurde, wird zur Berechnung der Übertragungsfunktion $H(f)$ des Wiener-Filters die Leistungsdichte $S_{uu}(f)$ des unverzerrten Messsignals $u(x)$ benötigt. Diese Größe ist in der Regel nicht verfügbar, weshalb sie im vorliegenden Fall aus dem Messsignal $y(x)$ geschätzt wurde. Dafür wurde das Periodogramm $\hat{S}_{yy}(f)$ berechnet und geglättet; anschließend wurde davon die bekannte Leistungsdichte des Rauschprozesses subtrahiert. Die damit nach (6.246) berechnete Übertragungsfunktion des Wiener-Filters ist in Abb. 6.30 (d) zu sehen. Das Ergebnis der Wiener-Filterung ist in Abb. 6.30 (b) dargestellt.

Zur besseren Interpretation des Wiener-Filters zeigt Abb. 6.30 (e) zwei horizontale Schnitte der Amplitudengänge von $G(f)$ und $H(f)$ durch den Ursprung des Ortsfrequenzbereiches. Beide Verläufe weisen zwar die gleichen Nullstellen auf. Allerdings führt das Filter eine höhere Verstärkung bei Frequenzen herbei, bei denen das Nutzsignal stark gegenüber der (im vorliegenden Fall konstanten) Störung ist. Entsprechend „schließt“ das Filter bei Frequenzen, bei denen die Störung stark gegenüber dem Nutzsignal ist.

Schließlich sei noch angemerkt, dass bei der Bewertung des Ergebnisses Abb. 6.30 (b) der Wiener-Filterung das visuelle Güteempfinden nicht mit dem Kriterium des quadratischen Fehlers konform ist, weshalb die Ergebnisse dieses Optimalfilters nicht zwangsläufig im Hinblick auf die visuelle Qualität optimal sein müssen. ■

(a) Realisierung des Messsignals $y(\mathbf{x})$ 

(b) Ergebnis der Wiener-Filterung

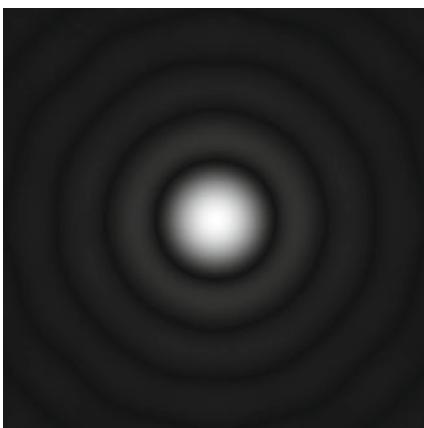
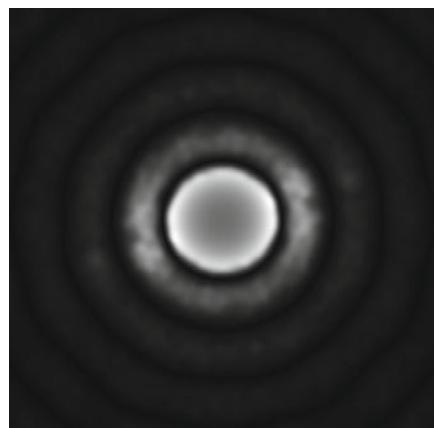
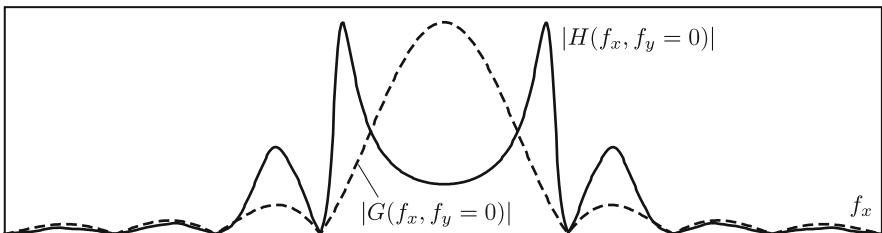
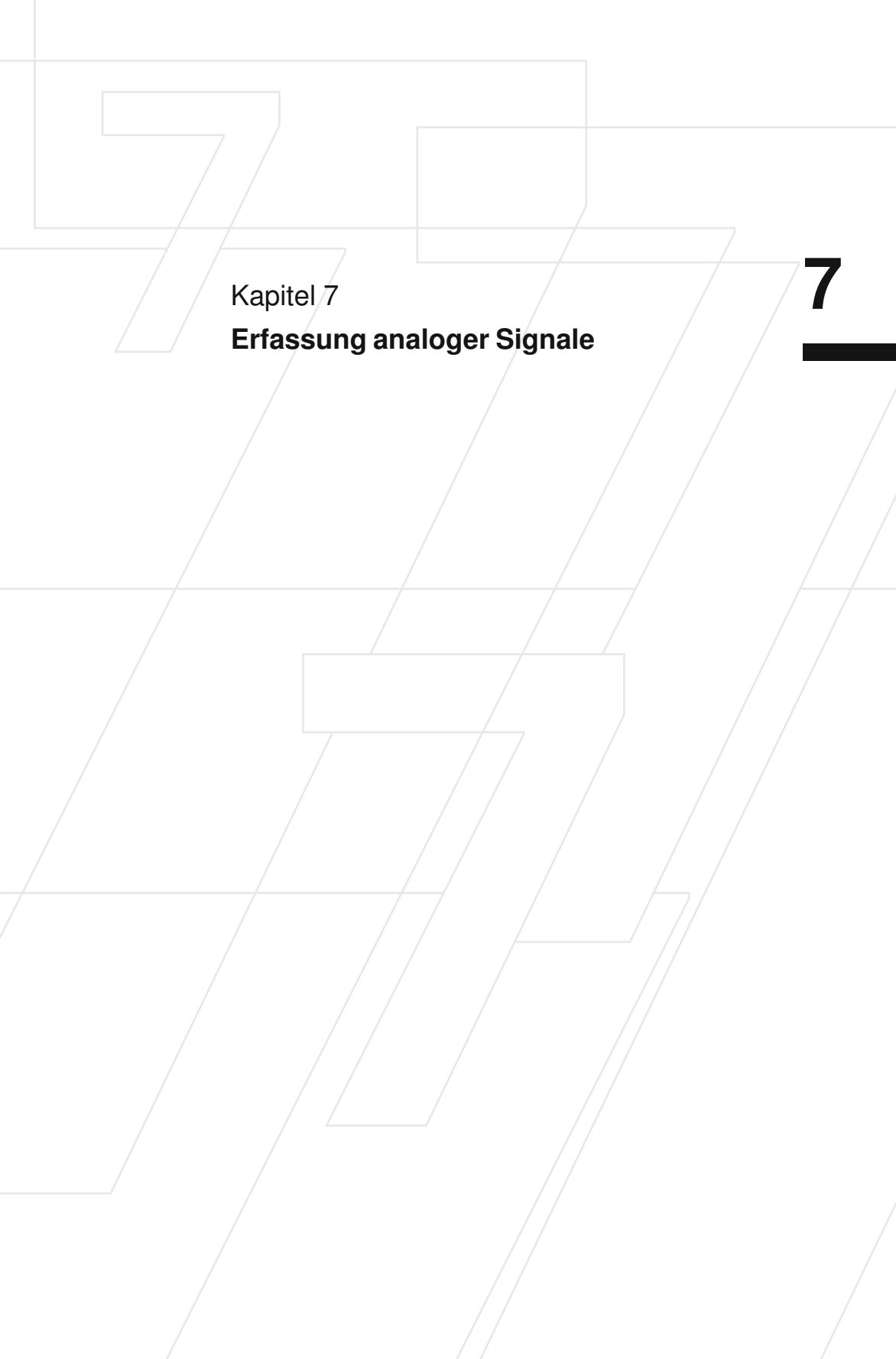
(c) Amplitudengang $|G(\mathbf{f})|$ (d) Amplitudengang $|H(\mathbf{f})|$ (e) Amplitudengänge von unscharfer Abbildung $|G(\mathbf{f})|$ und Wiener-Filter $|H(\mathbf{f})|$

Abbildung 6.30. Wiener-Filter zur Bildrestaurierung: (a) Realisierung des Messsignals $y(\mathbf{x})$ ($y(\mathbf{x}) = k \cdot \text{rect}(\|\mathbf{x}\|/9)$; $n(\mathbf{x})$: gaußverteiltes weißes Rauschen mit $\sigma^2 = 25$); (b) Ergebnis der Restaurierung mit dem Wiener-Filter; (c) Amplitudengang $|G(\mathbf{f})|$ der unscharfen optischen Abbildung; (d) Amplitudengang $|H(\mathbf{f})|$ des Wiener-Filters; (e) Vergleich der Amplitudengänge von unscharfer Abbildung $|G(\mathbf{f})|$ und Wiener-Filter $|H(\mathbf{f})|$.



Kapitel 7

Erfassung analoger Signale

7

7 Erfassung analoger Signale

| | | |
|---------|---|-----|
| 7.1 | Abtastung | 318 |
| 7.1.1 | Abtasttheorem | 318 |
| 7.1.2 | Bandüberlappungsfehler (<i>Aliasing</i>) | 320 |
| 7.1.3 | Anti-Aliasing-Filter | 322 |
| 7.1.4 | Mittelwertbildung bei endlicher Abtastdauer | 325 |
| 7.1.5 | Zeitliche Abtastfehler | 328 |
| 7.2 | Quantisierung | 333 |
| 7.2.1 | Wahrscheinlichkeitsdichte der Signalwerte | 336 |
| 7.2.2 | Amplitudendichte der Fourier-Reihe | 337 |
| 7.2.3 | Quantisierungstheorem | 340 |
| 7.2.4 | Wahrscheinlichkeitsdichte des Quantisierungsfehlers . | 345 |
| 7.2.5 | Signal-Rausch-Verhältnis infolge der Quantisierung . | 347 |
| 7.2.5.1 | Sinusförmige Signale | 348 |
| 7.2.5.2 | Signale mit gleichverteilten Amplituden | 348 |
| 7.2.6 | Optimale Quantisierung | 349 |
| 7.2.7 | Minimierung des relativen Quantisierungsfehlers . | 349 |
| 7.2.8 | Dithering | 352 |
| 7.3 | Analog-Digital-Umsetzer | 357 |
| 7.3.1 | A/D-Nachlaufumsetzer | 357 |
| 7.3.2 | A/D-Umsetzer mit sukzessiver Approximation | 359 |
| 7.3.3 | Integrierende A/D-Umsetzer | 360 |
| 7.3.4 | Delta-Sigma-Umsetzer | 363 |
| 7.3.4.1 | Delta-Sigma-Modulator | 363 |
| 7.3.4.2 | Rauschformung (<i>noise shaping</i>) | 365 |
| 7.3.4.3 | Digitalfilter | 368 |
| 7.3.4.4 | Stationäres Verhalten | 369 |
| 7.3.5 | Ratiometrische Messung | 372 |
| 7.4 | Digital-Analog-Umsetzer | 372 |
| 7.4.1 | Parallele D/A-Umsetzer | 373 |
| 7.4.1.1 | D/A-Umsetzer mit dyadisch gestuften Widerständen | 373 |
| 7.4.1.2 | D/A-Umsetzer mit $R/2R$ -Kettenleiternetzwerk | 373 |
| 7.4.2 | Serielle D/A-Umsetzer | 375 |

7 Erfassung analoger Signale

In der modernen Messtechnik werden Signale meist rechnergestützt ausgewertet. Dazu müssen die zeitkontinuierlichen Signale eines analogen Messwertaufnehmers in eine für den Rechner geeignete Form überführt werden. Bei diesem Prozess der sogenannten *digitalen Messdatenerfassung* werden die Signale implizit einer Signalverarbeitung unterzogen. Der mit einer Messdatenerfassung zwangsläufig verbundenen Signalverarbeitung ist sich der Ingenieur zumeist weniger bewusst, obwohl gerade die Bereitstellung der Messdaten das gesamte Informationsverarbeitungssystem in entscheidender Weise beeinflusst. Im Folgenden werden die wichtigsten Grundbausteine dieser Signalverarbeitung behandelt. Für eingehendere Darstellungen sei auf die Literatur verwiesen [34, 42, 49].

Da im Rechner nur eine endliche Anzahl von Messwerten gespeichert werden kann, muss das zeitkontinuierliche Signal des Messwertaufnehmers zu diskreten Zeitpunkten abgetastet werden (engl. *sampling*). Der wahre Signalverlauf wird im Rechner, ähnlich wie bei der Darstellung einer Kennlinie anhand von Messpunkten in Kap. 2, durch eine endliche Anzahl zeitdiskreter Stützstellen repräsentiert (Abb. 7.1). Durch die **Abtastung** selbst oder durch integrierende Verfahren bei der Messwertaufnahme werden die Messsignale verfälscht. Die dabei auftretenden Effekte werden im folgenden Abschnitt besprochen. Daneben können die Amplitudenwerte der abgetasteten Stützstellen nicht beliebig genau erfasst werden. Vielmehr sind tech-

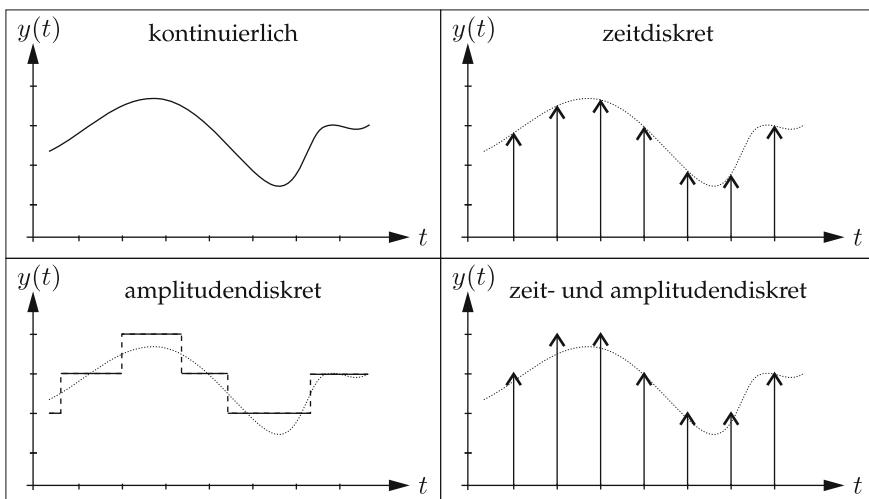


Abbildung 7.1. Kontinuierliche und diskrete Signale.

nisch nur diskrete Amplitudenstufen realisierbar. Auch im Rechner kann nur mit diskreten Zahlen gearbeitet werden, die meist dualcodiert werden. Dies führt bei der Messwerterfassung zu einer **Quantisierung** der Signalamplituden. Die damit einhergehenden Signalverfälschungen werden in Abschn. 7.2 diskutiert. Technisch wird die digitale Signalverarbeitung mit Hilfe von **Analog-Digital-Umsetzern** realisiert, die sowohl die Zeitdiskretisierung als auch die Amplitudenquantisierung verwirklichen. Die technische Realisierung dieses Umsetzverfahrens wird in Abschn. 7.3 erläutert. Umgekehrt werden **Digital-Analog-Umsetzer** eingesetzt, um digitale Zahlenwerte, wie sie z. B. in einem Rechner abgelegt werden, in analoge Signale zu überführen (Abschn. 7.4).

7.1

7.1 Abtastung

Ein kontinuierliches Signal zeichnet sich dadurch aus, dass man zu jedem beliebigen Zeitpunkt die Signalamplitude angeben kann. Bei technisch realisierbaren Systemen sind die Signalamplituden stetige Funktionen.

Bei der Abtastung werden nur die Signalwerte zu diskreten Zeitpunkten erfasst, d. h. die Signalwerte zwischen den Zeitpunkten gehen verloren. Unter bestimmten Voraussetzungen, die im **Abtasttheorem** festgelegt sind, kann man die Signalwerte zwischen den festen Abtastzeitpunkten wieder vollständig rekonstruieren [42]. Nur in diesem Fall entstehen durch die Abtastung keine Signalfehler.

7.1.1 Abtasttheorem

Bei der **idealen Abtastung** wird ein zeitkontinuierliches Signal $y(t)$ zu äquidistanten Zeitpunkten $n t_A$ in das zeitdiskrete Signal $y_*(t)$ überführt. Dies geschieht durch eine Multiplikation mit einer Impulsreihe:

$$y_*(t) = y(t) \cdot \sum_{n=-\infty}^{\infty} \delta(t - n t_A). \quad (7.1)$$

Die Fourier-Transformierte $Y_*(f)$ des abgetasteten Signals $y_*(t)$ wird durch eine Faltung mit der Fourier-Transformierten der Impulsreihe – die ebenfalls eine Impulsreihe ergibt – im Frequenzbereich beschrieben:

$$Y_*(f) = Y(f) * \frac{1}{t_A} \sum_{k=-\infty}^{\infty} \delta\left(f - \frac{k}{t_A}\right) \quad (7.2)$$

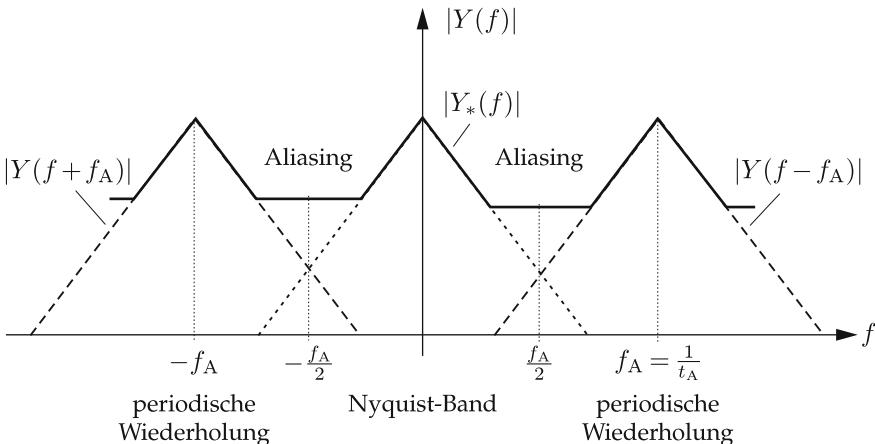


Abbildung 7.2. Spektrum einer abgetasteten Zeitfunktion bei Nichteinhaltung des Abtasttheorems (Aliasing).

$$= \int_{-\infty}^{\infty} Y(\nu) \frac{1}{t_A} \sum_{k=-\infty}^{\infty} \delta\left(f - \nu - \frac{k}{t_A}\right) d\nu \quad (7.3)$$

$$= \frac{1}{t_A} \cdot \sum_{k=-\infty}^{\infty} Y\left(f - \frac{k}{t_A}\right). \quad (7.4)$$

Man erkennt sofort, dass das Spektrum $Y(f)$ des zeitkontinuierlichen Signals $y(t)$ periodisch in Abständen von $1/t_A = f_A$ wiederholt wird (Abb. 7.2). Um eine eindeutige Rekonstruktion des ursprünglichen Signals $y(t)$ zu ermöglichen, darf die Bandbreite B des Signals $y(t)$ nicht größer als die **Abtastfrequenz** f_A sein, damit es zu keinen Überlappungen (Aliasing) der periodischen Wiederholungen kommt. Dies ist die Aussage des Abtasttheorems.

Satz 7.1: Abtasttheorem

7.1

Ein bandbegrenztes kontinuierliches Signal

$$y(t) \quad \circ \bullet \quad Y(f) \quad \text{mit} \quad Y(f) = 0 \quad \text{für} \quad |f| \geq \frac{B}{2} \quad (7.5)$$

lässt sich fehlerfrei aus seinen äquidistanten Abtastwerten $y(n t_A)$, $n \in \mathbb{Z}$, rekonstruieren, falls für die Abtastfrequenz $f_A = 1/t_A$ die Bedingung

$$f_A \geq B \quad (7.6)$$

erfüllt ist. \diamond

Die **Rekonstruktion** des kontinuierlichen Signals erfolgt, indem man die periodischen Wiederholungen des Spektrums aus $Y_*(f)$ herausfiltert. Dies gelingt durch die Multiplikation des Fourier-Spektrums $Y_*(f)$ mit einem Rechteckfenster $r_{f_A}(f)$ der Breite f_A , was einer **idealen Tiefpassfilterung** entspricht. Die Multiplikation im Frequenzbereich korrespondiert mit einer Faltung im Zeitbereich, womit man das folgende rekonstruierte Zeitsignal erhält:

$$\hat{y}(t) = y_*(t) * \mathcal{F}^{-1}\{t_A \cdot r_{f_A}(f)\} \quad (7.7)$$

$$= y(t) \sum_{n=-\infty}^{\infty} \delta(t - n t_A) * \frac{\sin(\pi f_A t)}{\pi f_A t} \quad (7.8)$$

$$= \sum_{n=-\infty}^{\infty} y(n t_A) \cdot \frac{\sin(\pi f_A(t - n t_A))}{\pi f_A(t - n t_A)}. \quad (7.9)$$

Anmerkung

Für die praktische Anwendung ist zu beachten, dass das Abtasttheorem eine theoretische Aussage liefert. Daher müssen die Fehler abgeschätzt werden, die in realen Anwendungen entstehen. Diese sind im Wesentlichen:

- Bandüberlappungsfehler (*Aliasing*) infolge nicht bandbegrenzter Signale, da vollständig bandbegrenzte Signale in der Praxis nicht existieren (siehe Abschn. 7.1.2).
- Eine ideale (d. h. unendlich kurze) Abtastung mit Dirac-Impulsen ist technisch nicht realisierbar. Vielmehr wird eine endliche Zeit benötigt, um den Abtastwert zu gewinnen. Dies führt zu Fehlern im Signalspektrum (siehe Abschn. 7.1.4).
- Zeitliche Abtastfehler (*Jitter*) bewirken eine Abweichung von der äquidistanten Abtastung. Dadurch kommt es ebenfalls zu einer Verfälschung des Signals (siehe Abschn. 7.1.5).
- Der Fehler infolge der Verwendung eines endlichen Beobachtungsfensters T_0 , da es in jeder realen Anwendung nur gelingt, einen zeitlich begrenzten Ausschnitt des Signals zu berücksichtigen. Dieses als **Leckefekt** bezeichnete Phänomen wird ausführlich in [42] behandelt.

⌚ 7.1.2 Bandüberlappungsfehler (*Aliasing*)

Bei der Berechnung des Spektrums $Y_*(f)$ des abgetasteten Signals $y_*(t)$ in (7.4) wurde das Spektrum als Summe der um ganzzahlige Vielfache der Ab-

tastfrequenz f_A verschobenen Spektren $Y(f)$ dargestellt:

$$Y_*(f) = f_A \cdot \sum_{k=-\infty}^{\infty} Y(f - k \cdot f_A). \quad (7.10)$$

Besitzt das Spektrum $Y(f)$ Anteile für Frequenzen $|f| \geq f_A/2$, so kommt es in der Regel zu spektralen Überlappungen (*Aliasing*). Ist das Spektrum $Y(f)$ bandbegrenzt, so kann es zu Aliasing kommen, wenn die Abtastfrequenz f_A zu klein gewählt wurde. Aus Abb. 7.2 erkennt man, dass durch eine geeignete Wahl von f_A der Aliasing-Fehler vermieden werden könnte.

Ist das Abtasttheorem nicht erfüllt, d. h. ist das Spektrum des abzutastenden Signals $y(t)$ nicht bandbegrenzt, oder ist die Abtastfrequenz f_A nicht größer als das Doppelte der Frequenz des größten Spektralanteils in $y(t)$, so kommt es zu spektraler Überlappung. Eine eindeutige Zuordnung der Spektralanteile ist dann nicht mehr möglich.

Beispiel 7.1 (Aliasing bei einer harmonischen Schwingung): Gegeben sei eine harmonische Schwingung

$$y(t) = \sin(2\pi \cdot 250 \text{ Hz} \cdot t) \quad (7.11)$$

der Frequenz $f = 250 \text{ Hz}$, die mit einer Abtastfrequenz von

$$f_A = 200 \text{ Hz}, \quad f_A = \frac{1}{t_A}, \quad (7.12)$$

abgetastet werden soll (Abb. 7.3). Hierbei ist das Abtasttheorem verletzt, weshalb nach der Abtastung eine eindeutige Zuordnung der Spektralan-

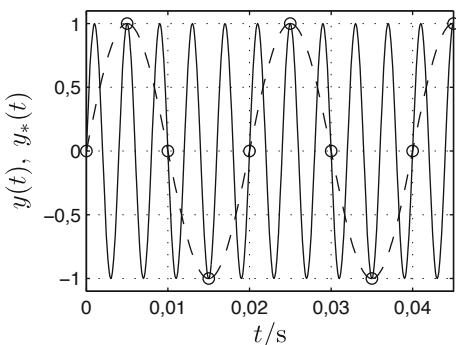


Abbildung 7.3. Harmonische Schwingung der Frequenz $f = 250 \text{ Hz}$. Bei Abtastung mit einer Abtastfrequenz von 200 Hz lässt sich das Abtastsignal $y_*(f)$ nicht von einer abgetasteten 50 Hz-Schwingung (gestrichelter Verlauf) unterscheiden.

teile nicht mehr möglich ist, wie die folgende Rechnung zeigt:

$$y_*(t) = y(n t_A) = \sin\left(2\pi \frac{250 \text{ Hz}}{200 \text{ Hz}} \cdot n\right) \quad (7.13)$$

$$= \sin\left(2\pi\left(1 + \frac{1}{4}\right)n\right) = \sin\left(\frac{\pi}{2}n + 2\pi n\right). \quad (7.14)$$

Wegen der 2π -Periodizität des Sinus lässt sich $y_*(t)$ nicht mehr von einer 50 Hz-Schwingung unterscheiden. Durch Verletzung des Abtasttheorems wird fälschlicherweise eine Schwingung

$$\sin\left(2\pi \frac{50 \text{ Hz}}{200 \text{ Hz}} \cdot n\right) = \sin\left(\frac{\pi}{2}n\right) = \sin\left(\frac{\pi}{2}n + 2\pi n\right) \quad (7.15)$$

der Frequenz 50 Hz vermutet. ■

Zur Verhinderung von Aliasing gibt es zwei Möglichkeiten:

- Bei nicht bandbegrenzten Signalen muss man vor der Abtastung mit Hilfe eines analogen Tiefpasses (Anti-Aliasing-Filter) Frequenzanteile größer oder gleich $f_A/2$ herausfiltern.
- Bei bandbegrenzten Signalen kann alternativ die Abtastfrequenz f_A größer als das Doppelte der höchsten im Signal vorkommenden Frequenz gewählt werden. Lässt sich die Abtastfrequenz nicht beliebig erhöhen, so muss man wie unter Punkt 1 mit Hilfe eines Tiefpasses Frequenzanteile größer oder gleich $f_A/2$ herausfiltern.

7.1.3 Anti-Aliasing-Filter

Zur Erfüllung des Abtasttheorems werden an ein praktisches Anti-Aliasing-Filter hohe Forderungen gestellt. Im Bereich bis zur halben Abtastfrequenz (Nyquist-Band) soll das Spektrum möglichst nicht verändert werden. Oberhalb von $f_A/2$ sollen alle Spektralanteile vollständig verschwinden. Der Amplitudengang eines derartigen idealen Tiefpasses ist in der Praxis nicht zu verwirklichen. Statt dessen muss man sich in der Praxis mit steilflankigen Filtern hoher Ordnung begnügen. Im Folgenden sollen die Aliasing-Fehler untersucht werden, die durch die Verwendung nichtideal er, aber dafür realisierbarer Anti-Aliasing-Filter entstehen.

Beispiel 7.2 (Aliasing-Fehler bei Filter 1. Ordnung): Gegeben sei ein zeitkontinuierliches, stochastisches Eingangssignal $x(t)$ mit konstanter Leistungsdichte (weißes Rauschen):

$$S_{xx}(f) = A^2 = \text{const}. \quad (7.16)$$

Das Signal wird mit der Abtastfrequenz f_A abgetastet. Als Anti-Aliasing-Filter kommt ein Filter 1. Ordnung mit einer Grenzfrequenz $f_g < f_A/2$ und der Übertragungsfunktion

$$G(f) = \frac{1}{1 + j \cdot f/f_g} \quad (7.17)$$

zum Einsatz. Der Nutzsignalbereich erstreckt sich auf den Bereich $|f| \leq f_g$ (Abb. 7.4). Die Störungen, die durch Aliasing entstehen, sind die spektralen Anteile bei $|f| \geq f_A/2$.

Zur Bewertung der Signalqualität nach der Abtastung sollen die Leistungen von Nutzsignal und Störsignal (Aliasing) verglichen werden. Mit der spektralen Leistungsdichte am Ausgang des Filters

$$S_{yy}(f) = |G(f)|^2 \cdot S_{xx}(f) = \frac{A^2}{1 + (f/f_g)^2} \quad (7.18)$$

erhält man die Signalleistungen durch Integration zu

$$P_{\text{Sig}} = 2 \int_0^{f_g} S_{yy}(f) df = 2 \int_0^{f_g} \frac{A^2}{1 + (f/f_g)^2} df \quad (7.19)$$

$$= 2 f_g A^2 \arctan\left(\frac{f}{f_g}\right) \Big|_0^{f_g} = \frac{\pi}{2} f_g A^2, \quad (7.20)$$

$$P_{\text{Stör}} = 2 \int_{f_A/2}^{\infty} S_{yy}(f) df = 2 f_g A^2 \arctan\left(\frac{f}{f_g}\right) \Big|_{f_A/2}^{\infty} \quad (7.21)$$

$$= f_g A^2 \left(\pi - 2 \arctan\left(\frac{f_A}{2 f_g}\right) \right). \quad (7.22)$$

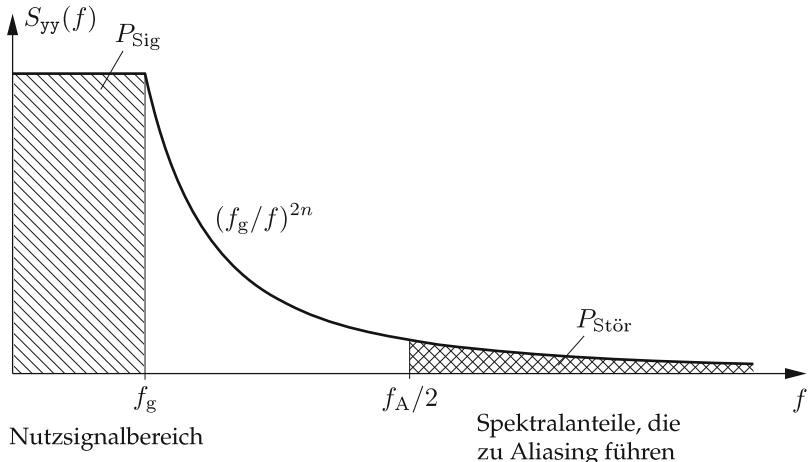
Das **Signal-Rausch-Verhältnis** (*signal to noise ratio*, SNR) wird damit

$$\text{SNR} = \frac{P_{\text{Sig}}}{P_{\text{Stör}}} = \frac{\pi/2}{\pi - 2 \arctan\left(\frac{f_A}{2 f_g}\right)}. \quad (7.23)$$

Legt man den Nutzsignalbereich auf ein Viertel der maximalen Bandbreite $f_g = f_A/(2 \cdot 4)$, so erhält man ein Signal-Rausch-Verhältnis von

$$\text{SNR} = \frac{P_{\text{Sig}}}{P_{\text{Stör}}} = \frac{\pi/2}{\pi - 2 \arctan(4)} \approx 3,205 \quad (7.24)$$

oder, in Dezibel, $\text{SNR} = 5,1 \text{ dB}$. Dieser Wert ist für viele Anwendungen unzureichend. ■

Abbildung 7.4. Aliasing-Fehler durch nicht-ideales Filter n -ter Ordnung.

Am letzten Beispiel erkennt man, wie wichtig der Einsatz steilflankiger Anti-Aliasing-Filter ist. Ein Filter 1. Ordnung ist in den allermeisten Fällen keinesfalls ausreichend. Man erkennt aber auch, dass es für ein großes Signal-Rausch-Verhältnis günstig ist, die Abtastfrequenz f_A deutlich größer als die Signalbandbreite B zu wählen.

Zur Untersuchung des Aliasing-Fehlers bei Tiefpassfiltern höherer Ordnung wird der Amplitudengang asymptotisch angenähert (Abb. 7.4). Für ein Filter n -ter Ordnung lautet dann das Quadrat des Amplitudenganges:

$$|G(f)|^2 \approx \begin{cases} 1 & \text{für } |f| \leq f_g, \\ \left(\frac{f_g}{f}\right)^{2n} & \text{für } |f| > f_g. \end{cases} \quad (7.25)$$

Die Signalleistungen erhält man dann für einen kontinuierlichen Eingangsprozess mit konstanter spektraler Leistungsdichte $S_{xx}(f) = A^2$:

$$P_{\text{Sig}} \approx 2 \int_0^{f_g} S_{yy}(f) df = 2 \int_0^{f_g} A^2 df = 2f_g A^2, \quad (7.26)$$

$$P_{\text{Stör}} \approx 2 \int_{f_A/2}^{\infty} S_{yy}(f) df = 2 \int_{f_A/2}^{\infty} \frac{A^2 f_g^{2n}}{f^{2n}} df \quad (7.27)$$

$$= -2A^2 \left. \frac{f_g^{2n}}{(2n-1) \cdot f^{2n-1}} \right|_{f_A/2}^{\infty} = A^2 \frac{(2f_g)^{2n}}{(2n-1) \cdot f_A^{2n-1}}. \quad (7.28)$$

Das Signal-Rausch-Verhältnis für ein Anti-Aliasing-Filter der Ordnung n ist:

$$\text{SNR} = \frac{P_{\text{Sig}}}{P_{\text{Stör}}} \approx (2n - 1) \cdot \left(\frac{f_A}{2f_g} \right)^{2n-1}. \quad (7.29)$$

Abbildung 7.5 zeigt das Signal-Rausch-Verhältnis für unterschiedliche Filterordnungen n und Grenzfrequenzen f_g . Selbst bei einem Filter 8. Ordnung sollte die maximale Nutzfrequenz f_g um den Faktor 4 kleiner als die Nyquist-Frequenz $f_A/2$ gewählt werden, um ein für manche Anwendungen ausreichend hohes Signal-Rausch-Verhältnis von 96 dB zu gewährleisten (vgl. z. B. Bsp. 7.5).

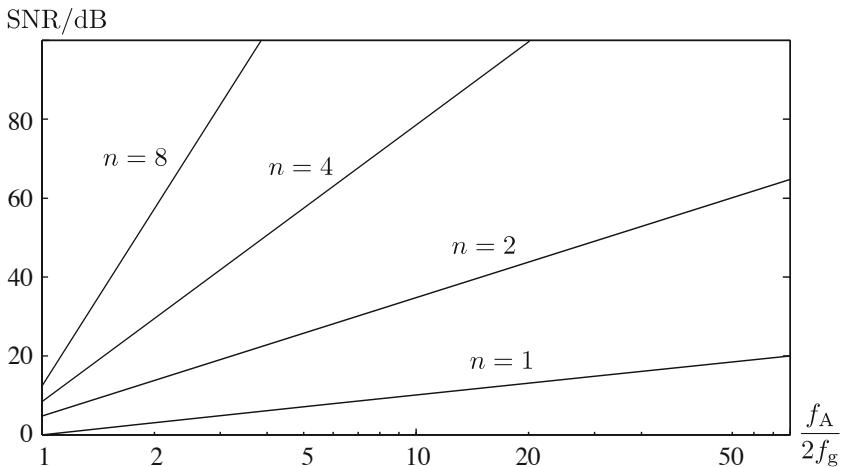


Abbildung 7.5. Signal-Rausch-Verhältnisse bei Anti-Aliasing-Filtern unterschiedlicher Ordnung n bei Annahme eines Eingangssignals mit konstanter Leistungsdichte.

7.1.4 Mittelwertbildung bei endlicher Abtastdauer

Eine Abtastung mit idealen Dirac-Impulsen ist in der Praxis nicht möglich. Vielmehr wird eine gewisse Zeit benötigt, bis die Abtastung durchgeführt ist. Man stelle sich beispielsweise einen Kondensator vor, der über die kontinuierliche Eingangsspannung $y(t)$ aufgeladen wird und der diesen Spannungswert für den Abtastvorgang kurz speichert. Dies entspricht implizit einer Integration der Eingangsspannung über die Ladezeit des Kondensators. Der Einfluss dieser Integration auf das Signal $y(t)$ soll nun untersucht werden.

Der Mittelwert eines Zeitsignals wird nach Kap. 6 über das Integral

$$\bar{y} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} y(t) dt \quad (7.30)$$

bestimmt. In der Praxis kann der Grenzübergang $T \rightarrow \infty$ nicht durchgeführt werden. Die Mittelung beschränkt sich daher auf ein Intervall der endlichen Breite T . Man kann nun für jeden Zeitpunkt t eine Mittelung über das Intervall $[t - T/2, t + T/2]$ durchführen und erhält so den **gleitenden Mittelwert** (engl. *moving average*, MA):

$$\bar{y}(t) = \frac{1}{T} \int_{t-T/2}^{t+T/2} y(\tau) d\tau. \quad (7.31)$$

Betrachtet man die Faltungsoperation

$$\bar{y}(t) = y(t) * g(t) = \int_{-\infty}^{\infty} y(\tau) g(t - \tau) d\tau \quad (7.32)$$

und vergleicht diese mit (7.31), so ergibt sich für die Impulsantwort der gleitenden Mittelung:

$$g(t) = \frac{1}{T} \cdot r_T(t) \quad \text{mit} \quad r_T(t) = \begin{cases} 1 & \text{für } |t| \leq T/2, \\ 0 & \text{für } |t| > T/2. \end{cases} \quad (7.33)$$

Die Mittelwertbildung, die einer mit $1/T$ gewichteten Integration über das Intervall T entspricht, kann durch eine Faltung des ursprünglichen Signals $y(t)$ mit einem Rechteckfenster $r_T(t)$ beschrieben werden. Im Frequenzbereich entspricht das einer Multiplikation des Fourier-Spektrums $Y(f)$ mit einer Sinc-Funktion [42]:

$$\bar{y}(t) = \frac{1}{T} y(t) * r_T(t) \quad (7.34)$$

○
●

$$\bar{Y}(f) = \frac{1}{T} Y(f) \cdot R_T(f) = Y(f) \cdot \frac{\sin(\pi T f)}{\pi T f} = Y(f) \cdot \text{sinc}(T f). \quad (7.35)$$

Die Filterwirkung der gleitenden Mittelung (MA-Filter) ist in Abb. 7.6 dargestellt. Das Eingangssignal wird also bereits vor der sich anschließenden idealen Abtastung mit Dirac-Impulsen verfälscht. Dabei können drei Frequenzbereiche unterschieden werden.

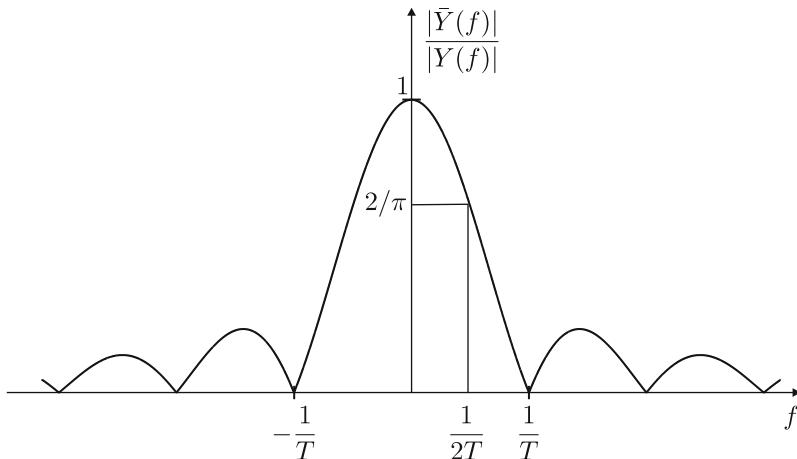


Abbildung 7.6. Amplitudengang der gleitenden Mittelung.

1. $|f| \leq 1/(2T)$:

Das Signal $y(t)$ wird nahezu ungefiltert durchgelassen:

$$\frac{2}{\pi} \leq \left| \frac{\bar{Y}(f)}{Y(f)} \right| \leq 1. \quad (7.36)$$

Das Nutzsignal sollte in diesem Frequenzintervall liegen.

2. $|f| \approx 1/T$:

Die Frequenz einer harmonischen Schwingung stimmt annähernd mit der Wiederholrate der Periode überein. Die Schwingung wird daher durch die Mittelwertbildung weitgehend aus dem Signal herausgefiltert.

3. $|f| > 1/T$:

Das Signal wird gedämpft.

Beispiel 7.3 (Mittelwertbildung als Anti-Aliasing-Filter): Bei vielen integrierenden A/D-Umsetzern erstreckt sich die Mittelwertbildung über die gesamte Abtastzeit t_A (vgl. Abschn. 7.3). Mit

$$T = t_A = \frac{1}{f_A} \quad (7.37)$$

erhält man das Spektrum des gemittelten Signals vor der Abtastung zu

$$\bar{Y}(f) = Y(f) \cdot \frac{\sin(\pi f/f_A)}{\pi f/f_A}. \quad (7.38)$$

Für die Nyquist-Frequenz $f = f_A/2$ ergibt sich daraus:

$$\bar{Y}\left(\frac{f_A}{2}\right) = Y\left(\frac{f_A}{2}\right) \cdot 2/\pi \approx 0,63 \cdot Y\left(\frac{f_A}{2}\right). \quad (7.39)$$

Da die Spektralanteile für $f \geq f_A/2$ teilweise nur schlecht gedämpft werden, ist bei breitbandigen Eingangssignalen $y(t)$ trotz Mittelung ein Anti-Aliasing-Filter vorzusehen. ■

Anmerkung

Die reale Abtastung ist durch das Aufladen von Kapazitäten immer mit einer Integration oder Mittelwertbildung verbunden. Dabei ist Folgendes zu beachten:

1. Ist das Mittelungsintervall T klein gegenüber der Abtastzeit t_A , so erhält man eine sehr breite Filtercharakteristik. Die spektrale Verfälschung im Nyquist-Band $|f| \leq f_A/2$ kann dann meist vernachlässigt werden.
2. Ist das Mittelungsintervall T gleich der Abtastzeit t_A , so kommt es wie im obigen Beispiel zu Verfälschungen im Nyquist-Band, die im Allgemeinen bei der anschließenden digitalen Signalverarbeitung kompensiert werden müssen. Die Filterwirkung ist allerdings für $|f| > f_A/2$ nicht ausreichend. Daher kann in der Regel auf ein Anti-Aliasing-Filter nicht verzichtet werden.

7.1.5 Zeitliche Abtastfehler

Bisher wurde davon ausgegangen, dass die Abtastung exakt – also ohne Zeitfehler – zu den Zeitpunkten $n t_A$ erfolgt. In realen Systemen kann es aber zu einem zeitlichen Fehler bei der Abtastung kommen, der auch als **Jitter** bezeichnet wird. Dies äußert sich in einer Signalverfälschung, die ähnliche Auswirkungen wie ein überlagertes Störsignal aufweist.

Abbildung 7.7 zeigt exemplarisch das Fehlersignal bei Abtastung einer harmonischen Schwingung mit Zeitfehlern. Es soll nun eine Abschätzung des Jitter-Fehlers vorgenommen werden [35, 38]. Für diese Abschätzung sei immer die Einhaltung des Abtasttheorems (Satz 7.1) vorausgesetzt.

Die ideal abgetastete Zeitfunktion sei durch

$$y_*(t) = y(t) \cdot \sum_{n=-\infty}^{\infty} \delta(t - n t_A) \quad \text{oder} \quad y_n = y(n t_A) \quad (7.40)$$

gekennzeichnet. Für die tatsächlichen, zufälligen Abtastzeitpunkte gelte

$$t_n = n t_A + \tau_n, \quad n = 0, \pm 1, \pm 2, \dots, \quad (7.41)$$

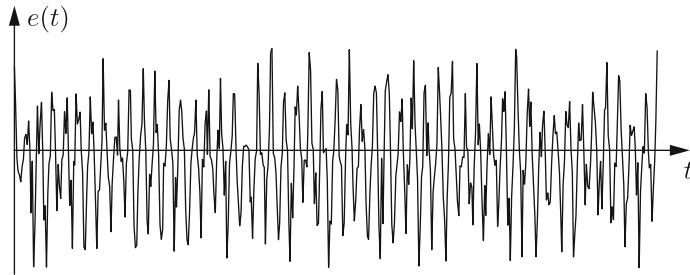


Abbildung 7.7. Fehlersignal infolge des Jitters bei der Abtastung einer harmonischen Schwingung.

wobei τ_n den Zeitfehler beim n -ten Abtastwert darstellt. Für die tatsächlich abgetastete Funktion gilt

$$\tilde{y}_*(t) = y(t) \cdot \sum_{n=-\infty}^{\infty} \delta(t - n t_A - \tau_n) \quad \text{oder} \quad \tilde{y}_n = y(n t_A + \tau_n). \quad (7.42)$$

Der Fehler durch zeitlich falsches Abtasten (Jitter-Fehler) ist die Differenz von gestörtem und ungestörtem Signal:

$$e_n = \tilde{y}_n - y_n. \quad (7.43)$$

Um den Jitter-Fehler quantitativ zu beschreiben, soll zunächst eine Abschätzung für die Ableitung eines bandbegrenzten Signals hergeleitet werden. Ausgangspunkt dafür ist der Differentiationsatz der Fourier-Transformation [42]:

$$\frac{dy(t)}{dt} = \mathcal{F}^{-1}\{(j2\pi f) Y(f)\} = \int_{-\infty}^{\infty} (j2\pi f) Y(f) e^{j2\pi f t} df. \quad (7.44)$$

Mit der Dreiecksungleichung [42] erhält man daraus die folgende Abschätzung für das erste Moment M_1 des auf f_g bandbegrenzten Spektrums $Y(f)$:

$$\left| \frac{dy(t)}{dt} \right| \leq \int_{-f_g}^{f_g} |2\pi f| |Y(f)| df = M_1. \quad (7.45)$$

Über den Differenzenquotienten im Abtastzeitpunkt t_n

$$|y(t_n) - y(n t_A)| \leq M_1 |t_n - n t_A| = M_1 |\tau_n| \quad (7.46)$$

lässt sich der Fehler nach (7.43) wie folgt abschätzen:

$$|\epsilon_n| \leq M_1 \cdot |\tau_n|. \quad (7.47)$$

Für ein gegebenes n hängt der Fehler ϵ_n nur von τ_n ab. Nimmt man an, die Zeitfehler τ_n seien voneinander unabhängige Zufallsgrößen mit

$$\mathbb{E}\{\tau_n\} = 0 \quad \text{und} \quad \mathbb{E}\{\tau_n^2\} = \sigma_\tau^2, \quad (7.48)$$

so folgt, dass die Amplitudenfehler ϵ_n ebenfalls statistisch unabhängig voneinander sind. Für die Fehlervarianz erhält man mit (7.47)

$$\mathbb{E}\{\epsilon_n^2\} \leq M_1^2 \sigma_\tau^2. \quad (7.49)$$

Betrachtet man wegen der Forderung des Abtasttheorems ein auf f_g bandbegrenztes Eingangssignal mit konstantem Spektrum

$$|Y(f)| = \begin{cases} A_{\max} & \text{für } |f| \leq f_g, \\ 0 & \text{sonst,} \end{cases} \quad (7.50)$$

so erhält man mit (7.45) für das erste Moment:

$$M_1 = \int_{-f_g}^{f_g} |2\pi f| \cdot A_{\max} df = 2\pi f_g^2 A_{\max}. \quad (7.51)$$

Damit wird der Erwartungswert des quadratischen Amplitudenfehlers, d.h. die Leistung des Fehlersignals, zu

$$\mathbb{E}\{\epsilon_n^2\} \leq M_1^2 \sigma_\tau^2 = 4\pi^2 f_g^4 A_{\max}^2 \sigma_\tau^2 \approx P_{\text{Stör}}. \quad (7.52)$$

Dies liefert eine Abschätzung für die Leistung des Störsignals infolge des Jitters. Um das Signal-Rausch-Verhältnis anzugeben, wird ferner die Leistung des Nutzsignals $y(t)$ benötigt, welches über das Amplitudenspektrum (7.50) definiert wurde. Gemäß (7.50) ist das Nutzsignal $y(t)$ jedoch ein Energiesignal, dessen Signalenergie über den Satz von Parseval ermittelt werden kann:

$$E_y = \int_{-\infty}^{\infty} |Y(f)|^2 df = \int_{-f_g}^{f_g} A_{\max}^2 df = 2 A_{\max}^2 f_g. \quad (7.53)$$

Berücksichtigt man nun, dass der wesentliche Teil der Energie des Signals (7.50) im Hauptgipfel enthalten ist, so kann die Signalleistung durch Division der Gesamtenergie E_y durch das Zeitintervall $2t_0$ zwischen den beiden Nullstellen links und rechts vom Hauptgipfel angegeben werden (Abb. 7.8).

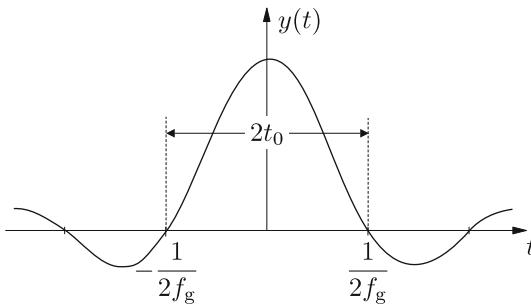


Abbildung 7.8. Zeitlicher Verlauf einer auf f_g bandbegrenzten konstanten Spektralfunktion.

Mit

$$t_0 = \frac{1}{2f_g} \quad (7.54)$$

erhält man die Signalleistung zu

$$P_{\text{Sig}} \approx \frac{E_y}{2t_0} = 2 A_{\max}^2 f_g^2. \quad (7.55)$$

Das Signal-Rausch-Verhältnis berechnet sich mit (7.52) zu

$$\text{SNR} = \frac{P_{\text{Sig}}}{P_{\text{Stör}}} \approx \frac{2 A_{\max}^2 f_g^2}{4\pi^2 f_g^4 A_{\max}^2 \sigma_\tau^2} = \frac{2}{(2\pi f_g \sigma_\tau)^2}. \quad (7.56)$$

Anmerkungen

1. Je näher die Signalfrequenzen an der Nyquist-Frequenz liegen, desto stärker macht sich der Jitter-Fehler bemerkbar, da sich hochfrequente Signale innerhalb des Zeitfehlers τ_n stärker verändern als tieffrequente Signalanteile. Für Signalfrequenzen, die klein gegenüber der halben Abtastfrequenz sind, kann der Fehler meist vernachlässigt werden.
2. Jitter-Fehler treten bei so genannten Peak-Hold-Schaltungen für schnelle A/D-Umsetzer auf, da innerhalb des Abtastfensters der maximale Signalwert dem idealen Abtastzeitpunkt zugeordnet wird.

Beispiel 7.4 (Harmonische Schwingung mit Jitter): Abbildung 7.9 zeigt eine Schätzung der spektralen Leistungsdichte einer mit Jitter-Fehlern abgetasteten harmonischen Schwingung der Frequenz $f = 200$ Hz. Das Signal wurde bei der Abtastung einem gleichverteilten Zeitfehler τ_n unterworfen mit $\tau_{\max} = 0,1 t_A$ bei $t_A = 1$ ms und somit $f_A = 1$ kHz. Neben dem Nutzsignal bei $f = 200$ Hz enthält das Signal \tilde{y}_n Störungen, die gleichmäßig im Nyquist-Band verteilt sind. ■

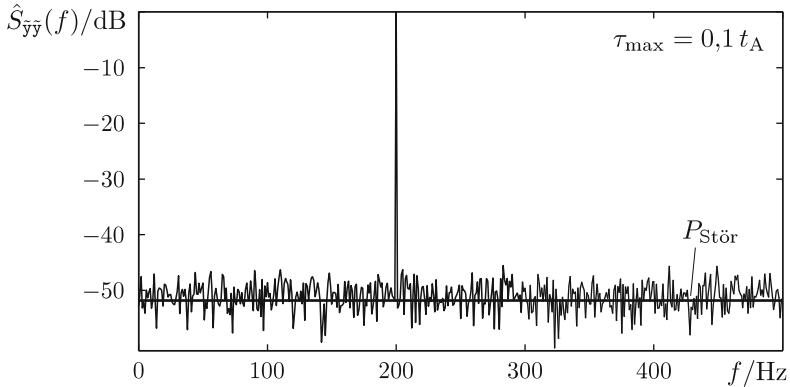


Abbildung 7.9. Spektrale Leistungsdichte einer harmonischen Schwingung mit Jitter.

Beispiel 7.5 (Jitter-Fehler bei der A/D-Umsetzung von Audiosignalen): Gegeben sei ein 16-Bit-A/D-Umsetzer ($N = 16$) für die CD-Produktion mit einer Abtastfrequenz von $f_A = 44100\text{ Hz}$. Es soll nun untersucht werden, wie hoch die Anforderungen an die zeitliche Genauigkeit der Abtastung sind. Dabei soll der Amplitudenfehler infolge des Abtast-Jitters kleiner sein als das Rauschen, das durch die Quantisierung entsteht (Abschn. 7.2). Das Signal-Rausch-Verhältnis bei einer Quantisierung mit 2^{16} Stufen ist nach (7.125)

$$\text{SNR}/\text{dB} = 10 \lg(2^{2N}) = 6,02 N \approx 96. \quad (7.57)$$

Die Musiksignale sind bandbegrenzt auf $f_g = 20\text{ kHz}$. Nach (7.56) erhält man bei vorgegebenem SNR die zulässige Varianz des Zeitfehlers zu

$$\sigma_\tau^2 = \frac{2}{(2\pi f_g)^2 \cdot \text{SNR}} \quad \text{mit} \quad \text{SNR} = 2^{2N}. \quad (7.58)$$

Nimmt man einen gleichverteilten Zeitfehler τ im Intervall

$$-\tau_{\max} \leq \tau \leq \tau_{\max} \quad \text{mit} \quad \sigma_\tau^2 = \frac{\tau_{\max}^2}{3} \quad (7.59)$$

an, so wird der maximal zulässige Zeitfehler

$$\tau_{\max} = \sqrt{3} \sigma_\tau = \frac{1}{2\pi f_g} \sqrt{\frac{6}{\text{SNR}}} \approx 297,4\text{ ps}. \quad (7.60)$$

Bezieht man nun noch den zulässigen Zeitfehler τ_{\max} auf die Abtastzeit t_A , so erhält man mit den oben gemachten Annahmen

$$\frac{\tau_{\max}}{t_A} \approx 1,31 \cdot 10^{-5}. \quad (7.61)$$

Das Ergebnis zeigt, dass bei A/D-Umsetzern mit hoher Auflösung eine sehr genaue Zeitbasis für die Abtastung notwendig ist, wenn der gewünschte Störabstand eingehalten werden soll. ■

7.2 Quantisierung

7.2

Bei der Quantisierung werden den kontinuierlichen Amplitudenwerten des Eingangssignals $x(t)$ diskrete Werte $x_q(t)$ zugeordnet. Die Quantisierung ist somit eine nichtdynamische, nichtlineare Operation mit einer Kennlinie nach Abb. 7.10. Jeder Eingangswert, welcher innerhalb eines Quantisierungsstreifens der Breite q liegt, ergibt am Ausgang eine konstante Amplitude, die der Mitte des zugehörigen Streifens entspricht (Rundungskennlinie). Für die Quantisierung sind aber auch andere Kennlinien denkbar, bei denen die Quantisierungsstufen nicht alle die gleiche Breite aufweisen [22]. Sie sollen aber hier nicht diskutiert werden.

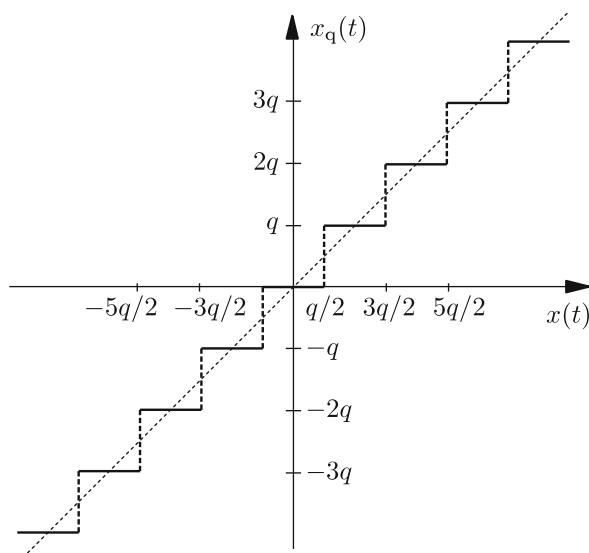


Abbildung 7.10. Kennlinie eines Quantisierers.

Bei äquidistanter Quantisierung ist der Betrag des Quantisierungsfehlers

$$|e_q(t)| = |x_q(t) - x(t)| \leq \frac{q}{2}. \quad (7.62)$$

Der Quantisierungsfehler wird durch den lokalen Verlauf des Signals und den Abstand zum jeweils nächsten Quantisierungsniveau bestimmt. Bei einem stetigen Zeitsignal weist der Quantisierungsfehler üblicherweise viele Sprünge auf. Das zugehörige Spektrum ist deshalb wesentlich breiter als jenes des ursprünglichen Zeitsignals $x(t)$. Die Bandbreite des Quantisierungsfehlers $e_q(t)$ ist näherungsweise proportional zur Steigung des Signals $|\dot{x}(t)|$ und umgekehrt proportional zur Breite einer Quantisierungsstufe q .

Beispiel 7.6 (Quantisierung einer harmonischen Schwingung): Als Eingangssignal eines Quantisierers soll eine harmonische Schwingung

$$x(t) = a \sin\left(2\pi \frac{t}{T_0}\right) \quad \text{mit} \quad f = \frac{1}{T_0} = 1,1 \text{ Hz} \quad (7.63)$$

betrachtet werden, deren Amplitude a größer als eine halbe Quantisierungsstufe q ist:

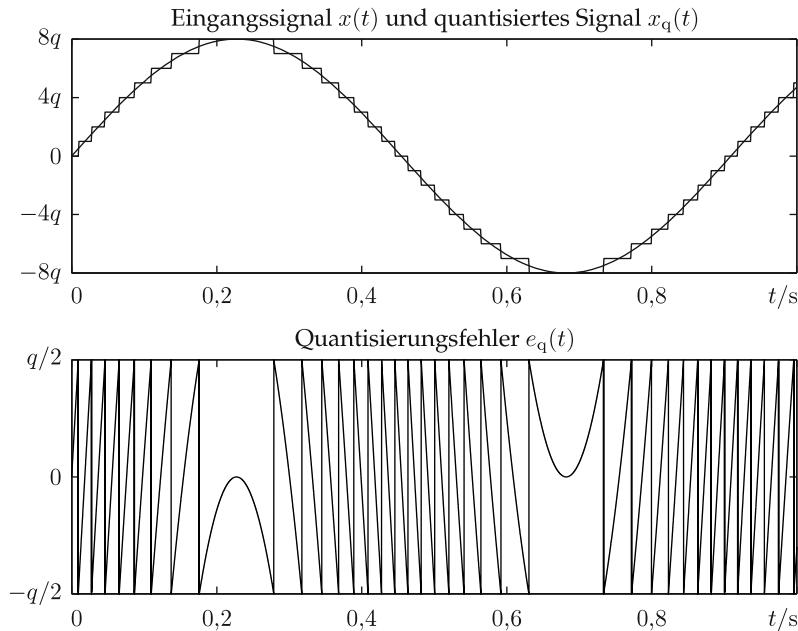
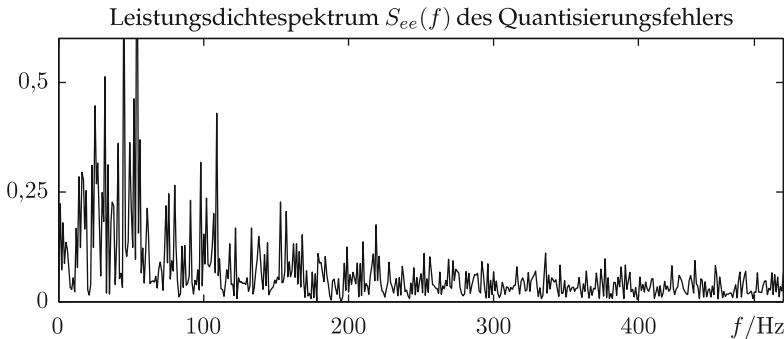
$$a > \frac{q}{2}. \quad (7.64)$$

Damit bleibt das Ausgangssignal nicht konstant in einer Quantisierungsstufe hängen, sondern springt zwischen verschiedenen diskreten Amplituden.

Abbildung 7.11 zeigt das Signal $x(t)$ vor und nach der Quantisierung sowie den Quantisierungsfehler $e_q(t)$. Man erkennt den treppenförmigen Verlauf des quantisierten Ausgangssignals $x_q(t)$ und sieht weiterhin, dass der Fehler viele Sprünge aufweist. Das zugehörige Leistungsdichtespektrum $S_{ee}(f)$ des Fehlersignals in Abb. 7.12 ist daher viel breitbandiger als jenes des Eingangssignals. ■

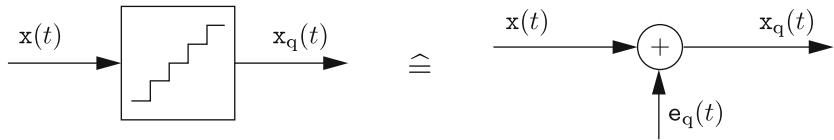
Aus diesen praktischen Beobachtungen heraus wird der Quantisierungsfehler näherungsweise als weißes Rauschen $e_q(t)$ angenommen. Er sei unkorreliert zum Eingangssignal $x(t)$ und habe innerhalb einer Quantisierungsstufe eine Gleichverteilung

$$f_{e_q}(e_q) = \begin{cases} \frac{1}{q} & \text{für } -\frac{q}{2} \leq e_q \leq \frac{q}{2}, \\ 0 & \text{sonst.} \end{cases} \quad (7.65)$$

Abbildung 7.11. Quantisiertes Signal und Quantisierungsfehler $e_q(t)$.Abbildung 7.12. Leistungsdichtespektrum $S_{ee}(f)$ des Quantisierungsfehlers.

Voraussetzung hierfür sind veränderliche Eingangssignale $x(t)$, die hinreichend viele Quantisierungsstufen überstreichen. Man erhält damit ein lineares Modell für die Quantisierung (Abb. 7.13), bei welchem dem stetigen Eingangssignal $x(t)$ das **Quantisierungsrauschen** $e_q(t)$ additiv überlagert wird. Die explizite Behandlung der Nichtlinearität des Systems wird durch diesen Ansatz vermieden.

Im Folgenden soll die theoretische Grundlage der eben gemachten Annahme durch Anwendung der Wahrscheinlichkeitstheorie überprüft wer-

Abbildung 7.13. Modellierung der Quantisierung mit überlagertem Rauschen $e_q(t)$.

den [54, 55]. Die Signale am Quantisierer wurden daher als stochastische Signale beschrieben. Die Zufallsvariable x repräsentiert dabei die möglichen Amplituden des Signals $x(t)$.

7.2.1 Wahrscheinlichkeitsdichte der Signalwerte

Zur statistischen Beschreibung der Verhältnisse am Quantisierer benötigt man die Wahrscheinlichkeitsdichte der Werte bzw. Amplituden des Eingangssignals. Diese sei im Folgenden kurz als „Amplitudendichte“ bezeichnet. Die Berechnung der Amplitudendichte lässt sich durch eine Rückführung auf eine Transformation von Zufallsvariablen (Satz 4.2) vereinfachen. Dazu wird der Zeit t im Beobachtungsintervall T_0 eine gleichverteilte Zufallsvariable t zugeordnet:

$$f_t(t) = \begin{cases} \frac{1}{T_0} & \text{für } 0 \leq t \leq T_0, \\ 0 & \text{sonst.} \end{cases} \quad (7.66)$$

Dieser Ansatz ist gerechtfertigt, da kein Zeitpunkt t_i einem anderen Zeitpunkt bevorzugt wird und daher jede Zeit t mit der gleichen Wahrscheinlichkeit auftritt. Die Amplitude zum Zeitpunkt t erhält man über eine Zeitfunktion

$$x = g(t). \quad (7.67)$$

Zur Zufallsvariablen für die Amplitude x gelangt man über die Transformation

$$x = g(t), \quad (7.68)$$

wobei die Funktion $g(\cdot)$ die Zufallsvariable t in die Zufallsvariable x für die Amplitude überführt. Die Bestimmung der Amplitudendichte entspricht also der Bestimmung der Wahrscheinlichkeitsdichte der transformierten Zufallsvariablen x nach (4.22):

$$f_x(x) = \sum_{i=1}^n f_t(t_i) \left| \frac{dg(t)}{dt} \right|_{t=t_i}^{-1}. \quad (7.69)$$

Beispiel 7.7 (Amplitudendichte einer harmonischen Schwingung): Gegeben sei eine allgemeine harmonische Schwingung der Form

$$x = g(t) = a_m \cos\left(2\pi \frac{m}{T_0} t\right). \quad (7.70)$$

Gesucht ist die Amplitudendichte $f_x(x)$ bei gegebener Gleichverteilung der Zeit t . Im Intervall $[0, T_0]$ hat die Schwingung m Perioden. Es gibt daher $2m$ Lösungen für die inverse Funktion:

$$t_n = g^{-1}(x) = \frac{T_0}{2\pi m} \arccos\left(\frac{x}{a_m}\right), \quad n = 1, \dots, 2m. \quad (7.71)$$

Einsetzen der $2m$ Lösungen (7.71) in die Ableitung

$$\frac{dg(t)}{dt} = -2\pi \frac{m}{T_0} a_m \sin\left(2\pi \frac{m}{T_0} t\right) \quad (7.72)$$

ergibt mit der Beziehung $\sin(\arccos x) = \sqrt{1 - x^2}$:

$$\left| \frac{dg(t)}{dt} \right|_{t=t_n} = 2\pi \frac{m}{T_0} \sqrt{a_m^2 - x^2}. \quad (7.73)$$

Bei einer Gleichverteilung der Zeit im Intervall $[0, T_0]$ mit

$$f_t(t) = \frac{1}{T_0} \quad (7.74)$$

erhält man nach (7.69) für die Amplitudendichte der harmonischen Schwingung:

$$f_x(x) = \sum_{n=1}^{2m} \frac{1}{T_0} \cdot \frac{T_0}{2\pi m \sqrt{a_m^2 - x^2}} \quad (7.75)$$

$$= \frac{1}{\pi \sqrt{a_m^2 - x^2}}, \quad |x| \leq a_m. \quad (7.76)$$

Abbildung 7.14 zeigt den Verlauf der Amplitudendichte $f_x(x)$. Das gleiche Ergebnis erhält man für sinusförmige Funktionen. ■

7.2.2 Amplitudendichte der Fourier-Reihe

Allgemein lassen sich periodische Signale in eine Fourier-Reihe entwickeln:

$$x(t) = \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k \cos\left(2\pi \frac{k}{T_0} t\right). \quad (7.77)$$

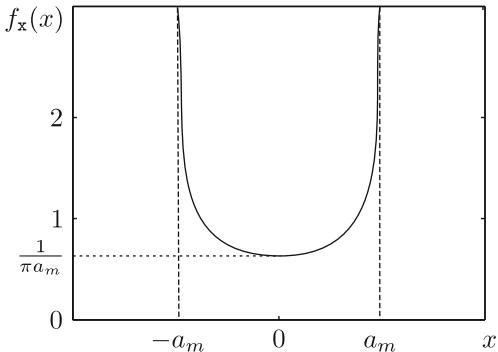


Abbildung 7.14. Amplitudendichte einer harmonischen Schwingung.

Dabei wurde in (7.77) wegen der Kausalität von Messsignalen Symmetrie zu $t = 0$ angenommen, so dass die Sinusterme herausfallen. Betrachtet man die Amplituden der einzelnen Schwingungen als voneinander unabhängige Zufallsgrößen x_k mit

$$x = x_0 + x_1 + \dots + x_n + \dots , \quad (7.78)$$

so erhält man die Amplitudendichte der Fourier-Reihe durch Faltung der Wahrscheinlichkeitsdichten der Teilsignale:

$$f_x(x) = f_{x_0}(x) * f_{x_1}(x) * \dots * f_{x_n}(x) * \dots . \quad (7.79)$$

Mit dem Ergebnis (7.76) aus obigem Beispiel und der Wahrscheinlichkeitsdichte des konstanten Gliedes

$$f_{x_0}(x) = \delta\left(x - \frac{a_0}{2}\right) \quad (7.80)$$

lässt sich die Amplitudendichte eines beliebigen periodischen Signals darstellen. Die explizite Berechnung von $f_x(x)$ ist in vielen Fällen nicht notwendig, wenn man sich den zentralen Grenzwertsatz (Satz 4.3) der Wahrscheinlichkeitsrechnung zu Nutze macht [37]. Dann kann man in den meisten Fällen die Summendichte durch eine Normalverteilung annähern.

Beispiel 7.8 (Amplitudendichte eines Summensignals): Es soll die Amplitudendichte einer Summe von vier harmonischen Signalen mit verschiedenen Amplituden und Frequenzen

$$x(t) = \sum_{k=1}^4 \left(\frac{6-k}{5} \right) \cos(2\pi kt) \quad (7.81)$$

mit einer Normalverteilung nach dem zentralen Grenzwertsatz verglichen werden. Nach dem zentralen Grenzwertsatz überlagern sich die Erwartungswerte und die Varianzen der Einzelsignale additiv.

Der Erwartungswert der einzelnen Amplitudendichten x_k verschwindet:

$$\mu_{x_k} = E\{x_k\} = \int_{-a_k}^{a_k} \underbrace{\frac{1}{\pi \sqrt{a_k^2 - x_k^2}}}_{\text{aus (7.76)}} x_k dx_k \quad (7.82)$$

$$= -\frac{\sqrt{a_k^2 - x_k^2}}{\pi} \Big|_{-a_k}^{a_k} = 0, \quad (7.83)$$

weshalb der Erwartungswert des Summensignals ebenfalls null ist. Die Varianz der einzelnen Amplitudendichten x_k berechnet sich zu

$$\sigma_{x_k}^2 = E\{(x_k - \mu_{x_k})^2\} = \int_{-a_k}^{a_k} \frac{x_k^2}{\pi \sqrt{a_k^2 - x_k^2}} dx_k \quad (7.84)$$

$$= -\frac{x_k}{2\pi} \sqrt{a_k^2 - x_k^2} + \frac{a_k^2}{2\pi} \arcsin\left(\frac{x_k}{a_k}\right) \Big|_{-a_k}^{a_k} = \frac{a_k^2}{2}. \quad (7.85)$$

Damit erhält man für die Varianz der Amplitudendichte des Summensignals:

$$\sigma_x^2 = \sum_{k=1}^4 \sigma_{x_k}^2 = \frac{1}{2} \sum_{k=1}^4 a_k^2. \quad (7.86)$$

Anmerkung

Die Standardabweichung σ_{x_k} stimmt erwartungsgemäß mit dem Effektivwert (quadratischen Mittelwert) eines sinusförmigen Signals überein:

$$\sigma_{x_k} = u_{\text{eff}} = \frac{a_k}{\sqrt{2}}. \quad (7.87)$$

Abbildung 7.15 zeigt die gute Übereinstimmung beider Verteilungen. ■

Das Beispiel zeigt, dass die Amplitudendichte vieler Signale durch eine Normalverteilung angenähert werden kann. Das erleichtert später die Abschätzung des Fehlers, der durch die Quantisierung entsteht.

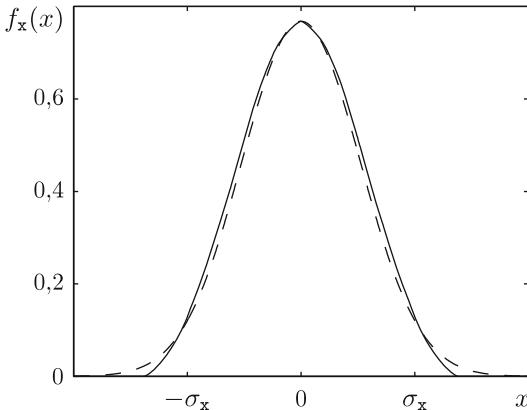


Abbildung 7.15. Wahrscheinlichkeitsdichte $f_x(x)$ des Summensignals (—) und der Normalverteilung (---).

7.2.3 Quantisierungstheorem

Ein Signal zu quantisieren bedeutet, den kontinuierlichen Signalamplituden diskrete Werte zuzuordnen. Die Diskretisierung der Amplituden entspricht gewissermaßen einer Abtastung der Amplitudendichte über dem Amplitudenbereich – analog zur Abtastung von Signalverläufen über dem Zeitbereich.

Ein Signal $x(t)$ mit der kontinuierlichen Amplitudendichte $f_x(x)$ durchlufe einen Quantisierer. Der Quantisierer erzeugt daraus ein Signal $x_q(t)$ mit den diskreten Amplituden $k \cdot q$ (Abb. 7.16). Die diskrete Wahrscheinlichkeit der quantisierten Amplituden muss gleich der Wahrscheinlichkeit sein, dass die Signalamplitude innerhalb des zugehörigen Quantisierungsstreifens liegt. Mit

$$P\left(x - \frac{q}{2} \leq x < x + \frac{q}{2}\right) = \int_{x - \frac{q}{2}}^{x + \frac{q}{2}} f_x(\alpha) d\alpha \quad (7.88)$$

ergibt sich die Wahrscheinlichkeitsdichte $f_{x_q}(x_q)$ des quantisierten Signals als Impulsreihe:

$$f_{x_q}(x_q) = \int_{x_q - \frac{q}{2}}^{x_q + \frac{q}{2}} f_x(\alpha) d\alpha \cdot \sum_{k=-\infty}^{\infty} \delta(x_q - kq). \quad (7.89)$$

Die Quantisierung kann demnach als eine **Flächenabtastung** der Amplitudendichte $f_x(x)$ des Eingangssignals angesehen werden. Dabei liegt der diskrete Abtastwert in der Mitte des Quantisierungintervalls und sein Gewicht ist gleich der Fläche unter der Wahrscheinlichkeitsdichte im Intervall.

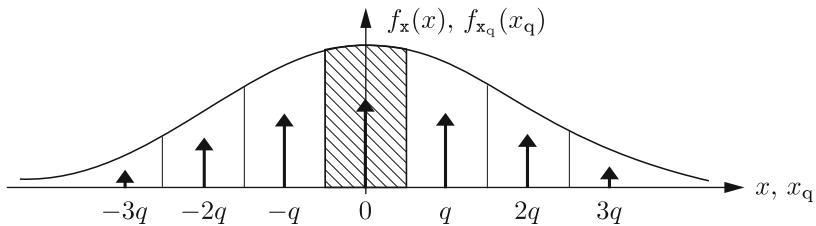


Abbildung 7.16. Wahrscheinlichkeitsdichten $f_x(x)$ und $f_{x_q}(x_q)$ eines amplitudenkontinuierlichen bzw. eines quantisierten Signals.

Beschreibt man (7.88) – wie bei der Mittelwertbildung in Abschn. 7.1.4 – als Faltung der Amplitudendichte mit einem Rechteckfenster der Breite q ,

$$\int_{x-\frac{q}{2}}^{x+\frac{q}{2}} f_x(\alpha) d\alpha = f_x(x) * r_q(x) \quad \text{mit} \quad r_q(x) = \begin{cases} 1 & \text{für } |x| \leq \frac{q}{2} \\ 0 & \text{sonst} \end{cases}, \quad (7.90)$$

so erhält man folgende Form der abgetasteten Wahrscheinlichkeitsdichte:

$$f_{x_q}(x) = (f_x(x) * r_q(x)) \cdot \sum_{k=-\infty}^{\infty} \delta(x - kq). \quad (7.91)$$

Dem entspricht im Frequenzbereich die charakteristische Funktion

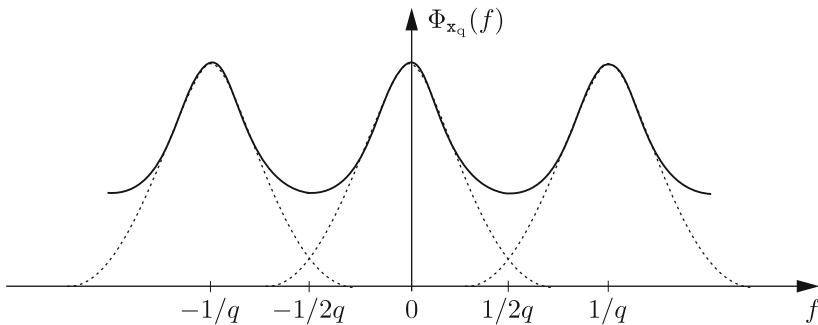
$$\Phi_{x_q}(f) = [\Phi_x(f) \cdot R_q(f)] * \frac{1}{q} \sum_{k=-\infty}^{\infty} \delta\left(f - \frac{k}{q}\right) \quad (7.92)$$

$$= \sum_{k=-\infty}^{\infty} \Phi_x\left(f - \frac{k}{q}\right) \cdot \frac{\sin(\pi q(f - \frac{k}{q}))}{\pi q(f - \frac{k}{q})}. \quad (7.93)$$

Die charakteristische Funktion $\Phi_{x_q}(f)$ des quantisierten Signals entspricht einer periodischen Wiederholung der mit $\sin(\pi q f)/(\pi q f)$ gefilterten charakteristischen Funktion $\Phi_x(f)$ des kontinuierlichen Signals (Abb. 7.17). Die Wiederholungsfrequenz $1/q$ wird dabei als Quantisierungsfrequenz bezeichnet. Dieses Ergebnis führt zum **Quantisierungstheorem**, welches besagt, wie grob Signale quantisiert werden dürfen, ohne wesentliche Signal-kennwerte zu verlieren.

Satz 7.2: Quantisierungstheorem

Ist die Wahrscheinlichkeitsdichte $f_x(x)$ eines amplitudenkontinuierlichen Signals $x(t)$ bandbegrenzt, d. h. besitzt die charakteristische Funktion $\Phi_x(f)$ nur Komponenten bei Frequenzen f , die kleiner sind als die halbe Quantisierungsfrequenz $1/q$, dann kann das Signal $x(t)$ mit einer Quantisierungsschrittweite q ohne wesentliche Verluste wiedergegeben werden.

Abbildung 7.17. Charakteristische Funktion $\Phi_{x_q}(f)$ eines quantisierten Signals.

sierungsfrequenz $1/2q$,

$$\Phi_x(f) = 0 \quad \text{für} \quad |f| \geq \frac{1}{2q}, \quad (7.94)$$

so lässt sich aus der Wahrscheinlichkeitsdichte $f_{x_q}(x_q)$ des quantisierten Signals die ursprüngliche Wahrscheinlichkeitsdichte des kontinuierlichen Signals fehlerfrei rekonstruieren.

Für die fehlerfreie Rekonstruktion der Momente der Wahrscheinlichkeitsdichte genügt anstelle von (7.94) die schwächere Bedingung

$$\Phi_x(f) = 0 \quad \text{für} \quad |f| \geq \frac{1}{q}, \quad (7.95)$$

da diese nach (4.62) für $f = 0$ berechnet werden. \diamond

Anmerkungen

1. Das Quantisierungstheorem weist eine große Ähnlichkeit mit dem Abtasttheorem auf. Bei der Zeitdiskretisierung werden die Signalamplituden über der Zeit abgetastet. Das resultierende Spektrum

$$Y_*(f) = \frac{1}{t_A} \sum_{k=-\infty}^{\infty} Y\left(f - \frac{k}{t_A}\right) \quad (7.96)$$

ist gleich dem mit $1/t_A$ gewichteten Spektrum des zeitkontinuierlichen Signals, welches mit der Frequenz $f_A = 1/t_A$ periodisch wiederholt wird. Bei der Quantisierung erfolgt hingegen keine ideale impulsförmige Abtastung, sondern eine Flächenabtastung der Amplitudendichte, wodurch sich im Unterschied zur Zeitdiskretisierung in (7.93) ein zusätzlicher Ge-

wichtungsfaktor

$$\frac{\sin(\pi q(f - \frac{k}{q}))}{\pi q(f - \frac{k}{q})} \quad (7.97)$$

ergibt.

2. Während das Abtasttheorem bei gegebener Abtastfrequenz f_A die Bandbreite eines Signals begrenzt und damit die Signalvariabilität zwischen zwei benachbarten Abtastzeitpunkten einschränkt, verlangt das Quantisierungstheorem dagegen nach einem – bezogen auf die Breite q der Quantisierungsstufen – möglichst stark veränderlichen Signal.
3. Zur vollständigen Rekonstruktion der kontinuierlichen Signale ist sowohl bei der Zeitdiskretisierung als auch bei der Amplitudenquantisierung auf die strikte Bandbegrenzung (7.5) bzw. (7.94) zu achten, damit es zu keinen spektralen Überlappungen (Aliasing) kommt.

In vielen Fällen genügt zur Beschreibung der statistischen Signaleigenschaften die Angabe der Momente. Im Falle der Quantisierung genügt dabei die Einhaltung von (7.95), da für die Berechnung der Momente aus der charakteristischen Funktion nach (4.62) nur Aliasing im Bereich um $f = 0$ vermieden werden muss.

4. Reale charakteristische Funktionen sind nicht bandbegrenzt. Da der Amplitudenspektrum von physikalischen Signalen begrenzt ist, ist auch die zugehörige Amplitudendichte beschränkt auf ein Intervall

$$-a_{\max} \leq x \leq a_{\max}. \quad (7.98)$$

Solche Dichten besitzen aber nach der Unschärferelation immer eine nicht bandbegrenzte charakteristische Funktion [27].

5. Es ist wichtig festzustellen, dass bei der Herleitung des Quantisierungstheorems keinerlei Voraussetzungen über den Quantisierungsfehler gemacht wurden. Die Eigenschaften dieses Fehlers werden später gesondert betrachtet.

Im Folgenden sollen der Erwartungswert und die Varianz des quantisierten Signals berechnet werden unter der Voraussetzung, dass das Quantisierungstheorem eingehalten wurde.

Der **Erwartungswert** wird mittels der charakteristischen Funktion $\Phi_{x_q}(f)$ ermittelt:

$$E\{x_q\} = \frac{1}{j2\pi} \left. \frac{d\Phi_{x_q}(f)}{df} \right|_{f=0} \quad (7.99)$$

$$= \frac{1}{j2\pi} \left. \frac{d}{df} \left(\Phi_x(f) \cdot \frac{\sin(\pi q f)}{\pi q f} \right) \right|_{f=0} \quad (7.100)$$

$$= \frac{1}{j2\pi} \left[\left. \frac{d\Phi_x(f)}{df} \frac{\sin(\pi q f)}{\pi q f} + \Phi_x(f) \frac{\pi q f \cos(\pi q f) - \sin(\pi q f)}{\pi q f^2} \right] \right|_{f=0} \quad (7.101)$$

$$= \frac{1}{j2\pi} \left. \frac{d\Phi_x(f)}{df} \right|_{f=0} = E\{x\}.$$

Das **2. Moment** berechnet sich zu

$$E\{x_q^2\} = \frac{1}{(j2\pi)^2} \left. \frac{d^2\Phi_{x_q}(f)}{df^2} \right|_{f=0} \quad (7.102)$$

$$= \frac{1}{(j2\pi)^2} \left. \frac{d^2}{df^2} \left(\Phi_x(f) \cdot \frac{\sin(\pi q f)}{\pi q f} \right) \right|_{f=0} \quad (7.103)$$

$$= \frac{1}{(j2\pi)^2} \left[\left. \frac{d^2\Phi_x(f)}{df^2} \frac{\sin(\pi q f)}{\pi q f} + 2 \frac{d\Phi_x(f)}{df} \frac{\pi q f \cos(\pi q f) - \sin(\pi q f)}{\pi q f^2} \right. \right. \\ \left. \left. + \Phi_x(f) \frac{[2 - (\pi q f)^2] \sin(\pi q f) - 2\pi q f \cos(\pi q f)}{\pi q f^3} \right] \right|_{f=0} \quad (7.104)$$

$$= \frac{1}{(j2\pi)^2} \left. \frac{d^2\Phi_x(f)}{df^2} \right|_{f=0} + \frac{q^2}{12} = E\{x^2\} + \frac{q^2}{12}. \quad (7.105)$$

Die **Varianz des quantisierten Signals** ist

$$\sigma_{x_q}^2 = E\{(x_q - E\{x_q\})^2\} = E\{x_q^2\} - (E\{x_q\})^2, \quad (7.106)$$

und mit der Varianz des stetigen Signals $\sigma_x^2 = E\{x^2\} - (E\{x\})^2$ ergibt sich daraus

$$\sigma_{x_q}^2 = \sigma_x^2 + \frac{q^2}{12}. \quad (7.107)$$

Für die **höheren Momente** erhält man [51]:

$$E\{x_q^3\} = E\{x^3\} - \frac{q^2}{4} E\{x\}, \quad (7.108)$$

$$E\{x_q^4\} = E\{x^4\} - \frac{q^2}{2} E\{x^2\} + \frac{7}{240} q^4. \quad (7.109)$$

Insgesamt ergibt sich für die geschätzten Momente bei Erfüllung des Quantisierungstheorems:

- Der Erwartungswert der quantisierten Zufallsvariablen $E\{x_q\}$ ist gleich dem Erwartungswert der stetigen Zufallsvariablen $E\{x\}$. Bei einer Mittelwertbildung als Approximation des Erwartungswertes kann man nach der Quantisierung den Erwartungswert der stetigen Größe rekonstruieren.
- Die Varianz der quantisierten Zufallsvariablen $E\{x_q^2\}$ ist gegenüber der des stetigen Signals um $q^2/12$ höher. Die Varianz wird somit bei Quantisierung zu groß geschätzt.
- Bei der Berechnung der Momente höherer Ordnung aus quantisierten Größen ist ebenfalls eine Korrektur erforderlich. In der Statistik-Literatur wird diese als Sheppard-Korrektur bezeichnet [51].

7.2.4 Wahrscheinlichkeitsdichte des Quantisierungsfehlers

Mit den Ergebnissen aus Bsp. 7.6 wurde durch Plausibilitätsüberlegungen der Quantisierungsfehler als additive Überlagerung eines gleichverteilten Rauschens angenommen (vgl. Abb. 7.13). Im Folgenden soll die Gültigkeit dieser Annahme nachgewiesen werden.

Ist das Quantisierungstheorem erfüllt, so kommt es zu keinen spektralen Überlappungen der charakteristischen Funktion des quantisierten Signals. Im Nyquist-Band $-\frac{1}{2q} \leq f \leq \frac{1}{2q}$ erhält man aus (7.93) für $k = 0$ die charakteristische Funktion als das folgende Produkt:

$$\Phi_{x_q}(f) = \Phi_x(f) \cdot \frac{1}{q} R_q(f) = \Phi_x(f) \cdot \frac{\sin(\pi q f)}{\pi q f}. \quad (7.110)$$

Nach (4.64) entspricht aber die Multiplikation zweier charakteristischer Funktionen gerade der Summation zweier unabhängiger Zufallsgrößen

$$x_q = x + e_q. \quad (7.111)$$

Dabei ist $R_q(f)$ bis auf einen Faktor q gerade die Fourier-Transformierte einer gleichverteilten Wahrscheinlichkeitsdichte $f_{e_q}(e_q)$. Dies bestätigt aber gerade die Modellvorstellung aus Abb. 7.13, wonach die Quantisierung als additive Überlagerung des stetigen Signals durch ein unabhängiges Quantisierungsrauschen $e_q(t)$ beschrieben werden kann. Aus (7.110) erhält man die charakteristische Funktion $\Phi_{e_q}(f)$ und die zugehörige **Wahrscheinlichkeitsdichte** $f_{e_q}(e_q)$ des Quantisierungsrauschens zu

$$\Phi_{e_q}(f) = \frac{1}{q} R_q(f) = \frac{\sin(\pi q f)}{\pi q f}, \quad (7.112)$$

$$f_{e_q}(e_q) = \begin{cases} \frac{1}{q} & \text{für } -\frac{q}{2} \leq e_q \leq \frac{q}{2}, \\ 0 & \text{sonst.} \end{cases} \quad (7.113)$$

7.3

Satz 7.3: Lineares Quantisierungsmodell

Ist das Quantisierungstheorem erfüllt, so lässt sich die Quantisierung durch ein lineares Modell beschreiben. Dem stetigen Eingangssignal $x(t)$ wird ein im Intervall $[-q/2, q/2]$ gleichverteiltes Quantisierungsrauschen $e_q(t)$ (Abb. 7.18) additiv überlagert:

$$x_q(t) = x(t) + e_q(t). \quad (7.114)$$

Die Momente

$$E\{x_q^k(t)\} = E\{x^k(t)\} + E\{e_q^k(t)\} \quad (7.115)$$

des quantisierten Signals ergeben sich aus der Summe der Momente des kontinuierlichen Signals und des Quantisierungsrauschen. \diamond

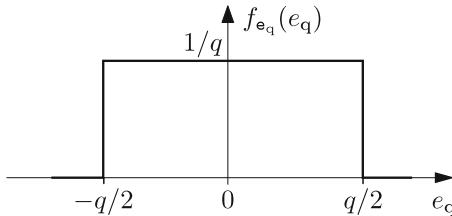


Abbildung 7.18. Wahrscheinlichkeitsdichte des Quantisierungsfehlers.

Mit dem Erwartungswert des Quantisierungsrauschen

$$E\{e_q(t)\} = \int_{-\infty}^{\infty} e_q \cdot f_{e_q}(e_q) de_q = \frac{1}{q} \int_{-\frac{q}{2}}^{\frac{q}{2}} e_q de_q = 0 \quad (7.116)$$

und der Varianz der zugehörigen Gleichverteilung

$$E\{e_q^2(t)\} = \int_{-\infty}^{\infty} e_q^2 \cdot f_{e_q}(e_q) de_q = \frac{1}{q} \int_{-\frac{q}{2}}^{\frac{q}{2}} e_q^2 de_q = \frac{q^2}{12} \quad (7.117)$$

erhält man sofort die Momente der quantisierten Größe nach (7.101) und (7.105).

In der Praxis ist das Quantisierungstheorem niemals exakt erfüllbar. Es zeigt sich aber, dass Satz 7.3 dennoch näherungsweise gilt, wenn die Eingangsamplituden mehrere Quantisierungsstufen durchlaufen.

Beispiel 7.9 (Grobe Quantisierung normalverteilter Größen): Es werden zwei Signale mit normalverteilten Amplituden nach Abb. 7.19 quantisiert. Die Breiten der gewählten Quantisierungsstufen betragen

$$q_1 = \sigma \quad \text{und} \quad q_2 = 3\sigma. \quad (7.118)$$

Dabei bezeichnet σ die Standardabweichung der Normalverteilung. Die Signalamplituden sind begrenzt auf die Amplitudenintervalle

$$|x_1| \leq 2,5\sigma \quad \text{bzw.} \quad |x_2| \leq 4,5\sigma. \quad (7.119)$$

Konstruiert man $f_{eq}(e_q)$ für die beiden Fälle, so gelangt man zu Abb. 7.20. Man erkennt, dass bei normalverteilten Amplituden bereits für 5 Quantisierungsstufen im Amplitudenbereich von $\pm 2,5\sigma$ die Annahme eines gleichverteilten Quantisierungsfehlers sehr gut erfüllt ist. Da nach Abschn. 7.2.2 die Amplitudendichte vieler Signale durch eine Normalverteilung beschrieben werden kann, ist der Ansatz eines gleichverteilten Quantisierungsfehlers in der Praxis gerechtfertigt. ■

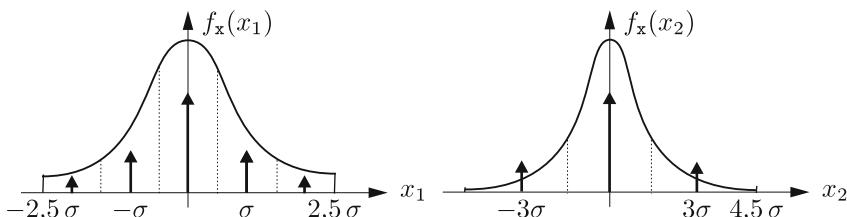


Abbildung 7.19. Quantisierung normalverteilter Amplituden.

7.2.5 Signal-Rausch-Verhältnis infolge der Quantisierung

Aus der Varianz des Quantisierungsrauschens lässt sich – bei gegebener Auflösung N des Quantisierers in Bit – das Signal-Rausch-Verhältnis (SNR) infolge der Quantisierung angeben. Das SNR ist das Verhältnis von Nutzsignalleistung P_{sig} zur Leistung des Quantisierungsrauschens $P_{stör}$.

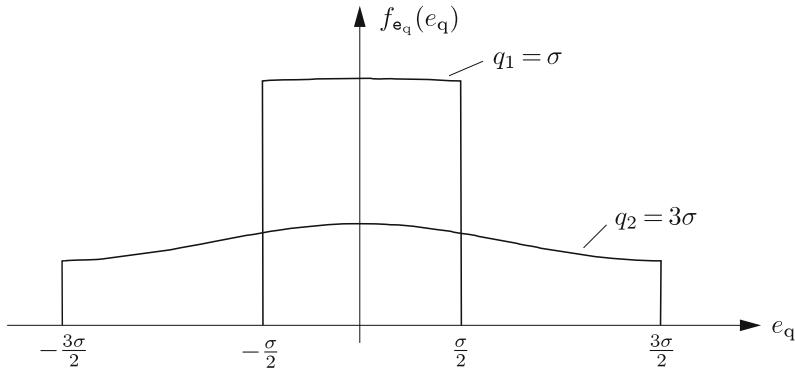


Abbildung 7.20. Wahrscheinlichkeitsdichte des Quantisierungsfehlers bei grober Quantisierung.

② 7.2.5.1 Sinusförmige Signale

Wird ein sinusförmiges Eingangssignal mit 2^N Stufen quantisiert, so erhält man den Effektivwert zu

$$u_{\text{eff}} = \frac{u_s}{\sqrt{2}} = \frac{2^{N-1} q}{\sqrt{2}}, \quad (7.120)$$

wobei u_s den Scheitelwert bezeichnet. Damit wird das Verhältnis der Signalleistungen zu

$$\text{SNR} = \frac{P_{\text{sig}}}{P_{\text{stör}}} = \frac{u_{\text{eff}}^2}{\sigma_{e_q}^2} = \frac{2^{2N-2} q^2 / 2}{q^2 / 12} = 2^{2N} \cdot 1,5. \quad (7.121)$$

Häufig wird das SNR in Dezibel angegeben:

$$\text{SNR}|_{\text{dB}} = 10 \lg(2^{2N} \cdot 1,5) = 6,02 \cdot N + 1,76. \quad (7.122)$$

② 7.2.5.2 Signale mit gleichverteilten Amplituden

Bei Signalen mit einer gleichverteilten Amplitudendichte beträgt der Effektivwert:

$$u_{\text{eff}} = \frac{2^{N-1} q}{\sqrt{3}}. \quad (7.123)$$

Damit erhält man verglichen mit (7.121) ein um den Faktor 1,5 kleineres Signal-Rausch-Verhältnis:

$$\text{SNR} = \frac{P_{\text{sig}}}{P_{\text{stör}}} = \frac{u_{\text{eff}}^2}{\sigma_{e_q}^2} = \frac{2^{2N-2} q^2 / 3}{q^2 / 12} = 2^{2N}. \quad (7.124)$$

In Dezibel ergibt sich:

$$\text{SNR}_{\text{dB}} = 10 \lg(2^{2N}) = 6,02 \cdot N. \quad (7.125)$$

7.2.6 Optimale Quantisierung

Bislang wurde bei der Quantisierung von äquidistanten Quantisierungsstufen ausgegangen (Abb. 7.10). Derartige Quantisierungskennlinien liefern jedoch meist suboptimale Ergebnisse bezüglich des Signal-Rausch-Verhältnisses, da die Amplituden des zu quantisierenden Signals nach Abschn. 7.2.2 in der Regel nicht gleichverteilt sind. Um Quantisierungsfehler zu reduzieren, ist es vorteilhaft, in Bereichen häufig vorkommender Signalamplituden eine dichtere Abtastung der Amplituden vorzunehmen; seltener vorkommende Amplituden dürfen dagegen größer quantisiert werden.

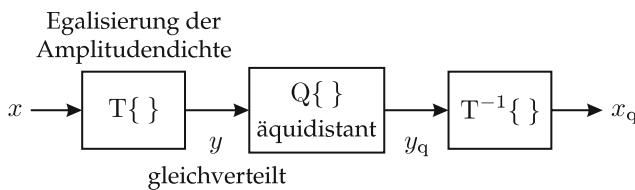


Abbildung 7.21. Optimale Quantisierung.

Bei Verwendung eines äquidistanten Quantisierers ist dafür vor der Quantisierung eine nichtlineare Transformation $T\{\}$ vorzunehmen (Abb. 7.21). Durch die nichtlineare Transformation wird die Amplitudendichte des Eingangssignals egalisiert, so dass die Signalamplituden im Quantisierungsbereich gleichverteilt sind. Nach der äquidistanten Quantisierung $Q\{\}$ wird die nichtlineare Transformation rückgängig gemacht. Durch diese Vorgehensweise erreicht man insgesamt eine Quantisierung, bei der die Quantisierungsstufen im Bereich hoher Amplitudendichte dichter zusammen liegen.

7.2.7 Minimierung des relativen Quantisierungsfehlers

Im Falle großer Variationsbereiche der Messgröße u und einer beschränkten Auflösung des Quantisierers ist bei Verwendung einer „linearen“ Quantisierungskennlinie (d. h. bei Verzicht auf die nichtlineare Transformation $T\{\}$) gemäß Abschn. 7.2.6) der relative Quantisierungsfehler

$$F_r = \frac{u_q - u}{u} = \frac{\Delta u}{u} \quad (7.126)$$

bei kleinen Messgrößen u sehr groß. Dabei bezeichnen u den wahren Messwert und u_q den angezeigten (quantisierten) Wert. Abhilfe schafft eine nicht-

lineare Transformation $T\{ \}$ vor der Quantisierung, die aus der Vorgabe nach einem konstanten relativen Fehler der Messgröße über dem Messbereich

$$\frac{\Delta u}{u} = F_r = \text{const.} \quad (7.127)$$

bei einer konstanten relativen Quantisierung

$$q = \frac{\Delta Z}{Z_{\max}} = \text{const.} \quad (7.128)$$

abgeleitet werden kann:

$$\frac{\Delta Z}{Z_{\max}} \cdot \frac{1}{q} = \frac{\Delta u}{u} \cdot \frac{1}{F_r} = 1, \quad (7.129)$$

$$\frac{Z}{Z_{\max}} = \frac{q}{F_r} \cdot \ln u + C. \quad (7.130)$$

Die absolute Quantisierung ist dabei ΔZ . Durch Abgleich der Kennlinie mit $u = u_{\text{ref}}$ bei $z = z_{\max}$ erhält man die Konstante

$$C = 1 - \frac{q}{F_r} \ln u_{\text{ref}} \quad (7.131)$$

und mit $u = u_{\min}$ bei $z = 0$ folgt das Verhältnis aus relativem Fehler und relativer Quantisierung:

$$\frac{F_r}{q} = \ln \frac{u_{\text{ref}}}{u_{\min}}. \quad (7.132)$$

Die Kennlinie der nichtlinearen Transformation wird damit die logarithmische Funktion

$$\frac{Z}{Z_{\max}} = \frac{q}{F_r} \ln \frac{u}{u_{\text{ref}}} + 1. \quad (7.133)$$

Beispiel 7.10 (Nichtlineare Quantisierung): Es sollen die Vorteile einer nichtlinearen Quantisierung verdeutlicht werden. Ein Messsignal soll bei einem Messbereich $[u_{\min}, u_{\max}]$ mit einem Dynamikumfang von

$$\frac{u_{\max}}{u_{\min}} = \frac{u_{\text{ref}}}{u_{\min}} = 100 \quad (7.134)$$

eine relative Genauigkeit von mindestens $F_r = 10^{-2}$ aufweisen. Bei einer linearen Kennlinie müsste die Quantisierung

$$q_{\text{lin}} \leq \frac{F_r}{u_{\text{ref}}/u_{\min}} = 10^{-4} \quad (7.135)$$

betragen, woraus sich eine Wortlänge von $\text{ld}(1/q_{\text{lin}}) \approx 13$ ergäbe. Bei der nichtlinearen logarithmischen Kennlinie ist nach (7.132) die Quantisierung lediglich

$$q_{\text{nl}} \leq \frac{F_r}{\ln(u_{\text{ref}}/u_{\text{min}})} = \frac{10^{-2}}{4,61} = 2,2 \cdot 10^{-3}, \quad (7.136)$$

was einer Wortlänge von $\text{ld}(1/q_{\text{nl}}) \approx 9$ entspricht. Durch die nichtlineare Kennlinie kann die Wortlänge um 4 Bit verkürzt werden. ■

Bei vergleichenden A/D-Umsetzern mit einem D/A-Umsetzer im Gegenkopplungszweig (Abschn. 7.3.1 und 7.3.2) kann die nichtlineare Kennlinie z. B. mit einem digitalen Speicher realisiert werden, der zwischen dem Ausgangswort Z und dem D/A-Umsetzer liegt, und in dem die Exponential-Kennlinie

$$\frac{u_b}{u_{\text{ref}}} = \left(\frac{u_{\text{min}}}{u_{\text{ref}}} \right)^{1-Z/Z_{\text{max}}} \quad (7.137)$$

abgelegt ist (Abb. 7.22). Eine Exponentielle-Kennlinie im Gegenkopplungszweig ergibt die gewünschte logarithmische Kennlinie des A/D-Umsetzers. Aufpassen muss man aber beim Anti-Aliasing-Filter am Eingang des A/D-Umsetzers. Durch die nichtlineare Kennlinie kann der Mittelwert der Ausgangsgröße Z verfälscht werden (Abschn. 4.2.5).

Beispiel 7.11 (KompaNDierung von Signalen): Ein weiteres Einsatzgebiet für nichtlineare Kennlinien ist die sogenannte **KompaNDierung** (engl. *compressing*, Kofferwort aus *compressing* und *expanding*) von Audiosignalen zur Datenreduktion. Bei der Kompression werden die Signale einer nichtlinearen Transformation $T\{\}$ unterzogen, bei der kleine Signalamplituden verstärkt werden, während große Amplituden weitgehend unverändert bleiben (Abb. 7.23). Auf diese Weise wird die Signalleistung erhöht, ohne den Wertebereich des Signals zu ändern. Nach der Kompression wird das Signal quantisiert, wodurch dem Signal ein additives Quantisierungsrauschen überlagert wird. Die anschließende Expansion $T^{-1}\{\}$ macht die Kompression rückgängig und verringert die Leistung des Signals, was sich auch auf das Quantisierungsrauschen auswirkt und somit zu einem höheren Signal-Rausch-Verhältnis führt als ohne KompaNDierung. ■

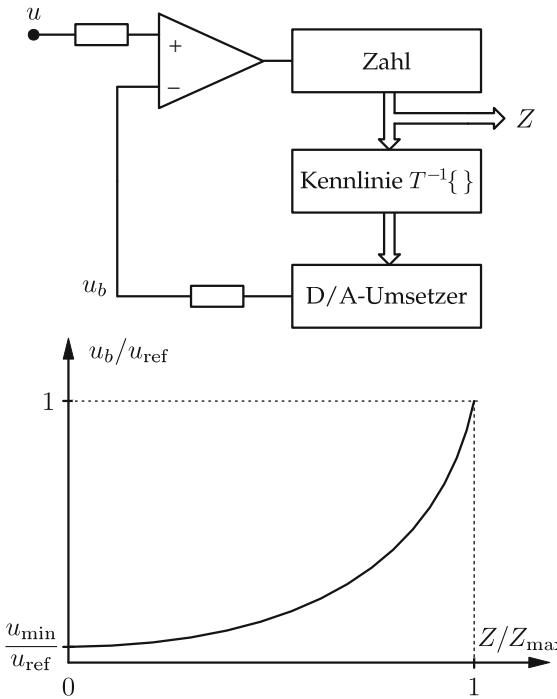


Abbildung 7.22. A/D-Umsetzer mit nichtlinearer inverser Kennlinie im Rückkopplungszweig.

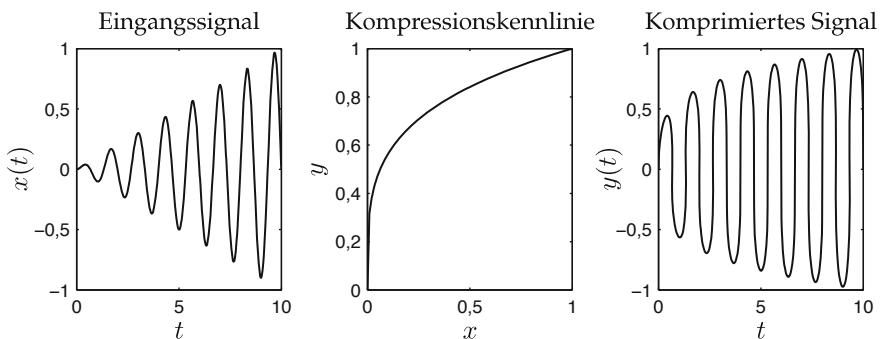


Abbildung 7.23. Kompression eines Signals.

7.2.8 Dithering

Voraussetzung für die approximative Erfüllung des Quantisierungstheorems ist, dass die Amplitude des kontinuierlichen Eingangssignals mehrere Quantisierungsstufen durchläuft. Dies setzt ein dynamisch veränderliches Signal voraus. Es gibt jedoch Fälle, in denen diese Voraussetzung verletzt ist – etwa bei langsam veränderlichen Signalen. Die Amplituden der zu messenden Signale durchlaufen dann zwar möglicherweise viele Quan-

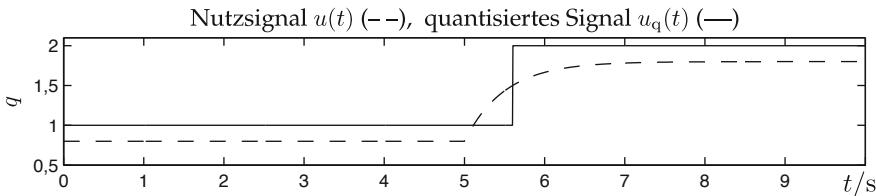


Abbildung 7.24. Verlauf eines langsam veränderlichen Signals (z. B. Temperaturverlauf).

tisierungsstufen, verharren aber während eines Beobachtungsintervalls innerhalb einer Stufe (vgl. Abb. 7.24).

Im Grenzfall eines konstanten Signalabschnittes

$$u(t) = u_0 = \text{const.} \quad (7.138)$$

sind die lokalen Amplitudendichten $f_u(u)$ daher impulsförmig. Die zugehörige charakteristische Funktion $\Phi_u(f)$ ist dann eine Harmonische:

$$f_u(u) = \delta(u - u_0) \quad \bullet\circ \quad \Phi_u(f) = e^{j2\pi f u_0}. \quad (7.139)$$

Durch die periodische Wiederholung von $\Phi_u(f)$ nach der Quantisierung entstehen wegen der fehlenden Bandbegrenzung starke spektrale Überlappungen der charakteristischen Funktion (ähnlich dem Aliasing bei der Zeitdiskretisierung). Dadurch lassen sich die ursprünglichen Signalmomente nicht mehr fehlerfrei bestimmen. Aus Abschn. 7.1.3 ist bekannt, dass spektrale Überlappungen mit Hilfe eines Anti-Aliasing-Filters verhindert werden können. Ein ähnliches Verfahren – das **Dithering** – ist auch bei der Amplitudenquantisierung anwendbar.

Das Dithering hat zum Ziel, spektrale Überlappungen infolge einer periodischen Fortsetzung der charakteristischen Funktion $\Phi_u(f)$ zu vermeiden. Dafür muss eine Bandbegrenzung von $\Phi_u(f)$ durch Multiplikation mit einer weiteren bandbegrenzten charakteristischen Funktion vorgenommen werden. Da die Multiplikation von charakteristischen Funktionen einer Addition von Zufallsvariablen entspricht, muss dem Nutzsignal $u(t)$ vor der Quantisierung ein Dithersignal $u_{\text{dith}}(t)$ addiert werden, welches eine näherungsweise bandbegrenzte charakteristische Funktion besitzt. Bei bekannten Momenten des Dithersignals lassen sich nach der Quantisierung die Momente des Nutzsignals durch Tiefpassfilterung rekonstruieren.

Beispiel 7.12 (Digitale Temperaturmessung): Die Temperatur eines chemischen Prozesses soll auf $\Delta T \leq 0,1^\circ\text{C}$ genau bestimmt werden. Dabei muss der Messbereich $0^\circ\text{C} \leq T < 400^\circ\text{C}$ erfasst werden. Die Tempera-

tur werde mittels eines A/D-Umsetzers erfasst. Wie hoch muss die Auflösung des Umsetzers sein?

Konventionell: Nach (7.62) ist der Quantisierungsfehler $e_q \leq q/2$. Daraus erhält man für die geforderte Auflösung die Breite q einer Quantisierungsstufe zu

$$\Delta T \leq 0,1^\circ\text{C} = \frac{q}{2} \quad \Rightarrow \quad q = 0,2^\circ\text{C}. \quad (7.140)$$

Für den Messbereich sind 2000 Stufen notwendig. Dies entspricht einem A/D-Umsetzer mit einer Auflösung von 11 Bit. Ohne überlagerte Störungen genügt ein einziger Abtastwert, um die Temperatur mit der geforderten Auflösung zu erfassen.

Dithering: Macht man sich die Eigenschaften des Dithering zunutze, so kann man mit einem A/D-Umsetzer weit geringerer Auflösung arbeiten. Dem Temperatursignal wird ein hochfrequentes, mittelwertfreies Dithersignal überlagert (Abb. 7.25). Das konstante Nutzsignal wird da-

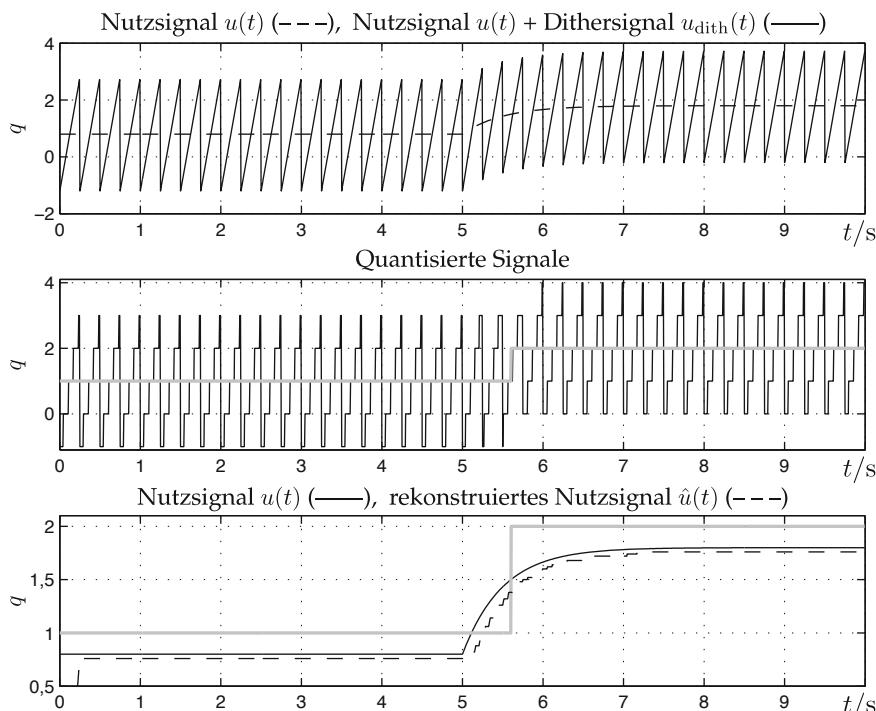


Abbildung 7.25. Verlauf des Messsignals ohne und mit Dithering (oben), quantisierte Signale (Mitte) und rekonstruiertes Signal durch Mittelwertbildung über eine Ditherperiode (unten).

durch über mehrere Quantisierungsstufen verteilt. Das Dithersignal sei eine hochfrequente Sägezahnschwingung der Amplitude $|u_{\max}| = 2q$. Die Amplitudendichte des Dithersignals ist eine Gleichverteilung im Intervall $-2q \leq u \leq 2q$, woraus für die charakteristische Funktion $\Phi_{\text{dith}}(f)$ eine Sinc-Funktion folgt (Abb. 7.26):

$$f_{\text{dith}}(u) = \begin{cases} \frac{1}{2u_{\max}} = \frac{1}{4q} & \text{für } -2q \leq u \leq 2q \\ 0 & \text{sonst} \end{cases} \quad (7.141)$$



$$\Phi_{\text{dith}}(f) = R_{4q}(f) = \frac{\sin(4\pi q f)}{4\pi q f} = \text{sinc}(4qf). \quad (7.142)$$

Durch die Abtastung der Amplitudendichte $f_u(u)$ wird die charakteristische Funktion $\Phi_u(f)$ mit der Quantisierungsfrequenz $1/q$ periodisch wiederholt. Ohne Dithering kommt es beim vorliegenden Temperatursignal $u(t)$ zu großen Aliasing-Fehlern (Abb. 7.27). Durch die Überlagerung des Dithersignals wird $\Phi_u(f)$ mit der charakteristischen Funktion des Dithersignals $\Phi_{\text{dith}}(f)$ multipliziert und somit bandbegrenzt. Nach Abb. 7.27 ist $\Phi_{\text{dith}}(f)$ sehr viel schmalbandiger als $\Phi_u(f)$. Dadurch werden hohe Frequenzen gedämpft und die Bandüberlappungen weitgehend vermieden.

Zur Bestimmung des Erwartungswertes der Temperatur benötigt man entsprechend (4.62) die gewichtete Ableitung der charakteristischen Funktion:

$$E\{u_q\} = \frac{1}{j2\pi} \left. \frac{d\Phi_{u_q}(f)}{df} \right|_{f=0}. \quad (7.143)$$

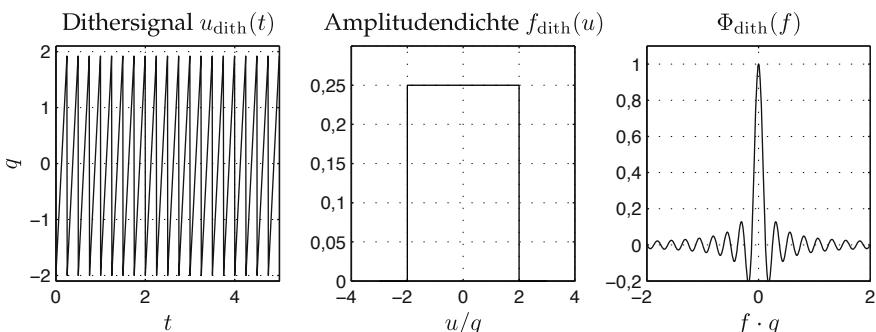


Abbildung 7.26. Sägezahnschwingung als Dithersignal, Amplitudendichte des Signals und charakteristische Funktion.

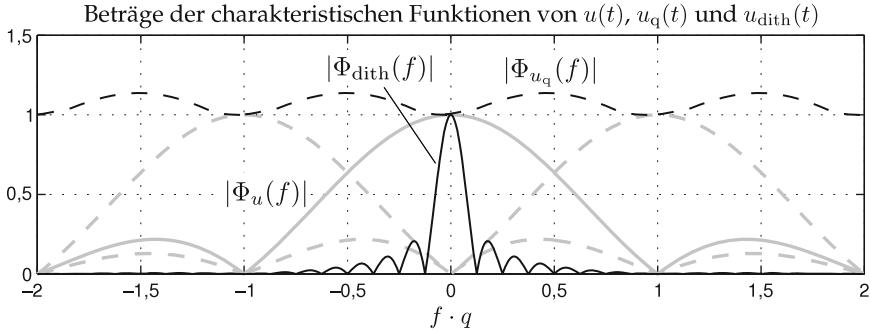


Abbildung 7.27. Betrag der charakteristischen Funktion des Nutzsignals $u(t)$ (grau) mit periodischen Wiederholungen und Aliasing infolge der Quantisierung (---) sowie zur Bandbegrenzung verwendete charakteristische Funktion des Dithersignals $u_{\text{dith}}(t)$ (—).

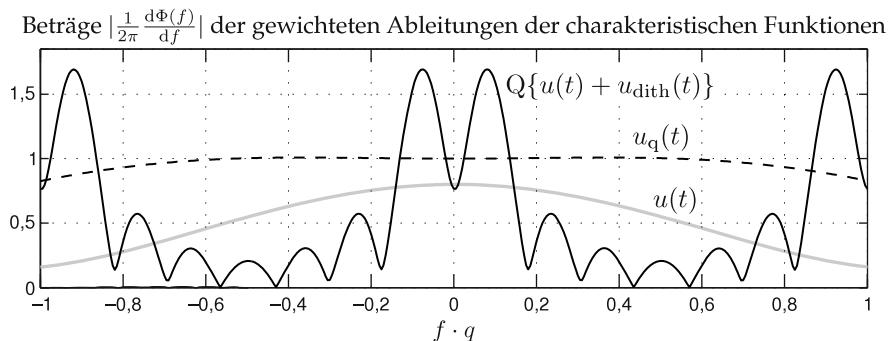


Abbildung 7.28. Betrag der gewichteten Ableitung der charakteristischen Funktion $\Phi(f)$ zur Bestimmung des Erwartungswertes für das kontinuierliche Signal $u(t)$ (grau) sowie für das mit (—) und ohne (---) Dithering quantisierte Signal $Q\{u(t) + u_{\text{dith}}(t)\}$ bzw. $u_q(t)$.

Da der Erwartungswert der Temperatur reell ist, verschwindet der Imaginärteil von (7.143) für $f = 0$. Abbildung 7.28 zeigt die Beträge der mit dem Faktor $1/(2\pi)$ gewichteten Ableitungen der charakteristischen Funktionen der mit und ohne Dithering quantisierten Signale $Q\{u(t) + u_{\text{dith}}(t)\}$ bzw. $u_q(t) = Q\{u(t)\}$ sowie des kontinuierlichen Signals $u(t)$. An der Stelle $f = 0$ kann der Erwartungswert (7.143) der Temperatur abgelesen werden. Bei Einsatz von Dithering zeigt sich eine gute Übereinstimmung mit dem Erwartungswert des kontinuierlichen Signals.

Eine anschließende Mittelwertbildung über zeitlich aufeinanderfolgende quantisierte Signale entfernt das Dithersignal und rekonstruiert den Erwartungswert des kontinuierlichen Nutzsignals. Die erzielbare Genauigkeit hängt vom Wertebereich des Dithersignals ab. Sie ist aber

höher als die halbe Quantisierungsstufe $q/2$ des verwendeten A/D-Umsetzers. ■

Anmerkungen

1. Als Dithersignale eignen sich sowohl periodische Signale als auch stochastische Signale wie weißes Rauschen. Es müssen nur die Momente des Dithersignals a priori bekannt sein. Vorteilhaft sind mittelwertfreie Signale, da sich dann die Bestimmung des Erwartungswertes des Nutzsignals durch eine gleitende Mittelung (MA-Filter) sehr einfach realisieren lässt.
2. Das Frequenzspektrum $U_{\text{dith}}(f)$ des Dithersignals sollte so beschaffen sein, dass keine spektralen Anteile im Frequenzbereich des Nutzsignals vorhanden sind. Vielmehr sollte der Frequenzbereich der beiden Signale weit auseinander liegen, um die anschließende Entfernung des Dithersignals durch eine Filterung zu erleichtern. Auf die Einhaltung des Abtasttheorems

$$f_{\text{Dither}} < f_A/2 \quad (7.144)$$

ist ebenfalls zu achten.

3. Den Gewinn an Genauigkeit erkauft man sich mit einer höheren Abtastrate und der anschließend notwendigen Filterung.

7.3 Analog-Digital-Umsetzer

7.3

Sollen Messwerte in einem Rechner verarbeitet werden, so müssen die stetigen, analogen Signale in zeitdiskrete, digitale Werte umgesetzt werden. Dies geschieht mit **Analog-Digital-Umsetzern**. Die Effekte, die allgemein bei der Abtastung und Quantisierung von Signalen auftreten, wurden bereits diskutiert. Im Folgenden wird deshalb die Arbeitsweise verschiedener Analog-Digital-Umsetzer (ADU) aufgezeigt. Dabei steht nicht die Schaltungstechnik im Vordergrund, die beispielsweise in [45, 52] nachgeschlagen werden kann. Vielmehr stehen die Effekte im Vordergrund, die mit einem spezifischen Umsetzungsverfahren einhergehen.

7.3.1 A/D-Nachlaufumsetzer

Abbildung 7.29 zeigt die Schaltung des **Nachlaufumsetzers**. Bei diesem einfachen ADU, der nach dem Kompensationsprinzip arbeitet, wird die umzu-setzende Eingangsspannung u mit der Ausgangsspannung u_v eines D/A-Umsetzers (DAU) verglichen. Der DAU setzt dabei den Zählerstand Z um, der die Ausgangsgröße darstellt, und wird mit der Referenzspannung u_{ref}

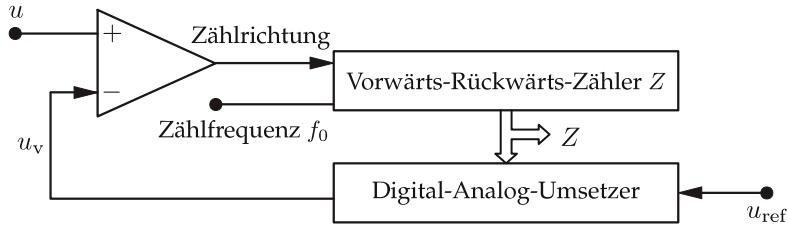


Abbildung 7.29. Schaltung des Nachlaufumsetzers.

gespeist. Für $u > u_v$ liegt am Komparatoriausgang für die Zähldurchrichtung ein 1-Signal an, sodass der Zählerstand Z mit der Zählfrequenz f_0 inkrementiert wird. Für $u < u_v$ wird rückwärts gezählt, weshalb Z mit der Frequenz f_0 dekrementiert wird.

Der Zähler pendelt sich bei konstanter Eingangsspannung u auf einen entsprechenden Zahlenwert Z ein, wobei er einen Grenzzyklus von einem Bit durchläuft. Diese Art der Umsetzung wird häufig auch als *Delta-Modulation* bezeichnet. Aus der Digital-Analog-Umsetzung

$$\frac{u_v}{u_{ref}} = \frac{Z}{Z_{\max}} \quad (7.145)$$

erhält man für konstante Eingangsspannungen $u = u_v$

$$Z = \frac{Z_{\max}}{u_{ref}} \cdot u . \quad (7.146)$$

Die Integrationszeitkonstante des Zählers

$$T_i = \frac{Z_{\max}}{f_0} = \frac{2^N}{f_0} \quad (7.147)$$

ist abhängig von der Auflösung 2^N und von der gewählten Zählfrequenz f_0 . Der A/D-Umsetzer reagiert auf Sprünge der Eingangsspannung mit einem rampenförmigen Nachlaufen (Abb. 7.30). Eine besonders einfache An-

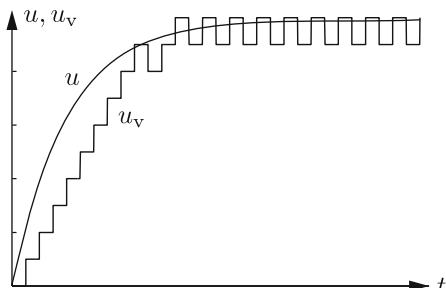


Abbildung 7.30. Rampenförmiges Nachlaufen des A/D-Umsetzers.

ordnung erhält man, wenn man für den D/A-Umsetzer einen sogenannten Rate-Multiplier (Abschn. 7.4.2) verwendet.

7.3.2 A/D-Umsetzer mit sukzessiver Approximation

Abbildung 7.31 zeigt rechts die Schaltung eines **A/D-Umsetzers mit sukzessiver Approximation**. Das Eingangssignal u wird auf einen Komparator gegeben und mit dem zurückgeführten Signal u_v verglichen, das sich aus einer schnellen D/A-Umsetzung (Abschn. 7.4.1) des Ausgangszahlenwertes Z ergibt. Der maximale Zahlenwert ist $Z_{\max} \approx 2^N$. Vor Beginn des Umsetzvorgangs werden alle N Bits a_i der Zahl Z auf null zurückgesetzt. Ausgehend vom höchstwertigen Bit (engl. *most significant bit*, MSB) a_{N-1} wird iterativ im i -ten Approximationsschritt das Bit a_{N-i} auf 1 gesetzt und die resultierende Spannungsdifferenz $u - u_v$ betrachtet.

- Für $u - u_v \geq 0$ bleibt das zuletzt gesetzte Bit a_{N-i} gesetzt,
- für $u - u_v < 0$ wird das zuletzt gesetzte Bit a_{N-i} auf 0 zurückgesetzt.

Somit ergibt sich die Zahl Z nach dem i -ten Approximationsschritt zu

$$Z(i) = \sum_{j=1}^i a_{N-j} \cdot 2^{N-j}. \quad (7.148)$$

Beispiel 7.13 (A/D-Umsetzung mit 6 Bit): Abbildung 7.31 zeigt links den Vorgang einer A/D-Umsetzung mit sukzessiver Approximation. Dabei soll die Eingangsspannung

$$\frac{u}{u_{\max}} = \frac{35}{64} \quad (7.149)$$

mit $N = 6$ Bit umgesetzt werden.

Schritt 1: Das höchstwertige Bit a_{N-1} wird probeweise auf 1 gesetzt. Die Zahl Z wird damit

$$Z(1) = a_{N-1} \cdot 2^{N-1} = 2^{N-1}. \quad (7.150)$$

Der Vergleich am Komparator ergibt, dass die Eingangsspannung u größer ist als die zurückgeführte Spannung $u_v(1)$:

$$\frac{u}{u_{\max}} \geq \frac{u_v(1)}{u_{\max}} = \frac{Z(1)}{Z_{\max}} = 2^{-1}, \quad (7.151)$$

weshalb das Bit a_{N-1} auf 1 gesetzt bleibt.

Schritt 2: Das Bit a_{N-2} wird probeweise auf 1 gesetzt. Die Zahl wird

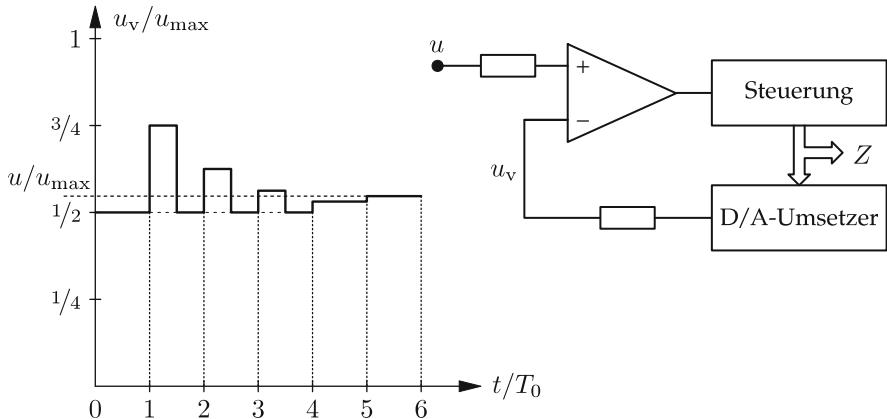


Abbildung 7.31. A/D-Umsetzer mit sukzessiver Approximation (rechts) und Umsetzung der Zahl $[100011]_2$ (links).

damit

$$Z(2) = 2^{N-1} + 2^{N-2}. \quad (7.152)$$

Der Vergleich am Komparator ergibt, dass die Eingangsspannung u kleiner ist als die zurückgeführte Spannung $u_v(2)$:

$$\frac{u}{u_{\max}} < \frac{u_v(2)}{u_{\max}} = \frac{Z(2)}{Z_{\max}} = 2^{-1} + 2^{-2}, \quad (7.153)$$

weshalb das Bit a_{N-2} auf 0 zurückgesetzt wird.

Die Eingangsspannung wird auf diese Weise in immer kleinere Vergleichsintervalle eingeschachtelt und verglichen. Nach N Schritten sind alle Stellen a_{N-i} ($i = 1, \dots, N$) und damit die Zahl Z bestimmt. ■

Für den gesamten Umsetzungsvorgang sind bei diesem A/D-Umsetzer N Schritte erforderlich. Damit der Wert der analogen Eingangsspannung u über den gesamten Verlauf der sukzessiven Approximation konstant bleibt, ist dem A/D-Umsetzer ein Abtast-Halte-Glied vorzuschalten [52].

7.3.3 Integrierende A/D-Umsetzer

Der **Zweirampen-A/D-Umsetzer** (engl. *dual-slope A/D converter*) ist ein integrierender ADU, bei dem die Umsetzung der Eingangsspannung auf eine Zeitintervallmessung zurückgeführt wird. Er ist in einer vereinfachten Ausführung in Abb. 7.32 dargestellt [43]. Ein Schalter Q schaltet abwechselnd die positive umzusetzende Eingangsspannung $u(t)$ und die feste negative Referenzspannung $-u_{\text{ref}}$ auf den Eingang des Integrators. Dessen Aus-

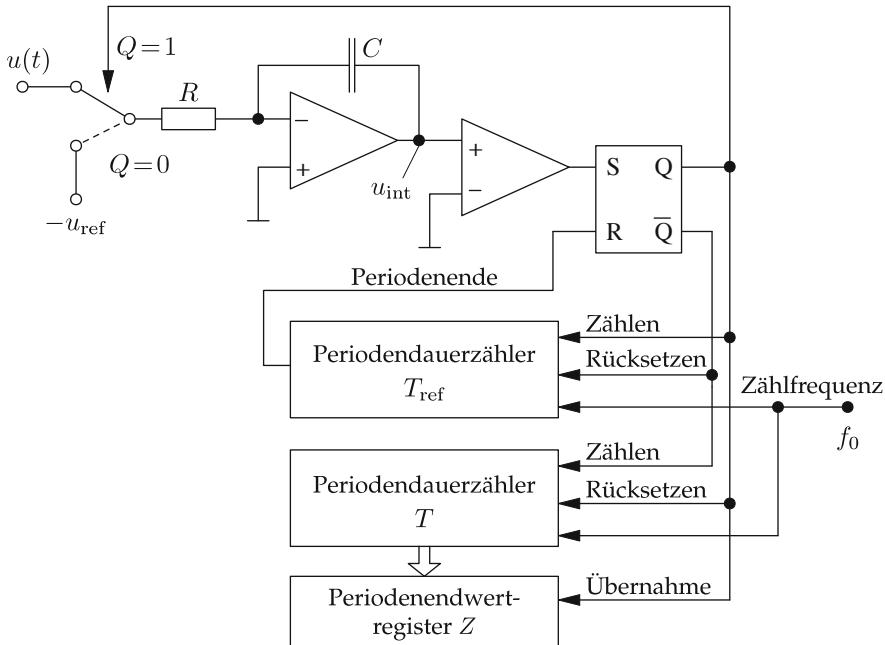


Abbildung 7.32. Vereinfachte schaltungstechnische Realisierung des integrierenden Zweirampen-A/D-Umsetzers.

Eingangsspannung

$$u_{\text{int}}(t) = -\frac{1}{RC} \int_0^t u(t) dt \quad (7.154)$$

verläuft während des Referenzintervalls T_{ref} linear abwärts (Abb. 7.33). Die Zeit T_{ref} wird dabei durch einen Zähler und einen Oszillatoren realisiert. Am Ende des Integrationsintervalls ist

$$u_{\text{int}}(T_{\text{ref}}) = -\frac{1}{RC} \int_0^{T_{\text{ref}}} u(t) dt = -\frac{T_{\text{ref}}}{RC} \bar{u}(t), \quad (7.155)$$

d. h. die Eingangsspannung wird durch die Integration im Referenzintervall gemittelt.

Am Ende der Periodendauer T_{ref} wird das Flipflop Q rückgesetzt und der Eingang auf die Referenzspannung $-u_{\text{ref}}$ umgeschaltet. Die Integrator-Ausgangsspannung verläuft im folgenden Zeitintervall $[T_{\text{ref}}, T_{\text{ref}} + T]$ ram-

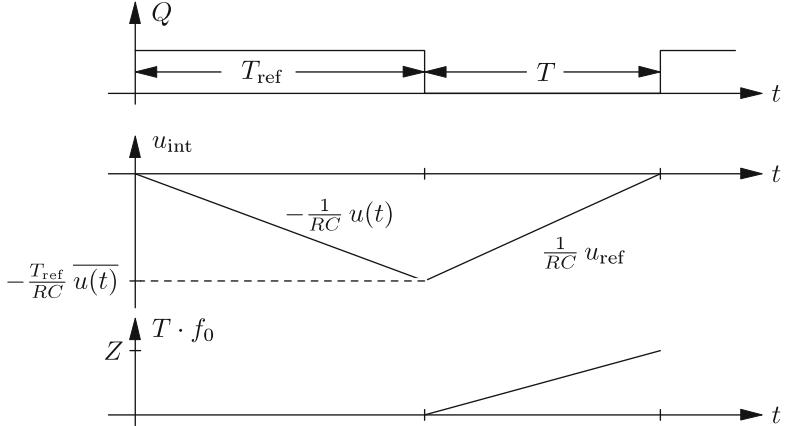


Abbildung 7.33. Verlauf der Integrator-Ausgangsspannung $u_{\text{int}}(t)$ und des Periodendauerzählers T während einer Umsetzungsperiode beim integrierender Zweirampen-A/D-Umsetzer.

penförmig bis null zurück:

$$u_{\text{int}}(T_{\text{ref}} + T) = -\frac{T_{\text{ref}}}{RC} \bar{u} + \frac{1}{RC} \int_0^T u_{\text{ref}} dt \quad (7.156)$$

$$= -\frac{T_{\text{ref}}}{RC} \bar{u} + \frac{T}{RC} u_{\text{ref}} = 0. \quad (7.157)$$

Mit dem Nulldurchgang der Spannung $u_{\text{int}}(t)$ setzt der Komparator das Flipflop Q . Der Zählerendwert $Z = T \cdot f_0$ wird am Ende der Periodendauer T in das Periodenendwertregister übernommen. Aus (7.157) erhält man

$$\frac{\bar{u}}{u_{\text{ref}}} = \frac{T}{T_{\text{ref}}} = \frac{T \cdot f_0}{T_{\text{ref}} \cdot f_0} = \frac{Z}{Z_{\text{ref}}} \quad \Rightarrow \quad \bar{u} = \frac{u_{\text{ref}}}{Z_{\text{ref}}} \cdot Z. \quad (7.158)$$

Die Zeitkonstante RC geht nicht in den Umsetzungsfehler ein. Auch die Zählfrequenz f_0 fällt heraus. Integrierende A/D-Umsetzer weisen eine sehr hohe stationäre Genauigkeit auf. Wenn der Mittelwert der Störungen während der Periodenzeitz T_{ref} null ist, gehen diese nicht in das Umsetzergebnis ein (Mittelwertfilter nach Abschn. 7.1.4). Dies wird mit einer sehr langen Umsetzungszeit erkauft. Signale mit höheren Frequenzkomponenten werden aber durch die implizite Mittelwertbildung gedämpft und das Ergebnis der A/D-Umsetzung dadurch verfälscht. Das Verfahren kann auf Eingangsspannungen mit beliebigem Vorzeichen erweitert werden.

7.3.4 Delta-Sigma-Umsetzer

Ein **Delta-Sigma-Umsetzer** besteht aus einem Modulator und einem Digitalfilter (Abb. 7.34). Das analoge Eingangssignal $u(t)$ wird im Modulator in eine binäre Impulsfolge $q(n)$ umgesetzt. Die Häufigkeit des hohen Spannungspegels ist dabei proportional zur Eingangsspannung. Das sich anschließende digitale Filter setzt die Impulsfolge in einen digitalen Zahlenwert $u_q(n)$ um, der dem abgetasteten und quantisierten Eingangssignal entspricht. Der Vorteil dieser Anordnung ist die einfache monolithische Integrierbarkeit. Durch die sogenannte **Rauschformung** (engl. *noise shaping*) dieses Umsetzungsprinzips können hohe Genauigkeiten erzielt werden. Delta-Sigma-Umsetzer werden häufig in Systemen mit dynamischen Signalen eingesetzt (Audio-Technik, Signalprozessoren). Aber auch für stationäre Signale in der Messtechnik finden diese Umsetzer Verwendung.



Abbildung 7.34. Struktur eines Delta-Sigma-Umsetzers.

7.3.4.1 Delta-Sigma-Modulator

Die Grundidee des **Delta-Sigma-Modulators** besteht darin, mit einem Integralregler $G_I(s)$ die Spannungs-Zeit-Flächendifferenz $e(t)$ von Eingangssignal $u(t)$ und zu modulierendem Ausgangssignal $q(t)$ zu null zu regeln (Abb. 7.35). Aus

$$E(s) = U(s) - E(s) \cdot G_I(s) \quad \text{mit} \quad G_I(s) = \frac{1}{T_i s} \quad (7.159)$$

erhält man die Laplace-Transformierte des Fehlers am Eingang des Integrators

$$E(s) = \frac{T_i s}{T_i s + 1} \cdot U(s), \quad (7.160)$$

wobei T_i die Integrationszeitkonstante des Reglers bezeichnet. Man kann mit Hilfe des Endwertsatzes der Laplace-Transformation zeigen, dass der Regelkreis stationär genau ist. Dazu wird angenommen, dass sich das Ein-

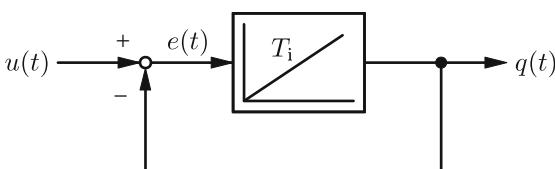


Abbildung 7.35. Integralregelkreis zwischen Eingangs- und Ausgangssignal.

gangssignal sprunghaft ändere,

$$U(s) = \frac{u_{\max}}{s}, \quad (7.161)$$

und während der A/D-Umsetzung konstant bleibe. Nach dem Endwertsatz verschwindet die stationäre Regeldifferenz $e(t)$ für $t \rightarrow \infty$:

$$\lim_{t \rightarrow \infty} e(t) = \lim_{s \rightarrow 0} s \cdot E(s) = \lim_{s \rightarrow 0} \frac{T_i s}{T_i s + 1} u_{\max} = 0. \quad (7.162)$$

Der Delta-Sigma-Modulator baut auf diesem Prinzip auf. Er besteht aus einem Integrator, der die Differenz zwischen Eingangssignal $u(t)$ und dem zurückgeführten Ausgangssignal integriert. Allerdings muss das Ausgangssignal nach Abb. 7.34 noch in eine Impulsfolge $q(n)$ umgesetzt werden. Dies erfolgt durch einen getakteten Komparator, der das Signal am Ausgang des Integrators mit 1 Bit quantisiert (Abb. 7.36).

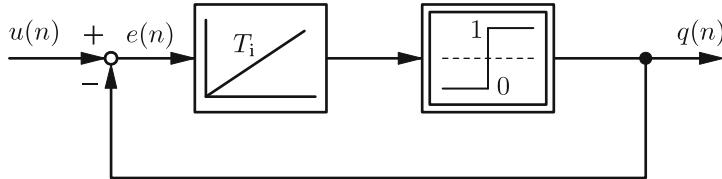


Abbildung 7.36. Strukturbild des Delta-Sigma-Modulators.

Durch die Taktung des Komparators mit der Periode t_A erhält man ein zeitdiskretes System, weshalb die Systemanalyse im z-Bereich erfolgen muss. Dazu beschreibt man die zeitdiskrete Integration durch die Rechteckregel vorwärts [42]:

$$y_n = y_{n-1} + t_A e_{n-1} \quad \circ \bullet \quad G_I(z) = \frac{t_A}{z - 1}. \quad (7.163)$$

Der Quantisierer wird entsprechend dem linearen Quantisierungsmodell (Satz 7.3) durch die Überlagerung des Ausgangs $q(n)$ mit einem Quantisierungsrauschen $n(n)$ beschrieben. Berücksichtigt man nun noch die Integrationszeitkonstante des Integrators T_i und eine zusätzliche Verstärkung V , so erhält man das in Abb. 7.37 dargestellte Modell des Delta-Sigma-Modulators in zeitdiskreter Form.

Die z-Transformierte des Ausgangssignals des geschlossenen Kreises ist einschließlich der Störung

$$Q(z) = \frac{G(z)}{1 + G(z)} U(z) + \frac{1}{1 + G(z)} N(z) \quad (7.164)$$

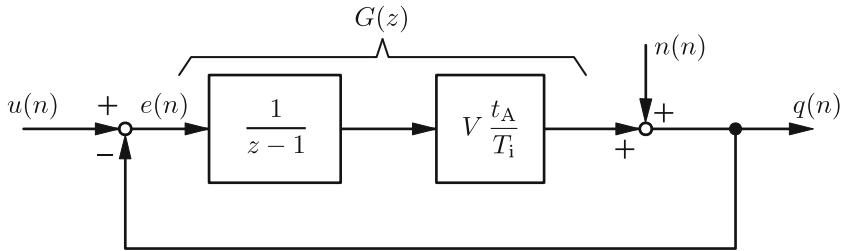


Abbildung 7.37. Strukturbild des Delta-Sigma-Modulators erster Ordnung.

$$= \frac{V \cdot t_A / T_i}{(z-1) + V \cdot t_A / T_i} U(z) + \frac{z-1}{(z-1) + V \cdot t_A / T_i} N(z). \quad (7.165)$$

Rückgekoppelte Systeme können instabil werden, weshalb die Stabilität der Anordnung überprüft werden muss. Zeitdiskrete Systeme sind stabil, wenn alle Polstellen der Übertragungsfunktion innerhalb des Einheitskreises in der komplexen z -Ebene liegen [42]:

$$|z_\infty| < 1 \quad \Leftrightarrow \quad |z_\infty| = |1 - V \cdot t_A / T_i| < 1. \quad (7.166)$$

Für die Stabilität des Regelkreises ergibt sich daraus die Bedingung

$$0 < V \cdot \frac{t_A}{T_i} < 2. \quad (7.167)$$

Die Integrationszeitkonstante T_i muss größer als das halbe Produkt von Verstärkung V und Abtastzeit t_A sein:

$$T_i > V \cdot t_A / 2. \quad (7.168)$$

⌚ 7.3.4.2 Rauschformung (*noise shaping*)

Setzt man $T_i = V \cdot t_A$ in (7.164) ein, so erhält man die Ausgangsfunktion

$$Q(z) = z^{-1} U(z) + (1 - z^{-1}) N(z) = G_u(z) U(z) + G_n(z) N(z) \quad (7.169)$$

mit der Führungsübertragungsfunktion $G_u(z) = z^{-1}$ und der Störübertragungsfunktion $G_n(z) = 1 - z^{-1}$. Mit $z = e^{j2\pi f t_A}$ ergeben sich aus $G_u(z)$ und $G_n(z)$ die zugehörigen Amplitudengänge durch Betragsbildung:

$$|G_u(f)| = |e^{-j2\pi f t_A}| = 1, \quad (7.170)$$

$$|G_n(f)| = |1 - e^{-j2\pi f t_A}| = 2 |\sin(\pi f / f_A)|. \quad (7.171)$$

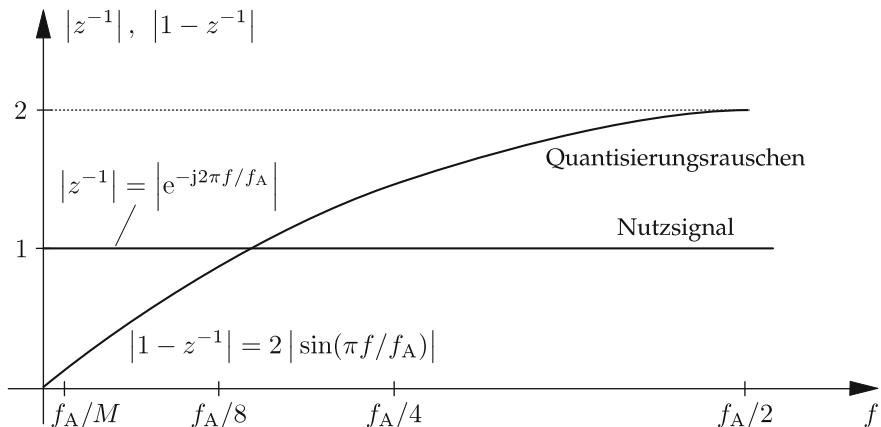


Abbildung 7.38. Amplitudengänge des Delta-Sigma-Modulators erster Ordnung für Nutzsignal $u(n)$ und Quantisierungsrauschen $n(n)$.

In Abb. 7.38 sind die beiden Amplitudengänge aufgetragen. Damit lässt sich das Übertragungsverhalten des Delta-Sigma-Modulators wie folgt zusammenfassen:

- Das Nutzsignal $u(t)$ wird nicht gedämpft, sondern lediglich um eine Abtastperiode t_A verzögert.
- Das Quantisierungsrauschen wird mit $1 - z^{-1}$ gefiltert, was einer Hochpassfilterung entspricht. Die Verlagerung der Rauschleistung zu hohen Frequenzen hin nennt man **Rauschformung** (engl. *noise shaping*).
- Für einen großen Störabstand ist es wichtig, dass die maximale Frequenz des Nutzsignals f_g wesentlich kleiner ist als die halbe Abtastfrequenz $f_A/2$. Eine anschließende Tiefpassfilterung isoliert das Nutzsignal und unterdrückt die Störanteile des Quantisierungsrauschen für Frequenzen oberhalb von f_g .

In der Praxis wird man das Eingangssignal daher M -fach überabtasten [42],

$$f_A = M f_g, \quad f_g = f_A/M, \quad (7.172)$$

wobei der **Überabtastfaktor** M (engl. *oversampling factor*) vom geforderten Signal-Rausch-Verhältnis abhängt. Als Maß für das Signal-Rausch-Verhältnis des Delta-Sigma-Modulators 1. Ordnung soll näherungsweise das qua-

drierte Verhältnis der Amplitudengänge für Nutz- und Störsignal

$$\text{SNR} = \frac{P_{\text{sig}}}{P_{\text{stör}}} = \frac{\int_0^{f_g} S_{uu}(f) |G_u(f)|^2 df}{\int_0^{f_g} S_{nn}(f) |G_n(f)|^2 df} \quad (7.173)$$

$$\approx \frac{|G_u(f_g)|^2}{|G_n(f_g)|^2} = \frac{1}{4 \sin^2(\pi f_g / f_A)} \quad (7.174)$$

bei der gewählten Grenzfrequenz f_g herangezogen werden, vgl. (6.191).

Handelsübliche Delta-Sigma-Umsetzer verwenden darüber hinaus auch Modulatoren höherer Ordnung. Die einfache Integration wird dabei durch Übertragungsfunktionen höherer Ordnung ersetzt. Man erreicht dadurch eine stärkere Rauschformung, d. h. eine noch größere Dämpfung des Quantisierungsrauschens für kleine Frequenzen. Abbildung 7.39 zeigt die erreichbaren Signal-Rausch-Verhältnisse für Modulatoren der Ordnungen eins bis drei abhängig vom Überabtastfaktor M .

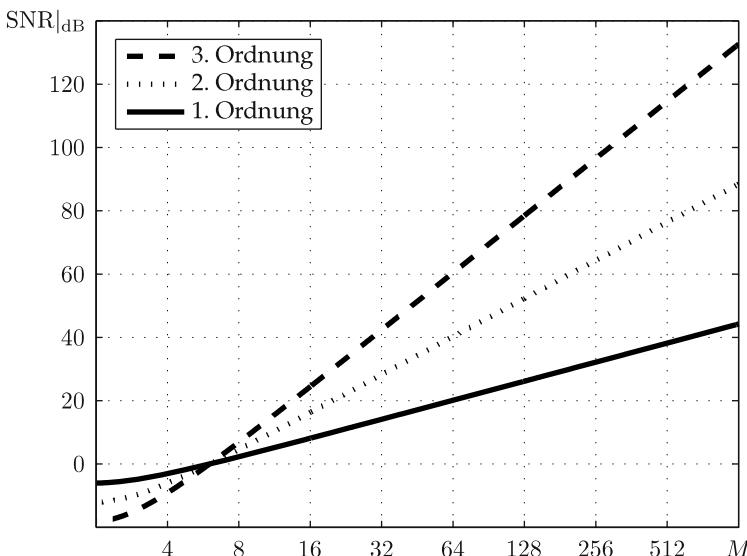


Abbildung 7.39. Signal-Rausch-Verhältnisse für Modulatoren verschiedener Ordnungen.

Beispiel 7.14 (Überabtastung bei Delta-Sigma-Umsetzern): Gegeben sei ein Delta-Sigma-Modulator 3. Ordnung. Wie groß muss die Oversampling-Rate M gewählt werden, damit der A/D-Umsetzer bei einem gleichverteilten Eingangssignal (vgl. Abschn. 7.2.5.2) einen Störabstand von

- 1.) $96 \text{ dB} \doteq 16 \text{ Bit Genauigkeit}$
 - 2.) $60 \text{ dB} \doteq 10 \text{ Bit Genauigkeit}$
- erreicht? Mit der Abb. 7.39 erhält man die Raten

$$M_1 = 256, \quad M_2 = 64. \quad (7.175)$$

Dies sind die Parameter gängiger, kommerziell erhältlicher Umsetzer. ■

② 7.3.4.3 Digitalfilter

Das **Digitalfilter** nach dem Modulator (Abb. 7.34) hat die folgenden Aufgaben:

1. Tiefpassfilterung zur Trennung von Nutzsignal $u(n)$ und Quantisierungsrauschen $n(n)$;
2. Umsetzung der binären Impulsfolge in eine digitale Zahl.

Zum Einsatz kommen **FIR-Filter** (FIR: *finite impulse response*) der Länge N mit folgender Struktur

$$y_n = \sum_{i=0}^{N-1} b_i \cdot q_{n-i}, \quad (7.176)$$

d. h. die letzten N Eingangswerte q_{n-i} werden mit Gewichtsfaktoren b_i versehen und aufsummiert. Da das System nicht rückgekoppelt ist, gibt es keine Stabilitätsprobleme. Um das hochfrequente Quantisierungsrauschen wirksam zu dämpfen, sind allerdings hohe Filterlängen notwendig ($N > 400$). Der Entwurf geeigneter Filter wird in der Literatur ausführlich diskutiert [15, 16, 24] und soll hier nicht besprochen werden.

Ein spezielles und sehr einfach zu implementierendes FIR-Filter soll jedoch kurz diskutiert werden. Setzt man alle Koeffizienten $b_i = 1/N$, so erhält man

$$y_n = \frac{1}{N} \sum_{i=0}^{N-1} q_{n-i}. \quad (7.177)$$

Nach Abschn. 6.1.7 ist y_n gerade der Mittelwert der letzten N Eingangswerte. Die Filtercharakteristik des MA-Filters (*moving average*) ist in Abschn. 7.1.4 beschrieben. Besonders einfach lässt sich dieses Filter mittels eines Zählers mit zyklisch anfallendem Zählerstand z implementieren, der die letzten N Werte des Bitstroms $q(n)$ aufsummiert und am Ende der Sum-

mation in ein Register übergibt:

$$\frac{Z}{Z_{\max}} = 2^{-k} \sum_{i=0}^{2^k-1} q_i = \frac{u}{u_{\text{ref}}} \quad \text{mit} \quad 2^k = N. \quad (7.178)$$

Nach der Summation über N Werte muss der Zähler zurückgesetzt werden. Man erreicht damit sowohl eine Tiefpassfilterung als auch eine Datenreduktion, da nur jeder N -te Wert ausgegeben wird. Die Überabtastung durch den Modulator wird dadurch wieder rückgängig gemacht.

Die reale Schaltung des Delta-Sigma-Umsetzers 1. Ordnung ist in Abb. 7.40 dargestellt. Der Integrationsverstärker und der Komparator werden beide mit f_A getaktet, um nach dem Zerhackerprinzip Offsetfehler zu eliminieren. Der Bitstrom $q(n)$ steuert die Rückkopplung der Referenzspannung $\pm u_{\text{ref}}$. Der Zähler, welcher ebenfalls mit f_A getaktet ist, summiert den Bitstrom über N Takte. Am Ende der Mittelungsperiode werden der Zählerstand in ein Register übernommen und der Zähler auf null zurückgesetzt.

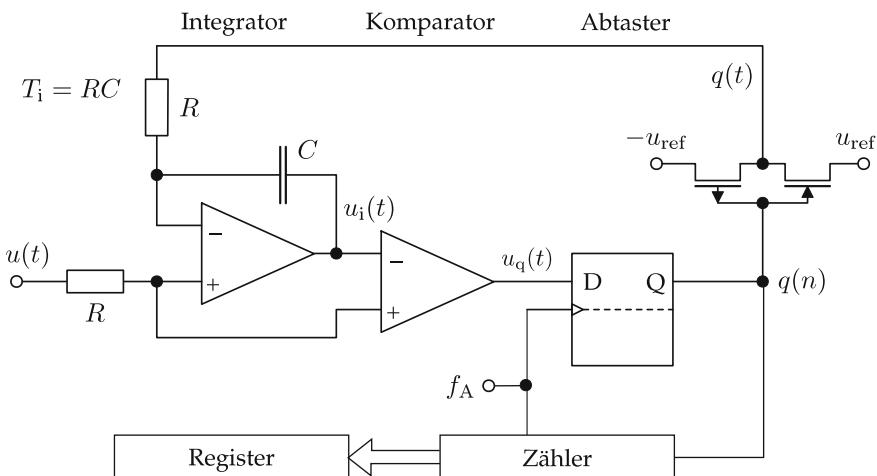


Abbildung 7.40. Schaltung des Delta-Sigma-Umsetzers 1. Ordnung mit MA-Filter.

▷ 7.3.4.4 Stationäres Verhalten

Bei dem Regelungsvorgang stellt sich ein Grenzzyklus ein, der im Folgenden analysiert werden soll. Die Impulsfolge $q(n)$ ist allgemein nicht-äquidistant. Zum einfacheren Verständnis sollen im Folgenden aber lediglich äquidistante Folgen betrachtet werden, die für die folgenden konstan-

ten Eingangsspannungen auftreten:

$$\frac{u}{u_{\text{ref}}} \in \left\{ \frac{1}{2^N}, \dots, \frac{1}{m}, \dots, \frac{1}{2}, \frac{3}{4}, \dots, \frac{m-1}{m}, \dots, \frac{2^N - 1}{2^N} \right\}. \quad (7.179)$$

Abbildung 7.41 zeigt ein Zeitdiagramm für die Funktion des Delta-Sigma-Umsetzers für $u/u_{\text{ref}} = 1/m$. Am Eingang des Integrators liegt das Differenzsignal $u(t)/u_{\text{ref}} - q(t)$, dessen Spannungs-Zeit-Fläche wegen der statio-nären Genauigkeit des Integralreglers gleich null ist. Die Integration ergibt im ersten Teilintervall $[0, t_A]$

$$u_i(t) = u_i(0) - \frac{1}{T_i} \int_0^t \frac{m-1}{m} \sigma(\tau) d\tau = u_i(0) - \frac{m-1}{m} \cdot \frac{t}{T_i}, \quad (7.180)$$

$$u_i(t_A) = u_i(0) - \frac{m-1}{m} \cdot \frac{t_A}{T_i}. \quad (7.181)$$

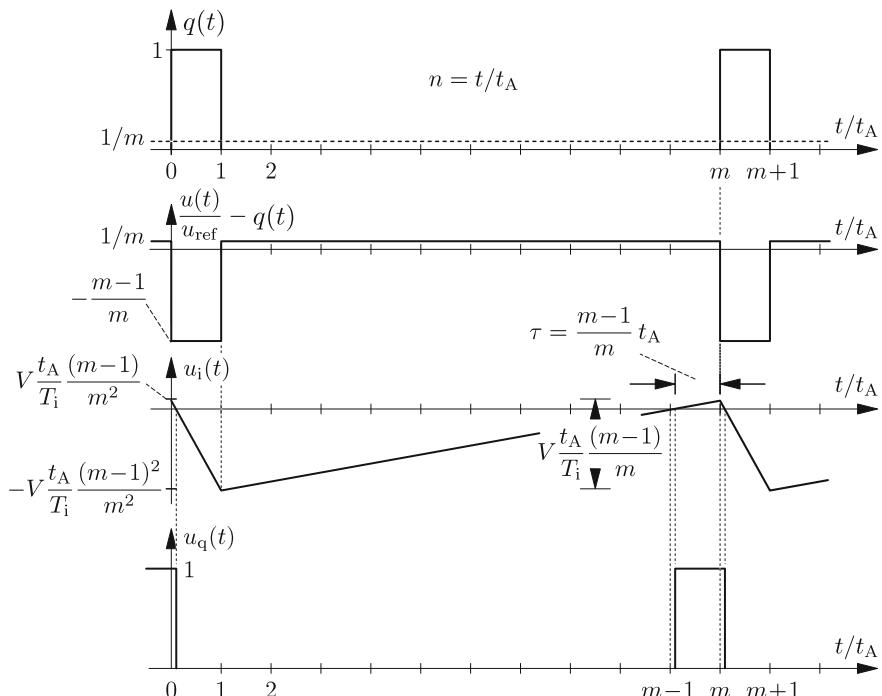


Abbildung 7.41. Grenzzyklus des Delta-Sigma-Umsetzers für äquidistante Impulsfolgen.

Im zweiten Teilintervall $[t_A, m t_A]$ erhält man

$$u_i(t) = u_i(t_A) + \frac{1}{T_i} \int_{t_A}^t \frac{1}{m} \sigma(\tau - t_A) d\tau = u_i(t_A) + \frac{1}{m} \cdot \frac{t - t_A}{T_i}, \quad (7.182)$$

$$u_i(m t_A) = u_i(0) = u_i(t_A) + \frac{m-1}{m} \cdot \frac{t_A}{T_i}. \quad (7.183)$$

Nach der Verstärkung um den Faktor V stellt sich ein Grenzzyklus mit folgender Amplitude und Frequenz ein:

$$A_{gr} = \frac{(m-1)}{m} \frac{t_A}{T_i} \cdot V, \quad f_{gr} = \frac{1}{m t_A}. \quad (7.184)$$

Damit lassen sich für das stationäre Verhalten folgende Aussagen gewinnen:

1. Für stationäre Eingangssignale ist $u_i(t)$ ein langsam veränderliches Signal. Das Quantisierungstheorem ist somit verletzt und die Annahme eines additiv überlagerten Quantisierungsrauschens gilt hier nicht mehr. Vielmehr erhält man für $q(n)$ eine periodische Impulsfolge. Das Störspektrum ist dann aber ein diskretes Linienspektrum!
2. Für große Werte von m , die bei kleinen Eingangsspannungen $|u(t)| \approx 0$ und bei großen Eingangsspannungen $|u(t)| \approx u_{ref}$ auftreten, erhält man sehr kleine Wiederholfrequenzen f_{gr} des Grenzzyklus. Der erste Impuls des diskreten Störspektrums kann deshalb im Nutzfrequenzbereich liegen.
3. Die effektiven Integrationszeitkonstanten im zweiten Teilintervall

$$\frac{(m-1) t_A}{A_g} = \frac{m T_i}{V} \quad (7.185)$$

sind bei großen m durch analoge Integrierer nur näherungsweise zu realisieren und bewirken deshalb einen mit m ansteigenden Mittelungsfehler.

In praktischen Anwendungen sollte man zur Vermeidung großer Fehler den nutzbaren Amplitudenbereich der umzusetzenden Eingangsspannung auf z. B.

$$\frac{1}{m} = 0,2 \leq \frac{u}{u_{ref}} = \frac{Z}{Z_{max}} \leq 0,8 = \frac{m-1}{m} \quad (7.186)$$

begrenzen, d. h. $m \leq 5$. In diesem Bereich sind die Fehler gering. Der Delta-Sigma-Umsetzer erster Ordnung hat bei Einhaltung dieser Randbedingungen den Vorteil einer sehr hohen stationären Genauigkeit. Die hohe Genauigkeit wird mit langen Umsetzzeiten erkauft.

7.3.5 Ratiometrische Messung

Die Genauigkeit vieler Messverfahren hängt von der Genauigkeit ab, mit der eine Speisespannung u_{ref} konstant gehalten werden kann. Dieser deformierende Fehlereinfluss kann eliminiert werden, wenn man für die Speisespannung des Sensors und die Referenzspannung des A/D-Umsetzers die gleiche Spannung u_{ref} verwendet. Das Messergebnis ist dann unabhängig von der Speisespannung u_{ref} . Man spricht von einer **ratiometrischen Messung**. Als Beispiel sei eine Widerstandsmessbrücke betrachtet, deren Speisespannung gleich der Referenzspannung des A/D-Umsetzers sein soll (Abb. 7.42). Der Zahlenwert

$$\frac{Z}{Z_{\max}} = \frac{\Delta u}{u_{\text{ref}}} = \frac{\Delta R}{2R} \cdot \frac{1}{1 + \frac{\Delta R}{2R}} \quad (7.187)$$

hängt ausschließlich von den *Widerstandsverhältnissen* der Brückenschaltung ab.

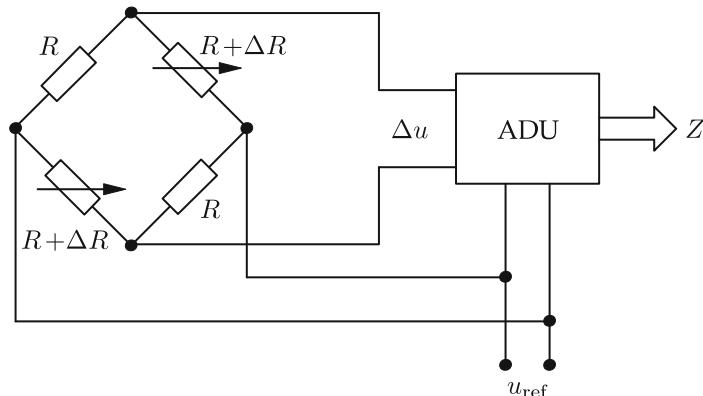


Abbildung 7.42. Ratiometrische Messung.

7.4 Digital-Analog-Umsetzer

Digital-Analog-Umsetzer werden immer dann eingesetzt, wenn digitale Zahlenwerte, die z. B. in einem Rechner abgelegt sind, in eine analoge Spannung umgesetzt werden müssen. Die Ausgangsspannung ist proportional zum Produkt aus digitalem Eingangssignal und einer Referenzspannung u_{ref} :

$$u = u_{\text{ref}} \sum_{k=0}^{N-1} a_k 2^{k-N}. \quad (7.188)$$

Dabei ist $a_k \in \{0, 1\}$ das k -te Bit des digitalen Signals. D/A-Umsetzer werden z. B. in der Automatisierungstechnik eingesetzt, wo beispielsweise eine zuvor im Rechner ermittelte Regelgröße auf ein System geschaltet werden muss. Wie im vorherigen Abschnitt bereits besprochen, benötigen viele in der digitalen Messwerterfassung eingesetzte Analog-Digital-Umsetzer prinzipiell ebenfalls einen Digital-Analog-Umsetzer.

7.4.1 Parallele D/A-Umsetzer

Das klassische Verfahren der Digital-Analog-Umsetzung besteht darin, mit Hilfe eines Widerstandsnetzwerkes (oder Kapazitätsnetzwerkes) entsprechend dem digitalen Eingangswort dyadisch gestufte Ströme i_k zu erzeugen und von einer Summationsschaltung zu addieren. Hierbei spricht man von einer **parallelen D/A-Umsetzung**.

7.4.1.1 D/A-Umsetzer mit dyadisch gestuften Widerständen

Abbildung 7.43 zeigt einen einfachen D/A-Umsetzer mit dyadisch gestuften Widerständen. Der kleinste Widerstand $2R$ besitzt den größten Leitwert und ist somit für die Darstellung des höchstwertigen Bits (MSB) zuständig. Die Ströme

$$i_k = a_k \cdot \frac{u_{\text{ref}}}{2^{N-k} R} \quad (7.189)$$

ergeben summiert und nach der i/u -Wandlung am Operationsverstärker die gewünschte Ausgangsspannung

$$u = -R i = -R \sum_{k=0}^{N-1} i_k = -u_{\text{ref}} \sum_{k=0}^{N-1} \frac{a_k}{2^{N-k}}. \quad (7.190)$$

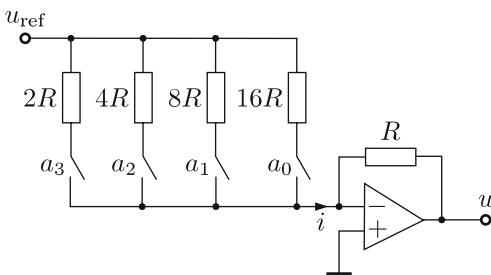
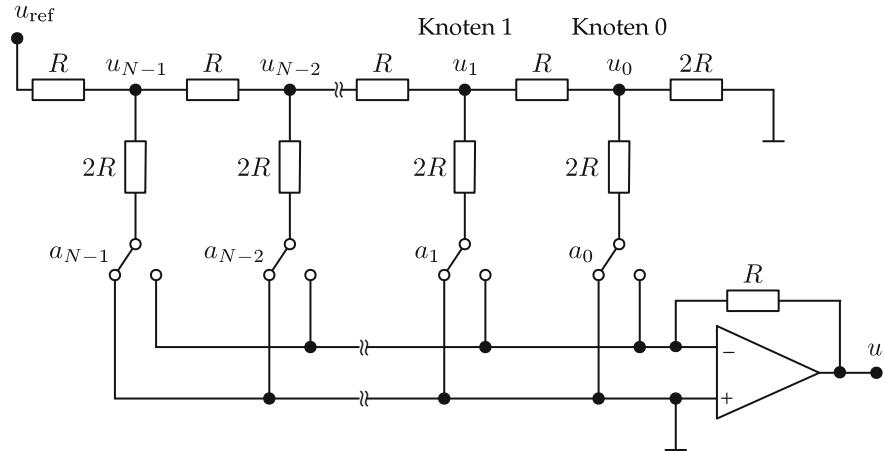


Abbildung 7.43. D/A-Umsetzer mit dyadisch gestuften Widerständen.

7.4.1.2 D/A-Umsetzer mit $R/2R$ -Kettenleiternetzwerk

In der Praxis lassen sich N dyadisch gestufte Widerstände für die gewünschte N -Bit-Auflösung nur mit begrenzter Genauigkeit herstellen, so

Abbildung 7.44. Schaltung eines D/A-Umsetzers mit $R/2R$ -Kettenleiternetzwerk.

dass sich dieses Verfahren nicht für hohe Auflösungen eignet. Abhilfe schafft ein **D/A-Umsetzer mit $R/2R$ -Kettenleiternetzwerk** (Abb. 7.44). Der D/A-Umsetzer kann durch ein Widerstandsnetzwerk oder durch ein Kapazitätsnetzwerk realisiert werden.

Das Verhältnis der Spannungen an den Knoten 0 und 1 ergibt sich mit der Spannungsteilerregel zu:

$$\frac{u_0}{u_1} = \frac{2R \| 2R}{R + 2R \| 2R} = \frac{1}{2} \quad \Rightarrow \quad u_1 = 2u_0 . \quad (7.191)$$

Der Widerstand des Restnetzwerkes vom Knoten 1 nach rechts ist

$$R_{\text{rest}} = R + 2R \| 2R = 2R , \quad (7.192)$$

d. h. im Knoten 1 findet man die gleichen Widerstandsverhältnisse wie im Knoten 0. Man kann deshalb verallgemeinert feststellen, dass die Spannung am nächsthöheren Knoten jeweils doppelt so hoch ist:

$$u_{k+1} = 2u_k . \quad (7.193)$$

Bei Speisung mit der Referenzspannung u_{ref} erhält man

$$u_k = \frac{1}{2^{N-k}} u_{\text{ref}} . \quad (7.194)$$

Die Summenspannung wird abhängig von den Stellenwerten a_k der Zahl Z gebildet:

$$-u = \sum_{k=0}^{N-1} a_k u_k = u_{\text{ref}} \sum_{k=0}^{N-1} a_k 2^{k-N} = u_{\text{ref}} \frac{Z}{Z_{\max}}. \quad (7.195)$$

Derartige D/A-Umsetzer werden häufig in schnellen A/D-Umsetzern mit sukzessiver Approximation (Abschn. 7.3.2) eingesetzt.

7.4.2 Serielle D/A-Umsetzer

Eine besonders einfach zu integrierende Anordnung erhält man, wenn man D/A-Umsetzer durch sogenannte **Rate-Multiplier** [46] realisiert. Rate-Multiplier sind Schaltungen, die ähnlich wie ein Frequenzteiler eine Verringerung der Frequenz¹ der Eingangsimpulsfolge bewirken. Allerdings sind Rate-Multiplier bei der Wahl des Teilungsverhältnisses nicht auf Zweierpotenzen beschränkt, sondern erlauben ebenfalls rationale Teilungsverhältnisse. Praktisch erfolgt die Frequenzteilung dadurch, dass aus der Eingangsimpulsfolge einzelne Impulse entfernt werden, wodurch die Impulse in der resultierenden Folge i. Allg. nicht äquidistant sind.

Abbildung 7.45 zeigt eine Schaltung, die das eingehende Taktsignal f_0 in eine Reihe nicht-koinzidenter Impulsfolgen unterschiedlicher Frequenzen zerlegt. Dabei läuft der Zähler $Q(n)$, der die gleiche Wortlänge N wie der umzusetzende Zahlenwert Z hat, zyklisch mit der Frequenz f_0 durch. Die Zählfreigabe C_0 liegt während eines Taktintervalls fest auf 0 oder 1; das Taktsignal f_0 hat demgegenüber in dieser Zeit zwei Flanken. Bei der D/A-Umsetzung liegt die Zählfreigabe C_0 fest auf 1. Zur Vereinfachung der Schreibweise wird im Folgenden von $C_0 = 1$ ausgegangen. Durch die Decodierung werden aus den Binärstellen Q_k des Zählers nicht-koinzidente Impulsfolgen $q_k(n)$ mit den dyadisch gestuften Frequenzen (Impulshäufigkeit pro Zeit)

$$f_{q_k} = f_0 2^{-(k+1)} \quad (7.196)$$

erzeugt, die sich nicht gegenseitig überlappen und somit zu einer Gesamtimpulsfolge

$$q(n) = \sum_{k=0}^{N-1} q_k(n) \quad (7.197)$$

¹Der Begriff „Frequenz“ wird hier verallgemeinert verwendet, und zwar als Anzahl der Impulse pro Umsetzungsintervall.

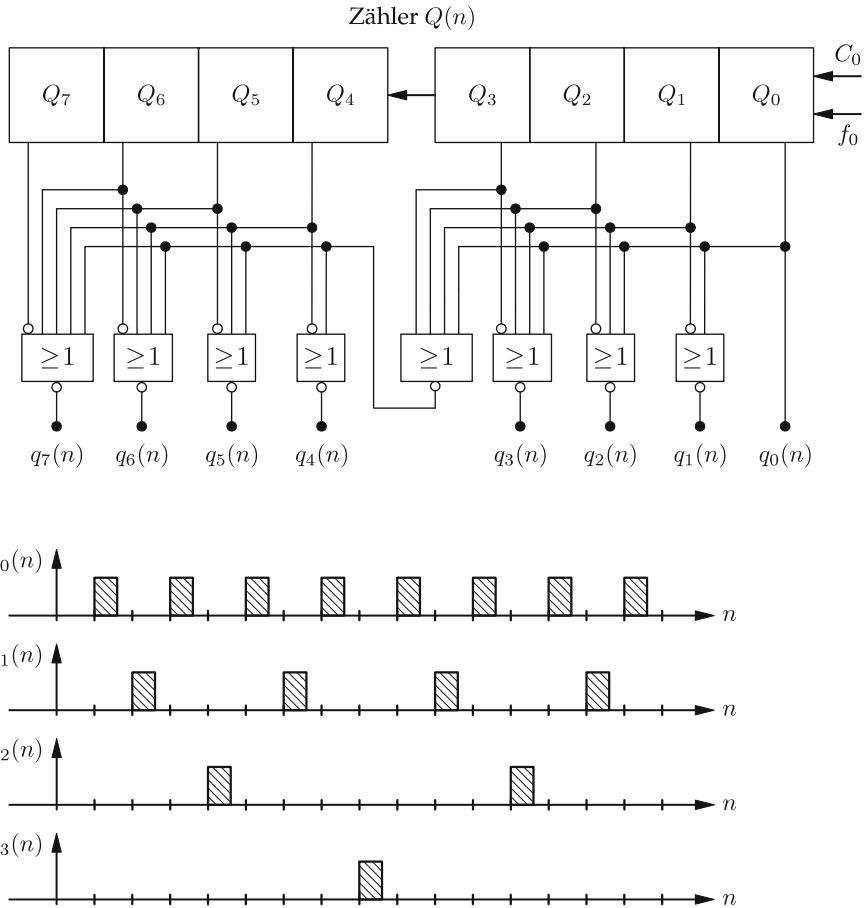


Abbildung 7.45. Erzeugung nicht-koinzidenter Impulsfolgen zur seriellen D/A-Umsetzung.

der Frequenz

$$f = \sum_{k=0}^{N-1} f_{q_k} = f_0 \sum_{k=0}^{N-1} 2^{-(k+1)} \quad (7.198)$$

summiert werden können. Bei Multiplikation der Impulsfolgen $q_k(n)$ mit den jeweiligen Bits a_{N-k-1} des umzusetzenden Zahlenwertes Z gemäß Abb. 7.46 ergibt sich eine Gesamtempulsfolge

$$q(n) = \sum_{k=0}^{N-1} a_{N-k-1} q_k(n), \quad (7.199)$$

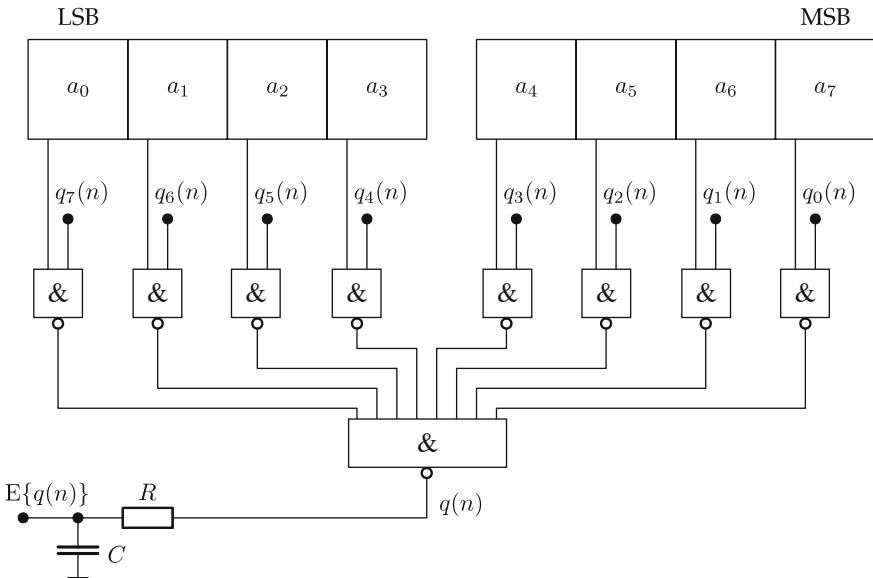


Abbildung 7.46. Digital-Analog-Umsetzung über die mittlere Häufigkeit von Impulsen.

deren Frequenz

$$f = f_0 \sum_{k=0}^{N-1} a_{N-k-1} 2^{-(k+1)} = f_0 \sum_{k=1}^N a_{N-k} 2^{-k} \quad (7.200)$$

proportional zum Zahlenwert

$$Z = Z_{\max} \sum_{k=1}^N a_{N-k} \cdot 2^{-k} \quad (7.201)$$

ist.

Zur D/A-Umsetzung des Zahlenwertes Z ist der Erwartungswert der Gesamtimpulsfolge $q(n)$ zu bilden,

$$u \propto E\{q(n)\}, \quad (7.202)$$

was beispielsweise durch eine Tiefpassfilterung erfolgen kann. In Abb. 7.46 wird dies über einen RC-Tiefpass realisiert.

Wie beim Delta-Sigma-Umsetzer werden auch bei diesem seriellen D/A-Umsetzer die Fehler zu den Grenzen des Messbereiches größer, da dort der Erwartungswert nicht mehr ausreichend genau durch den Mittelwert approximiert werden kann.

D/A-Umsetzer basierend auf dem Rate-Multiplier-Prinzip werden häufig in A/D-Nachlaufumsetzern (Abschn. 7.3.1) eingesetzt. Ihr Vorteil liegt in einer leichten monolithischen Integration der gesamten Schaltung. Aufgrund einer eventuell ungleichförmigen Impulsverteilung kann allerdings der Grenzzyklus des Vorwärts-Rückwärts-Zählers beim Nachlaufumsetzer größer als 1 Bit werden. Abhilfe erreicht man durch zusätzliche Mittelung des Impulszuges $q(n)$ in einem Untersetzergenerator. Der Rate-Multiplier muss dann aber mit einer um den Untersetzungsfaktor höheren Frequenz f_0 betrieben werden.

Kapitel 8

Erfassung frequenzanaloger Signale

8

8 Erfassung frequenzanaloger Signale

| | | |
|---------|--|-----|
| 8.1 | Allgemeiner Frequenzbegriff | 382 |
| 8.2 | Digitale Drehzahlmessung | 389 |
| 8.2.1 | Periodendauermessung | 390 |
| 8.2.2 | Frequenzmessung | 391 |
| 8.2.3 | Maximaler Quantisierungsfehler für einen Zählvorgang | 392 |
| 8.2.3.1 | Quantisierungsfehler bei der Periodendauermessung | 393 |
| 8.2.3.2 | Quantisierungsfehler bei der Frequenzmessung | 393 |
| 8.2.4 | Mittelwertbildung bei der Drehzahlmessung | 395 |
| 8.2.5 | Abtastung bei der Drehzahlmessung | 397 |
| 8.2.6 | Quantisierung bei fortlaufenden Periodendauermessungen | 398 |
| 8.2.7 | Leistungsdichte des Quantisierungsfehlers | 402 |
| 8.2.8 | Kompensation mechanischer Fehler des Sensorrades | 405 |
| 8.2.8.1 | Stochastische Zahnflankenfehler | 405 |
| 8.2.8.2 | Schätzung der Zahnflankenfehler | 407 |
| 8.3 | Kontinuierliche Frequenzmessung | 411 |
| 8.3.1 | Phasenregelkreis | 411 |
| 8.3.2 | Frequenzregelkreis | 416 |
| 8.4 | Positions- und Richtungserkennung | 418 |
| 8.4.1 | Drehrichtungserkennung | 418 |
| 8.4.2 | Positionsbestimmung | 420 |

8 Erfassung frequenzanaloger Signale

In der bisherigen Betrachtung diente meist die Amplitude eines Signals $u(t)$ als Informationsträger. Man spricht daher auch von **amplitudenanalogen Signalen** (Abschn. 1.3.4). Zur digitalen Signalverarbeitung wurde daher auch der Amplitudenverlauf des Signals abgetastet und in diskrete, digitale Werte umgesetzt. In der Messtechnik spielt aber noch eine andere Klasse von Signalen eine Rolle: die der **frequenzanalogen Signale**. Hier dient die Frequenz der Signale als Informationsparameter, den es zu erfassen oder zu verarbeiten gilt. Hierbei stößt man bei verschiedenen Signalen zunächst auf eine unterschiedliche Interpretation des Informationsparameters **Frequenz** [33] (Abb. 8.1).

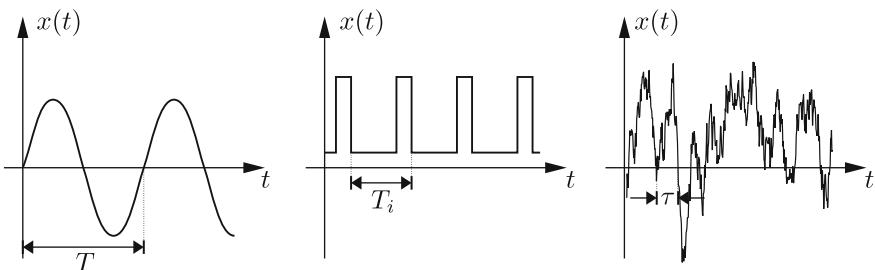


Abbildung 8.1. Interpretationen des Informationsparameters Frequenz: (links) harmonisches Signal; (Mitte) impulsförmiges Signal; (rechts) bandbegrenztes stochastisches Signal.

- **Harmonische Signale:** Für ein harmonisches Signal ist die Kreisfrequenz

$$\omega(t) = \frac{d\varphi(t)}{dt} \quad (8.1)$$

gleich der **Momentanfrequenz**, die als zeitliche Ableitung der Phase definiert ist. Die Momentanfrequenz ist allerdings messtechnisch schwer zugänglich. Bei konstanter Kreisfrequenz $\omega = \text{const.}$ über $\varphi = 2\pi$ gilt

$$\omega = 2\pi f. \quad (8.2)$$

Mit der Periodendauer T kann die Frequenz der harmonischen Schwingung im Zeitbereich ermittelt werden:

$$f = \frac{1}{T} = \frac{\omega}{2\pi}. \quad (8.3)$$

- **Impulsförmige Signale:** Für impulsförmige Signale lässt sich nur die sogenannte Folgefrequenz

$$f_F = \frac{1}{T_i} \quad (8.4)$$

durch die Periodendauer T_i zwischen den Impulsen angeben – eine Momentanfrequenz im Sinne von (8.1) existiert nicht.

- **Stochastische Signale:** Schließlich wird bei bandbegrenzten stochastischen Signalen meist lediglich ein Mittelwert der Frequenz angegeben, der sich aus der Häufigkeit $N_0(T)$ der Nulldurchgänge innerhalb eines Zeitintervalls der Länge T ermitteln lässt. Mit der mittleren Zeitdauer $\bar{\tau}$ zwischen zwei Nulldurchgängen

$$\bar{\tau} = \frac{1}{n} \sum_{i=1}^n \tau_i \quad (8.5)$$

erhält man die zugehörige Frequenz:

$$f = \frac{1}{2\bar{\tau}}. \quad (8.6)$$

Die mittlere Zeitdauer $\bar{\tau}$ entspricht aber dem Beobachtungsintervall T geteilt durch die Anzahl $E\{N_0\}$ der erwarteten Nulldurchgänge in diesem Intervall. Damit erhält man

$$\bar{\tau} = \frac{T}{E\{N_0(T)\}} \quad \Rightarrow \quad f = \frac{1}{2\bar{\tau}} = \frac{E\{N_0(T)\}}{2T}. \quad (8.7)$$

Bislang wurden für verschiedene Signaltypen entsprechende Ad-hoc-Interpretationen der Frequenz angegeben. Allerdings ist diese Klassifikation von Signalen weder systematisch noch vollständig; außerdem sind die Übergänge zwischen den Signaltypen fließend. Daraus wird die Notwendigkeit deutlich, einen allgemeinen, signalunabhängigen Frequenzbegriff einzuführen.

8.1 Allgemeiner Frequenzbegriff

Die Einführung eines allgemeinen Frequenzbegriffs erfolgt in Analogie zur Wahrscheinlichkeitsrechnung. Mittels der Wahrscheinlichkeitsdichte wurden dort der Mittelwert und die Varianz von Zufallsvariablen definiert. Entsprechend wird für eine allgemeine Frequenzdefinition eine „Frequenzverteilungsfunktion“ $f_x(f)$ benötigt, welche die relative Verteilung der Signalanteile über der Frequenz beschreibt. Bei Leistungssignalen ist dies die auf

die Leistung P_x normierte Leistungsdichte $S_{xx}(f)$

$$f_x(f) = \frac{S_{xx}(f)}{P_x} \quad \text{mit} \quad P_x = \int_{-\infty}^{\infty} S_{xx}(f) \, df. \quad (8.8)$$

Bei Energiesignalen $x(t)$ ergibt sich die Frequenzverteilungsfunktion, indem man die spektrale Energiedichte $S_{xx}^E(f)$ auf die Energie E_x bezieht:

$$f_x^E(f) = \frac{S_{xx}^E(f)}{E_x} = \frac{|X(f)|^2}{\|X(f)\|^2}, \quad (8.9)$$

wobei die Energiedichte $S_{xx}^E(f)$ in (8.9) direkt über die Fourier-Transformierte $X(f) = \mathcal{F}\{x(t)\}$ des Signals berechnet werden kann.

Im Folgenden werden die Frequenzverteilungsfunktionen von Leistungs- und Energiesignalen einheitlich als normierte Leistungsdichten $f_x(f)$ bezeichnet. Eine normierte Leistungsdichte hat die gleichen Eigenschaften wie eine Wahrscheinlichkeitsdichte. Insbesondere gilt:

$$f_x(f) \geq 0 \quad \text{und} \quad \int_{-\infty}^{\infty} f_x(f) \, df = 1. \quad (8.10)$$

Analog zur Wahrscheinlichkeitsrechnung können nun durch Erwartungswertbildung Momente berechnet werden. Wichtig sind hierbei das erste Moment und das zweite zentrale Moment.

Definition 8.1: Mittlere Frequenz

8.1

Die **mittlere Frequenz** eines Signals $x(t)$ ist als das erste Moment

$$\bar{f} = \int_{-\infty}^{\infty} f \cdot f_x(f) \, df = \int_{-\infty}^{\infty} f \cdot \frac{S_{xx}(f)}{P_x} \, df \quad (8.11)$$

der normierten Leistungsdichte $f_x(f)$ des Signals definiert. \diamond

Entsprechend der Standardabweichung wird die Wurzel des zweiten zentralen Moments der normierten Leistungsdichte als Bandbreite interpretiert.

Definition 8.2: Bandbreite

8.2

Die **Bandbreite** Δ_f eines Signals $x(t)$ berechnet sich als die Wurzel des zweiten zentralen Moments der normierten Leistungsdichte:

ten zentralen Moments

$$\Delta_f^2 = \int_{-\infty}^{\infty} (f - \bar{f})^2 \cdot f_x(f) df = \int_{-\infty}^{\infty} (f - \bar{f})^2 \cdot \frac{S_{xx}(f)}{P_x} df \quad (8.12)$$

der normierte Leistungsdichte $f_x(f)$ des Signals. \diamond

Insbesondere für reelle Signale ergeben die Definitionen 8.1 und 8.2 keinen Sinn, da ihre Leistungsdichtespektren gerade sind und die mittleren Frequenzen somit zu null werden. In solchen Fällen sollten die mittlere Frequenz \bar{f} und die Bandbreite Δ_f vom zugehörigen analytischen Signal berechnet werden, dessen Spektrum für negative Frequenzen verschwindet [42]. Alternativ lassen sich in (8.11) und (8.12) die unteren Integrationsgrenzen zu null setzen, wenn das Ergebnis der Integration entsprechend um dem Faktor 2 korrigiert wird.

Zu Beginn des Kapitels wurde bereits angedeutet, dass die Frequenz in manchen Anwendungen durch die Häufigkeit der Nulldurchgänge eines Signals bestimmt wird (Abb. 8.1). Zur Abschätzung der Frequenz wird ein mittelwertfreies stochastisches Signal mit normalverteilter Amplitudendichte angenommen. Dies ist deshalb gerechtfertigt, weil die Lage der Nulldurchgänge von $x(t)$ durch überlagerte Störungen verschoben wird.

8.1

Satz 8.1: Effektivfrequenz

Die mittlere Häufigkeit $E\{N_0(T)\}$ der Nulldurchgänge eines mittelwertfreien Signals $x(t)$ innerhalb eines Intervalls T ist bei Annahme normalverteilter Amplituden gleich der **Effektivfrequenz**

$$f_{\text{eff}} = \frac{E\{N_0(T)\}}{2T} = \sqrt{\int_{-\infty}^{\infty} f^2 \cdot f_x(f) df} = \sqrt{\int_{-\infty}^{\infty} f^2 \cdot \frac{S_{xx}(f)}{P_x} df} \quad (8.13)$$

$$= \sqrt{\bar{f}^2 + \Delta_f^2}, \quad (8.14)$$

d.h. der Wurzel des zweiten Momentes der normierten Leistungsdichte $f_x(f)$ (Abb. 8.2). \diamond

Beweis 8.1 (Effektivfrequenz): Für den Beweis von Satz 8.1 muss zunächst die Anzahl $N_0(T)$ der Nulldurchgänge des stationären Prozesses $x(t)$ in einem Intervall der Länge T bestimmt werden, also die Anzahl von Ereigniszeitpunkten t_i in diesem Intervall, für die $x(t_i) = 0$ gilt. Man

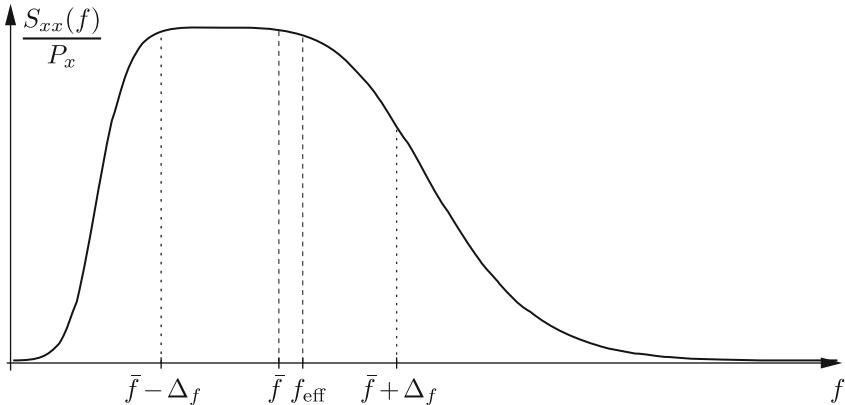


Abbildung 8.2. Kenngrößen der Frequenz im normierten Leistungsdichtespektrum.

spricht daher auch von einem Ereignisprozess. Aufgrund der stochastischen Natur dieses Prozesses muss schließlich der Erwartungswert $E\{N_0(T)\}$ gebildet werden.

Für den Erwartungswert der Ereignisanzahl gilt folgende Gleichung:

$$E\{N_0(T)\} = T \cdot f_x(x=0) \cdot E\{|\dot{x}(t)| \mid x(t)=0\}, \quad (8.15)$$

wobei $f_x(x)$ die Amplitudendichte des Prozesses $x(t)$ bezeichnet. Zum Beweis und zur gleichzeitigen Veranschaulichung von (8.15) verwendet man die folgende Eigenschaft der Dirac'schen Impulsfunktion, die in Abb. 8.3 illustriert wird [36]:

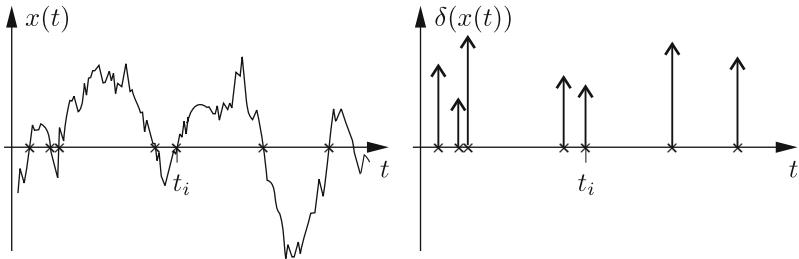
$$\delta(x(t)) = \sum_i \frac{\delta(t - t_i)}{|\dot{x}(t_i)|} = \frac{1}{|\dot{x}(t)|} \sum_i \delta(t - t_i), \quad (8.16)$$

$$t_i = \{t \mid x(t) = 0\}, \quad (8.17)$$

$$\sum_i \delta(t - t_i) = \delta(x(t)) |\dot{x}(t)| = \zeta(t). \quad (8.18)$$

Demnach wird mittels der δ -Funktion aus der Funktion $x(t)$ eine Impulsreihe erzeugt, die nur an den Stellen der Nulldurchgänge von $x(t)$ von null verschieden ist. Wegen der Eigenschaft der Dirac'schen Impulsfunktion

$$\int_{-\infty}^{\infty} \delta(t) dt = 1 \quad (8.19)$$

Abbildung 8.3. Impulsfunktion $\delta(x(t))$ zur Erzeugung von Ereignissen.

erhält man durch Integration des Ereignisprozesses $\zeta(t)$ über T die Anzahl der Ereignisse $N_0(T)$ im Intervall T :

$$N_0(T) = \int_0^T \zeta(t) dt \quad \Rightarrow \quad E\{N_0(T)\} = T \cdot E\{\zeta(t)\}. \quad (8.20)$$

Nach (8.18) ist $\zeta(t)$ eine Funktion der Zufallsgrößen x und \dot{x} . Mit der Verbundwahrscheinlichkeitsdichte $f_{x\dot{x}}(x, \dot{x})$ berechnet sich der Erwartungswert zu

$$E\{\zeta(t)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\dot{x}| \delta(x) f_{x\dot{x}}(x, \dot{x}) dx d\dot{x} \quad (8.21)$$

$$= \int_{-\infty}^{\infty} |\dot{x}| f_{x\dot{x}}(x=0, \dot{x}) d\dot{x}. \quad (8.22)$$

Mit der bedingten Wahrscheinlichkeitsdichte (Def. 4.8)

$$f_{\dot{x}|x}(\dot{x} | x=0) = \frac{f_{x\dot{x}}(x=0, \dot{x})}{f_x(x=0)} \quad (8.23)$$

folgt der Erwartungswert der Nulldurchgänge nach (8.15):

$$E\{\zeta(t)\} = f_x(x=0) \int_{-\infty}^{\infty} |\dot{x}| f_{\dot{x}|x}(\dot{x} | x=0) d\dot{x} \quad (8.24)$$

$$= f_x(x=0) E\{|\dot{x}(t)| \mid x(t)=0\}. \quad (8.25)$$

Damit ist (8.15) bewiesen.

Bezieht man den Erwartungswert noch auf die Intervalldauer T , so erhält man die mittlere Ereignisrate, die in der Markov-Theorie eine

wichtige Rolle spielt [25]:

$$\lambda = \frac{E\{N_0(T)\}}{T} = f_x(x=0) \cdot E\{|\dot{x}(t)| \mid x(t)=0\}. \quad (8.26)$$

Im nächsten Schritt soll gezeigt werden, dass für normalverteilte Zufallsprozesse

$$E\{N_0(T)\} = T \cdot f_x(x=0) \cdot E\{|\dot{x}(t)|\} \quad (8.27)$$

gilt, d. h. dass der bedingte Erwartungswert $E\{|\dot{x}(t)| \mid x(t) = 0\}$ in den Erwartungswert $E\{|\dot{x}(t)|\}$ übergeht. Für den Beweis benötigt man die Korrelationsfunktionen differenzierter Signale. Betrachtet man die Differentiation als lineares System (Abb. 8.4) mit der Übertragungsfunktion

$$G(f) = j2\pi f, \quad (8.28)$$

so folgt nach (6.189) für die Leistungsdichthespektren bzw. Korrelationsfunktionen

$$S_{x\dot{x}}(f) = -j2\pi f S_{xx}(f) \quad \bullet\circ \quad r_{x\dot{x}}(\tau) = -\dot{r}_{xx}(\tau), \quad (8.29)$$

$$S_{\dot{x}\dot{x}}(f) = (-j2\pi f)(j2\pi f) S_{xx}(f) \quad \bullet\circ \quad r_{\dot{x}\dot{x}}(\tau) = -\ddot{r}_{xx}(\tau). \quad (8.30)$$

Da die Autokorrelationsfunktion allgemein eine gerade Funktion mit dem Maximum an der Stelle $\tau = 0$ ist, folgt daraus für die Ableitung

$$\dot{r}_{xx}(\tau=0) = 0 = r_{x\dot{x}}(0) = E\{x(t) \dot{x}(t)\} = 0. \quad (8.31)$$

Das bedeutet, dass die Zufallsgrößen $x(t)$ und $\dot{x}(t)$ unkorreliert sind. Für normalverteilte Prozesse folgt daraus auch die statistische Unabhängigkeit, so dass für die bedingte Wahrscheinlichkeitsdichte

$$f_{\dot{x}|x}(\dot{x} \mid x=0) = f_{\dot{x}}(\dot{x}) \quad (8.32)$$

gilt und (8.15) in den Erwartungswert (8.27) übergeht. Dieser berechnet sich für mittelwertfreie Gaußprozesse zu

$$E\{|\dot{x}(t)|\} = \int_{-\infty}^{\infty} |\dot{x}| \frac{1}{\sigma_{\dot{x}}\sqrt{2\pi}} \exp\left(-\frac{\dot{x}^2}{2\sigma_{\dot{x}}^2}\right) d\dot{x} \quad (8.33)$$

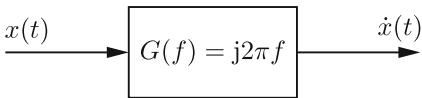


Abbildung 8.4. System zur Differentiation.

$$= \frac{2}{\sigma_{\dot{x}} \sqrt{2\pi}} \int_0^\infty \dot{x} \exp\left(-\frac{\dot{x}^2}{2\sigma_{\dot{x}}^2}\right) d\dot{x} \quad (8.34)$$

$$= \frac{2\sigma_{\dot{x}}}{\sqrt{2\pi}} \quad \text{mit} \quad -\ddot{r}_{xx}(0) = r_{\dot{x}\dot{x}}(0) = \sigma_{\dot{x}}^2 \quad (8.35)$$

$$= \sqrt{\frac{-2\ddot{r}_{xx}(0)}{\pi}}. \quad (8.36)$$

wobei zur Vereinfachung von (8.34) der Zusammenhang

$$\int_0^\infty x e^{-ax^2} dx = \frac{1}{2a} \quad (8.37)$$

verwendet wurde. Setzt man nun noch

$$f_x(x=0) = \frac{1}{\sigma_x \sqrt{2\pi}} = \frac{1}{\sqrt{2\pi r_{xx}(0)}}, \quad (8.38)$$

so folgt schließlich für die mittlere Anzahl von Ereignissen bei normalverteilten Prozessen aus (8.27)

$$E\{N_0(T)\} = T \sqrt{\frac{-\ddot{r}_{xx}(0)}{\pi^2 r_{xx}(0)}} = T \frac{\sigma_{\dot{x}}}{\pi \sigma_x}. \quad (8.39)$$

Damit erhält man schließlich die Effektivfrequenz

$$f_{\text{eff}} = \frac{E\{N_0(T)\}}{2T} = \sqrt{\frac{-\ddot{r}_{xx}(0)}{4\pi^2 r_{xx}(0)}} \quad \circ \bullet \quad \sqrt{\frac{\int_{-\infty}^{\infty} f^2 S_{xx}(f) df}{\int_{-\infty}^{\infty} S_{xx}(f) df}}, \quad (8.40)$$

womit (8.13) bewiesen ist. ♦

Der Name Effektivfrequenz röhrt von seinem Analogon Effektivwert bei amplitudenanalogen Signalen her:

$$u_{\text{eff}} = \sqrt{\mu_u^2 + \sigma_u^2} \approx \sqrt{\frac{1}{2T} \int_{-T}^T u(t)^2 dt}. \quad (8.41)$$

Das in Abschn. 8.2.2 behandelte Frequenzmessverfahren ermittelt gerade die Effektivfrequenz f_{eff} durch das Zählen von Nulldurchgängen innerhalb eines Referenzintervalls. Die Effektivfrequenz besitzt jedoch einen systematischen Fehler, der von der Bandbreite Δ_f des Signals abhängt (Abb. 8.2).

Für kleine Fehler (d. h. für eine kleine relative Bandbreite $\Delta f/\bar{f}$) folgt mit der Näherung $\sqrt{1+x} \approx 1 + \frac{x}{2}$:

$$f_{\text{eff}} = \sqrt{\bar{f}^2 + \Delta_f^2} \approx \bar{f} \left(1 + \frac{1}{2} \frac{\Delta_f^2}{\bar{f}^2} \right) = \bar{f} + \frac{1}{2} \frac{\Delta_f^2}{\bar{f}} > \bar{f}. \quad (8.42)$$

8.2 Digitale Drehzahlmessung

8.2

Die digitale Erfassung frequenzanaloger Signale soll am Beispiel der **Drehzahlmessung** vorgestellt werden. Hierfür gibt es prinzipiell mehrere Möglichkeiten. So kann die Winkelgeschwindigkeit einer Welle z. B. mit einem Tachogenerator in ein *amplitudenanaloges* Signal umgeformt werden. Weniger störanfällig und leichter zu realisieren sind demgegenüber Verfahren, die das Drehzahlsignal mittels Inkrementalgeber in ein periodisches Signal umformen, so dass die Information als *frequenzanaloges* Signal codiert ist. Die eigentliche Erfassung der Drehzahl beruht dann auf einer Zeitintervallmessung, welche durch die Verwendung von Quarzoszillatoren um ein Vielfaches genauer realisiert werden kann als dies bei einer Spannungs auswertung mit vertretbarem Aufwand der Fall wäre [33]. Die dafür notwendigen Zähler können ohne großen Aufwand realisiert werden und sind in der Peripherie von handelsüblichen Mikrocontrollersystemen mehrfach vorhanden.

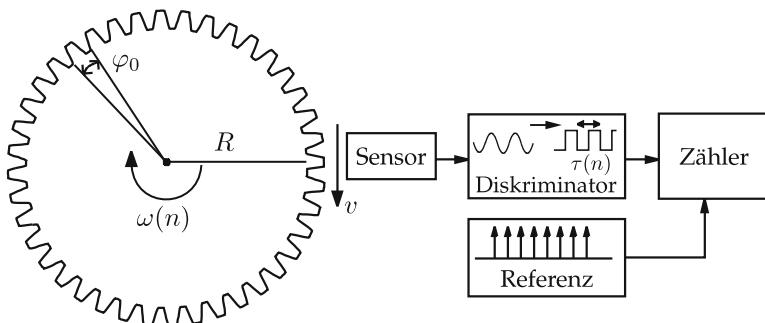


Abbildung 8.5. Messsystem zur frequenzanalogen Drehzahlmessung.

Abbildung 8.5 zeigt das Prinzip der Drehzahlmessung. Ein auf der Welle mitrotierendes Zahnrad bildet den Modulator zur Umformung der amplitudenanalogen Eingangsgröße Winkelgeschwindigkeit ω in ein frequenzanaloges Signal. Die Zahnhöhe wird z. B. über einen induktiven Sensor oder einen Hall-Sensor erfasst und in ein periodisches Signal umgewandelt. Das

dabei pro Zahn überstrichene Winkelinkrement beträgt

$$\varphi_0 = \frac{2\pi}{Z}, \quad (8.43)$$

wobei Z die Anzahl aller Zähne auf dem Zahnrad bezeichnet. Das entstehende periodische Signal wird über einen Schmitt-Trigger in ein Rechtecksignal gleicher Periodendauer umgewandelt, wobei die Periodendauer anschließend in einem Mikrocontroller mit Hilfe eines Referenzzählers ausgezählt, d. h. digitalisiert, wird. Die gemessene Winkelgeschwindigkeit ω_m berechnet sich mit Hilfe des Differenzenquotienten aus dem Winkelintervall $\varphi_2 - \varphi_1$ und der dafür benötigten Durchlaufzeit $t_2 - t_1$ zu

$$\omega_m = 2\pi n = \frac{\varphi_2 - \varphi_1}{t_2 - t_1}, \quad (8.44)$$

wobei n die Drehzahl bezeichnet. Für die Auswertung von (8.44) gibt es die beiden grundsätzlichen Möglichkeiten,

- das frequenzanaloge Signal winkelsynchron zu erfassen und die Periodendauer zu messen (Abschn. 8.2.1) oder
- das Signal zeitsynchron zu erfassen und eine Frequenzmessung vorzunehmen (Abschn. 8.2.2).

Abhängig vom Betrag der Messgröße weisen beide Verfahren eine unterschiedliche Genauigkeit auf (Abschn. 8.2.3).

8.2.1 Periodendauermessung

Abbildung 8.6 zeigt das Strukturbild der **Periodendauermessung**. Bei diesem Verfahren wird die kontinuierliche Zeit $t_2(i) - t_1(i) = T_m(i)$ für einen diskreten Winkelschritt

$$\varphi_0 = \varphi_2 - \varphi_1 = \frac{2\pi}{Z}, \quad (8.45)$$

der durch die Winkelteilung des Zahnrades fest vorgegeben ist, gemessen:

$$\omega_m(i) = \frac{2\pi}{T_m(i) \cdot Z}, \quad (8.46)$$

wobei i den Zahnindex bezeichnet. Die Periodendauer $T_m(i)$, die für das Überstreichen des Winkelinkrements φ_0 benötigt wird, wird durch das Auszählen von Impulsen einer Referenzfrequenz f_0 ermittelt (Abb. 8.7). Die Abtastung erfolgt hierbei **winkelsynchron**, da mit jeder neuen Zahnflanke φ_0 ein digitaler Messwert in den Rechner übernommen wird. Aufgrund der Diskretisierung des Referenzzeitintervalls $1/f_0$ weist die gemessene Periodendauer einen Quantisierungsfehler auf.

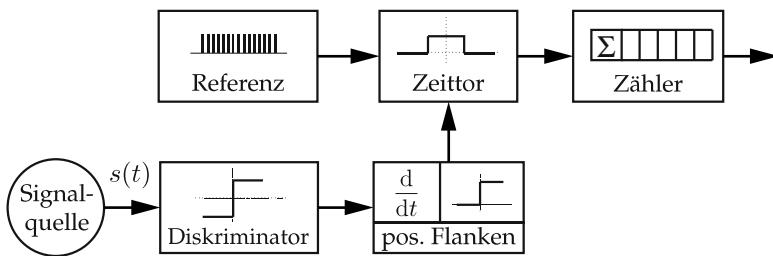


Abbildung 8.6. Strukturbild der Periodendauermessung.

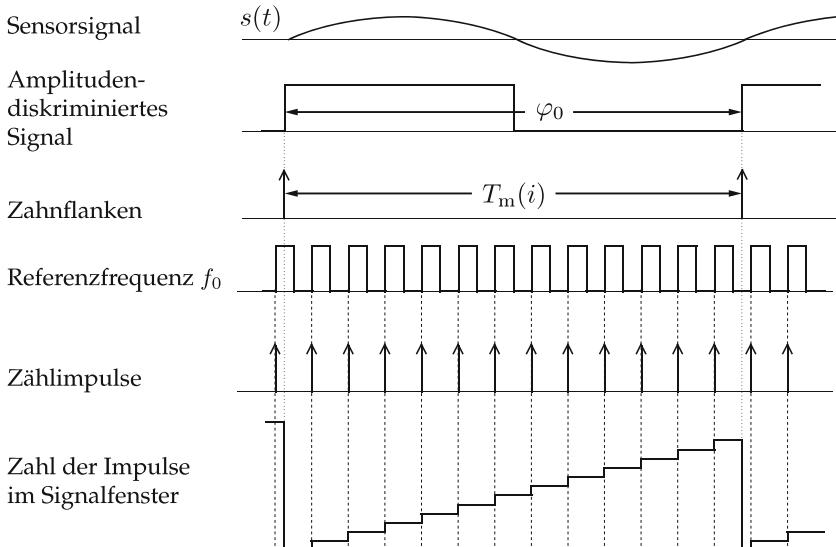


Abbildung 8.7. Digitale Drehzahlmessung durch Periodendauermessung (winkelsynchrone Abtastung).

8.2.2 Frequenzmessung

Abbildung 8.8 zeigt das Strukturbild der **Frequenzmessung**. Das Frequenzmessverfahren gibt dabei das Zeitintervall $T_{\text{ref}} = t_2 - t_1$ fest vor und misst den überstrichenen Winkel $\varphi_2(i) - \varphi_1(i) = \varphi_m(i)$, um daraus die Winkelgeschwindigkeit zu berechnen:

$$\omega_m(i) = \frac{\varphi_m(i)}{T_{\text{ref}}} . \quad (8.47)$$

Der überstrichene kontinuierliche Winkel $\varphi_m(i)$ wird durch das Auszählen der Nulldurchgänge des Sensorsignals $s(\varphi)$ innerhalb der diskreten Messperiode T_{ref} ermittelt (Abb. 8.9). Die Abtastung erfolgt bei der Frequenzmessung **zeitsynchron**, da nach jedem Zeitintervall T_{ref} ein Messwert er-

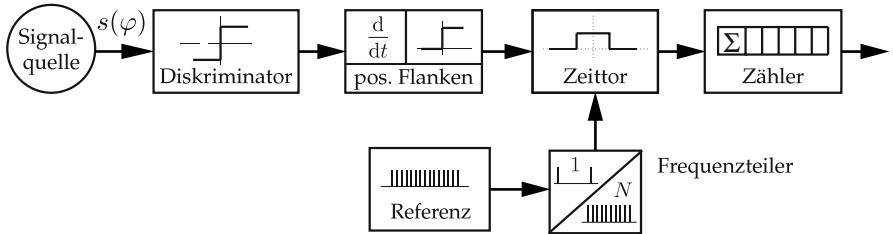


Abbildung 8.8. Strukturbild der Frequenzmessung.

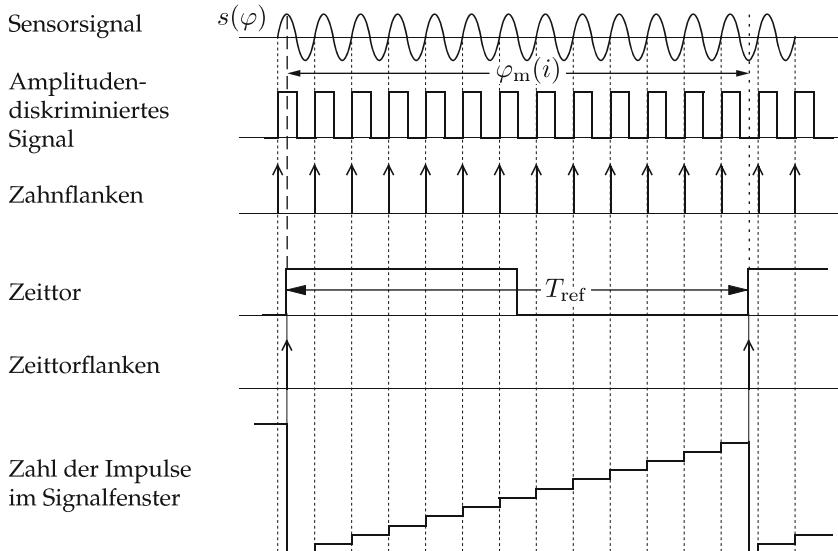


Abbildung 8.9. Digitale Drehzahlmessung durch Frequenzmessung (zeitsynchrone Abtastung).

fasst wird. Durch die feste Zahnteilung weist der ermittelte Winkel einen Quantisierungsfehler auf.

8.2.3 Maximaler Quantisierungsfehler für einen Zählvorgang

Bei den beiden vorgestellten Verfahren wird die Drehzahl ω entweder über das Zählen der Impulse eines Referenztaktes $f_0 = 1/T$ zwischen zwei Flanken des Sensorsignals (winkelsynchrone Abtastung) oder über die Anzahl der in einer Referenzperiode T_{ref} einlaufenden Flankenimpulse des Sensorsignals (zeitsynchrone Abtastung) bestimmt. In beiden Fällen erfolgt eine Quantisierung mit entsprechenden Quantisierungsfehlern. Nun soll der relative Quantisierungsfehler F_r in einer *einzig* Periode T_m bzw. in einem *einzig* Referenzintervall T_{ref} abgeschätzt werden. Aufgrund der statischen Fehlerbetrachtung können die Zeitindizes im Folgenden weggelassen wer-

den. Die wahre stationäre Winkelgeschwindigkeit ist

$$\omega = \frac{\varphi}{T}. \quad (8.48)$$

▷ 8.2.3.1 Quantisierungsfehler bei der Periodendauermessung

Im ungünstigsten Fall kann hier die quantisierte Periodendauer $N \cdot T_0$ durch die endliche Auflösung des Zählers um ein Quantisierungsintervall des Referenztaktes $T_0 = 1/f_0$ von der kontinuierlichen Periodendauer T_m zwischen zwei Flanken des Sensorsignals abweichen:

$$|T_m - NT_0| \leq T_0. \quad (8.49)$$

Mit der quantisierten Winkelgeschwindigkeit $\omega_q = \frac{\varphi_0}{NT_0}$ ergibt sich der relative Quantisierungsfehler bei der Periodendauermessung betragsmäßig zu:

$$F_r = \frac{|\omega_q - \omega_m|}{\omega_m} = \left| \frac{\frac{\varphi_0}{NT_0} - \frac{\varphi_0}{T_m}}{\frac{\varphi_0}{T_m}} \right| = \left| \frac{T_m}{NT_0} - 1 \right|. \quad (8.50)$$

Bei einer maximalen Abweichung von einem Zählintervall T_0 erhält man mit (8.49) den **relativen Quantisierungsfehler**

$$F_r \leq \frac{1}{N} = \frac{1}{NT_0 f_0} = \frac{\omega_q}{\varphi_0 f_0}. \quad (8.51)$$

Bei der Periodendauermessung ist somit der maximale relative Quantisierungsfehler proportional zur Drehzahl.

▷ 8.2.3.2 Quantisierungsfehler bei der Frequenzmessung

Beim Frequenzmessverfahren erhält man in der Referenzperiode T_{ref} gerade N Sensorflankenimpulse: $T_{ref} = N \cdot \varphi_0 / \omega_q$. Jeder Flankenimpuls entspricht dem Durchlauf des Sensorrades durch das Winkelinkrement $\varphi_0 = 2\pi/Z$. Der in der Messzeit T_{ref} tatsächlich überstrichene Winkel φ_m kann bis zu einem Winkelinkrement φ_0 vom gemessenen Wert $N \cdot \varphi_0$ abweichen:

$$|\varphi_m - N \cdot \varphi_0| \leq \varphi_0. \quad (8.52)$$

Mit der quantisierten Winkelgeschwindigkeit $\omega_q = \varphi_0 \frac{N}{T_{ref}}$ folgt für den Betrag des relativen Quantisierungsfehlers:

$$F_r = \frac{|\omega_q - \omega_m|}{\omega_m} = \left| \frac{\varphi_0 \frac{N}{T_{ref}} - \frac{\varphi_m}{T_{ref}}}{\frac{\varphi_m}{T_{ref}}} \right| = \left| \frac{\varphi_0 \cdot N - \varphi_m}{\varphi_m} \right| \quad (8.53)$$

$$\approx \left| \frac{\varphi_0 \cdot N - \varphi_m}{\varphi_0 \cdot N} \right|. \quad (8.54)$$

Bei einer maximalen Abweichung von einem Winkelinkrement φ_0 erhält man mit (8.52) den **relativen Quantisierungsfehler** zu

$$F_r \leq \frac{1}{N} = \frac{\varphi_0}{\omega_q T_{\text{ref}}} . \quad (8.55)$$

Bei der Frequenzmessung ist der maximale relative Quantisierungsfehler also umgekehrt proportional zur Drehzahl.

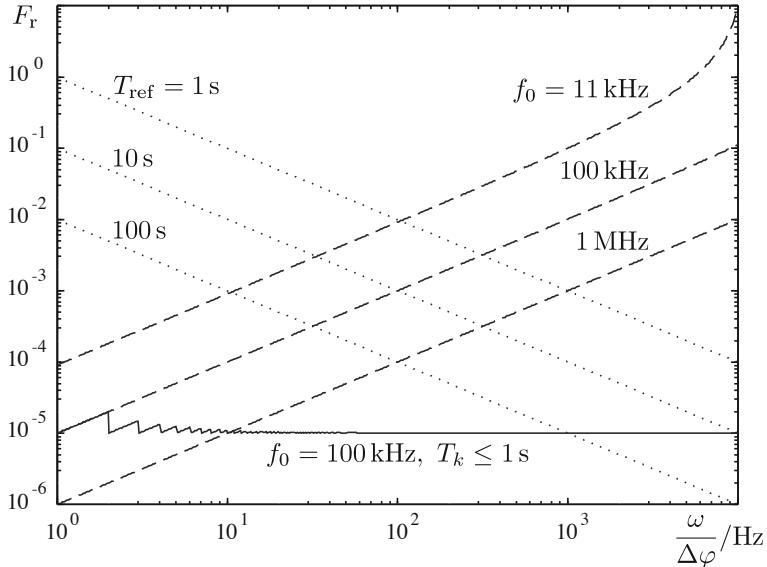


Abbildung 8.10. Maximaler relativer statischer Quantisierungsfehler (\cdots Frequenzmessung, $-$ Periodendauermessung, — Mischverfahren: Messung über mehrere Perioden).

Abbildung 8.10 zeigt den maximalen relativen statischen Quantisierungsfehler für verschiedene Referenzperiodendauern und Referenzzählfrequenzen in doppeltlogarithmischem Maßstab. Der Knick der oberen Kurve zur Periodendauermessung ist nach (8.51) bedingt durch eine Division durch den Zählerstand $N = 0$. Darüber hinaus zeigt die Abbildung auch den Quantisierungsfehler für das quasizeitdiskrete Mischverfahren über eine ganzzahlige Anzahl $k = (T_k \omega_q)/\varphi_0$ von Perioden bei einer maximalen Messzeit von $T_k \leq 1$ s. Dabei wird die Periode T_k für den Durchlauf des Sensorrades durch den Winkelschritt $k \cdot \varphi_0$ gemessen. Es ergibt sich der gleiche Fehler wie bei der Periodendauermessung, wenn man dort den festen, einfachen Winkelschritt φ_0 durch den k -fachen Winkelschritt $k \cdot \varphi_0$ ersetzt:

$$F_r = \frac{|\omega_q - \omega_m|}{\omega_m} \leq \frac{\omega_q}{k \varphi_0 f_0} = \frac{\omega_q}{\frac{T_k \omega_q}{\varphi_0} \varphi_0 f_0} = \frac{1}{T_k f_0} . \quad (8.56)$$

8.2.4 Mittelwertbildung bei der Drehzahlmessung

Alle Verfahren zur digitalen Drehzahlmessung besitzen einen mittelwertbildenden Charakter. Ändert sich die Drehzahl ω innerhalb der Messzeit, so wird sie mit dieser Mittelungsfunktion verfälscht. Entsprechend der Approximation der Winkelgeschwindigkeit durch den Differenzenquotienten (8.44) wird dem jeweiligen Abtastwert eine mittlere Winkelgeschwindigkeit zugewiesen.

Bei der Periodendauermessung erfolgt die Mittelung zwischen zwei Zahnflanken im Abstand φ_0 . Daraus folgt für die messbare Winkelgeschwindigkeit

$$\hat{\omega}(\varphi) = \frac{1}{\varphi_0} \int_{\varphi - \frac{\varphi_0}{2}}^{\varphi + \frac{\varphi_0}{2}} \omega(\varphi) d\varphi. \quad (8.57)$$

Die Frequenzmessung bewirkt eine Mittelung über das Referenzintervall T_{ref} :

$$\hat{\omega}(t) = \frac{1}{T_{\text{ref}}} \int_{t - \frac{T_{\text{ref}}}{2}}^{t + \frac{T_{\text{ref}}}{2}} \omega(t) dt. \quad (8.58)$$

Da die Periodendauermessung winkelsynchron erfolgt, eignet sich zur Untersuchung des Fehlers infolge der Mittelung der Winkelfrequenzbereich. Das zeitsynchrone Frequenzmessverfahren lässt sich analog im Zeitfrequenzbereich beschreiben.

Nach Abschn. 7.1.4 kann die Mittelwertbildung als Faltung der Eingangsgröße mit einem Rechteckfenster $r_{\varphi_0}(\varphi)$ bzw. $r_{T_{\text{ref}}}(t)$ interpretiert werden. Damit kann man für die Periodendauermessung

$$\hat{\omega}(\varphi) = \omega(\varphi) * r_{\varphi_0}(\varphi) \quad \circ \bullet \quad \hat{\Omega}(f_\varphi) = \Omega(f_\varphi) \cdot R_{\varphi_0}(f_\varphi) \quad (8.59)$$

und für die Frequenzmessung

$$\hat{\omega}(t) = \omega(t) * r_{T_{\text{ref}}}(t) \quad \circ \bullet \quad \hat{\Omega}(f) = \Omega(f) \cdot R_{T_{\text{ref}}}(f) \quad (8.60)$$

schreiben. Dabei bezeichnet f_φ die Winkelfrequenz. Abbildung 8.11 zeigt die Fourier-Transformierte eines Rechteckfensters im Winkelbereich. Infolge der Mittelwertbildung wird das Originalspektrum durch Multiplikation mit einer Sinc-Funktion [42] verfälscht. Schnell veränderliche Drehzahländerungen werden dabei unterdrückt; die mittlere Drehzahl bleibt jedoch erhalten.

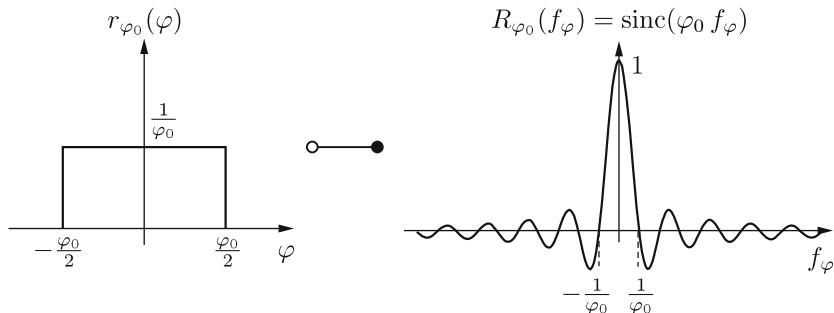


Abbildung 8.11. Rechteckfenster der Breite φ_0 im Winkelbereich und Fourier-Transformierte davon im Winkelfrequenzbereich.

Beispiel 8.1 (Drehzahlmessung an einem Verbrennungsmotor): Die Drehzahl n eines Verbrennungsmotors mit $N_{\text{Zyl}} = 6$ Zylindern wird durch Periodendauermessung zwischen zwei Zahnflanken bestimmt. Der Drehzahl sind durch die Verbrennungsvorgänge in den einzelnen Zylindern Störungen mit einer Winkelfrequenz von

$$f_{\varphi, \text{stör}} \approx \frac{N_{\text{Zyl}}}{2 \text{ rad}} = 3 \frac{1}{\text{rad}} \quad (8.61)$$

überlagert (3 Störhübe pro Umdrehung). Die Abtastfrequenz des Sensors ist bei einem Zahnrad mit $Z = 60$ Zähnen

$$f_{\varphi, A} = 60 \frac{1}{\text{rad}}, \quad (8.62)$$

also 60 Zahnflankenimpulse pro Umdrehung. Wegen

$$f_{\varphi, \text{stör}} \leq \frac{f_{\varphi, A}}{2} \quad (8.63)$$

werden die Verbrennungsmomente in den einzelnen Zylindern durch die inhärente Mittelwertbildung der Motordrehzahl bei der Periodendauermessung kaum unterdrückt, sondern nahezu ungedämpft durchgelassen. Aus der gemessenen Motordrehzahl kann deshalb das Verbrennungsmoment geschätzt werden [26]. ■

Die Mittelwertbildung soll nun speziell für das Frequenzmessverfahren untersucht werden. Gegeben sei eine konstante Drehzahl $\bar{\omega}$, die von einer harmonischen Störung der Frequenz f_1 überlagert ist:

$$\omega(t) = \bar{\omega} + \Delta\omega e^{j2\pi f_1 t}. \quad (8.64)$$

Die Mittelwertbildung nach (8.58) ergibt

$$\hat{\omega}(t) = \frac{1}{T_{\text{ref}}} \int_{t - \frac{T_{\text{ref}}}{2}}^{t + \frac{T_{\text{ref}}}{2}} (\bar{\omega} + \Delta\omega e^{j2\pi f_1 t}) dt \quad (8.65)$$

$$= \bar{\omega} + \frac{\Delta\omega}{j2\pi f_1 T_{\text{ref}}} \cdot \left[e^{j2\pi f_1 (t + T_{\text{ref}}/2)} - e^{j2\pi f_1 (t - T_{\text{ref}}/2)} \right] \quad (8.66)$$

$$= \bar{\omega} + \Delta\omega e^{j2\pi f_1 t} \cdot \frac{\sin(\pi f_1 T_{\text{ref}})}{\pi f_1 T_{\text{ref}}}. \quad (8.67)$$

Die überlagerte Störung wird durch die Sinc-Funktion gefiltert:

$$\hat{\omega}(t) = \bar{\omega} + \Delta\bar{\omega} e^{j2\pi f_1 t} \quad \text{mit} \quad \Delta\bar{\omega} = \Delta\omega \cdot \frac{\sin(\pi f_1 T_{\text{ref}})}{\pi f_1 T_{\text{ref}}}. \quad (8.68)$$

8.2.5 Abtastung bei der Drehzahlmessung

Nach der Mittelwertbildung wird der Wert $\hat{\omega}$ abgetastet. Bei der Frequenzmessung erfolgt die Abtastung zeitäquidistant mit der Referenzperiode T_{ref} :

$$\hat{\omega}(n T_{\text{ref}}) = \hat{\omega}(t) \cdot \sum_{n=-\infty}^{\infty} \delta(t - n T_{\text{ref}}). \quad (8.69)$$

Im Frequenzbereich resultiert daraus nach Abschn. 7.1 eine Wiederholung des Spektrums mit der Periode $1/T_{\text{ref}}$:

$$\hat{\Omega}_*(f) = \frac{1}{T_{\text{ref}}} \sum_{k=-\infty}^{\infty} \hat{\Omega}(f - k/T_{\text{ref}}). \quad (8.70)$$

Es soll erneut der Fall (8.64) einer konstanten Drehzahl $\bar{\omega}$ betrachtet werden, die von einer harmonischen Störung der Frequenz f_1 überlagert ist:

$$\hat{\omega}(t) = \bar{\omega} + \Delta\bar{\omega} e^{j2\pi f_1 t} \quad (8.71)$$

$$\hat{\Omega}(f) = \overset{\bullet}{\underset{\circ}{\hat{\Omega}}}(f) = \bar{\omega} \delta(f) + \Delta\bar{\omega} \delta(f - f_1). \quad (8.72)$$

Nach dem Abtasttheorem (Satz 7.1) lassen sich zwei Fälle unterscheiden:

1. $f_1 \leq \frac{1}{2T_{\text{ref}}}$: es liegt keine spektrale Überlappung (*Aliasing*) vor, da die Störung im Nyquist-Band liegt.
2. $f_1 > \frac{1}{2T_{\text{ref}}}$: es kommt zu einer spektralen Überlappung, d.h. zu einer zusätzlichen Störung infolge von *Aliasing*.

Bei der Periodendauermessung erfolgt die Abtastung winkeldiskret:

$$\hat{\omega}(n\varphi_0) = \hat{\omega}(\varphi) \cdot \sum_{n=-\infty}^{\infty} \delta(\varphi - n\varphi_0). \quad (8.73)$$

Nur für den Spezialfall $\omega(t) = \text{const.}$ ist auch die Abtastzeit konstant. Im Spektrum ergibt sich eine in $1/\varphi_0$ periodisch wiederholte Funktion:

$$\hat{\Omega}_*(f_\varphi) = \frac{1}{\varphi_0} \sum_{k=-\infty}^{\infty} \hat{\Omega}(f_\varphi - k/\varphi_0). \quad (8.74)$$

Das Abtasttheorem gilt hier analog. Allerdings benötigt man hier den Frequenzbegriff „pro Umdrehung“ mit der Einheit $1/\text{rad}$, der vom physikalischen Frequenzbegriff abweicht. Für das Beispiel 8.1 der Drehzahlmesung beim Verbrennungsmotor tritt kein Aliasing auf, da die Störfrequenz kleiner als die halbe Abtastfrequenz ist:

$$f_{\varphi,\text{stör}} \leq \frac{f_{\varphi,A}}{2}. \quad (8.75)$$

8.2.6 Quantisierung bei fortlaufenden Periodendauermessungen

Der durch die Zahnräder des Sensorrades gegebene Winkel φ_0 wird bei einer Winkelgeschwindigkeit ω_m im i -ten Zählintervall in der Zeit $T_m(i)$ durchlaufen. Der Quantisierungsfehler der Periodendauermessung soll fortlaufend über viele Intervalle i betrachtet und statistisch beschrieben werden.

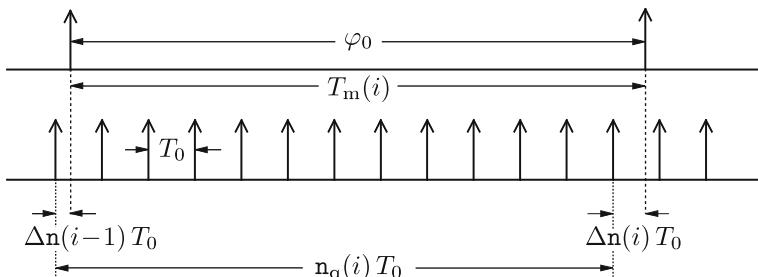


Abbildung 8.12. Quantisierungsfehler im i -ten Intervall am Anfang und am Ende einer Periode $T_m(i)$.

Das Intervall wird bei durchlaufendem Zähler am Anfang und am Ende quantisiert (Abb. 8.12). Mit dem ersten in die Zeit $T_m(i)$ fallenden Zählimpuls wird eine ganze Zählimpulsperiode T_0 gezählt. Die Teilperiode $\Delta n(i-1) T_0$ ist der Quantisierungsfehler am Anfang des Zählvorgangs. Am Ende wird die Teilperiode $\Delta n(i) T_0$ nicht mehr gezählt, was den Quantisie-

rungsfehler am Ende darstellt. Das Zählergebnis ist die Differenz zwischen dem Zählerstand am Ende der Periode $T_m(i)$ und dem Zählerstand am Anfang von $T_m(i)$.

Nachdem die von den Sensorflanken abgeleitete Zeit $T_m(i)$ und die Zählperiode T_0 statistisch unabhängig voneinander in Zeitdauer und Phasenlage sind, kann für beide Quantisierungsfehler Δn eine Gleichverteilung angenommen werden:

$$f_{\Delta n}(n) = \begin{cases} 1 & \text{für } 0 \leq n \leq 1, \\ 0 & \text{sonst.} \end{cases} \quad (8.76)$$

Das diskrete Zählergebnis ist bei Quantisierung am Intervallanfang und -ende

$$n_q(i) = n(i) + \Delta n_q(i) \quad (8.77)$$

mit dem kontinuierlichen Verhältnis

$$n(i) = \frac{T_m(i)}{T_0} \quad (8.78)$$

und dem resultierenden Quantisierungsfehler

$$\Delta n_q(i) = \Delta n(i-1) - \Delta n(i). \quad (8.79)$$

Die Wahrscheinlichkeitsdichte $f_{\Delta n_q}(n)$ des resultierenden Quantisierungsfehlers entspricht der Faltung der beiden Gleichverteilungen (8.76):

$$f_{\Delta n_q}(n) = f_{\Delta n}(n) * f_{\Delta n}(n). \quad (8.80)$$

Es ergibt sich die Dreiecksverteilung nach Abb. 8.13.

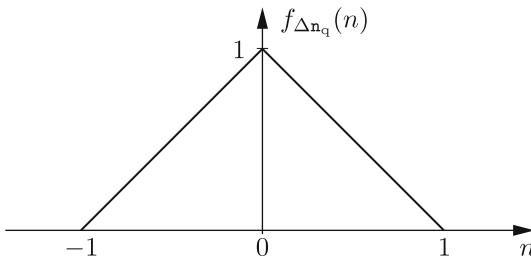
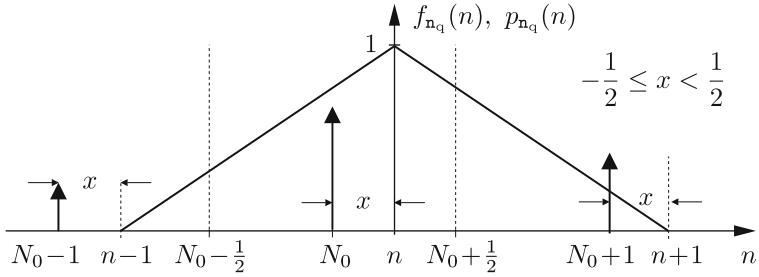


Abbildung 8.13. Wahrscheinlichkeitsdichte des resultierenden Quantisierungsfehlers $\Delta n_q(i)$.

Im Folgenden wird die Quantisierung erneut durch das lineare Modell (Satz 7.3) beschrieben, bei dem das kontinuierliche Verhältnis $n(i)$ additiv von einem Quantisierungsrauschen $\Delta n_q(i)$ mit der Dichte $f_{\Delta n_q}(n)$ überlagert ist. Aus dem deterministischen Signal $n(i)$ wird durch das überlagerte

Abbildung 8.14. Wahrscheinlichkeitsdichte des diskreten Zählergebnisses $n_q(i)$.

Rauschen ein stochastisches Signal $n_q(i)$, das eine Dreiecksdichte aufweist (Abb. 8.14).

Das Zählergebnis $n_q(i)$ kann nur diskrete Werte annehmen, die für ein betrachtetes Intervall i mit unterschiedlichen Wahrscheinlichkeiten auftreten. Die Diskretisierung der Wahrscheinlichkeitsdichte entspricht einer Flächenabtastung. Die Einzelwahrscheinlichkeiten erhält man durch Integration der Dichte $f_{\Delta n_q}(n)$ über das Quantisierungsintervall um das diskrete Zählergebnis. Es sei

$$n(i) = N_0 + x \quad \text{mit} \quad -\frac{1}{2} \leq x < \frac{1}{2}. \quad (8.81)$$

Damit werden die Einzelwahrscheinlichkeiten (Abb. 8.15):

$$p(N_0) = \int_{N_0 - \frac{1}{2}}^{N_0 + \frac{1}{2}} f_{n_q}(n) dn = \int_{-x - \frac{1}{2}}^{-x + \frac{1}{2}} f_{\Delta n_q}(n) dn = \frac{3}{4} - x^2, \quad (8.82)$$

$$p(N_0 - 1) = \int_{N_0 - \frac{3}{2}}^{N_0 - \frac{1}{2}} f_{n_q}(n) dn = \int_{-1}^{-x - \frac{1}{2}} f_{\Delta n_q}(n) dn = \frac{1}{2} \left(\frac{1}{2} - x \right)^2, \quad (8.83)$$

$$p(N_0 + 1) = \int_{N_0 + \frac{1}{2}}^{N_0 + \frac{3}{2}} f_{n_q}(n) dn = \int_{-x + \frac{1}{2}}^1 f_{\Delta n_q}(n) dn = \frac{1}{2} \left(\frac{1}{2} + x \right)^2. \quad (8.84)$$

Die diskreten Zählergebnisse treten mit der diskreten Wahrscheinlichkeitsverteilung

$$\begin{aligned} p_{n_q}(n) &= \frac{1}{2} \left(\frac{1}{2} - x \right)^2 \delta(n - (N_0 - 1)) + \left(\frac{3}{4} - x^2 \right) \delta(n - N_0) \\ &\quad + \frac{1}{2} \left(\frac{1}{2} + x \right)^2 \delta(n - (N_0 + 1)) \end{aligned} \quad (8.85)$$

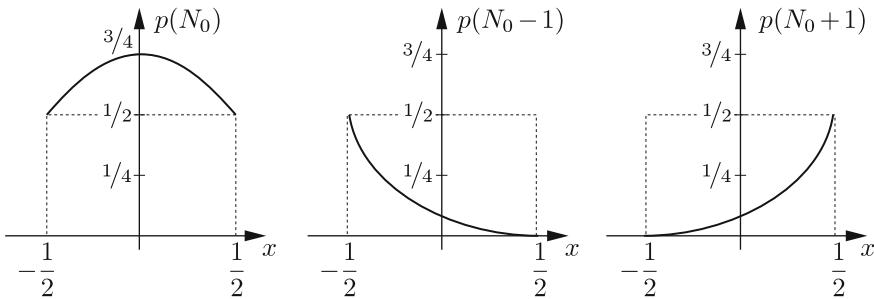


Abbildung 8.15. Einzelwahrscheinlichkeiten der diskreten Zählergebnisse.

auf. Damit lässt sich der Erwartungswert des quantisierten Zählergebnisses berechnen:

$$E\{n_q\} = \sum_{n=-\infty}^{\infty} n \cdot p_{n_q}(n) = \sum_{i=-1}^1 (N_0 + i) \cdot p(N_0 + i) \quad (8.86)$$

$$= N_0 + x = n. \quad (8.87)$$

Die Schätzung der Periode $T_m(i)$ zwischen zwei Zahnrändern über die Auszählung mit der Zählperiode T_0 ist erwartungstreu:

$$E\{n_q \cdot T_0\} = n T_0 = T_m(i). \quad (8.88)$$

Aus dem 2. Moment

$$E\{n_q^2\} = \sum_{n=-\infty}^{\infty} n^2 \cdot p_{n_q}(n) \quad (8.89)$$

$$\begin{aligned} &= (N_0 - 1)^2 \frac{1}{2} \left(\frac{1}{2} - x \right)^2 + N_0^2 \left(\frac{3}{4} - x^2 \right) \\ &\quad + (N_0 + 1)^2 \frac{1}{2} \left(\frac{1}{2} + x \right) \end{aligned} \quad (8.90)$$

$$= n^2 + \frac{1}{4} \quad (8.91)$$

erhält man die Varianz des quantisierten Zählergebnisses:

$$\sigma_{n_q}^2 = E\{n_q^2\} - E^2\{n_q\} = \frac{1}{4}. \quad (8.92)$$

Nun sollen diese Ergebnisse auf die Messung der Winkelgeschwindigkeit ω_m nach dem Periodendauerverfahren angewendet werden. Nach (8.46) gilt für den kontinuierlichen Wert der Winkelgeschwindigkeit $\omega_m(i)$ ohne

Quantisierung

$$\omega_m(i) = \frac{\varphi_0}{T_m(i)} = \frac{\varphi_0}{n(i) \cdot T_0} . \quad (8.93)$$

Der Schätzwert nach der Quantisierung lautet

$$\hat{\omega}_m(i) = \frac{\varphi_0}{n_q(i) \cdot T_0} = \frac{\varphi_0}{[n(i) + \Delta n_q(i)] T_0} . \quad (8.94)$$

Da die Zufallsgröße $\Delta n_q(i)$ nichtlinear in den Schätzwert eingeht, wird $\hat{\omega}_m(i)$ zur Bestimmung der Erwartungswerte in eine Taylor-Reihe um den kontinuierlichen Wert $\omega_m(i)$ entwickelt:

$$\hat{\omega}_m(i) \approx \omega_m(i) \left(1 - \frac{\Delta n_q(i)}{n(i)} + \frac{\Delta n_q^2(i)}{n^2(i)} \right) . \quad (8.95)$$

Mit $E\{\Delta n_q\} = 0$ und (8.92) erhält man für den Erwartungswert der Schätzung

$$E\{\hat{\omega}_m(i)\} \approx \omega_m(i) \left(1 + \frac{\sigma_{n_q}^2}{n^2(i)} \right) = \omega_m(i) \left(1 + \frac{1}{4n^2(i)} \right) . \quad (8.96)$$

Mit wachsendem n geht der zweite Term gegen null. Aufgrund der nichtlinearen Beziehung ist die Schätzung der Winkelgeschwindigkeit des Sensorrades lediglich asymptotisch (für $n \rightarrow \infty$) erwartungstreu. Die Schätzfehlervarianz lautet

$$\sigma_{\hat{\omega}_m}^2 = E\{(\hat{\omega}_m - \omega_m)^2\} \approx \omega_m^2 \frac{\sigma_{n_q}^2}{n^2(i)} \left(1 - 2 \frac{E\{\Delta n_q\}}{n} + \frac{\sigma_{n_q}^2}{n^2(i)} \right) \quad (8.97)$$

$$= \omega_m^2 \frac{1}{4n^2(i)} \left(1 + \frac{1}{4n^2(i)} \right) . \quad (8.98)$$

Die Schätzung ist somit konsistent:

$$\lim_{n \rightarrow \infty} \sigma_{\hat{\omega}_m}^2 = 0 . \quad (8.99)$$

8.2.7 Leistungsdichte des Quantisierungsfehlers

Im Folgenden sollen bei der Periodendauermessung die Zählergebnisse zwischen zwei Zahnflanken

$$n_q(i) = \underbrace{\frac{\varphi_0}{\omega_m(i) \cdot T_0}}_{n(i)} + \Delta n(i-1) - \Delta n(i) \quad (8.100)$$

in vielen aufeinanderfolgenden Intervallen betrachtet werden. Die einzelnen Quantisierungsfehler $\Delta n(i)$ in unterschiedlichen Intervallen i sind statistisch unabhängig. Die Autokorrelationsfolge des einzelnen Quantisierungsfehlers $\Delta n(i)$ lautet dann:

$$r_{\Delta n \Delta n}(k) = E\{\Delta n(i+k) \Delta n(i)\} = r_{\Delta n \Delta n}(0) \cdot \delta(k), \quad (8.101)$$

wobei $r_{\Delta n \Delta n}(0)$ dem zweiten Moment der gleichverteilten Wahrscheinlichkeitsdichte $f_{\Delta n}(n)$ entspricht:

$$r_{\Delta n \Delta n}(0) = \int_{-\infty}^{\infty} n^2 f_{\Delta n}(n) dn = \int_0^1 n^2 dn = \frac{1}{3}. \quad (8.102)$$

Die resultierenden Quantisierungsfehler $\Delta n_q(i) = \Delta n(i-1) - \Delta n(i)$ in aufeinanderfolgenden Intervallen sind dagegen statistisch voneinander abhängig. Ein einzelner Quantisierungsfehler $\Delta n(i)$ verkürzt z. B. nicht nur das Intervall i , sondern verlängert in gleichem Maße das Intervall $i+1$. Die Autokorrelationsfolge des resultierenden Quantisierungsfehlers $\Delta n_q(i)$ ist

$$r_{\Delta n_q \Delta n_q}(k) = E\{\Delta n_q(i+k) \Delta n_q(i)\} \quad (8.103)$$

$$= E\{(\Delta n(i+k-1) - \Delta n(i+k))(\Delta n(i-1) - \Delta n(i))\} \quad (8.104)$$

$$= E\{\Delta n(i+k-1) \Delta n(i-1)\} - E\{\Delta n(i+k-1) \Delta n(i)\} \\ - E\{\Delta n(i+k) \Delta n(i-1)\} + E\{\Delta n(i+k) \Delta n(i)\} \quad (8.105)$$

$$= \frac{1}{3}[2\delta(k) - \delta(k-1) - \delta(k+1)]. \quad (8.106)$$

Das Leistungsdichtespektrum ist die z-Transformierte der Autokorrelationsfolge

$$S_{\Delta n_q \Delta n_q}(z) = \sum_{k=-\infty}^{\infty} r_{\Delta n_q \Delta n_q}(k) \cdot z^{-k} = \frac{1}{3}(2 - z^{-1} - z). \quad (8.107)$$

Allgemein ist die Angabe der z-Transformierten bei der Periodendauermessung nicht möglich, da die Abtastung winkelsynchron erfolgt. Nimmt man aber an, dass die Winkelgeschwindigkeit $\omega_m(i)$ im Beobachtungszeitraum näherungsweise konstant sei, so erhält man ebenfalls konstante Abtastperioden für die Zählergebnisse $T_m(i)$. Dann kann die z-Transformierte und damit das Leistungsdichtespektrum des Quantisierungsfehlers angegeben werden. Setzt man

$$T_m = n T_0 = \frac{\varphi_0}{\omega_m} = \frac{2\pi}{Z \omega_m} = \frac{1}{Z f_m} \quad (8.108)$$

und wertet die z-Transformierte entlang des Einheitskreises aus,

$$z = \exp(j2\pi f T_m) = \exp\left(j2\pi f \frac{2\pi}{Z \omega_m}\right), \quad (8.109)$$

so erhält man die Leistungsdichte zu

$$S_{\Delta n_q \Delta n_q}(f) = \frac{1}{3} \left[2 - \exp\left(-j2\pi f \frac{2\pi}{Z \cdot \omega_m}\right) - \exp\left(j2\pi f \frac{2\pi}{Z \cdot \omega_m}\right) \right] \quad (8.110)$$

$$= \frac{4}{3} \sin^2\left(\frac{\pi f}{Z f_m}\right). \quad (8.111)$$

Mit der Abtastfrequenz $f_A = 1/T_m$ erhält man das Nyquist-Band, das sich über das folgende Intervall erstreckt:

$$-\frac{1}{2T_m} \leq f \leq \frac{1}{2T_m}. \quad (8.112)$$

Mit der Drehzahl $f_m = (Z T_m)^{-1}$ ergibt sich daraus:

$$-Z \frac{f_m}{2} \leq f \leq Z \frac{f_m}{2}. \quad (8.113)$$

Das Nyquist-Band ist damit proportional zur Winkelgeschwindigkeit ω_m der sich drehenden Welle. Je mehr Zähne Z das Sensorrad besitzt, desto breiter ist das Nyquist-Band und desto höherfrequent ist das Quantisierungsrauschen. Ähnlich wie beim Delta-Sigma-Modulator (Abschn. 7.3.4) ist das Spektrum zu höheren Frequenzen hin verschoben (Abb. 8.16).

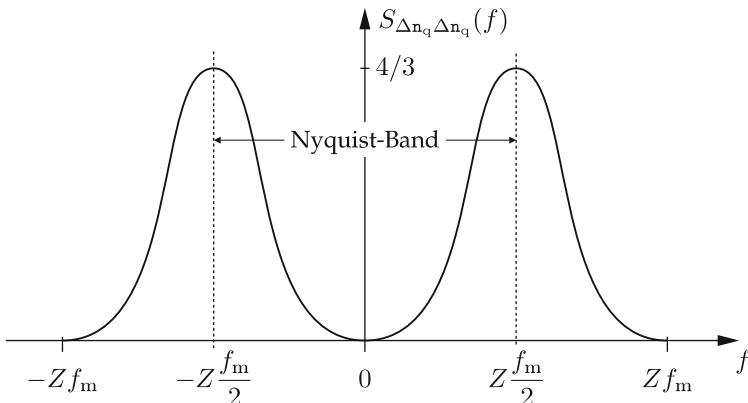


Abbildung 8.16. Leistungsdichtespektrum des Quantisierungsfehlers bei der Periodendauermessung für konstantes $\omega_m(i)$.

Mittels eines Tiefpasses könnte das Quantisierungsrauschen wirkungsvoll unterdrückt werden – allerdings auf Kosten eines Informationsverlustes bei den höherfrequenten Spektralanteilen von ω_m .

Das Quantisierungsrauschen ist unabhängig von der Zählperiode T_0 . Mit höheren Auszählfrequenzen $f_0 = 1/T_0$ steigt allerdings das absolute Zählergebnis n , so dass die relative Quantisierung abnimmt (vgl. Abschn. 8.2.3).

⌚ 8.2.8 Kompensation mechanischer Fehler des Sensorrades

Die Periode T_m wird durch Auszählung gemessen und dann durch Invertierung die Winkelgeschwindigkeit

$$\hat{\omega} = \frac{\varphi_0}{T_m} \quad (8.114)$$

bestimmt (Abb. 8.7). Die Abtastung erfolgt an den Flanken des Sensorrades, d. h. an festen Winkeln. Diese weisen in der Praxis mechanische Toleranzen $\Delta\varphi(i)$ auf, die sich durch die Ungenauigkeiten des mechanischen Herstellungsprozesses ergeben (Abb. 8.17). Aufgrund der periodischen Wiederholung mit jeder Umdrehung des Sensorrades ist der Erwartungswert aller Winkelfehler über dem Umfang des Sensorrades gleich null:

$$E\{\Delta\varphi(i)\} = 0, \quad i \in \{1, \dots, Z\}. \quad (8.115)$$

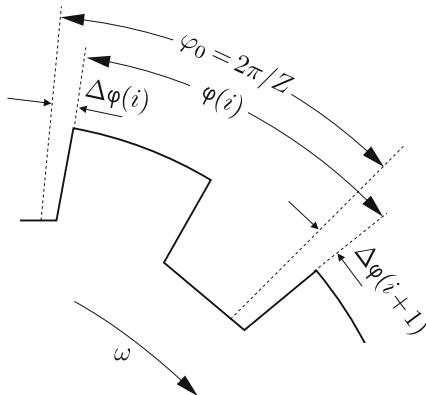


Abbildung 8.17. Winkelfehler durch mechanische Toleranzen der Zahnflanken.

⌚ 8.2.8.1 Stochastische Zahnflankenfehler

Zur weiteren Betrachtung sei zunächst angenommen, dass die mechanischen Winkelfehler $\Delta\varphi(i)$ der Zahnflanken zufällig und voneinander sta-

tistisch unabhängig seien:

$$\mathrm{E}\{\Delta\varphi(i) \Delta\varphi(l)\} = \sigma_\varphi^2 \delta_{il}, \quad i, l \in \{1, \dots, Z\}. \quad (8.116)$$

Die Periode T_m wird über den fehlerhaften Winkel

$$\varphi(i) = \frac{2\pi}{Z} + \Delta\varphi(i+1) - \Delta\varphi(i) \quad (8.117)$$

ausgezählt. Der exakte Winkel ist durch den Winkel 2π für eine Umdrehung geteilt durch die Anzahl der Zähne Z gegeben. Der Erwartungswert des fehlerhaften Winkels

$$\mathrm{E}\{\varphi(i)\} = \mathrm{E}\left\{\frac{2\pi}{Z} + \Delta\varphi(i+1) - \Delta\varphi(i)\right\} \quad (8.118)$$

$$= \frac{2\pi}{Z} + \mathrm{E}\{\Delta\varphi(i+1)\} - \mathrm{E}\{\Delta\varphi(i)\} = \frac{2\pi}{Z} \quad (8.119)$$

ist gerade gleich dem exakten Winkel φ_0 zwischen den Zahnräumen. Die Kovarianz des Winkels ist eine in der Zähnezahl Z periodische Funktion, entsprechend der periodischen Wiederholung mit jeder Umdrehung des Sensorrades:

$$C_{\varphi\varphi}(k) = \mathrm{E}\{(\varphi(i+k) - 2\pi/Z)(\varphi(i) - 2\pi/Z)\}. \quad (8.120)$$

Einsetzen des fehlerhaften Winkels (8.117) ergibt

$$C_{\varphi\varphi}(k) = \mathrm{E}\{[\Delta\varphi(i+1+k) - \Delta\varphi(i+k)] \cdot [\Delta\varphi(i+1) - \Delta\varphi(i)]\} \quad (8.121)$$

$$\begin{aligned} &= \underbrace{\mathrm{E}\{\Delta\varphi(i+1) \Delta\varphi(i+1+k)\}}_{\sigma_\varphi^2 \delta(k)} - \underbrace{\mathrm{E}\{\Delta\varphi(i+1) \Delta\varphi(i+k)\}}_{\sigma_\varphi^2 \delta(k+1)} \\ &\quad - \underbrace{\mathrm{E}\{\Delta\varphi(i) \Delta\varphi(i+1+k)\}}_{\sigma_\varphi^2 \delta(k-1)} + \underbrace{\mathrm{E}\{\Delta\varphi(i) \Delta\varphi(i+k)\}}_{\sigma_\varphi^2 \delta(k)} \quad (8.122) \end{aligned}$$

$$= \sigma_\varphi^2(2\delta(k) - \delta(k+1) - \delta(k-1)). \quad (8.123)$$

Die Kovarianz ist daher nicht vom absoluten Winkelintervall i abhängig. Außerdem ergibt sich eine statistische Abhängigkeit zwischen jeweils zwei benachbarten Intervallen i . Der Flankenfehler $\Delta\varphi(i+1)$ verlängert das Intervall $\varphi(i)$ und verkürzt gleichzeitig das Intervall $\varphi(i+1)$. Ebenso verkürzt der Flankenfehler $\Delta\varphi(i)$ das Intervall $\varphi(i)$ und verlängert das Intervall $\varphi(i-1)$.

Die z-Transformierte der periodischen Kovarianzfunktion $C_{\varphi\varphi}(k)$ ergibt ein diskretes Leistungsdichtespektrum:

$$S_{\varphi\varphi}(z) = \sum_{k=-\infty}^{\infty} C_{\varphi\varphi}(k) z^{-k} \quad (8.124)$$

$$= \sum_{k=-\infty}^{\infty} \sigma_{\varphi}^2 [2\delta(k) - \delta(k+1) - \delta(k-1)] z^{-k} \quad (8.125)$$

$$= \sigma_{\varphi}^2 (2 - z - z^{-1}). \quad (8.126)$$

Da die Kovarianzfunktion $C_{\varphi\varphi}(k)$ periodisch ist mit einer Periodendauer von einer Zahnradumdrehung $Z T_m$, ergibt sich ein diskretes Leistungsdichtespektrum, das nur bei den diskreten Frequenzen $f = i/(Z T_m)$ definiert ist. Wertet man die z-Transformierte mit $z = e^{j2\pi f T_m} = e^{j2\pi i/Z}$ entlang des Einheitskreises aus, so erhält man das folgende diskrete, periodische Leistungsdichtespektrum der Zahnfehler (Abb. 8.18)

$$S_{\varphi\varphi}(i) = 4 \sigma_{\varphi}^2 \sin^2(\pi i/Z). \quad (8.127)$$

Die als stochastisch angenommenen Zahnflankenfehler können – genauso wie das Quantisierungsrauschen – durch Mittelung aufeinanderfolgender Winkelgeschwindigkeiten $\omega_m(i)$ unterdrückt werden. Allerdings führt dies zu einem Verlust der höherfrequenten Spektralanteile von ω_m .

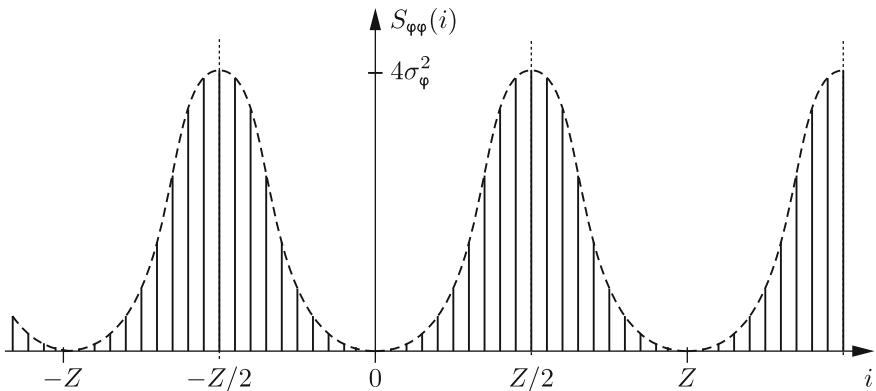


Abbildung 8.18. Leistungsdichtespektrum des stochastischen Zahnflankenfehlers bei Periodendauerauszählung.

④ 8.2.8.2 Schätzung der Zahnflankenfehler

Bislang wurden die Zahnflankenfehler als zufällig angenommen. Bei genauerer Betrachtung verursachen die mechanischen Zahnflankenfehler

einen deterministischen Fehler $\Delta T(i)$ der gemessenen Periodendauer im i -ten Intervall. Die gemessene Periodendauer zwischen zwei Zahnflanken ist

$$T_m(i) = \frac{2\pi}{Z \omega_m(i)} + \Delta T(i) + e(i), \quad i \in \{0, \dots, Z-1\}. \quad (8.128)$$

Dabei bezeichnen i den Zählindex für die Zähne des Sensorrades und $e(i)$ einen zufälligen¹ Messfehler, z. B. aufgrund des Quantisierungsrauschens. Alternativ zur Mittelwertbildung sollen im Folgenden die Zählergebnisse dadurch korrigiert werden, dass in jedem Intervall zwischen zwei Zahnflanken eine Kompensation erfolgt. Die Zahnflankenfehler können durch den Ansatz

$$\bar{T}_m(i) = c_i \cdot T_m(i) \quad (8.129)$$

mit den Korrekturfaktoren c_i kompensiert werden. Jede Messzeit $T_m(i)$ ist im i -ten Intervall durch den Zahnflankenfehler $\Delta T(i)$ verfälscht, der durch einen konstanten Korrekturterm c_i behoben wird. Zur Bestimmung der konstanten Korrekturfaktoren c_i in den Intervallen i ist es notwendig, die wahre Winkelgeschwindigkeit $\omega_m(i)$ zu kennen. Da der Zahnflankenfehler über eine Umdrehung des Sensorrades mittelwertfrei ist, ist es möglich, die wahre Winkelgeschwindigkeit $\omega_m(i)$ bzw. die wahren Durchlaufzeiten zwischen zwei Zahnflanken mit Hilfe eines gleitenden Mittelwerts $\bar{T}_m(i)$ über eine volle Umdrehung zu bestimmen. Die Voraussetzung hierfür ist, dass die Winkelgeschwindigkeit über eine Umdrehung näherungsweise konstant ist:

$$\omega_m(i) \approx \text{const.} \quad (8.130)$$

Das Gedächtnis des akausalen Mittelwertfilters ist symmetrisch über die vergangene halbe und die zukünftige halbe Umdrehung des Sensorrades verteilt (Abb. 8.19):

$$\begin{aligned} \bar{T}_m(i) &= \frac{1}{Z} \sum_{j=i-\frac{Z}{2}+1}^{i+\frac{Z}{2}-1} T_m(j) + \frac{1}{2} \left[T_m\left(i - \frac{Z}{2}\right) + T_m\left(i + \frac{Z}{2}\right) \right], \quad Z \text{ gerade}, \\ \bar{T}_m(i) &= \frac{1}{Z} \sum_{j=i-\frac{Z-1}{2}}^{i+\frac{Z-1}{2}} T_m(j), \quad Z \text{ ungerade}. \end{aligned} \quad (8.131)$$

¹Strenghalten ist die Periodendauer $T_m(i)$ aufgrund des überlagerten Fehlers $e(i)$ ebenfalls eine Zufallsgröße. Da hier die deterministischen Einflüsse im Vordergrund stehen, wird auf die Kennzeichnung als Zufallsvariable bewusst verzichtet.

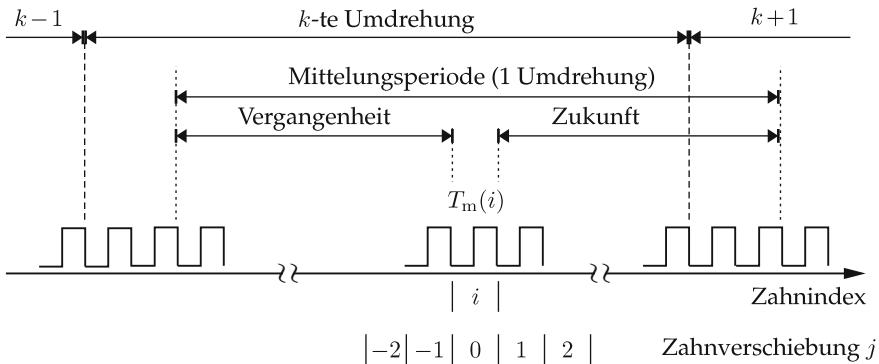


Abbildung 8.19. Akausale Mittelung zur Bestimmung der wahren Durchlaufzeit $\bar{T}_m(i)$.

Nun könnte man jedes c_i mit Hilfe eines Quotienten aus gemessener Zeit $T_m(i)$ und wahrer Zeit $\bar{T}_m(i)$ bestimmen. Es wird die k -te volle Umdrehung betrachtet. Durch Multiplikation mit dem Kompensationsfaktor c_i wird der im i -ten Intervall gemessenen Zeit $T_m(kZ + i)$ der gleitende Mittelwert über eine Umdrehung

$$\bar{T}_m(kZ + i) = c_i \cdot T_m(kZ + i) \quad (8.132)$$

zugeordnet, der frei von Zahnräderfehlern ist. Für alle Intervalle i einer Umdrehung lautet die Beziehung

$$\begin{bmatrix} \bar{T}_m(kZ) \\ \bar{T}_m(kZ + 1) \\ \vdots \\ \bar{T}_m(kZ + Z - 1) \end{bmatrix} = \begin{bmatrix} T_m(kZ) & 0 & \mathbf{0} \\ 0 & T_m(kZ + 1) & \ddots \\ & & T_m(kZ + Z - 1) \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_{Z-1} \end{bmatrix},$$

$$\bar{\mathbf{T}}_m(k) = \Phi_m(k) \cdot \mathbf{c}, \quad (8.133)$$

aus der die Faktoren c_i geschlossen berechnet werden könnten. Den Messungen ist allerdings noch ein Messfehler $e(i)$ überlagert. Erstreckt man die Messung über mehrere Umdrehungen k , so kann man mit einem Least-Squares-Schätzer den Einfluss des Messfehlers $e(i)$ bei der Bestimmung der Korrekturfaktoren c_i minimieren. Das Signalmodell für mehrere Umdrehungen ist:

hungen lautet:

$$\bar{\mathbf{T}} = \begin{bmatrix} \bar{\mathbf{T}}_m(1) \\ \bar{\mathbf{T}}_m(2) \\ \vdots \\ \bar{\mathbf{T}}_m(K) \end{bmatrix} = \begin{bmatrix} \Phi_m(1) \\ \Phi_m(2) \\ \vdots \\ \Phi_m(K) \end{bmatrix} \cdot \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_{Z-1} \end{bmatrix} = \Phi \mathbf{c}. \quad (8.134)$$

Der gesuchte Vektor der Kompensationsfaktoren ist die Pseudoinverse

$$\mathbf{c} = (\Phi^T \Phi)^{-1} \Phi^T \cdot \bar{\mathbf{T}}. \quad (8.135)$$

Man erhält damit Z Werte c_i , die das Sensorrad „kalibrieren“. Später ist nur noch jeder gemessene Wert $T_m(i)$ mit dem entsprechenden konstanten Faktor c_i zu multiplizieren, um die Fehler zu kompensieren. Abbildung 8.20 zeigt den Effekt der Korrektur am Beispiel der Kurbelwellendrehzahl eines Motors. Dargestellt ist die zur Drehzahl umgekehrt proportionale Periodendauer.

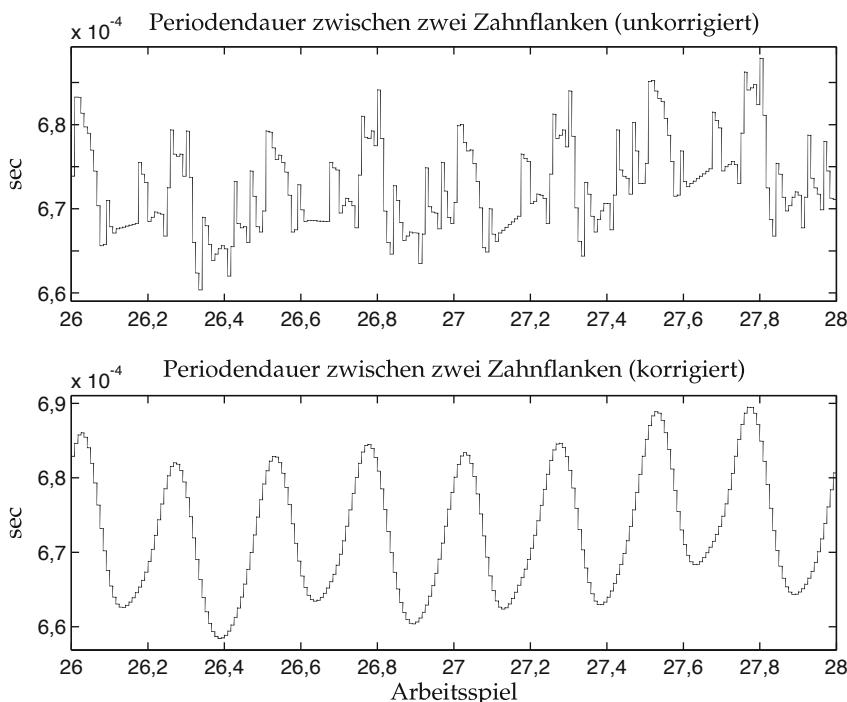


Abbildung 8.20. Korrektur mechanischer Zahnräderfehler durch Least-Squares-Schätzung.

In praktischen Anwendungen können Winkelfehler auch durch eine exzentrische Lagerung des Sensorrades verursacht werden (Abb. 8.21).

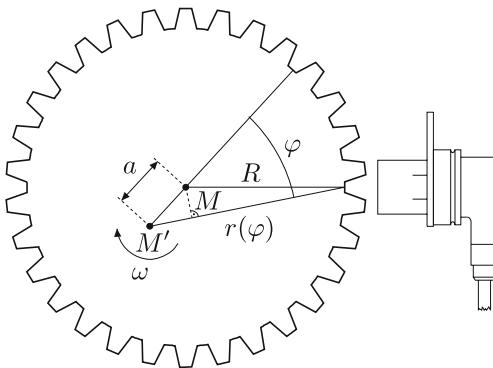


Abbildung 8.21. „Zahnräderflankenfehler“ durch exzentrische Lagerung des Sensors.

Das Fehlerbild ist ähnlich wie beim Zahnräderflankenfehler, aber abhängig von der aktuellen Drehzahl des Rades. Die Kompensation kann ebenfalls mit dem vorgestellten Least-Squares-Schätzer erfolgen. Bei aufeinanderfolgenden Schätzungen ergeben sich in Abhängigkeit von der momentanen Rad-drehzahl abweichende Kompensationsfaktoren c_i .

8.3 Kontinuierliche Frequenzmessung

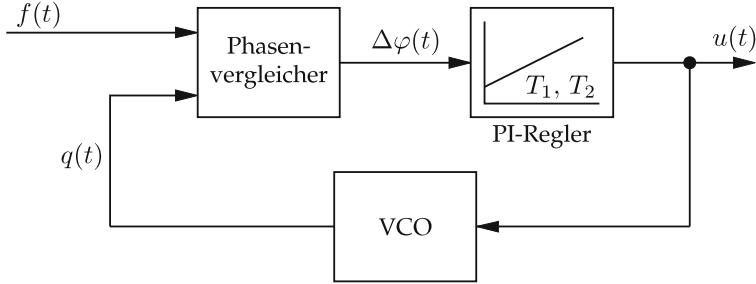
8.3

Bisher wurde die digitale Erfassung frequenzanaloger Signale besprochen. Daneben gibt es auch kontinuierliche Methoden zur Frequenzmessung. Die vorgestellten Verfahren basieren beide auf Regelkreisen, die es ermöglichen, eine Frequenz in eine proportionale Spannung (Phasenregelkreis, Abschn. 8.3.1) oder in einen quasikontinuierlichen Zählerstand (Frequenzregelkreis, Abschn. 8.3.2) umzusetzen.

8.3.1 Phasenregelkreis

Der **Phasenregelkreis** (engl. *phase-locked loop*, PLL) setzt das frequenzanaloge Eingangssignal $f(t)$ in eine frequenzproportionale Spannung $u(t)$ durch Ausregelung des Phasenfehlers um. Abbildung 8.22 zeigt den Regelkreis.

Um die Frequenz des Eingangssignals $f(t)$ zu messen, wird mittels eines spannungsgesteuerten Oszillators (engl. *voltage-controlled oscillator*, VCO) ein Vergleichssignal $q(t)$ aus der Ausgangsspannung $u(t)$ erzeugt und durch den Regelkreis die Phase der beiden Signale angeglichen. Der digitale Phasenvergleicher erfasst die Phasendifferenz und leitet diese als pulsweitenmoduliertes Signal $\Delta\varphi(t)$ an den Regler weiter. Um daraus die Phasendifferenz zu erhalten, muss das pulsweitenmodulierte Signal $\Delta\varphi(t)$ zunächst tiefpassgefiltert werden, was durch den I-Anteil des Reglers erreicht wird. Außer dieser Filterwirkung sorgt der PI-Regler für stationäre Genauigkeit.

Abbildung 8.22. Phasenregelkreis (*phase-locked loop, PLL*).

Somit ist im stationären Zustand die Ausgangsspannung $u(t)$ proportional zur Frequenz des Eingangssignals $f(t)$, die im Folgenden als $f_f(t)$ bezeichnet wird.

Das dynamische Verhalten des Phasenregelkreises wird durch den Phasenvergleicher und den Regler bestimmt. An dieser Stelle sollen nur Phasenvergleicher betrachtet werden, die den Phasenregelkreis aus jedem beliebigen Anfangszustand in den korrekten stationären Zustand überführen. Dies gilt nicht für Vergleicher, bei denen nur ein sogenannter eingeschränkter Fangbereich existiert, z. B. wenn ein einfaches XOR-Gatter als Phasenvergleicher verwendet wird. Der hier verwendete Phasenvergleicher arbeitet ausschließlich mit Rechtecksignalen (Abb. 8.23). Das Phasendifferenzsignal $\Delta\varphi(t)$ wird durch Differenz der beiden Ausgangssignale $u_{\text{up}}(t)$ und $u_{\text{down}}(t)$ gebildet:

$$\Delta\varphi(t) = u_{\text{up}}(t) - u_{\text{down}}(t) . \quad (8.136)$$

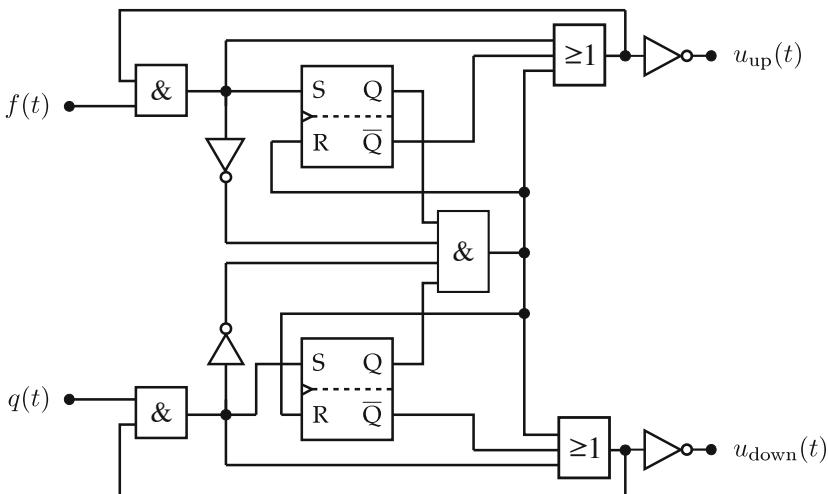
Bei einem Betrag der Phasendifferenz zwischen Eingangssignal $f(t)$ und Vergleichssignal $q(t)$ von mehr als π entstehen Mehrdeutigkeiten. In diesem Fall wird um eine oder evtl. mehrere ganze Perioden versetzt eingeregelt.

Bei einer konstanten Differenz der Frequenz der Signale $f(t)$ und $q(t)$ steigt das mittlere Phasendifferenzsignal $\overline{\Delta\varphi(t)}$ mit der Zeit an (Abb. 8.24). Im kontinuierlichen Fall gilt für die Phasendifferenz nach (8.1)

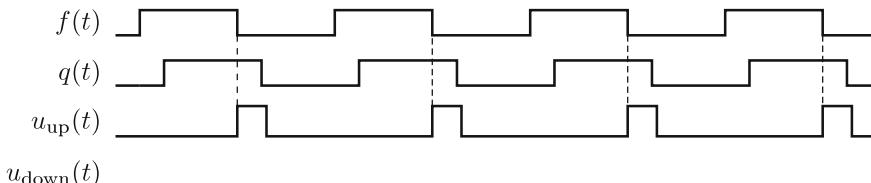
$$\Delta\varphi(t) = \int_0^t \Delta\omega(t) dt \quad \xrightarrow{\mathcal{L}} \quad \Delta\Phi(s) = \frac{1}{s} \Delta\Omega(s) . \quad (8.137)$$

Mit der Normierung auf die Maximalwerte

$$\frac{\omega_{\max}}{\varphi_{\max}} = f_{\max} = \frac{1}{T_{\min}} \quad (8.138)$$



a) $q(t)$ eilt $f(t)$ nach



b) $q(t)$ eilt $f(t)$ vor

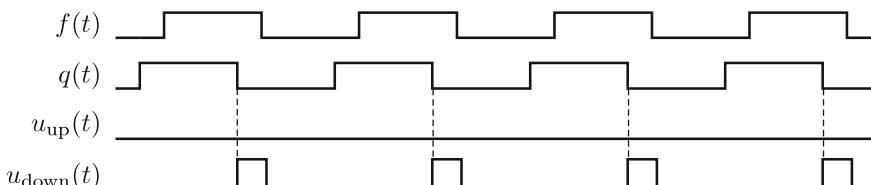


Abbildung 8.23. Digitaler Phasenvergleicher, Vergleichssignale bei konstanter Phasendifferenz.

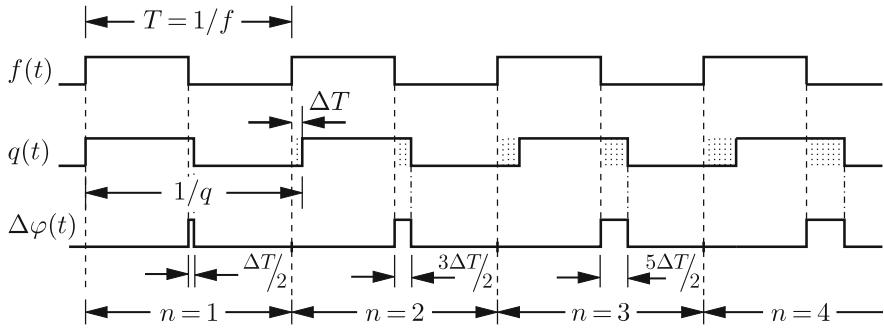


Abbildung 8.24. Phasenvergleichssignal bei konstanter Frequenzdifferenz.

erhält man aus (8.137) die Beziehung

$$\frac{\Delta\Phi(s)}{\varphi_{\max}} = \frac{1}{s T_{\min}} \frac{\Delta\Omega(s)}{\omega_{\max}} = \frac{1}{s T_{\min}} \frac{\Delta F(s)}{f_{\max}}. \quad (8.139)$$

Der Phasenvergleicher ist ein Integrator mit der Zeitkonstante $T_{\min} = 1/f_{\max}$. Der Ausgang $\Delta\varphi(t)$ des Phasenvergleichers ist zwar ein pulsweitenmoduliertes Signal, aber wegen der nachfolgenden Mittelung wird mit einem kontinuierlichen Vergleichssignal gerechnet. Das PI-Glied stabilisiert den Phasenregelkreis. Mit der Übertragungsfunktion des PI-Gliedes erhält man:

$$\frac{U(s)}{u_{\max}} = \frac{1 + s T_1}{s T_2} \frac{\Delta\Phi(s)}{\varphi_{\max}} = \frac{1 + s T_1}{s^2 T_2 T_{\min}} \frac{\Delta F(s)}{f_{\max}} \quad (8.140)$$

$$= \frac{1 + s T_1}{s^2 T_2 T_{\min}} \left[\frac{F_f(s)}{f_{\max}} - \frac{F_q(s)}{f_{\max}} \right], \quad (8.141)$$

wobei $F_f(s)$ und $F_q(s)$ die Laplace-Transformierten der Frequenzen der Signale $f(t)$ bzw. $q(t)$ bezeichnen. Das Verhalten des spannungsgesteuerten Oszillators (VCO) kann als verzögerungsfrei angenommen werden:

$$\frac{f_q(t)}{f_{\max}} = \frac{u(t)}{u_{\max}} \quad \text{mit} \quad f_q(t) = \mathcal{L}^{-1}\{F_q(s)\}. \quad (8.142)$$

Die Übertragungsfunktion des geschlossenen Phasenregelkreises ist damit

$$\frac{U(s)}{F_f(s)} = \frac{u_{\max}}{f_{\max}} \cdot \frac{1 + s T_1}{1 + s T_1 + s^2 T_{\min} T_2}. \quad (8.143)$$

Der Nenner der Übertragungsfunktion besitzt PT₂-Verhalten. Das System ist schwingungsfähig; Frequenzschwankungen oberhalb der Kennkreisfre-

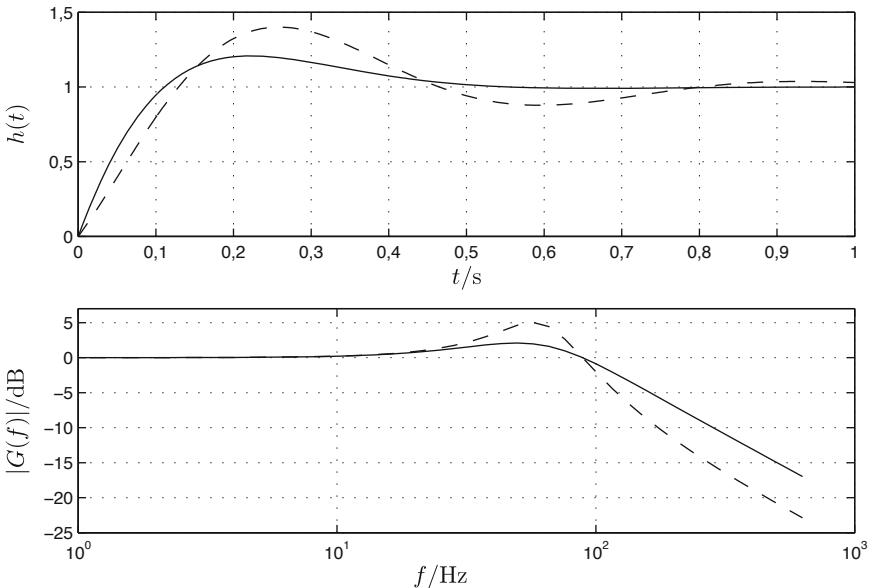


Abbildung 8.25. Sprungantwort $h(t)$ und Amplitudengang $|G(f)|$ eines Phasenregelkreises für $f_{\max} = 100 \text{ Hz}$, $f_{\min} = 1/T_2 = 1 \text{ Hz}$ und $T_1 = 1/\sqrt{2} \text{ s}$ (—) bzw. $T_1 = 1/(2\sqrt{2}) \text{ s}$ (---).

quenz

$$\omega_0 = \frac{1}{\sqrt{T_{\min} T_2}} \quad (8.144)$$

werden stark gedämpft. Abbildung 8.25 zeigt die Sprungantwort und den Amplitudengang des Phasenregelkreises für zwei verschiedene Reglereinstellungen. Zur geeigneten Parametrierung des Reglers wird die Dämpfung δ des PT₂-Gliedes eingeführt:

$$\delta = \frac{1}{2} \frac{T_1}{\sqrt{T_{\min} T_2}} = \frac{1}{2} T_1 \omega_0. \quad (8.145)$$

Nach Bsp. 5.8 erhält man mit $\delta = 1/\sqrt{2}$ ein günstiges Einschwingverhalten, woraus sich die Bedingung

$$T_1^2 = 2 T_{\min} T_2 \quad (8.146)$$

ergibt (durchgezogene Linie in Abb. 8.25). Als Beispiel für eine Dimensionierung des Regelkreises wird die Integrationszeitkonstante T_2 gleich der Periode der minimalen Eingangsfrequenz f_{\min} gesetzt: $f_{\min} = 1/T_2$. Die Kennkreisfrequenz (8.144) des Phasenregelkreises wird dann mit (8.138) gleich

dem geometrischen Mittel

$$\omega_0 = \sqrt{f_{\min} \cdot f_{\max}} \quad (8.147)$$

aus maximaler und minimaler Eingangsfrequenz. Die Schwingfrequenz des Einschwingvorgangs ist mit dieser Dimensionierung

$$\frac{1}{T_1} = \frac{\omega_0}{\sqrt{2}} = \frac{\sqrt{f_{\min} \cdot f_{\max}}}{\sqrt{2}} \quad (8.148)$$

und die Abklingzeitkonstante T_A des Einschwingvorgangs ist gerade gleich der Zeitkonstanten T_1 :

$$T_A = \frac{1}{\omega_0 \delta} = T_1. \quad (8.149)$$

8.3.2 Frequenzregelkreis

Bei der Frequenzmessung nach Abschn. 8.2.2 wurde die Frequenz als Folge von Impulsereignissen abschnittsweise über ein festes Zeitintervall T_{ref} gemittelt. Alternativ dazu ist eine quasikontinuierliche Frequenzmessung durch Umsetzung in einen Zählerstand $z(n)$ möglich. Verwendet man einen Vorwärts-Rückwärts-Zähler als Integrator und den in Abschn. 7.4.2 vorgestellten Rate-Multiplier als Frequenzgenerator, so erhält man mit dem Frequenzvergleicher

$$\Delta f(n) = f(n) - q(n) \quad (8.150)$$

der binären Impulsfolgen $f(n)$ und $q(n)$ einen digitalen Regelkreis (Abb. 8.26). Der Zählerstand des Vorwärts-Rückwärts-Zählers

$$z(n) = t_A \sum_{j=-\infty}^n \Delta f(j) = \frac{1}{f_{\max}} \sum_{j=-\infty}^n [f(j) - q(j)] \quad (8.151)$$

summiert mit der Abtastrate $f_{\max} = 1/t_A$ die Differenz zwischen Eingangsfrequenz $f(n)$ und zurückgeführter Frequenz $q(n)$. Letztere leitet sich wie folgt vom Zählerstand ab:

$$q(n) = f_{\max} \frac{z(n)}{z_{\max}}. \quad (8.152)$$

Durch Einsetzen von (8.152) in (8.151) und z-Transformation resultiert der Zählerstand im Frequenzbereich:

$$Z(z) = \frac{1}{f_{\max}} \frac{1}{1 - z^{-1}} F(z) - \frac{1}{z_{\max}} \frac{1}{1 - z^{-1}} Z(z). \quad (8.153)$$

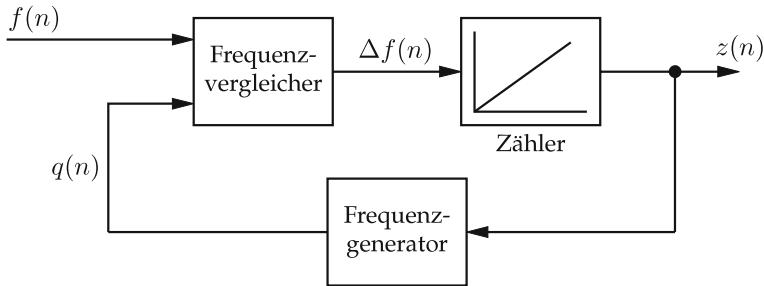


Abbildung 8.26. Quasikontinuierliche Mittelung der Eingangsfrequenz.

Damit erhält man schließlich die Übertragungsfunktion des Regelkreises

$$G(z) = \frac{Z(z)}{F(z)} = \frac{1/f_{\max}}{1 - z^{-1} + 1/z_{\max}} = \frac{1/f_{\max}}{k - z^{-1}} \quad (8.154)$$

mit

$$k = 1 + \frac{1}{z_{\max}}. \quad (8.155)$$

Mit der Sprungfolge

$$f(n) = f_{\max} \cdot \sigma(n) \quad \circ \bullet \quad F(z) = \mathcal{Z}\{f(n)\} = f_{\max} \frac{z}{z - 1} \quad (8.156)$$

ergibt sich die Sprungantwort des Systems im z-Bereich zu

$$Z(z) = \frac{z}{z - 1} \cdot \frac{z}{kz - 1} = \frac{z}{k} \left[\frac{z}{(z - 1)(z - 1/k)} \right]. \quad (8.157)$$

Zur Rücktransformation benötigt man die Partialbruchzerlegung des Klammerausdrucks. Damit erhält man

$$Z(z) = \underbrace{\frac{1}{k-1}}_{=z_{\max}} \frac{z}{z-1} - \frac{1}{k(k-1)} \frac{z}{z-1/k} \quad (8.158)$$

$$z(n) = z_{\max} (1 - k^{-(n+1)}) \sigma(n), \quad (8.159)$$

d. h. der Zählerstand $z(n)$ schwingt exponentiell gegen den Maximalwert ein. Mit der Umformung

$$k^{-(n+1)} = e^{-(n+1) \ln k}, \quad (8.160)$$

der Näherung

$$\ln k = \ln(1 + 1/z_{\max}) \approx 1/z_{\max} \quad (8.161)$$

und dem Übergang auf die kontinuierliche Zeit $t = n t_A$ kann man den kontinuierlichen Verlauf der Sprungantwort rekonstruieren:

$$z(t) \approx z_{\max} (1 - e^{-(t+t_A)/T}) \sigma(t) \quad \text{mit} \quad T = z_{\max} t_A. \quad (8.162)$$

Der Frequenzregelkreis besitzt also PT₁-Verhalten und damit Tiefpasscharakter. Frequenzschwankungen oberhalb der Eckfrequenz $\omega_e = 1/T$ werden gedämpft.

8.4

8.4 Positions- und Richtungserkennung

8.4.1 Drehrichtungserkennung

In den Abschnitten 8.2.1 und 8.2.2 wurden zwei Methoden vorgestellt, die es ermöglichen, die Winkelgeschwindigkeit ω einer Welle mittels eines Sensorsrades zu bestimmen. Zusätzlich soll nun die Drehrichtung der Welle bestimmt werden. Zur Erkennung der Drehrichtung wird der Aufbau in Abb. 8.27 eingesetzt.

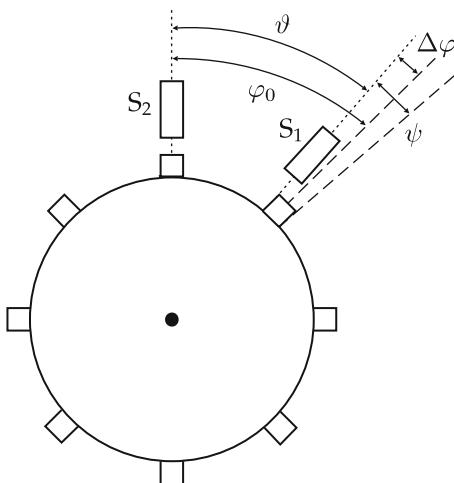


Abbildung 8.27. Einrichtung zur Drehrichtungserkennung.

Die beiden Sensoren S₁ und S₂ sind dabei so ausgerichtet, dass sie nicht gleichzeitig genau über beiden Sensorzähnen stehen, sondern der eine leicht

versetzt dazu. Der Winkel zwischen zwei Zähnen ist durch

$$\varphi_0 = \frac{2\pi}{Z} \quad (8.163)$$

gegeben, wobei Z die Anzahl der Zähne bezeichnet. Durch ψ wird der Winkel definiert, den das Rad während der Dauer eines Impulses im Sensor zurücklegt (Zahnbreite). Für den Winkelabstand zwischen den Sensoren S_1 und S_2 gilt:

$$\vartheta = k \cdot \varphi_0 \pm \Delta\varphi \quad \text{mit} \quad 0 < \Delta\varphi < \frac{\psi}{2} \quad \text{und} \quad k \in \mathbb{N}. \quad (8.164)$$

Durch die Bedingung $\Delta\varphi < \psi/2$ ist es gewährleistet, dass sich die Impulse in den Sensoren zeitlich überschneiden. In Abb. 8.27 und in der weiteren Betrachtung gilt $\vartheta = \varphi_0 - \Delta\varphi$ mit $\Delta\varphi = \psi/2$. Geht man von einer Rechtsdrehung aus, so ergibt sich das Impulsdiagramm in Abb. 8.28 für die Sensorsignale. Die Impulse erscheinen bei einer Rechtsdrehung am Sensor S_1 früher als am Sensor S_2 .

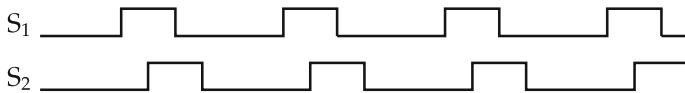


Abbildung 8.28. Sensorsignale bei einer Rechtsdrehung.

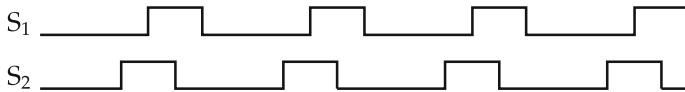


Abbildung 8.29. Sensorsignale bei einer Linksdrehung.

Bei einer Linksdrehung ergibt sich ein entsprechendes Impulsdiagramm nach Abb. 8.29. Hier erscheinen die Impulse am Sensor S_2 früher. So kann man aus der Information, an welchem Sensor die Impulse früher erscheinen, mittels einer einfachen logischen Schaltung (Abb. 8.30) auf die Drehrichtung schließen.

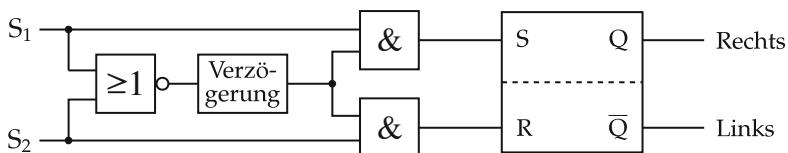


Abbildung 8.30. Gatter zur Erkennung der Drehrichtung.

Drehrichtungsänderungen während eines Impulses werden von dieser Schaltung jedoch nicht erkannt. Hierzu benötigt man ein komplizierteres Netzwerk. Das Prinzip der Drehrichtungserkennung wird beispielsweise bei mechanischen Computermäusen und *Trackballs* benutzt, um die Bewegungsrichtung zu bestimmen.

8.4.2 Positionsbestimmung

In den vorherigen Abschnitten wurde durch eine äquidistante Anbringung von Zähnen an einer Welle die Drehbewegung als frequenzanaloges Signal erfasst und daraus sowohl die Winkelgeschwindigkeit als auch die Drehrichtung abgeleitet. Analog lassen sich durch Einteilung eines linearen Maßstabs durch gleich große Gitterelemente mit abwechselnden physikalischen Eigenschaften Wege und Geschwindigkeiten messen (Abb. 8.31). Beispielsweise kommen lineare **Inkrementalgeber** bei Tintenstrahldruckern zum Einsatz, um die Position des Druckwagens zu bestimmen, sowie bei Werkzeugmaschinen oder bei bildgebenden Messsystemen in der Medizin.

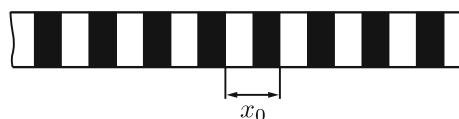


Abbildung 8.31. Inkrementalgeber mit der Gitterkonstante x_0 .

Allerdings erlauben Inkrementalgeber lediglich die Angabe von Positionsänderungen und sind somit nicht geeignet, die absolute Position zu liefern. Letztere Eigenschaft ist z. B. bei der Bestimmung des Öffnungswinkels der Drosselklappe in einem Automotor oder bei Positioniersystemen im Falle von Betriebsunterbrechungen durch Stromausfall wichtig.

Zur absoluten Codierung der Position werden mehrspurige Codelineale verwendet, die parallel abgetastet werden müssen, wobei die Gesamtheit der binären Abtastwerte ein Codewort ergibt (Abb. 8.32). Entsprechend den vielen parallelen Spuren und Sensoren ist der Aufwand deutlich höher als

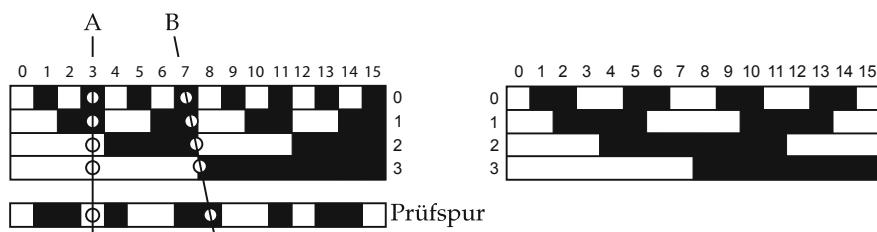


Abbildung 8.32. Links: vierspuriges, dualcodiertes Codelineal mit Prüfspur (A und B: korrekt bzw. fehlerhaft justiertes Abtastlineal); rechts: vierspuriges Codelineal mit Gray-Codierung.

bei Inkrementalgebern. So benötigt man bei einer Auflösung von 256 Schritten 8 Spuren und ebenfalls 8 Sensoren. Zur Codierung der Position bieten sich mehrere Möglichkeiten an:

- **Dualcode:** Der Dualcode ist vor allem aus dem Einsatz in Rechnersystemen bekannt. Dabei weist das i -te Bit die feste Stellenwertigkeit 2^i auf. Zur Codierung von Weg- und Winkelgebern ist der Dualcode jedoch ungeeignet, da bei fehlerhafter Justierung des Abtastlineals an den Übergangsstellen große Fehler auftreten können (siehe Abb. 8.32 links, Fall B). Durch Hinzufügen einer Prüfspur lassen sich Abtastfehler erkennen. Das Prüf- oder Paritätsbit ist genau dann eins, wenn die Quersumme der tatsächlichen Spuren ungerade ist, sonst ist es null. Durch Vergleich von Quersumme und Paritätsbit kann somit jede ungerade Fehlerzahl erkannt werden [3].
- **Gray-Code:** Der Gray-Code ist ein einschrittiger binärer Code, d. h. aufeinander folgende Gray-Zahlen unterscheiden sich jeweils nur in einem Bit (Abb. 8.32 rechts). Anders als der Dualcode weist der Gray-Code keine feste Stellenwertigkeit auf. In den Übergängen von einer Stufe zur nächsten ergeben sich wenige Fehlermöglichkeiten, da sich nur ein Bit ändert. Daher wird dieser Code bevorzugt bei weg- und winkelgebenden Messungen eingesetzt.

Zur Messung der absoluten Winkelposition können ebenfalls **Absolutwinkelgeber** mit Gray-Codierung eingesetzt werden (Abb. 8.33). Aufgrund des hohen Realisierungsaufwands werden jedoch Methoden bevorzugt, bei denen bereits aus wenigen Spuren die absolute Position bestimmt werden kann. Der einfachste Fall basiert darauf, auf der Codescheibe neben der In-

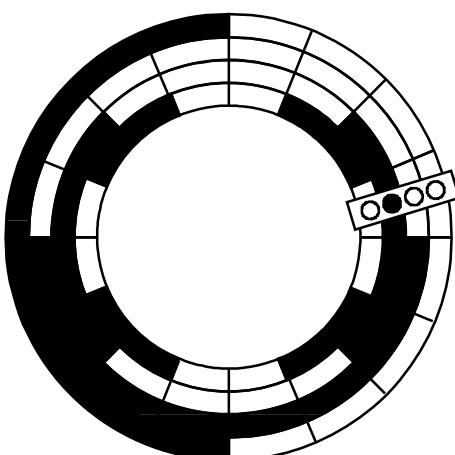


Abbildung 8.33. Absolutwinkelgeber mit Gray-Codierung.

inkrementspur eine weitere Spur mit einer Indexmarke aufzubringen. Die Indexmarke definiert dabei die Nullposition der Scheibe. Diese Methode hat allerdings den entscheidenden Nachteil, dass sich die Welle im ungünstigsten Fall (z. B. bei Verzählen) um 360° drehen muss, bis der Detektor die Indexmarke erfasst und damit die Nullposition feststellt. Erst ab diesem Zeitpunkt kann die absolute Position nach einem Fehler erneut bestimmt werden.

Die Nachteile der einfachen Scheibe mit 2 Spuren können mit einer „abstandscodierten“ Scheibe behoben werden. Dazu befinden sich weitere Indexmarken auf der zweiten Spur der Scheibe. Diese Markierungen unterteilen die Codespur in einzelne Segmente. Die Indexmarken werden so angeordnet, dass sich aus dem Abstand dazwischen die Segmentnummer und damit die absolute Position errechnen lässt. Die Marken 1 und 2 sind d Inkremente voneinander entfernt, die Marken 2 und 3 ($d + 1$) Inkremente, die Marken 3 und 4 ($d + 2$) Inkremente und so weiter. Die Abstände zwischen den Indexmarken sind damit jeweils eindeutig einer bestimmten Position zugeordnet. Als Beispiel zeigt Abb. 8.34 den Aufbau einer abstandscodierten Scheibe mit $d = 4$ und insgesamt 72 Inkrementmarken. Nach spätestens 12 Impulsen (entspricht 60°) kann die absolute Position erkannt werden.

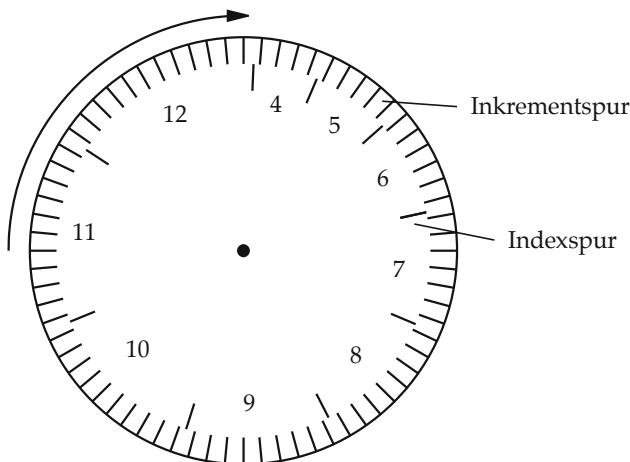
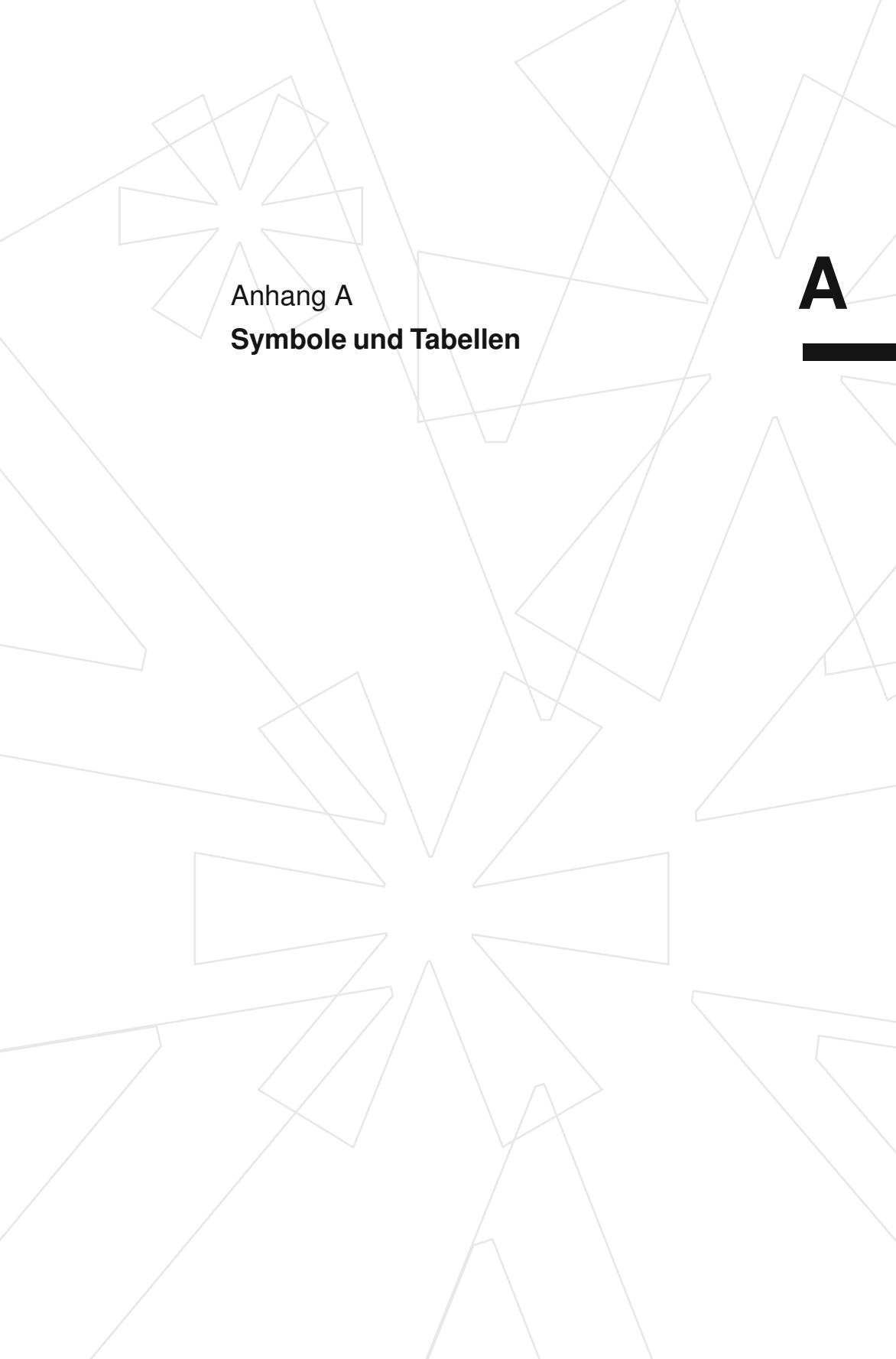


Abbildung 8.34. Beispiel einer abstandscodierten Scheibe.



Anhang A

Symbole und Tabellen

A

A

A Symbole und Tabellen

| | | |
|-------|-------------------------------|-----|
| A.1 | Symbolverzeichnis | 425 |
| A.1.1 | Konventionen | 425 |
| A.1.2 | Operatoren | 425 |
| A.1.3 | Lateinische Symbole | 425 |
| A.1.4 | Griechische Symbole | 429 |
| A.2 | Tabellen | 430 |

A Symbole und Tabellen

A.1

A.1 Symbolverzeichnis

Die Auflistung der Symbole spiegelt die Verwendung in diesem Buch wider. Dabei wurde die internationale Schreibweise beachtet. Es muss jedoch darauf hingewiesen werden, dass sich in anderen Quellen die Schreibweise unterscheiden kann.

⌚ A.1.1 Konventionen

| | |
|---------------------------------------|---|
| a, b, \dots | Konstante, Bezeichner |
| x, y, X, Y, \dots | skalare Variable (kursiv, Kleinbuchstaben, gelegentlich Großbuchstaben) |
| $\mathbf{x}, \mathbf{y}, \dots$ | Vektoren (fett, Kleinbuchstaben) |
| $\mathbf{A}, \mathbf{B}, \dots$ | Matrizen (fett, Großbuchstaben) |
| a, b, \dots | Zufallsvariable (Schreibmaschinenschrift) |
| $x(t), y(t), \dots$ | Zeitsignale |
| $\mathbf{x}(t), \mathbf{y}(t), \dots$ | stochastische Prozesse (Schreibmaschinenschrift) |

⌚ A.1.2 Operatoren

| | |
|-----------------------|--|
| Δ | Differenzenoperator |
| $E\{\}$ | Erwartungswertoperator |
| $\Im\{\}$ | Imaginärteil einer komplexen Zahl |
| $\mathcal{L}\{x(t)\}$ | Laplace-Transformierte der Funktion $x(t)$ |
| $\mathcal{F}\{x(t)\}$ | Fourier-Transformierte der Funktion $x(t)$ |
| $\Re\{\}$ | Realteil einer komplexen Zahl |
| $\mathcal{S}\{\}$ | System |
| $\mathcal{Z}\{x(t)\}$ | z-Transformierte der Funktion $x(t)$ |

⌚ A.1.3 Lateinische Symbole

| | |
|-------------------|--------------------------------------|
| a, \mathbf{a} | Beschleunigung |
| a_1, a_2, \dots | Koeffizienten |
| \mathbf{a} | Koeffizientenvektor, Parametervektor |
| A | Fläche |
| B | Bandbreite eines Signals |

| | |
|-------------------------------------|--|
| c | Federkonstante, Wärmekapazität |
| C | Kapazität, Kondensator |
| C_{xx} | Autokovarianz |
| C_{xy} | Kreuzkovarianz |
| \mathbb{C} | Menge der komplexen Zahlen |
| d | Durchmesser (Zufallsgröße) |
| e | Euler'sche Zahl |
| \mathbf{e} | Einheitsvektor |
| $e(t), \mathbf{e}(t)$ | Fehlersignal allgemein |
| $e_q(t), \mathbf{e}_q(t)$ | Quantisierungsfehler |
| $\mathbf{e}_1, \mathbf{e}_2, \dots$ | Einheitsvektoren |
| E | Energie |
| E_x | Energie des Signals $x(t)$ |
| $E(s), E(f)$ | Fehlersignal im Frequenzbereich |
| f | Frequenz |
| \bar{f} | mittlere Frequenz |
| f_A | Abtastfrequenz |
| f_0 | Zählfrequenz |
| f_{eff} | Effektivfrequenz |
| f_g | Grenzfrequenz |
| $f_x(x)$ | Wahrscheinlichkeitsdichte von x |
| $F_x(x)$ | Wahrscheinlichkeitsverteilung von x |
| F | absoluter Fehler, Kraft |
| F_r | relativer Fehler, relativer Kennlinienfehler bezogen auf die Anzeigespanne |
| F_{rA} | relativer Kennlinienfehler bezogen auf den Anzeigebereich |
| F_{rS} | relativer Kennlinienfehler bezogen auf den Sollwert |
| F_{rE} | relativer Kennlinienfehler bezogen auf den Endwert |
| $F_{r,\text{sup}}$ | relativer superponierender Fehler (additiv) |
| $F_{r,\text{def}}$ | relativer deformierender Fehler (multiplikativ) |
| g | Erdbeschleunigung |
| $G(s), G(f)$ | Übertragungsfunktion |
| h | Hysterese |
| h | Höhe (Zufallsgröße) |
| H | Hurwitz-Determinante |
| i | Stromstärke |

| | |
|------------------|---|
| $i(u)$ | Impulsantwort eines Interpolationsfilters |
| $I(f)$ | Übertragungsfunktion eines Interpolationsfilters |
| j | imaginäre Einheit |
| K, K' | Konstante |
| ℓ | Länge |
| $L_i(u)$ | Lagrange-Polynom |
| m, m | Masse |
| n | Index der Abtastwerte, allgemeiner Zahlenindex, Drehzahl |
| $n(t)$ | stochastisches Signal, Rauschen |
| n_q | quantisiertes Zählergebnis |
| N | allgemeiner Zahlenindex $\in \mathbb{N}$ |
| \mathbb{N} | Menge der natürlichen Zahlen |
| p | Druck |
| p_i | Wahrscheinlichkeit, dass $x_i \leq x < x_{i+1}$ |
| $P\{x \leq x\}$ | Wahrscheinlichkeit für $x \leq x$ |
| P_x | Leistung des Signals $x(t)$ |
| q | Quantisierungsstufe |
| Q | Gütekriterium allgemein |
| r, r | Radius |
| $r_T(t)$ | Rechteckfenster der Breite T |
| $r_{xx}(\tau)$ | Autokorrelationsfunktion |
| $r_{xy}(\tau)$ | Kreuzkorrelationsfunktion |
| $r_{xy}^E(\tau)$ | Korrelation für Energiesignale |
| R | elektrischer Widerstand |
| $R_q(f)$ | Fourier-Transformierte der gleichverteilten Wahrscheinlichkeitsdichte |
| \mathbb{R} | Menge der reellen Zahlen |
| s | komplexe Frequenz $s = \sigma + j\omega$ im Laplace-Bereich |
| $s(u)$ | Spline |
| $s_{xx}(t)$ | Energiedichte im Zeitbereich |
| s_x^2 | Stichprobenvarianz |
| S | Empfindlichkeit (engl. <i>sensitivity</i>) |
| S_i | Empfindlichkeit, ideal, linear |
| $S_{xx}(f)$ | Energie- bzw. Leistungsdichtespektrum |
| $S_{ee}(f)$ | Leistungsdichte eines allgemeinen Fehlersignals |
| $S_{nn}(f)$ | Leistungsdichte eines Rauschsignals |

| | |
|--------------------|--|
| t | Zeit, Zeitpunkte |
| t_A | Abtastzeit |
| T | Periodendauer, Temperatur |
| T_i | Integrationszeitkonstante |
| T_m | Periodendauer zwischen zwei Zahnflanken bei Periodendauermessung |
| T_{ref} | Referenzperiode bei Frequenzmessung |
| u | Messgröße, elektrische Spannung |
| $u(t)$ | Eingangssignal, Messgröße |
| $u_q(t)$ | quantisiertes Signal |
| u_0 | Arbeitspunkt |
| u_a | Messanfang |
| u_e | Messende |
| u_{ref} | Referenzspannung |
| U | elektrische Spannung |
| v, \mathbf{v} | Geschwindigkeit |
| V | Verstärkungsfaktor |
| \mathbf{V} | Vandermonde-Matrix |
| x | Umkehrspanne |
| \mathbf{x} | Zustandsvektor |
| \bar{x} | zeitlicher Mittelwert |
| \hat{x} | Schätzwert allgemein |
| $\hat{\mathbf{x}}$ | Stichprobenmittelwert |
| y | Messwert |
| \mathbf{y} | Messwertvektor |
| $y(t)$ | Ausgangssignal, kontinuierliches Zeitsignal |
| $y_*(t)$ | abgetastetes Signal |
| $y_n, y(nT)$ | zeitdiskretes Signal |
| y_a | angezeigter Wert, Anzeigeanfang |
| y_e | Anzeigeende |
| y_i | Sollwert, Ausgangswert |
| y_w | wahrer Wert |
| $Y(f)$ | Spektrum allgemein, Frequenzgang |
| $Y_*(f)$ | Spektrum eines zeitdiskreten Signals |
| \mathbf{z} | Störgrößenvektor |
| z_0 | feste Störgröße bei Normalbedingungen |

- z komplexe Variable im z-Bereich
- Z Zählerstand, digitaler Zahlenwert
- Z Anzahl der Zähne eines Zahnrades
- \mathbb{Z} Menge der ganzen Zahlen

⌚ A.1.4 Griechische Symbole

- α Signifikanzniveau, Irrtumswahrscheinlichkeit
- β Dämpfungskonstante
- δ Dämpfung eines PT₂-Gliedes
- δ_i^j Kronecker-Delta
- $\delta(t)$ Dirac-Impuls
- Δ_f Bandbreite eines Signals
- Δy Abweichung vom wahren Wert/Arbeitspunkt
- $\Delta\varphi$ Winkelfehler am Zahnrad
- $\zeta(t)$ Ereignisprozess
- θ Winkel
- ϑ Frequenzverschiebung
- κ Krümmung einer Kurve
- λ Ausfallrate, Übergangsrate
- $\Lambda(x)$ Dreieckfunktion
- μ Mittelwert, Scharmittelwert
- ρ Dichte (Zufallsgröße)
- ρ_{xy} Korrelationskoeffizient
- $\varrho_{XY}(\vartheta)$ spektrale Korrelationsfunktion
- $\sigma(t)$ Sprungfunktion (Heaviside-Funktion)
- σ_x^2 Varianz der Zufallsgröße x
- Σ_x Kovarianzmatrix des Zufallsvektors \mathbf{x}
- τ Zeitverschiebung
- ϕ, φ Winkel allgemein
- φ_0 Winkelinkrement bei Sensorzahnrad
- $\varphi_i(u), \varphi_j(u)$ Basisfunktionen
- $\Phi(f)$ charakteristische Funktion
- Φ Beobachtungsmatrix des linearen Signalmodells
- χ^2 χ^2 -verteilte Zufallsgröße
- ψ Winkel
- ω Winkelgeschwindigkeit, Kreisfrequenz

- $\omega_m(t)$ kontinuierliche Winkelgeschwindigkeit
 $\Omega(f)$ Fourier-Transformierte der Winkelgeschwindigkeit

A.2**A.2 Tabellen**

Tabelle A.1. Gauß'sche Fehlerfunktion $\text{erf}(c)$ und statistische Sicherheit $P(c)$ bzw. $P_n(c)$ bei bekannter sowie bei geschätzter Standardabweichung (Wahrscheinlichkeiten in Prozent).

| c | $\text{erf}(c)$ | $P(c)$ | $P_2(c)$ | $P_3(c)$ | $P_4(c)$ | $P_5(c)$ | $P_{10}(c)$ | $P_{20}(c)$ |
|-----|-----------------|---------|----------|----------|----------|----------|-------------|-------------|
| 0,5 | 0,52050 | 38,2925 | 33,3333 | 34,8552 | 35,6670 | 36,1701 | 37,2106 | 37,7468 |
| 0,6 | 0,60386 | 45,1494 | 39,0567 | 40,9199 | 41,9159 | 42,5340 | 43,8145 | 44,4756 |
| 0,7 | 0,67780 | 51,6073 | 44,3607 | 46,5673 | 47,7500 | 48,4851 | 50,0112 | 50,8010 |
| 0,8 | 0,74210 | 57,6289 | 49,2366 | 51,7801 | 53,1473 | 53,9986 | 55,7700 | 56,6889 |
| 0,9 | 0,79691 | 63,1880 | 53,6895 | 56,5549 | 58,0994 | 59,0629 | 61,0721 | 62,1170 |
| 1,0 | 0,84270 | 68,2689 | 57,7350 | 60,8998 | 62,6099 | 63,6783 | 65,9107 | 67,0743 |
| 1,1 | 0,88021 | 72,8668 | 61,3960 | 64,8317 | 66,6916 | 67,8549 | 70,2893 | 71,5603 |
| 1,2 | 0,91031 | 76,9861 | 64,6997 | 68,3738 | 70,3649 | 71,6109 | 74,2204 | 75,5838 |
| 1,3 | 0,93401 | 80,6399 | 67,6753 | 71,5532 | 73,6548 | 74,9699 | 77,7234 | 79,1616 |
| 1,4 | 0,95229 | 83,8487 | 70,3526 | 74,3993 | 76,5899 | 77,9596 | 80,8235 | 82,3165 |
| 1,5 | 0,96611 | 86,6386 | 72,7607 | 76,9416 | 79,2000 | 80,6096 | 83,5493 | 85,0764 |
| 1,6 | 0,97635 | 89,0401 | 74,9269 | 79,2095 | 81,5151 | 82,9505 | 85,9318 | 87,4723 |
| 1,7 | 0,98379 | 91,0869 | 76,8767 | 81,2309 | 83,5645 | 85,0123 | 88,0031 | 89,5369 |
| 1,8 | 0,98909 | 92,8139 | 78,6334 | 83,0320 | 85,3762 | 86,8242 | 89,7948 | 91,3035 |
| 1,9 | 0,99279 | 94,2567 | 80,2181 | 84,6368 | 86,9761 | 88,4137 | 91,3378 | 92,8052 |
| 2,0 | 0,99532 | 95,4500 | 81,6497 | 86,0674 | 88,3883 | 89,8061 | 92,6612 | 94,0734 |
| 2,1 | 0,99702 | 96,4271 | 82,9450 | 87,3435 | 89,6347 | 91,0247 | 93,7923 | 95,1382 |
| 2,2 | 0,99814 | 97,2193 | 84,1191 | 88,4828 | 90,7347 | 92,0906 | 94,7559 | 96,0271 |
| 2,3 | 0,99886 | 97,8552 | 85,1852 | 89,5012 | 91,7061 | 93,0228 | 95,5746 | 96,7653 |
| 2,4 | 0,99931 | 98,3605 | 86,1550 | 90,4126 | 92,5644 | 93,8379 | 96,2684 | 97,3751 |
| 2,5 | 0,99959 | 98,7581 | 87,0388 | 91,2293 | 93,3233 | 94,5510 | 96,8553 | 97,8766 |
| 2,6 | 0,99976 | 99,0678 | 87,8459 | 91,9624 | 93,9952 | 95,1751 | 97,3509 | 98,2873 |
| 2,7 | 0,99987 | 99,3066 | 88,5841 | 92,6214 | 94,5906 | 95,7216 | 97,7687 | 98,6222 |
| 2,8 | 0,99992 | 99,4890 | 89,2607 | 93,2147 | 95,1188 | 96,2006 | 98,1205 | 98,8943 |
| 2,9 | 0,99996 | 99,6268 | 89,8820 | 93,7498 | 95,5882 | 96,6209 | 98,4166 | 99,1146 |
| 3,0 | 0,99998 | 99,7300 | 90,4534 | 94,2331 | 96,0058 | 96,9901 | 98,6656 | 99,2924 |
| 3,1 | 0,99999 | 99,8065 | 90,9799 | 94,6704 | 96,3779 | 97,3147 | 98,8749 | 99,4355 |
| 3,2 | 0,99999 | 99,8626 | 91,4659 | 95,0668 | 96,7099 | 97,6005 | 99,0508 | 99,5504 |
| 3,3 | 1,00000 | 99,9033 | 91,9152 | 95,4267 | 97,0067 | 97,8524 | 99,1987 | 99,6424 |
| 3,4 | 1,00000 | 99,9326 | 92,3313 | 95,7539 | 97,2723 | 98,0749 | 99,3229 | 99,7159 |
| 3,5 | 1,00000 | 99,9535 | 92,7173 | 96,0519 | 97,5104 | 98,2716 | 99,4273 | 99,7745 |

Literaturverzeichnis

- [1] W. J. Bartz und H. Brenneke (Hrsg.). *Frühdiagnose von Schäden an Maschinen und Maschinenanlagen*. Expert-Verlag, Ehningen, 1988.
- [2] J. Beyerer, F. Puente León und C. Frese. *Automatische Sichtprüfung – Grundlagen, Methoden und Praxis der Bildgewinnung und Bildauswertung*. Springer, Berlin, Heidelberg, 2012.
- [3] M. Bossert. *Kanalcodierung*. Teubner, Stuttgart, 2. Auflage, 1998.
- [4] Y. Cao. Correcting the minimum ITAE standard forms of zero-displacement-error systems. *Journal of Zhejiang University (Natural Science)*, 23(4):550–559, 1989.
- [5] Deutsches Institut für Normung. *Grundlagen der Messtechnik – Teil 1: Grundbegriffe*. DIN 1319-1, 1995.
- [6] R. C. Dorf und R. H. Bishop. *Modern control systems*. Pearson Prentice Hall, Upper Saddle River, 11. Auflage, 2008.
- [7] L. Fahrmeir, R. Künstler, I. Pigeot und G. Tutz. *Statistik*. Springer, Berlin, Heidelberg, New York, 5. Auflage, 2007.
- [8] O. Fiedler. *Strömungs- und Durchflußmeßtechnik*. Oldenbourg Verlag, München, 1992.
- [9] L. Finkelstein. Widely, strongly and weakly defined measurement. *Measurement*, 34:39–48, 2003.
- [10] O. Föllinger. *Nichtlineare Regelungen I: Grundbegriffe, Anwendung der Zustandsebene, Direkte Methode*. Oldenbourg Verlag, München, 1998.
- [11] O. Föllinger. *Regelungstechnik: Einführung in die Methoden und ihre Anwendung*. Hüthig, Heidelberg, 2008.
- [12] M. Giacinta und S. Hildebrandt. *Calculus of variations Vol 1. and 2*. Springer, Berlin, Heidelberg, New York, 1996.
- [13] F. D. Graham und R. C. Lathrop. Synthesis of optimum transient response: Criteria and standard forms. *AIEE Transactions*, 72(9):273–286, 1953.
- [14] F. H. Harmuth. *Transmission of information by orthogonal functions*. Springer, Berlin, Heidelberg, New York, 1972.
- [15] H. Heppner und I. Benecke. Entwurf optimaler Filter für einen Sigma-Delta-ADU bei statischen Eingangssignalen. *Wissenschaftliche Zeitschrift der TU Magdeburg*, 35, Heft 2:59–65, 1991.
- [16] W. Hess. *Digitale Filter*. Teubner, Stuttgart, 2. Auflage, 1993.
- [17] P. G. Hoel. *Introduction to mathematical statistics*. J. Wiley and Sons, New York, 1984.
- [18] A. Hurwitz. Über die Bedingungen, unter welchen eine Gleichung nur Wurzeln mit negativen reellen Teilen besitzt. *Math. Annalen*, 46:273–284, 1895.
- [19] E. Hänsler. *Statistische Signale: Grundlagen und Anwendungen*. Springer, Berlin, Heidelberg, New York, 3. Auflage, 2001.

- [20] International Organization for Standardization. *Quantities and units – Part 1: General*. ISO 80000-1, 2009.
- [21] R. Isermann. *Identifikation dynamischer Systeme, Band 1 und 2*. Springer, Berlin, Heidelberg, New York, 1992.
- [22] N. S. Jayant und P. Noll. *Digital coding of waveforms*. Prentice-Hall, Englewood Cliffs, New Jersey, 1984.
- [23] F. Jondral und A. Wiesler. *Grundlagen der Wahrscheinlichkeitsrechnung und stochastischer Prozesse für Ingenieure*. Teubner, Stuttgart, Leipzig, 2000.
- [24] K. Kammeyer und K. Kroschel. *Digitale Signalverarbeitung: Filterung und Spektralanalyse*. Vieweg + Teubner, Wiesbaden, 2009.
- [25] U. Kiencke. *Ereignisdiskrete Systeme: Modellierung und Steuerung verteilter Systeme*. Oldenbourg Verlag, München, 2. Auflage, 2006.
- [26] U. Kiencke und L. Nielsen. *Automotive control systems*. Springer, Berlin, Heidelberg, New York, 2. Auflage, 2005.
- [27] U. Kiencke, M. Schwarz und T. Weickert. *Signalverarbeitung – Zeit-Frequenz-Analyse und Schätzverfahren*. Oldenbourg Verlag, München, 2008.
- [28] A. Leonhard. Ein neues Verfahren zur Stabilitätsuntersuchung. *Archiv für Elektrotechnik*, 38:17–29, 1944.
- [29] L. Ljung. *System identification*. Prentice-Hall, Englewood Cliffs, New Jersey, 1987.
- [30] F. Mesch. Geschwindigkeits- und Durchflußmessung mit Korrelationsverfahren. *Regelungstechnische Praxis*, 24(3):73–82, 1982.
- [31] K. Meyberg und P. Vachenauer. *Höhere Mathematik 2*. Springer, Berlin, Heidelberg, New York, 4. Auflage, 2001.
- [32] G. C. Newton, L. A. Gould und J. F. Kaiser. *Analytical design of linear feedback controls*. John Wiley & Sons, New York, London, Sidney, 1957.
- [33] P. V. Novickij, V. G. Knorring und V. S. Gutnikov. *Frequenzanaloge Meßeinrichtungen*. VEB Verlag Technik, Berlin, 1975.
- [34] A. V. Oppenheim, R. W. Schafer und J. R. Buck. *Zeitdiskrete Signalverarbeitung*. Pearson Studium, München, 2. Auflage, 2004.
- [35] A. Papoulis. Error analysis in sampling theory. *Proceedings of the IEEE*, 54(7):947–955, 1966.
- [36] A. Papoulis. *Systems and transformations with applications in optics*. Krieger Publishing, Malabar, Florida, 1981.
- [37] A. Papoulis und S. Unnikrishna Pillai. *Probability, random variables and stochastic processes*. McGraw-Hill, New York, 4. Auflage, 2002.
- [38] M. Paul. Einige Abschätzungen der Fehler infolge zeitlicher Abtastfehler bei der Diskretisierung kontinuierlicher Signale. *Regelungstechnik*, 12:409–413, 1975.
- [39] K. R. Popper. *Logik der Forschung*. Akademie Verlag, Berlin, 3. Auflage, 2007.

- [40] F. Puente León. *Automatische Identifikation von Schußwaffen*. Fortschritt-Berichte VDI, Reihe 8, Vol. 787. VDI Verlag, Düsseldorf, 1999.
- [41] F. Puente León. Automated comparison of firearm bullets. *Forensic Science International*, 156(1):40–50, 2006.
- [42] F. Puente León, U. Kiencke und H. Jäkel. *Signale und Systeme*. Oldenbourg Verlag, München, 5. Auflage, 2011.
- [43] E. Schrüfer. *Elektrische Messtechnik*. Hanser, München, 9. Auflage, 2007.
- [44] D. W. Scott. On optimal and data-based histograms. *Biometrika*, 66(3):605–610, 1979.
- [45] M. Seifart. *Analoge Schaltungen*. Verlag Technik, Berlin, 6. Auflage, 2003.
- [46] T. R. H. Sizer. *The digital differential analyzer*, Seiten 21–23. Chapman and Hall, London, 1968.
- [47] M. R. Spiegel und J. Liu. *Mathematical handbook of formulas and tables*. McGraw-Hill, New York, 2. Auflage, 1999.
- [48] M. D. Springer. *The algebra of random variables*. J. Wiley and Sons, New York, 1979.
- [49] S. D. Stearns und D. R. Hush. *Digitale Verarbeitung analoger Signale*. Oldenbourg Verlag, München, Wien, 1999.
- [50] A. Steland. *Basiswissen Statistik*. Springer, Berlin, Heidelberg, New York, 2. Auflage, 2010.
- [51] A. Stuart und J. K. Ord. *Kendall's advanced theory of statistics*, volume 1, Distribution theory. Arnold, London, 6. Auflage, 2004.
- [52] U. Tietze und C. Schenk. *Halbleiter-Schaltungstechnik*. Springer, Berlin, Heidelberg, 13. Auflage, 2010.
- [53] VDE/VDI 2183. *Meßumformer für Differenzdruck*, Oktober 1964.
- [54] B. Widrow. A study of rough amplitude quantization by means of Nyquist sampling theory. *IRE Transactions on Circuit Theory*, CT-3:266–276, Dec. 1956.
- [55] B. Widrow. Statistical analysis of amplitude quantized sampled data systems. *AIEE Transactions on Applications and Industry*, 81:555–568, Jan. 1961.
- [56] N. Wiener. *The Extrapolation, Interpolation and Smoothing of Stationary Time Series with Engineering Applications*. J. Wiley, New York, 1949.
- [57] H. Wolf. *Lineare Systeme und Netzwerke*. Springer, Berlin, Heidelberg, New York, 2. Auflage, 1989.

Index

- Abgleich, 56
- Abschirmung, 98
- Absolutdruckmesser, 90
- absoluter Fehler, 16
- Absolutwinkelgeber, 421
- abstandscodierte Scheibe, 422
- Abtastfrequenz, 319, 321
- Abtasttheorem, 318, 319
- Abtastung, 305, 317, 318
 - bei Drehzahlmessung, 397
 - winkelsynchrone ~, 390
 - winkelsynchrone ~, 391
 - zeitsynchrone ~, 391
 - zeitsynchrone ~, 392
- Addition unabhängiger Zufallsvariabler, 126
- Additivität, 182
- Aliasing, 319, 321
- Aliasing-Fehler, 322
- Allpass, 204
- Alternativhypothese, 155
- AM, *siehe* Amplitudenmodulation
- Ampere, 7
- amplitudenanaloges Signal, 14, 381
- Amplitudengang, 110
- Amplitudendichte, 336
 - der Fourier-Reihe, 337
 - harmonische Schwingung, 337
- Amplitudengang, 190
 - konstanter ~, 202
- Amplitudenmodulation, 14
- Analog-Digital-Umsetzer, 318, 357
 - Delta-Sigma-Umsetzer, 363
 - dual slope converter, 360
 - integrierender ~, 360
 - mit sukzessiver Approximation, 359
 - Nachlaufumsetzer, 357
 - verzögert nachlaufender ~, 357
 - Zweirampen-A/D-Umsetzer, 360
- Anpassungstest, 148, 157
- Anti-Aliasing-Filter, 322
- Anzeigeanfang, 55
- Anzeigebereich, 55
- Anzeigeende, 55
- Anzeigespanne
 - ideale ~, 56
- Anzeigewert
 - idealer ~, 56
- aperiodische Einstellung, 187
- aperiodischer Grenzfall, 184
- Approximation, 25
- Approximationsfehlerquadrate, 26, 31
- Aufnehmer, 11
- Ausfallrate, 164, 165
- Ausfallwahrscheinlichkeit, 165
- Ausgangsgleichung, 12
- Ausgangsgröße, 12
- Ausschussrate, 163
- äußere Störgrößen, 18
- Autokorrelationsfunktion, 252
 - einer PRBS-Folge, 293
 - eines ergodischen Prozesses, 258
 - von weißem Rauschen, 289
- Autokovarianzfunktion, 252
 - eines ergodischen Prozesses, 259
- Autoleistungsdichtespektrum, 286
- Bandbreite, 383
- Basis
 - orthonormale ~, 27
- Basiseinheit, 6
- Basisgröße, 6
- Bauelemente-Qualität, 164
- Bauelementestunden, 165
- Bayes-Theorem, 115, 116
- Beobachtungsfehler, 19
- Bernoulli'sches Gesetz der großen Zahlen, 134
- Beschleunigungssensor, 80
- Beschleunigungstreue, 200, 217
- Beschreibungsfunktion, 191
- Bessel'sche Ungleichung, 29
- Bildrestaurierung, 313
- bilineare Interpolation, 48
- binäre Impulsfolge, 363

- Binärfolge
 - pseudostochastische ~, *siehe*
 - PRBS-Folge
- Binomialverteilung, 158, 164
- Bode-Diagramm, 190
- Brückenschaltung, 68
- Butterworth-Filter, 204
- Candela, 7
- Charakteristische Funktion
 - quantisierter Signale, 341
- charakteristische Funktion, 125
- Chi-Quadrat-Anpassungstest, 157
- Chi-Quadrat-Verteilung, 142
- Closed-loop-Korrelation, 279
- Dämpfung, 183, 196, 225
- deformierende Störgrößen, 87, 90
- Dehnungsmessstreifen, 66
- Dekade, 191
- Delta-Sigma-Modulator, 363
- Delta-Sigma-Umsetzer, 363
 - Digitalfilter, 368
 - Modulator, 363
 - Rauschformung, 365
 - stationäres Verhalten, 369
- Determinante, 140
- deterministisches Signal, 110
- Dichte, *siehe* Wahrscheinlichkeitsdichte
- Differentialgleichung
 - lineare, 182
- Differenz
 - erste ~, 38
 - zweite ~, 39
- Differenzen, 38
- Differenzenoperator, 38
- Differenzenschema, 40
- Differenzmethode, 77
- Digital-Analog-Umsetzer, 318, 372
 - mit $R/2R$ -Kettenleiternetzwerk, 374
 - mit dyadisch gestuften Widerständen, 373
 - parallele ~, 373
 - serielle ~, 375
- digitale Messdatenerfassung, 317
- digitales Signal, 14
- Digitalfilter, 368
- Dirac-Impuls, 248
- direktes Messverfahren, 10
- Dithering, 352, 353
- DMS, *siehe* Dehnungsmessstreifen
- dominante Pole, 227
- Doppler-Effekt, 282
- Dopplerfrequenzmessung, 281
- Drehrichtungserkennung, 418
- Drehzahlmessung, 389
- Dreieckfunktion, 47
- Druck-Messumformer, 85
- dual-slope converter, 360
- Dualcode, 421
- Durchflussmessung
 - Wirkdruckverfahren zur ~, 86
- dynamische Fehler, 19
- Effektivfrequenz, 384, 388
- Effektivwert, 339, 348, 388
- Effizienz, 130
- Eichen, 58
- Eigenvektor, 140
- Eigenwert, 141
- Eingangsgröße, 12
- Einheit, 5
- Einheitensystem, 6
 - kohärentes ~, 7
- Elementarereignis, 110, 245
- Elle, 6
- Ellipsoid, 140
- Empfindlichkeit, 55
 - ideale ~, 56
 - mittlere ~, 62
- Empfindlichkeitsdifferenz, 62
- empirisches relationales System, 3
- Endwertsatz, 192
- Energie, 262, 267
- Energiedichte
 - im Frequenzbereich, 268
 - im Zeitbereich, 268
- Energiedichtespektrum, 287, 295
- Energiesignal, 262, 295
 - Innenprodukt zweier ~e, 264

- Norm eines ~s, 264
Entropie, 4
Ergebnismenge eines Zufallsexperimentes, 110
Ergodizität, 257
schwache ~, 257
ERS, *siehe* empirisches relationales System
erste Differenz, 38
Erwartungstreue, 129
Erwartungswert, 117, 118, 250
des Quantisierungsrauschen, 346
quantisierter Signale, 343, 344
Extrapolation, 32
Exzess, 119, 133

Falschalarm, 155
Faltung mit Rechteckfenster, 326
farbiges Rauschen, 290
Feder-Masse-Dämpfer-System, 80, 195
Fehler, 16
 1. Art, 155
 2. Art, 155
 absoluter ~, 16
 Aliasing-~, 322
 dynamische ~, 186
 Jitter-~, 328
 Quantisierungs-~, 333
 relativer ~, 17
 systematischer ~, 18, 109, 129
 Winkel~, 405
 zufälliger ~, 18, 109
Fehlerfortpflanzung, 173
Fehlerfunktion, 150
Fehlergrenze, 175
Fehlerquadrate
 Minimierung der ~, 26
Fehlerursachen, 18
Fensterfunktion, 262, 268
Filter
 FIR-~, 368
 moving average ~, 169, 326, 368
 Optimal-~, 32
 Wiener-~, 306
Filter-Optimal, 306

FIR-Filter, 368
Fixpunktjustierung, 57
Flächenabtastung, 340, 400
FM, *siehe* Frequenzmodulation
Folgefrequenz, 382
Fourier-Reihe, 29
Frühausfälle, 166
Frequenz, 381
 Effektiv-~, 384, 388
 eines harmonischen Signals, 381
 eines impulsförmigen Signals, 382
 eines stochastischen Signals, 382
 Folge-~, 382
 Kreis-~, 381
 mittlere ~, 383
 Momentan-~, 381
frequenzanaloges Signal, 15, 381
Frequenzbegriff, 382
Frequenzgang, 183
Frequenzgangmessung, 188
Frequenzmessung, 391, 411
 relativer Quantisierungsfehler der ~, 393
Frequenzmodulation, 15
Frequenzregelkreis, 416
Frequenzvergleicher, 416
Frequenzverschiebung, 282
Fühler, 11
Funktionenraum, 263
 unitärer ~, 263, 264
Funktionensystem
 orthonormales ~, 26

Gammafunktion, 144
Gauß'sche Fehlerfunktion, 150, 151
Gauß'sches Fehlerfortpflanzungsgesetz, 154, 173, 176
Gegenkopplung, 83, 93, 194, 231
Geschwindigkeitstreue, 200, 217
Gesetz der großen Zahlen, 134
gleitender Mittelwert, 169, 326
Gray-Code, 421
Gütemaß, 26, 71, 214–216, 219, 223–225

Häufigkeitsverteilung, 127

- Hauptachse, 140
 Histogramm, 127
 höchstwertiges Bit, 359
 höhere Differenz, 38
 höhere Momente, 119
 quantisierter Signale, 344
 Homogenität, 182
 Homomorphismus, 3
 Hurwitz-Determinante, 185
 Hurwitz-Kriterium, 184
 Hypothese, 155
 Hysterese, 59
- ideale Abtastung, 318
 ideale Anzeigespanne, 56
 ideale Empfindlichkeit, 56
 idealer Anzeigewert, 56
 idealer Messwert, 56
 idealer Tiefpass, 320
 ideales Übertragungsverhalten, 189
 Identifikation, 278
 Impulsantwort, 198, 262, 295, 326
 eines Interpolationsfilters, 46
 Impulskorrelation, 267, 298
 Impulsreihe, 318
 indirektes Messverfahren, 10
 Inkrementalgeber, 389, 420
 Innenprodukt, 263
 zweier Energiesignale, 264
 zweier Leistungssignale, 264
 innere Störgrößen, 18
 Integration der Eingangsspannung, 325
 Interpolation, 25, 34
 bilineare \sim , 48
 Deutung der \sim , 46
 Kennfeld- \sim , 47
 Lagrange- \sim , 36
 Newton- \sim , 38
 Polynom- \sim , 34
 Spline- \sim , 41
 zweidimensionale \sim , 47
 Intervallskala, 4
 inverses Filter, 309
 Irrtumswahrscheinlichkeit, 149, 155
 ITAE-Kriterium, 194, 214
- Jitter, 328
 Justierung, 56
- Kardinalskala, 5
 Kausalität, 182
 Kelvin, 7
 Kennfeldinterpolation, 47
 Kennlinie
 ideale \sim , 55
 Kennlinienfehler, 58
 relativer \sim , 58, 59, 64
 Kennlinieninterpolation, 34
 Kilogramm, 6
 Kompaßierung, 351
 Kompensation, 18, 97, 194
 des Zeitverhaltens, 226
 mechanischer Zahnräderfehler, 405
 Kompensationsverfahren, 188
 Konfidenzintervall, 149, 154
 Konfidenzniveau, 166
 Konsistenz, 129
 konstanter Amplitudengang, 193
 konstanter Realteil, 194
 Korrelation
 Eigenschaften der \sim , 269
 eines Summensignals, 294
 Impuls- \sim , 267
 Kurzzeit- \sim , 272
 Maximum der \sim , 272
 von Energiesignalen, 267
 von Leistungssignalen, 265
 Korrelationsfunktion
 eines ergodischen Prozesses, 258
 Messung der \sim , 271
 Korrelationskoeffizient, 4, 122, 140, 273
 Korrelationslänge, 290
 Korrelationsmesstechnik, 271
 Kovarianz, 120, 121
 Kovarianzfunktion
 eines ergodischen Prozesses, 258
 Kovarianzmatrix, 140
 Kreisfrequenz, 381
 Kreuzkorrelation
 Differentiation, 387
 Differentiation der \sim , 280

- Kreuzkorrelationsfunktion, 253
Abschätzung der ~, 273
zweier ergodischer Prozesse, 258
- Kreuzkovarianzfunktion, 253
zweier ergodischer Prozesse, 258
- Kreuzleistungsdichtespektrum, 286
- Kronecker-Delta, 27
- Krümmung, 42, 68, 78
- Lageparameter, 4, 118
- Lagrange-Polynom, 36
- Laufzeitkorrelation, 277
- Laufzeitmessung, 272
- Least-Squares-Schätzer, 31
- Lebensdauer
mittlere ~, 165
- Leckeffekt, 320
- Leistung, 266, 286, 287
mittlere ~, 262, 287
- Leistungsdichtespektrum, 285
Eigenschaften des ~, 288
in LTI-Systemen, 295
von weißem Rauschen, 289
- Leistungssignal, 262
Innenprodukt zweier ~e, 264
Norm eines ~s, 264
- Lichtgeschwindigkeit, 6
- lineare Regression, 32
- Linearität, 182
- Lose, 60
- LS-Schätzer, 31, 171, 409
- LTI-System, 182
- MA-Filter, 169, 170, 326, 357, 368
- Marginalisierung, 115
- Maßeinheit, 5
- mathematisches relationales System, 3
- Median, 4, 5
- Merkmal, 3
begriffliches ~, 4
kategoriales ~, 4
komparatives ~, 4
metrisches ~, 5
nominales ~, 4
ordinales ~, 4
- qualitatives ~, 5
quantitatives ~, 5
Rangmerkmal, 4
- Messabweichung, 16
- Messanfang, 55
- Messbereich, 55
günstigster ~, 71
herabsetzen, 64
- Messen, 3, 5
- Messende, 55
- Messfehler, 16
- Messgröße, 5, 10, 11
- Messkennlinie, 13, 55
ideale ~, 55
Krümmung der ~, 78
stationäre ~, 13
- Messkette, 11, 98
- Messskala, 3, 4
- Messspanne, 56
- Messumformer, 11
- Messung
ratiometrische ~, 372
- Messunsicherheit, 149, 154
der Einzelmessung, 154
Stichprobenmittelwert, 152, 154
- Messvektor, 35
- Messverfahren
direktes ~, 10
indirektes ~, 10
- Messwert
idealer ~, 56
- Meter, 6, 8
- metrische Skala, 5
- metrisches Merkmal, 5
- Minimalphasensystem, 204
- Mittelung, 136
- Mittelwert, 4, 5, 118
 χ^2 -Verteilung, 146
arithmetischer ~, 5
gleitender ~, 169, 259, 326
- Mittelwertbildung
bei A/D-Umsetzung, 362
bei Drehzahlmessung, 395
bei endlicher Abtastdauer, 325

- Mittelwertsatz der Integralrechnung, 136
- mittlere Frequenz, 383
- mittlere Lebensdauer, 165
- mittlere Leistung, 287
- Modalwert, 4, 5
- Modellbildung, 11, 25, 184
- Modulation, 100, 282
 - AM, 14
 - FM, 15
 - PAM, 14
 - PCM, 14
 - PWM, 15
- Modus, *siehe* Modalwert
- Moivre-Laplace-Theorem, 158
- Mol, 7
- Moment, 118
 - der Impulsantwort, 193, 198, 199
 - einer Zufallsvariablen, 118
 - eines ergodischen Prozesses, 258
 - eines stochastischen Prozesses, 251
 - gemeinsames ~, 120
 - zentrales ~, 118
- Momentanfrequenz, 15, 381, 382
- Momentanleistung, 287
- MRS, *siehe* mathematisches relationales System
- MSB, *siehe* höchstwertiges Bit
- MTTF, *siehe* mittlere Lebensdauer
- Musterfunktion, 245–247, 250, 252, 256–258
- Nachlaufkorrelator, 279
- Nachlaufumsetzer, 357, 378
- Naturkonstanten, 8
- Newton-Polynom, 38
- Nickel-Thermometer, 70
- Noise shaping, 363, 365
- Nominalskala, 3
- Norm, 263
 - eines Energiesignals, 264
 - eines Leistungssignals, 264
- Normal, 5, 17
- Normalbedingungen, 56, 64
- Normalverteilung, 138, 139
- mehrdimensionale ~, 140
- Transformation der ~, 140
- Nullhypothese, 155
- Nullpunktfehler, 58, 64
- Nyquist-Band, 322
- Offsetfehler, 58
- optimale Quantisierung, 349
- Optimalfilter, 32, 305, 306
- Ordinalskala, 4
- Orthogonalität
 - stochastischer Prozesse, 254
 - von Funktionen, 28
 - von Vektoren, 27
- orthonormale Basis, 27
- Oszillographendämpfung, 206, 218
- P-Regler, 232, 237
- PAM, *siehe* Pulsamplitudenmodulation
- Parameteroptimierung, 193, 195, 304
- Parameterschätzung, 32
- Parametertest, 156
- Parseval'sches Theorem, 219, 225, 268, 282, 330
- PCM, *siehe* Pulscodemodulation
- Periodendauer, 381
- Periodendauermessung, 390
 - Quantisierungsrauschen, 398
 - relativer Quantisierungsfehler der ~, 393
- periodisches Signal, 284
- Periodogramm, 300
- Phase-locked loop, 411
- Phasengang, 190
- Phasenregelkreis, 411
- Phasenvergleicher, 412
- physikalische Messkennlinie, 13
- PI-Regler, 196, 236, 238
- Poisson'sches Theorem, 164
- Poisson-Verteilung, 165
- Polaritätskorrelator, 280
- Pole
 - dominante ~, 227
- Polynom, 34
 - kubisches ~, 43

- Lagrange-~, 36
Newton-~, 38
Spline-~, 43
Polynomansatz, 34
Polynominterpolation, 34
Positionsbestimmung, 420
PRBS-Folge, 291–293, 298, 304
Prozessbrauchbarkeitsindex, 162
Prozessfähigkeitsindex, 162
Pseudoinverse, 31, 172
pseudostochastische Binärfolge, *siehe*
 PRBS-Folge, *siehe* PRBS-Folge
PT₁-Verhalten, 190
PT₂-Verhalten, 191
Pulsamplitudenmodulation, 14
Pulscodemodulation, 14
Pulsweitenmodulation, 15
PWM, *siehe* Pulsweitenmodulation
- quadratisches Fehlerintegral, 194, 219
Quantil, 4
Quantisierung, 14, 305, 318, 333
 nichtlineare Kennlinie, 349
 optimale ~, 349
 relativer Quantisierungsfehler, 349
Quantisierungsfehler, 333
 Gleichverteilung, 334, 345
 Leistungsdichte bei
 Periodendauermessung, 402
Quantisierungsfrequenz, 341, 342
Quantisierungsmodell, 346
Quantisierungrauschen, 335
Quantisierungstheorem, 340, 341
- Rückwirkung, 103
Raffungsfaktor, 167
Randdichte, 114, 115
Rangkorrelation, 4
Rate-Multiplier, 375
ratiometrische Messung, 372
Rauschabstand, *siehe*
 Signal-Rausch-Verhältnis
Rauschen, 18, 247, 288
 farbiges ~, 290, 298
 weißes ~, 289
- weißes Gauß'sches ~, 289
Rauschformung, 363, 366, 367
Rauschunterdrückung, 309
Realteil des Frequenzganges, 208
Rechteckfunktion, 47
Regression
 lineare ~, 32
Regressionsanalyse, 25, 32
Regressionsgerade, 33
Reibung, 60
Rekonstruktion, 304, 320, 342
relativer Fehler, 17
relativer Kennlinienfehler, 58, 59, 64
relativer Quantisierungsfehler
 der Frequenzmessung, 394
 der Periodendauermessung, 393
Rückwirkung, 19
- Satz von Bayes, *siehe* Bayes-Theorem
Scharmittelwert, 250, 251
Schätzer, 129
Schätzfunktion, *siehe* Schätzer
Schätzung
 der Korrelation, 272
 der Periodendauer, 401
 der Zahnflankenfehler, 407
 des Leistungsdichtespektrums, 300
 effiziente ~, 130
 erwartungstreue ~, 129
 konsistente ~, 129
Schiefe, 119, 133
schwache Ergodizität, 257
Schwarz'sche Ungleichung, 123, 264,
 266, 267
Schwerpunkt, 118
schwingende Einstellung, 186
Sekunde, 7
Sensitivität, *siehe* Empfindlichkeit
Sensor, 11
Sheppard-Korrektur, 345
SI-System, 6
Signal, 181
 Ähnlichkeit zweier ~, 272
 amplitudanaloges ~, 14, 317, 381
 digitales ~, 14

- Distanz zweier ~, 272
- Energie~, 262
- frequenzanaloges ~, 15, 381
- harmonisches ~, 14, 381
- impulsförmiges ~, 14, 382
- Leistung eines ~, 266, 287
- Leistungs~, 262
- periodisches ~, 284
- Rekonstruktion eines ~, 305
- stochastisches ~, 382
- zeitanaloges ~, 15
- Signal-Rausch-Verhältnis
 - Aliasing, 323, 325
 - Delta-Sigma-Modulator, 366
 - infolge der Quantisierung, 347
 - Jitter, 331
 - Kompandierung, 351
 - Signalenergie, 267, 268
 - Signalklassen, 261
 - Signalleistung, 266
 - Signalmodell, 32, 171, 172
 - Signalrekonstruktion, 305
 - Signalschätzung, 305
 - Signalverarbeitung, 11, 317
 - Signifikanzniveau, 155
 - Signifikanztest, 148, 156
 - Sinc-Funktion, 47, 326, 355, 395, 397
 - Skalenniveau, 3
 - Skalentyp, 3
 - SNR, *siehe* Signal-Rausch-Verhältnis
 - sonstige Signale, 263
 - spannungsgesteuerter Oszillator, 281
 - Spannungsmessung, 105
 - spektrale Überlappungen, 321
 - Spektrum, 47, 101
 - Ähnlichkeit zweier ~, 281
 - Distanz zweier ~, 282
 - Leistungsdichte~, 285
 - Spezifikation, 20
 - Spline
 - Interpolation, 41
 - kubischer ~, 43
 - Sprungantwort, 183, 186
 - Kennwerte der ~, 186
- Stabilität, 184
- Stabilitätskriterium, 184
 - Hurwitz, 184
- Standardabweichung, 5, 119, 131
- Standardnormalverteilung, 139, 143
- Stationarität, 13, 254
 - schwache ~, 255
 - strenge ~, 254
 - verbundene ~, 255
- statische Fehler, 20
- Statistik, 148
 - Statistik n -ter Ordnung, 117
 - statistische Prozessüberwachung, 169
 - statistische Sicherheit, 149
 - statistische Unabhängigkeit, 115, 290
 - eines stochastischen Prozesses, 249
 - statistischer Test, 148
- Stichprobe, 127
- Stichprobenmittelwert, 128
- Stichprobenvarianz, 131
- stochastische Prozesse, 245
- stochastischer Prozess, 245
 - ergodischer ~, 256
 - schwach ergodischer ~, 257
 - schwach stationärer ~, 255
 - stationärer ~, 254
- Störabstand, *siehe*
 - Signal-Rausch-Verhältnis
- Störgroße, 12
- Störgrößen
 - deformierende ~, 90
 - in Messketten, 98
 - superponierende ~, 87, 89
- Störgroßenvektor, 13, 64
- Störungsunterdrückung, 136
- strenge Stationarität, 254
- Streuung, 5
- Streuungsmaß, 4, 119
- Strommessung, 105
- Strukturänderung, 194, 226
- Student-t-Verteilung, 147
- superponierende Störgrößen, 87
- Superpositionsprinzip, 182
- System, 181

- Eigenschaften, 181
kausales ~, 182
Klassifikation, 181
lineares ~, 182
nichtlineares ~, 188
schwingungsfähiges ~, 184
zeitinvariantes ~, 182
systematischer Fehler, 109, 129
Systemidentifikation, 298, 309
Systemparameter, 183, 192
Optimierung der ~, 193
- t-Verteilung, 147, 153
Taylor-Reihe, 48, 63
Test
statistischer ~, 155
Testsignal, 184
Testverfahren
Anpassungstest, 157
Signifikanztest, 156
Theorem von Bayes, *siehe*
Bayes-Theorem
Thermostatisierung, 98, 99
Toleranzbandjustierung, 57
Toleranzfeld, 151, 162
Toleranzgrenze, 161
Totzeit, 170, 181–183, 187, 226
Transinformation, 4
Tschebyscheff'sche Ungleichung, 135
- Überabtastfaktor, 366
Übertragungsfunktion, 183
ideale ~, 193
Übertragungsverhalten
ideales ~, 189, 191
Umkehrspanne, 60
Ungleichung
Bessel'sche ~, 29
Schwarz'sche ~, 123, 264
Tschebyscheff'sche ~, 135
Unkorreliertheit, 121, 290, 294
stochastischer Prozesse, 254
Unschärferelation, 343
Urmeter, 8, 17
- Vandermonde-Matrix, 35
Varianz, 4, 119
 χ^2 -Verteilung, 146
des Quantisierungsrauschens, 346
quantisierter Signale, 343, 344
Variationskoeffizient, 4
VCO, *siehe* voltage-controlled oscillator
Verbunddichte, *siehe*
Verbundwahrscheinlichkeitsdichte
Verbundwahrscheinlichkeitsdichte, 114
vereinfachtes Modell, 18
Verhältnisskala, 5
Verstärkungsprinzip, 182
Verteilung, *siehe*
Wahrscheinlichkeitsverteilung
Verteilungsdichte, *siehe*
Wahrscheinlichkeitsdichte
Vertrauensintervall, 149
voltage-controlled oscillator, 281, 411
- wahrer Wert, 17
Wahrscheinlichkeit, 110
Wahrscheinlichkeitsdichte, 113
bedingte ~, 115
des Quantisierungsfehlers, 345
des Quantisierungsrauschens, 345
diskretes Zählergebnis, 400
eines stochastischen Prozesses, 248
gemeinsame ~, 114
quantisierter Signale, 340
- Wahrscheinlichkeitstheorie, 110
Wahrscheinlichkeitsverteilung, 112
eines stochastischen Prozesses, 247
eingipflige ~, 119
flachgipflige ~, 120
gemeinsame ~, 114
leptokurtische ~, 120
linksschiefe ~, 119
platykurtische ~, 120
rechtsschiefe ~, 119
steilgipflige ~, 120
symmetrische ~, 119
unimodale ~, 119, 121
- Walsh-Funktionen, 30
Wechselstrombrücke, 82

- Wegmessung, 66
mit DMS, 66
- weißes Gauß'sches Rauschen, 289
- weißes Rauschen, 288, 289
- Wendetangentenverfahren, 187
- Widerstandsthermometer, 68
- Wiener-Filter, 305, 306, 309
 akausales ~, 311
 kausales ~, 311
- Wiener-Khintchine-Theorem, 286
- Winkelfehler, 405
- winkelsynchrone Abtastung, 390, 391
- Wirkdruckverfahren
 zur Durchflussmessung, 86
- Wirksamkeit, *siehe* Effizienz
- Wölbung, 133
- zeitanaloges Signal, 15
- Zeitinvarianz, 182
- Zeitliche Abtastfehler, 328
- Zeitmittelwert, 250
- zeitsynchrone Abtastung, 391, 392
- Zentraler Grenzwertsatz, 338
- zentraler Grenzwertsatz, 141
- zentrales Moment, 118
- Zerhackerverstärker, 100
- zufälliger Fehler, 109
- zufälliges Signal, 110
- Zufallsexperiment, 110, 245
- Zufallsprozess, *siehe* stochastischer
 Prozess
- Zufallsvariable, 111, 246
- Zustandsgleichung, 12
- Zustandsgrößen, 12
- Zustandsraum, 12
- zweidimensionale Interpolation, 47
- Zweirampen-A/D-Umsetzer, 360
- zweite Differenz, 39