

Identifikation dynamischer Systeme

Eric Lenz

Skript – Wintersemester 2014/15



TECHNISCHE
UNIVERSITÄT
DARMSTADT

REGELUNGSTECHNIK
UND MECHATRONIK **rtm**

Identifikation dynamischer Systeme

Skript – Wintersemester 2014/15

Eric Lenz

REGELUNGSTECHNIK
UND MECHATRONIK **rtm**

Technische Universität Darmstadt
Institut für Automatisierungstechnik
Fachgebiet Regelungstechnik und Mechatronik
Prof. Dr.-Ing. U. Konigorski

Landgraf-Georg-Straße 4
64283 Darmstadt
Telefon 06151/16-3014
www.rtm.tu-darmstadt.de



Inhaltsverzeichnis

1 Einführung	7
1.1 Darstellungsformen dynamischer Systeme	7
1.1.1 Modellgleichungen	7
1.2 Einflussfaktoren	10
1.3 Ziele der Identifikation	12
1.4 Durchführung einer Systemidentifikation	14
1.5 Inhalt der Vorlesung	14
1.6 Beispielsystem	16
2 Kenngrößenermittlung	18
2.1 PT1-Glied	18
2.2 DT1-Glied	20
2.3 IT1-Glied	23
2.4 PT2-Glied	26
2.4.1 Schwingfähiges PT2-Glied	27
2.4.2 Nichtschwingfähiges PT2-Glied	33
2.5 PTn-Glied	34
2.6 Totzeitglied	36
3 Minimierung des Ausgangsfehlers	41
3.1 Ansatz	41
3.2 Beispiele	42
3.2.1 PT2-Glied	42
3.2.2 Drei-Massen-Schwinger	44
3.2.3 Doppelter Drei-Massen-Schwinger	47
3.3 Fazit	48
4 Grundlagen Stochastik	49
4.1 Zufallsvariablen	49
4.2 Ausgewählte Verteilungen	56
4.2.1 Normalverteilung	56
4.2.2 Gleichverteilung	58
4.2.3 Simulation und Messung von Zufallsvariablen	59
4.3 Zufallsprozesse	60
4.3.1 Korrelation und Leistungsdichtespektren	63
4.4 Übertragung stochastischer Eigenschaften bei linearen Systemen	67
4.4.1 Übertragung durch Gewichtsfunktion und Frequenzgang	67
4.4.2 Zustandsraumdarstellung	70
4.5 Stochastische Konvergenz	72
4.5.1 Konvergenzbegriffe	73
4.5.2 Eigenschaften von Schätzern	74
4.6 Anwendungen	75
4.6.1 Eigenschaften von Schätzern	75
4.6.2 Optimales gewichtetes Mittel	80

4.6.3	Quotientenbildung von Zufallsprozessen	84
5	Grundlagen Signalverarbeitung	91
5.1	Fourierreihe (FR)	91
5.2	Fouriertransformation (FT)	94
5.3	Laplacetransformation (LT)	102
5.4	Zeitdiskrete Fouriertransformation (DTFT)	103
5.5	Diskrete Fouriertransformation (DFT)	113
5.6	Berechnung der Transformationen	123
5.6.1	Fourierreihe und Fouriertransformation	123
5.6.2	DFT und DTFT	126
6	Signale	128
6.1	Signalleistung- und energie	128
6.2	Stochastische und deterministische Störungen	130
6.2.1	Rauschen	130
6.2.2	Konstante Störung	136
6.2.3	Rampenförmige Störung	138
6.2.4	Sinusförmige Störung	139
6.3	Anregungssignale	141
6.3.1	Signale zur Anregung einer Frequenz	141
6.3.2	Impulsförmige Anregungssignale	142
6.3.3	Breitbandige Anregungssignale	145
6.3.4	Kombination von Signalen	156
7	Nichtparametrische Identifikation	159
7.1	Frequenzgang	159
7.2	Messaufbau	160
7.3	Frequenzgangsschätzung über Amplitudenspektren (H0-Schätzung)	162
7.3.1	Ablesen aus dem Zeitschrieb	162
7.3.2	Schätzung über Fourierkoeffizienten	164
7.3.3	Schätzung über Fouriertransformierte	171
7.3.4	Einfluss von Störungen	182
7.3.5	Umgang mit periodischen Störungen	190
7.4	Frequenzgangsschätzung über Impuls-Leistungsdichten (H1- und H2-Schätzung)	192
7.4.1	H0-Schätzung	192
7.4.2	H1- und H2-Schätzung	195
7.5	Glättung des geschätzten Frequenzgangs	201
7.6	Mehrgrößensysteme	202
7.7	Weitere Verfahren zur nichtparametrischen Identifikation	205
8	Methode der kleinsten Quadrate	208
8.1	Beispiele	208
8.2	Allgemeine Formulierung	212
8.2.1	Prozess	212
8.2.2	LS-Problem	214
8.3	Lösung des LS-Problems	216
8.3.1	Schätzgleichung	216
8.3.2	Eigenschaften	220
8.3.3	Schätzung der Varianz des Messrauschen	226

8.4	Varianz der Schätzwerte	227
8.4.1	Monotonie der Varianz der Schätzwerte	228
8.4.2	Einflussfaktoren auf die Varianz	229
8.4.3	Konsistenz	235
8.5	Erweiterungen	238
8.5.1	Verfahren der gewichteten kleinsten Quadrate (Weighted Least Squares)	238
8.5.2	Exponentielles Vergessen	240
8.6	Iterative und rekursive Berechnung	240
8.6.1	Iterative Berechnung	240
8.6.2	Rekursives Least-Squares-Verfahren	243
8.7	Numerische Berechnung	245
8.7.1	Berechnung des Schätzwertes bei der Methode der Normalgleichungen	246
8.7.2	Orthogonalisierungsverfahren	248
8.7.3	Vergleich der Methoden	250
8.8	Diskussion von Anregung, Kondition und Varianz	255
9	Parametrische Identifikation dynamischer Systeme	258
9.1	Schätzung dynamischer Zusammenhänge aus Zeitbereichsdaten	258
9.1.1	Zeitdiskrete Übertragungsfunktionen	258
9.1.2	Stochastische Eigenschaften	261
9.1.3	Methode der Instrumentalvariablen	269
9.1.4	Wahl der Abtastzeit und der Anregungsgröße	276
9.1.5	Zeitvariante Systeme	283
9.1.6	Zeitkontinuierliche Übertragungsfunktionen	284
9.2	Schätzung von Übertragungsfunktionen aus Frequenzbereichsdaten	293
9.3	Mehrgrößensysteme	305
9.3.1	Systemmodelle	305
9.3.2	Identifikation aus Zeitbereichsdaten	307
9.3.3	Identifikation aus Frequenzbereichsdaten	309
9.4	Identifikation im geschlossenen Regelkreis	311
9.5	Anmerkungen zum Aufstellen von LS-Problemen	313
10	Kalmanfilter	317
10.1	Parameteridentifikation über Zustandsschätzung	317
10.2	Hinführendes Beispiel	318
10.3	Diskretes Kalmanfilter	323
10.3.1	Korrekturschritt	323
10.3.2	Prädiktion	326
10.3.3	Zusammenfassung	326
10.3.4	Beispiele	332
10.4	Kontinuierlich-diskretes Kalmanfilter	336
10.5	Erweitertes Kalmanfilter	338
10.5.1	Zeitdiskretes erweitertes Kalmanfilter	338
10.5.2	Kontinuierlich-diskretes erweitertes Kalmanfilter	339
10.6	Beispiel	340
A	Stochastik	343
A.1	Varianz Schätzung des quadratischen Mittels	343

B Signalverarbeitung	345
B.1 atan2	345
B.2 Verbindung zwischen der DFT und der Fourier-Reihe bei periodischen Signalen	346
B.3 Goertzel-Algorithmus	347
B.3.1 Goertzel-Algorithmus für nicht-ganzzahlige l	349
B.4 Nebenrechnungen Fourierreihe	350
C Signale	354
C.1 Zur Berechnung der Varianz bei Rauschen	354
D Methode der kleinsten Quadrate	357
D.1 Rekursives LS-Verfahren	357
D.1.1 Gewichtetes LS-Verfahren	357
D.1.2 LS-Verfahren mit exponentiellem Vergessen	358
D.2 QR-Zerlegung durch Householder-Transformationen	358
D.3 DSFI-Verfahren	360
E Parametrische Identifikation	366
E.1 Zeitkontinuierliches Butterworth-Filter	366
E.2 Alternatives Aufstellen der Matrizen bei Schätzung aus Frequenzbereichsdaten	366
E.3 Überlagerung von harmonischen Schwingungen	369
F Kalmanfilter	370
F.1 Betrachtung des Korrektur- und Prädiktionsschritt bezüglich des MSE	370
F.2 Parallele Messungen zusammen oder einzeln verarbeiten	372
Literatur	375

1 Einführung

In den bisherigen Vorlesungen zur Regelungstechnik ist zur Systemanalyse und dem Reglerentwurf immer das Modell des betrachteten Systems vorausgesetzt. In der Regel ist das System durch eine Übertragungsfunktion oder eine Zustandsraumdarstellung angegeben, für manche Entwurfsverfahren aus SDRT1 genügt auch die grafische Angabe des Frequenzganges.

Für einfache Systeme sind Übertragungsfunktionen und Zustandsraumdarstellungen dabei über bekannte physikalischen Zusammenhänge abgeleitet. Dieses Vorgehen wird als theoretische Modellbildung bezeichnet. Die notwendigen Kenntnisse dazu sind in den verschiedenen Grundlagenvorlesungen zur Elektrik, Mechanik und Strömungslehre vermittelt und die Vorlesung „Modellbildung und Simulation“ zeigt, wie diese genutzt werden können, um auch bei komplexeren Systemen auf methodische Weise für die Regelungstechnik nutzbare Modelle zu erhalten.

Dieses Vorgehen setzt aber voraus, dass ein vollständiges Wissen über das betrachtete System vorliegt, d. h. sowohl die physikalischen Zusammenhänge als auch deren Parameter müssen bekannt sein.

Diese Vorlesung „Identifikation dynamischer Systeme“ befasst sich dagegen mit der Aufgabe, ohne oder mit nicht vollständigem Wissen über das System aus Messdaten geeignete mathematischen Modelle realer Systeme abzuleiten.

1.1 Darstellungsformen dynamischer Systeme

Schwerpunkt dieser Vorlesung ist die Identifikation von linearen, zeitinvarianten Modellen. Damit werden hier auch nur Darstellungsformen für die Modellklasse der LTI-Systeme betrachtet.

Prinzipiell wird zwischen parametrischen und nicht-parametrischen Modellen unterschieden.

Unter nicht-parametrischen Modellen werden beispielsweise der an (vielen) einzelnen Frequenzstützstellen gegebene Frequenzgang oder der Zeitschrieb der Impuls- oder Sprungantwort verstanden. Also Darstellungsformen, bei denen das Systemverhalten punktweise, ohne inneren Zusammenhang gegeben ist.

Parametrische Modelle sind beispielsweise Übertragungsfunktionen oder Zustandsraumdarstellungen, bei denen das gesamte Systemverhalten durch eine endliche Anzahl von Parametern, z. B. den Koeffizienten des Zähler- und Nennerpolynoms der Übertragungsfunktion oder den Einträgen in Ein-, System- und Ausgangsmatrix, beschrieben ist. So ist auch die Impulsantwort eines System eine parametrische Darstellung, sofern diese nicht als eine Reihe von Zeit- und Ausgangswerten, sondern als Funktion gegeben ist, z. B. $g(t) = K \cdot e^{-\frac{t}{T}}$. Hier wären K und T die Parameter.

1.1.1 Modellgleichungen

In diesem Abschnitt werden die allgemeinen Gleichungen der verschiedenen Darstellungsformen, die im Weiteren verwendet werden, in Erinnerung gerufen.

Frequenzgang und Gewichtsfunktion

Zu den nicht-parametrischen Verfahren können prinzipbedingt keine allgemeinen Gleichungen angegeben werden. Daher werden hier beispielhaft die Frequenzgänge (Abbildung 1.1) und Gewichtsfunktionen bzw. Impulsantworten (Abbildung 1.2) für ein MIMO-System gezeigt.

Wesentlich ist, dass für jeden Übertragungspfad, d. h. jedes Paar aus Eingangsgröße u_j und Ausgangsgröße y_i ein einzelner Frequenzgang $H_{ij}(\omega)$ bzw. eine einzelne Gewichtsfunktion $g_{ij}(t)$ existiert und angegeben werden muss.

Den hier gezeigten Frequenzgängen (Abbildung 1.1) und Gewichtsfunktionen (Abbildung 1.2) gehören zu einem MIMO-System mit zwei Ein- und drei Ausgängen.

Die Frequenzgänge werden in der Regel als Bode-Diagramm, d. h. getrennt in Betrag und Phase dargestellt.

Übertragungsfunktion

Allgemein ist die Übertragungsfunktion eines LTI-SISO-Systems durch

$$G(s) = \frac{b_m s^m + b_{m-1} s^{m-1} + \dots + b_1 s + b_0}{s^n + a_{n-1} s^{n-1} + \dots + a_1 s + a_0}, \quad m \leq n,$$

bzw. bei Vorhandensein einer Totzeit T_t durch

$$G(s) = \frac{b_m s^m + b_{m-1} s^{m-1} + \dots + b_1 s + b_0}{s^n + a_{n-1} s^{n-1} + \dots + a_1 s + a_0} \cdot e^{-T_t s}, \quad m \leq n,$$

gegeben, wobei in aller Regel bei technischen Systemen auch $m < n$ gilt.

Im MIMO-Fall tritt eine Übertragungsmatrix

$$\mathbf{G}(s) = \begin{bmatrix} G_{11}(s) & \dots & G_{1p}(s) \\ \vdots & & \vdots \\ G_{q1}(s) & \dots & G_{qp}(s) \end{bmatrix}$$

anstelle einer einzelnen Übertragungsfunktion. Die Einträge dieser Übertragungsmatrix sind Übertragungsfunktionen $G_{ij}(s)$, die das Übertragungsverhalten vom j -ten Eingang auf den i -ten Ausgang beschreiben, $Y_i(s) = G_{ij}(s)U_j(s)$. Wenn das betrachtete System nicht völlig entkoppelte Teilsysteme beinhaltet, dann sind die Nenner der Übertragungsfunktion nicht unabhängig voneinander, sondern besitzen – bis auf ggf. durch Nullstellen kompensierte Pole – die gleichen Pole.

Im Zeitdiskreten ist die Übertragungsfunktion allgemein durch

$$G(z) = \frac{b_m + b_{m-1} z^{-1} + \dots + b_0 z^{-m}}{1 + a_{n-1} z^{-1} + \dots + a_0 z^{-n}} \cdot z^{-(n-m)}, \quad m \leq n,$$

gegeben, wobei $n - m$ die Totzeit in Abtastperioden ist. Wird ein diskretisiertes kontinuierliches System betrachtet und enthält dieses keine echten Totzeiten, dann gilt in aller Regel $n - m = 1$.

Zustandsraumdarstellung

Die Zustandsraumdarstellung lautet im Zeitkontinuierlichen

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t) \end{aligned}$$

und im Zeitdiskreten

$$\begin{aligned} \mathbf{x}_{k+1} &= \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k \\ \mathbf{y}_k &= \mathbf{C}\mathbf{x}_k + \mathbf{D}\mathbf{u}_k. \end{aligned}$$

Die Durchgriffsmatrix \mathbf{D} kann dabei für technische Systeme meist als Nullmatrix angesetzt werden.

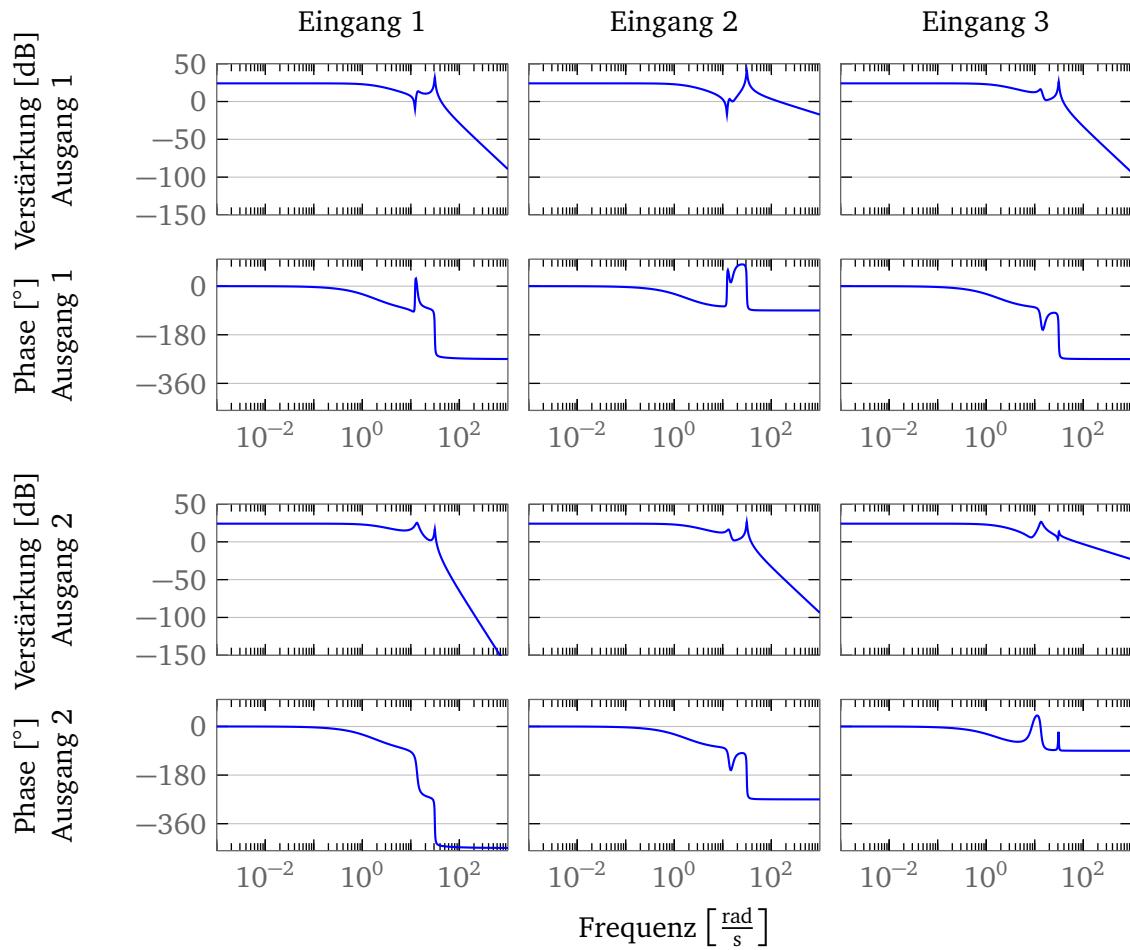


Abbildung 1.1: Frequenzgangdarstellung eines MIMO-Systems mit zwei Ein- und drei Ausgängen (Beispiel)

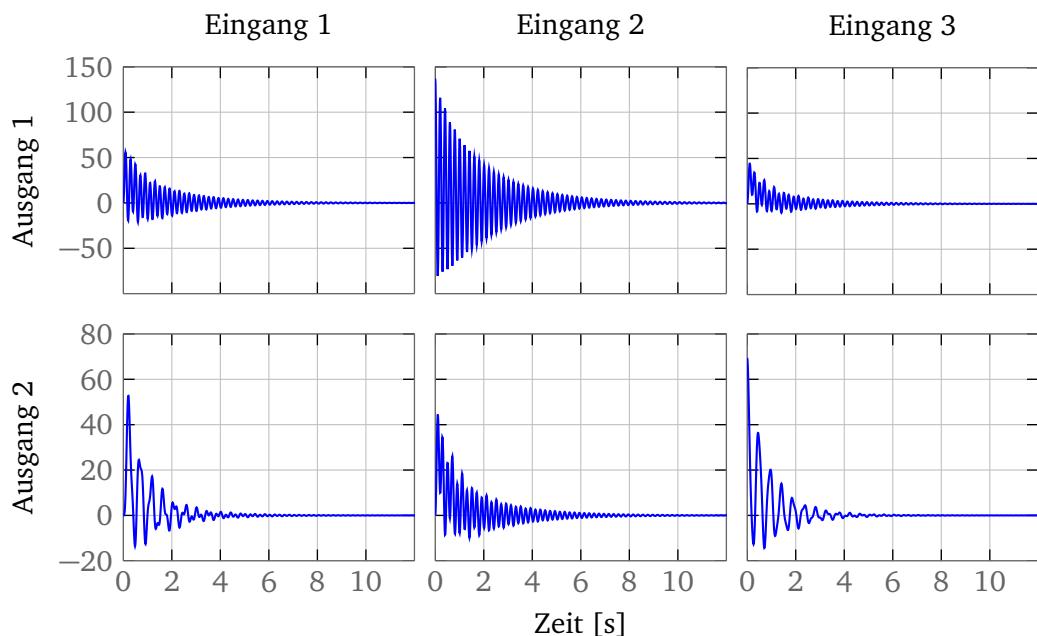


Abbildung 1.2: Gewichtsfunktionen (Impulsantworten) eines MIMO-Systems mit zwei Ein- und drei Ausgängen (Beispiel)

1.2 Einflussfaktoren

Bei der Wahl des Identifikationsverfahrens spielen viele Faktoren eine Rolle, die hier zunächst nur kurz angerissen werden. Diese Diskussion wird am Ende der Vorlesung wieder aufgenommen, wenn die einzelnen Verfahren bekannt sind.

Welche Systemklasse liegt vor?

Im Rahmen dieser Vorlesung wird im Wesentlichen die Identifikation von LTI-Systemen behandelt. Bei den parametrischen Verfahren werden dazu noch (langsame) Parameteränderungen und damit – mit Einschränkungen – LZV-Systeme zugelassen, sowie an verschiedenen Stellen auf die Identifikation von nichtlinearen Systemen eingegangen.

Ist das System nichtlinear, so stellt sich die Frage, ob es für die Identifikation an einzelnen Arbeitspunkten sinnvoll linearisiert werden kann.

Bei den LTI-Systemen ist noch zu unterscheiden, ob es sich um ein zeitkontinuierliches oder zeitdiskretes System handelt. Dies hängt neben dem System an sich auch von den realisierbaren Abtastzeiten ab, da die Identifikation – wenn von Analogrechnerimplementierungen abgesehen wird – immer auf Digitalrechnern erfolgt.

Welche Art an Störungen liegt vor?

In der Praxis werden immer Messfehlern und Prozessrauschen auftreten. In Abbildung 1.3 sind nur rauschartige bzw. externe Störungen bei der Messung des Eingangs- und Ausgangssignals sowie Prozessstörungen eingezeichnet. Die Identifikation kann naturgemäß nur auf den gemessenen Größen $\tilde{u}(t)$ und $\tilde{y}(t)$ beruhen, so dass diese in gewisser Weise robust gegenüber den Störungen sein muss.

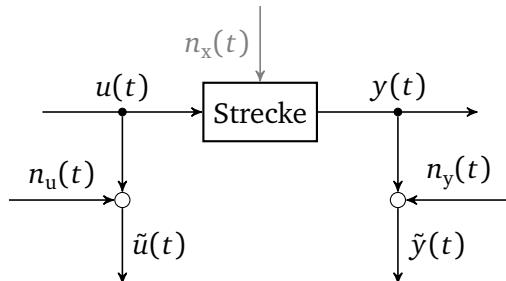


Abbildung 1.3: Messfehler

Dabei ist zu unterscheiden, ob die Störungen deterministischer oder stochastischer Natur sind, und welche Eigenschaften diese haben.

Werden die Fehler $n_u(t)$ der gemessenen Eingangsgröße explizit berücksichtigt, so spricht man auch von einem „Error-in-Variables“-Ansatz. In der Regel wird im Rahmen dieser Vorlesung davon ausgegangen, dass kein Fehler im Eingang vorliegt.

Auch wenn der Eingang nicht gemessen, sondern vorgegeben wird, kann dabei ein Fehler auftreten. So kann das Stellglied auch Rauschen oder Störungen aufweisen. Dies ist in Abbildung 1.4a dargestellt. Dort stellt $u'(t)$ den vorgegebenen Eingangsgrößenverlauf dar, der durch die Störung $n_a(t)$ verfälscht wird, so dass tatsächlich die Größe $u(t)$ auf das System wirkt. Dies kann dann aber wie in Abbildung 1.4b wieder dahingehend interpretiert werden, dass die vorgegebene Größe $u'(t)$ die Messgröße ist, die entsprechend um $-n_a(t)$ von der tatsächlichen Eingangsgröße abweicht.

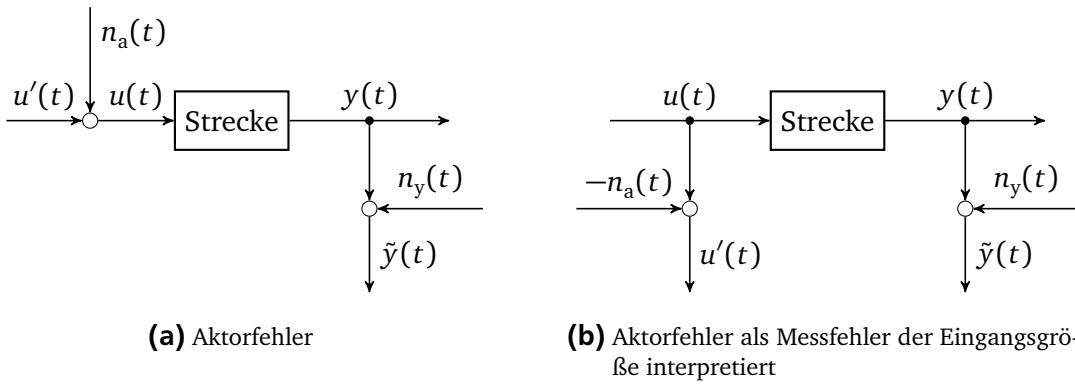


Abbildung 1.4: Aktorfehler

In den hier gezeigten Abbildungen Abbildung 1.3 und 1.4 sind nur rauschartige bzw. externe Störungen eingezeichnet. Daneben kann auch eine Dynamik des Messgliedes und Aktors vorliegen.

Randbedingungen der Versuchsdurchführungen

Zunächst stellt sich die Frage, ob wirklich gezielte Versuche durchgeführt werden können, oder ob die Identifikation ausschließlich auf Daten basieren muss, die im normalen Betrieb des betrachteten Prozesses anfallen. Zwischen den Extrema „völlige Freiheit bei der Wahl der Anregungssignale“ und „es muss mit Betriebsdaten gearbeitet werden“ existieren natürlich auch Zwischenstufen.

So ist es möglich, dass der Prozess weiter geregelt betrieben werden muss, man aber zusätzlich zum Regler eine additive Stellgröße vorgeben kann, oder dass man dem Regler in gewissem Rahmen freie Sollgrößenverläufe vorgeben kann. Dabei ist zu beachten, dass die Identifikation im geschlossenen Regelkreis aufgrund dessen, dass die Ausgangsgrößen über den Regler wieder auf den Eingang zurückwirken aus stochastischen Gründen kritisch sein kann.

Auch wenn die Stellgröße des Systems im Prinzip frei vorgegeben werden kann, so können Beschränkungen bezüglich maximaler Stellamplituden oder Änderungen der Stellgröße existieren. Diese können entweder aus Sicherheitsgründen oder durch eine begrenzte Aktordynamik gegeben sein.

Es ist auch zu klären, wie das Eingangssignal aufgebracht werden kann. Idealerweise besteht eine Schnittstelle zu einem Rechner und einem Messsystem, mit dem beliebige Eingangsgrößenverläufe vorgegeben werden können. Wenn eine solche Schnittstelle nicht existiert und das Eingangssignal per Hand eingestellt werden muss, so ergeben sich dabei natürlich Einschränkungen. So ist es schwierig, dem System manuell eine saubere Sinusschwingung oder einen Gleitsinus vorzugeben.

Neben dem Aufbringen von Eingangsgrößen müssen zumindest die Ausgangsgrößen (ggf. auch die tatsächlich wirksame Stellgröße) geeignet gemessen werden. Die möglichen Abtastraten und maximale speicherbare zusammenhängende Länge stellen dabei Bedingungen an das Identifikationsverfahren.

Zuletzt ist die Frage zu klären, wie viel Zeit für die Messungen zur Verfügung steht.

Einsatzzweck

Hierbei liegt das Hauptaugenmerk auf der zur Verfügung stehenden Rechenleistung zur Identifikation. Soll beispielsweise eine Identifikationsmethode auf einem Mikrocontroller implementiert werden, so ist dies mit erheblichen Einschränkungen bezüglich der Rechenzeit sowie der Genauigkeit der Zahlendarstellung verbunden, und macht eine sorgfältigere Betrachtung der numerischen Eigenschaften der Verfahren notwendig. Unabhängig von der Rechenleistung ist zwischen Verfahren zur Offline- und Online-Identifikation zu unterscheiden. Bei der Online-Identifikation sollen Schätzwerte möglichst kon-

tinuierlich, während der laufenden Messung bestimmt werden. Bei der Offline-Identifikation werden zunächst alle Daten gesammelt und die Messungen abgeschlossen, und dann anhand der gespeicherten Daten die Identifikation durchgeführt.

1.3 Ziele der Identifikation

Kurz gesagt ist das Ziel der Identifikation, zu einem gegebenen realen System ein Modell zu finden, dessen Ausgangsgrößenverlauf $\hat{y}(t)$ bei einem beliebigen Eingangsgrößenverlauf $u(t)$ immer dem Ausgangsgrößenverlauf $y(t)$ des realen Systems entspricht. Der Fehler $e(t)$ in Abbildung 1.5 sollte also identisch Null sein.

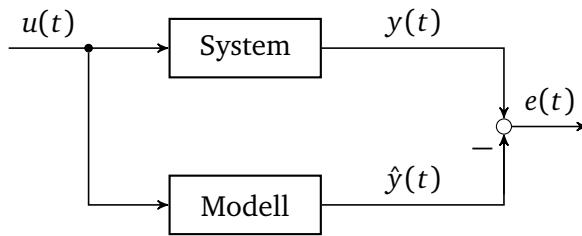


Abbildung 1.5: Ausgangsfehler (ohne Störungen)

Für LZI-Systeme kann dies auch im Frequenzbereich formuliert werden. Gesucht ist ein Systemmodell, dessen Frequenzgang $\hat{H}(\omega)$ genau dem tatsächlichen Frequenzgang $H(\omega)$ entspricht, d. h. der Fehler

$$E(\omega) = H(\omega) - \hat{H}(\omega)$$

sollte für alle Frequenzen ω verschwinden.

Dieses Ziel kann in der Praxis nicht erreicht werden. So sind zum einen Rauschen und Störungen zu beachten. Auch wenn eigentlich der wahre Systemausgang $y(t)$ für die Bewertung des identifizierten Modells relevant ist, so kann in der Praxis dafür nur der gemessene Ausgang $\tilde{y}(t)$ herangezogen werden, wie in Abbildung 1.6 dargestellt. (Das Eingangsmessrauschen ist dort vernachlässigt.)

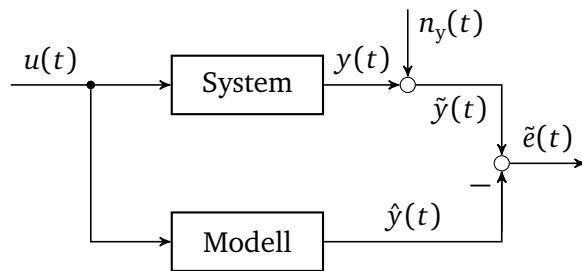


Abbildung 1.6: Ausgangsfehler

Darüber hinaus werden stets Annahmen über das System getroffen, d. h. es wird eine gewisse, weiter oder enger gefasste Modellklasse für das identifizierte Modell vorgegeben. So ist auch bei der nicht-parametrischen Identifikation, d. h. der Schätzung des Frequenzganges oder der Gewichtsfunktion, implizit schon die Annahme enthalten, dass es sich um ein LZI-System handelt. Bei der parametrischen Identifikation sind darüber hinaus noch weitere Annahmen, wie beispielsweise die Systemordnung, zu tätigen.

Zuletzt gibt es auch Gründe, nicht den Ausgangsfehler zu betrachten und das Modell daraufhin zu optimieren, sondern die Identifikation anhand einer Struktur wie in Abbildung 1.7 dargestellt durchzuführen. Dies wird im Kapitel zur parametrischen Identifikation wieder aufgenommen. Dort zeigt sich, dass

bei der Schätzung von Übertragungsfunktionen von LZI-Systemen der sogenannte Gleichungsfehler auf ein quadratisches Optimierungsproblem führt, dessen Lösung direkt angegeben werden kann.

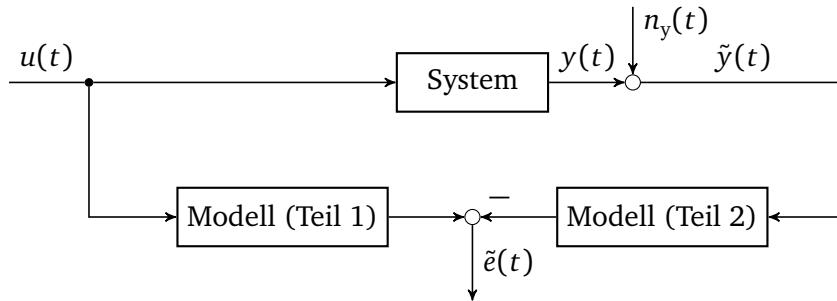


Abbildung 1.7: Gleichungsfehler

Überlegungen zur Wahl der zu identifizierenden Darstellungsform

Aus der Zustandsraumdarstellung lässt sich einfach die Übertragungsfunktion, und aus einer Übertragungsfunktion der Frequenzgang ermitteln. Aus dieser Sichtweise stellt die Zustandsraumdarstellung, die im Gegensatz zur Übertragungsfunktion auch Informationen über innere Zusammenhänge des Systems beinhaltet, ein „höherwertiges“ Modell dar. Der Vorteil der Darstellung der inneren Zusammenhänge geht aber verloren, wenn die physikalische Bedeutung der einzelnen Zustände nicht mehr bekannt ist. Dies ist jedoch der Normalfall bei der Identifikation von Zustandsraummodellen.

Auf der anderen Seite haben die nicht-parametrischen Modelle den Vorteil, dass keinerlei Annahmen über die Modellstruktur (abgesehen von der LZI-Eigenschaft) getroffen werden müssen. So sollte auch bei der Ermittlung der Übertragungsfunktionen zumindest die Systemordnung und die Anzahl der Nullstellen bekannt sein. (Es zeigt sich leider, dass man nicht einfach eine relativ hohe Systemordnung ansetzen kann, und die überzähligen Koeffizienten automatisch als null identifiziert werden.) Die Darstellung des Frequenzgangs hat auch den Vorteil, dass – bei geeigneter Wahl der Messzeit und ggf. weiteren Maßnahmen – periodische Störungen den identifizierten Frequenzgang nur an den entsprechenden Stellen stark beeinflussen, während beispielsweise bei Übertragungsfunktionen jeder Koeffizient von allen Frequenzstützstellen beeinflusst wird.

Es können auch Verfahren der nicht-parametrischen und parametrischen Identifikation kombiniert werden. So kann zunächst der Frequenzgang identifiziert werden, und dann basierend auf den Frequenzbereichsdaten eine Übertragungsfunktion oder eine Zustandsraumdarstellung angepasst werden.

In Abbildung 1.8 sind die bisher genannten parametrischen und nicht-parametrischen Modellformen genannt und darüber hinaus die Identifikationswege dargestellt, die in dieser Vorlesung behandelt werden.

White-, Gray- und Black-Box-Modelle

In der Literatur zur Identifikation treten auch die Begriffe White-, Gray- und Black-Box-Modelle auf. Diese bewerten das betreffende Modell anhand des Grades, welche A-priori-Informationen zur Identifikation des Systems einfließen. White-Box-Modelle sind dabei Modelle, die vollständig aus der theoretischen Modellbildung entstehen, d. h. es sind die Struktur sowie die Parameterwerte bekannt. Bei Black-Box-Modellen ist kein Vorwissen vorhanden. Bei linearen Systemen entspricht dies den nicht-parametrischen Modellen. Innerhalb der Gray-Box-Modelle, die das Spektrum dazwischen abdecken, sind weitere Unterabstufungen möglich. So ist bei Modellen, die sich aus den Gleichungen bekannter physikalischer Zusammenhänge ergeben, aber die Parameterwerte nicht bekannt sind noch sehr viel Vorwissen vorhanden. Kennt man dagegen nur die Ordnung und die Anzahl der Nullstellen eines LZI-Systems, so steckt

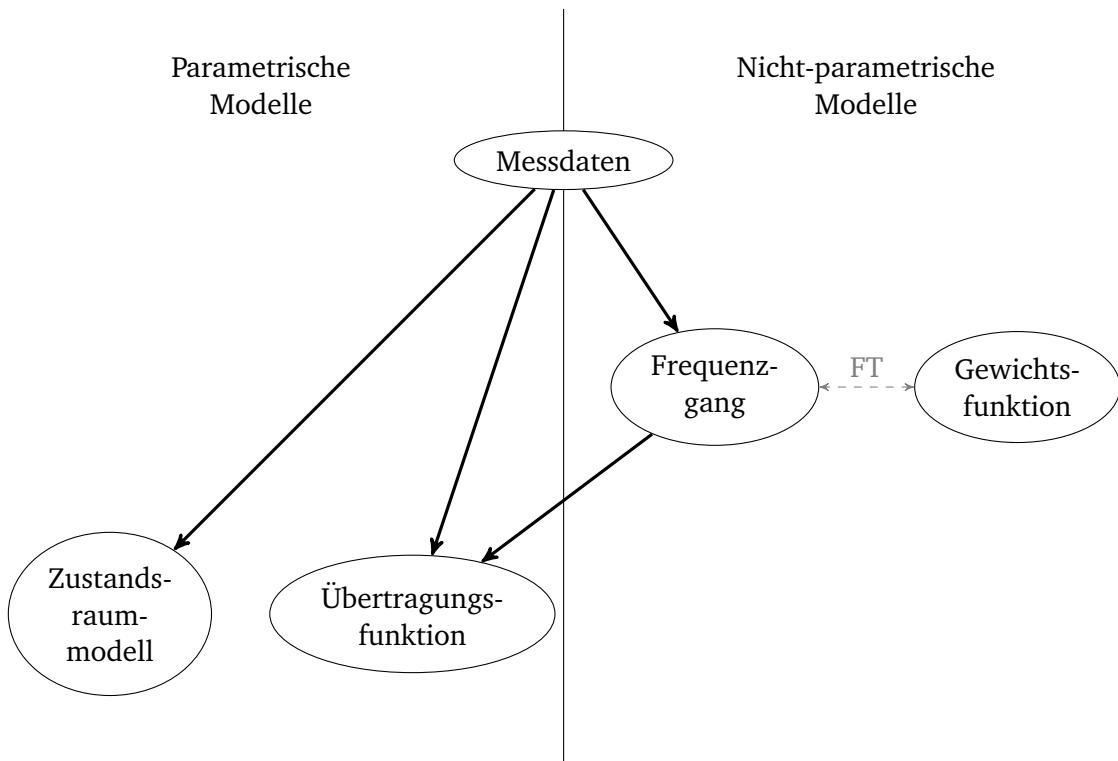


Abbildung 1.8: Parametrische und nicht-parametrische Modelle mit den in dieser Vorlesung behandelten Identifikationswegen

in diesem Vorwissen deutlich weniger Information. (Diese Abstufung zwischen Gray- und Black-Box-Modellen ist nicht einheitlich definiert.)

1.4 Durchführung einer Systemidentifikation

In Abbildung 1.9 ist ein prinzipieller, kompletter Ablauf einer Systemidentifikation dargestellt. In diesem finden sich viele Punkte wieder, die bis hier schon kurz angesprochen wurden. Wesentlich ist der iterative Charakter der Identifikation. D. h. das erhaltene Modell wird validiert, um dann entsprechende Anpassungen im Identifikationsablauf vorzunehmen, um in einem folgenden Schritt zu einem besseren Modell zu kommen.

1.5 Inhalt der Vorlesung

Im folgenden letzten Teil dieses Einführungskapitels wird ein Beispielsystem, welches des Öfteren herangezogen wird, vorgestellt.

In Kapitel 2 werden zunächst die elementaren Übertragungsglieder, die schon aus SDRT1 bekannt sind, wiederholt. Dabei wird darauf eingegangen, mit welchen Größen diese in der Regel beschrieben werden, und wie diese Größen anhand von Sprungantworten oder Frequenzgängen bestimmt werden können.

Daraufhin wird in Kapitel 3 die Minimierung des in Abbildung 1.6 dargestellten Ausgangsfehlers betrachtet. Da dies jedoch auf ein nichtlineares Optimierungsproblem führt, ist die praktische Umsetzung kritisch. Dies wird an ein paar Beispielen diskutiert.

Die weiteren in dieser Vorlesung behandelten Verfahren werden entgegen der nichtlinearen Optimierung auch in theoretischer Hinsicht ausführlicher diskutiert. Die dazu notwendigen Grundlagen der Stochastik sowie der Signalverarbeitung werden in den Kapiteln 4 und 5 behandelt.

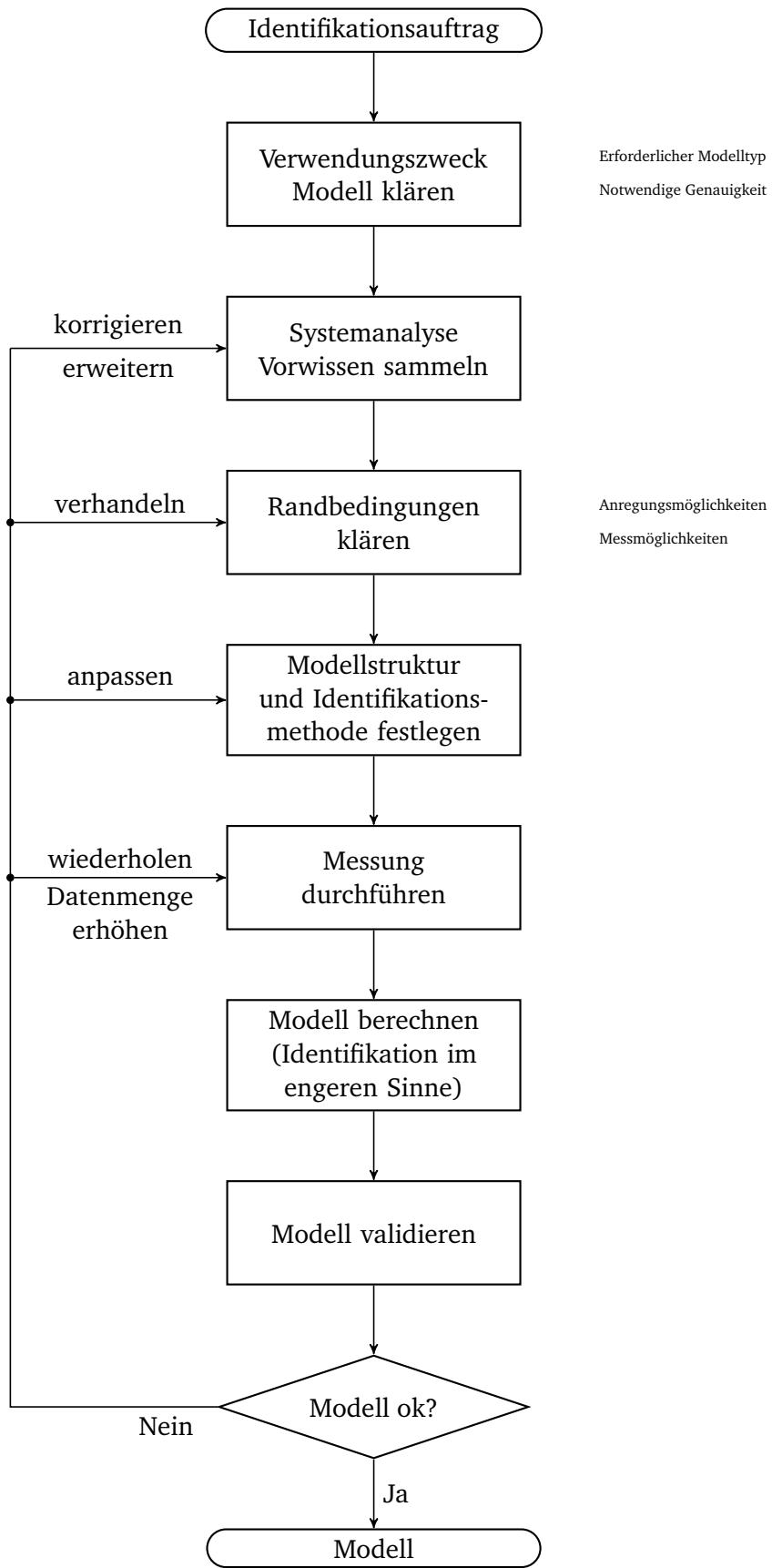


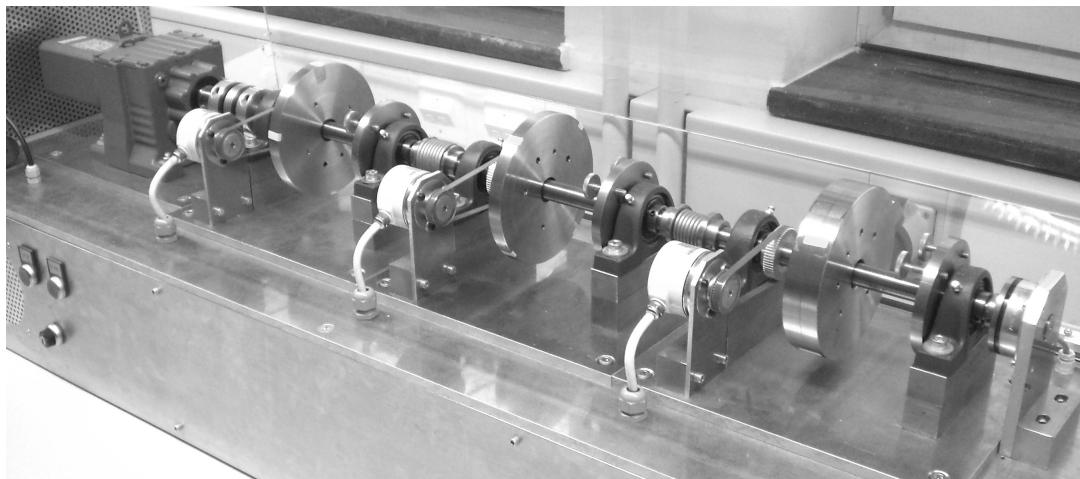
Abbildung 1.9: Ablauf einer Systemidentifikation

Dann werden zunächst die Verfahren zur nicht-parametrischen Identifikation besprochen, wobei hier die Verfahren zur Schätzung von Frequenzgängen zeitkontinuierlicher Systeme mit deterministischen Signalen im Mittelpunkt stehen.

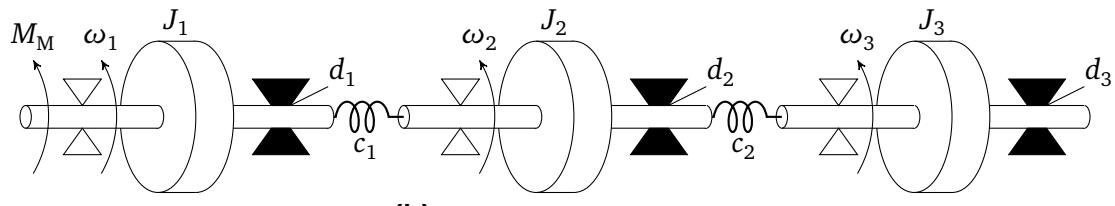
Die Behandlung der parametrischen Verfahren beginnt mit der Schätzung von zeitdiskreten und zeitkontinuierlichen Übertragungsfunktionen mit der Methode der kleinsten Quadrate. Des weiteren wird auf die Ermittlung von parametrischen Modellen aus Frequenzgängen eingegangen. Weitere behandelte Methoden sind das Kalman-Filter sowie Subspace-Methoden.

1.6 Beispielsystem

Als Beispielsystem dient der in Abbildung 1.10a gezeigte Drei-Massen-Schwinger. Dieser besteht aus drei Wellen mit je einer Trägheit, die über Drehfedern miteinander verbunden sind. Die erste Welle wird durch einen Motor angetrieben. Die Trägheit des Motors wird (entsprechend mit der Getriebeübersetzung auf die erste Welle umgerechnet) der Trägheit der ersten Welle zugeschlagen. Der Motor wird im Momentenbetriebsmodus betrieben und es wird davon ausgegangen, dass die Strom- bzw. Momentenregelung so schnell ist, dass deren Dynamik gegenüber den dominanten, mechanischen Anteilen vernachlässigt werden kann. Als Eingangsgröße kann damit direkt das Motormoment M_M (wieder auf die erste Welle bezogen) vorgegeben werden. Für jede Welle ist noch die Lagerdämpfung zu berücksichtigen. In Abbildung 1.10b ist die damit entstehende schematische Darstellung des Drei-Massen-Schwingers gezeigt.



(a) Realer Aufbau



(b) Schematische Darstellung

Abbildung 1.10: Beispielsystem Drei-Massen-Schwinger

Wird der Zustandsvektor zu

$$\mathbf{x}^T = [\varphi_1 \quad \varphi_2 \quad \varphi_3 \quad \omega_1 \quad \omega_2 \quad \omega_3]$$

Tabelle 1.1: Parameter des Drei-Massen-Schwingers

$J_1 = 1,4 \cdot 10^{-2} \text{ kg m}^2$	$d_1 = 5,6 \cdot 10^{-2} \text{ Nm/(rad/s)}$	$c_1 = 3,44 \text{ Nm/rad}$
$J_2 = 7,3 \cdot 10^{-3} \text{ kg m}^2$	$d_2 = 2,5 \cdot 10^{-3} \text{ Nm/(rad/s)}$	$c_2 = 2,14 \text{ Nm/rad}$
$J_3 = 1,44 \cdot 10^{-2} \text{ kg m}^2$	$d_3 = 4,5 \cdot 10^{-3} \text{ Nm/(rad/s)}$	

festgelegt, so lautet die Zustandsmatrix

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ -\frac{c_1}{J_1} & \frac{c_1}{J_1} & 0 & -\frac{d_1}{J_1} & 0 & 0 \\ \frac{c_1}{J_2} & -\frac{c_1+c_2}{J_2} & \frac{c_2}{J_2} & 0 & -\frac{d_2}{J_2} & 0 \\ 0 & -\frac{c_2}{J_3} & -\frac{c_2}{J_3} & 0 & 0 & -\frac{d_3}{J_3} \end{bmatrix}.$$

Für das Motormoment als Eingangsgröße ergibt sich

$$\mathbf{B}^T = \begin{bmatrix} 0 & 0 & 0 & \frac{1}{J_1} & 0 & 0 \end{bmatrix}.$$

Wird die Winkelgeschwindigkeit ω_3 der dritten Masse als Ausgangsgröße gewählt, so gilt

$$\mathbf{C} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Mit den in Tabelle 1.1 aufgelisteten Werten für die Parameter ergibt sich die Übertragungsfunktion

$$G(s) = \frac{\mathcal{L}(\omega_3(t))|_s}{\mathcal{L}(M_M(t))|_s} = \frac{5,0022 \cdot 10^6}{s^5 + 4,6550 \cdot s^4 + 1161 \cdot s^3 + 4103 \cdot s^2 + 1,7976 \cdot 10^5 \cdot s^1 + 3,1514 \cdot 10^5}. \quad (1.1)$$

Die genaue Abhängigkeit der Koeffizienten b_0 und a_0, \dots, a_4 von den physikalischen Parametern ist etwas länglich. Wesentlich ist aber sich hier bewusst zu machen, dass die Übertragungsfunktion durch sechs Koeffizienten bestimmt wird, aber acht physikalische Parameter vorliegen. D. h. aus der Übertragungsfunktion, bzw. allgemeiner aus allen Darstellungen, die lediglich das Ein-/Ausgangsverhalten dieses einen Übertragungspfades beschreiben, wie z. B. auch der Frequenzgang, können von vorneherein nicht alle physikalischen Parameter eindeutig bestimmt werden.

2 Kenngrößenermittlung

In diesem Kapitel werden die grundlegenden dynamischen Übertragungsglieder wiederholt und darauf eingegangen, wie man die Kenngrößen, die diese beschreiben aus Messungen von z. B. Sprungantworten oder anhand des Frequenzgangs bestimmen kann. Methoden zur messtechnischen Ermittlung des Frequenzgangs eines Systems werden später vorgestellt.

Auf die Herleitungen der Übertragungsfunktionen und der Zeitbereichslösungen wird nicht detailliert eingegangen sondern auf SDRT1 verwiesen.

2.1 PT₁-Glied

Beispiele

Das erste Beispiel eines Systems mit PT₁-Verhalten ist das RL-Glied, welches in Abbildung 2.1a dargestellt ist. Die Eingangsgröße ist hier die Spannung $u(t)$ zwischen den linken Klemmen, Ausgangsgröße der Strom $i(t)$. Dieses System wird durch die Differentialgleichung

$$\frac{di(t)}{dt} = -\frac{R}{L}i(t) + \frac{1}{L}u(t)$$

beschrieben, woraus die Übertragungsfunktion

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

folgt.

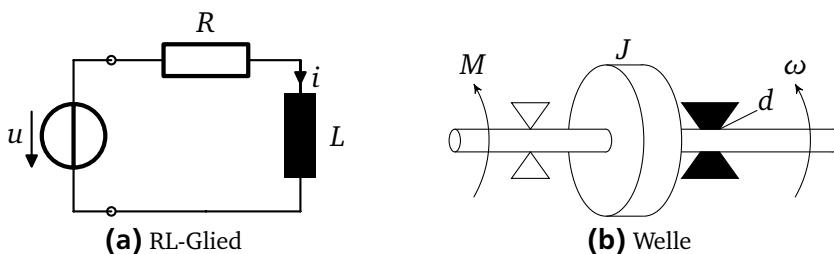


Abbildung 2.1: Beispiele für PT₁-Glieder

Ein weiteres Beispiel ist die Welle mit der Massenträgheitsmoment J aus Abbildung 2.1b mit dem Antriebsmoment $M(t)$ als Eingangsgröße und der Winkelgeschwindigkeit $\omega(t)$ als Ausgangsgröße. Es gilt

$$J \cdot \dot{\omega}(t) = -d \cdot \omega(t) + M(t)$$

bzw.

$$G(s) = \frac{1}{Js + d} .$$

Übertragungsfunktion

Die Übertragungsfunktion eines PT₁-Gliedes lautet

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

(Bzw. kann diese immer in die angegebene Form gebracht werden.) Die beiden zu bestimmenden Parameter sind die stationäre Verstärkung K sowie die Zeitkonstante T .

Das PT₁-Glied besitzt einen Pol bei

$$\lambda = -\frac{1}{T}.$$

Bestimmung der Parameter aus der Sprungantwort

Die Zeitlösung der Sprungantwort ($u(t) = u_0 \cdot \sigma(t)$) lautet

$$y(t) = K \cdot \left(1 - e^{-\frac{1}{T} \cdot t}\right) \cdot u_0.$$

Für $t \rightarrow \infty$ gilt $y_\infty = K \cdot u_0$, so dass der Parameter K einfach anhand des Endwerts der Sprungantwort abgelesen werden kann, siehe Abbildung 2.2.

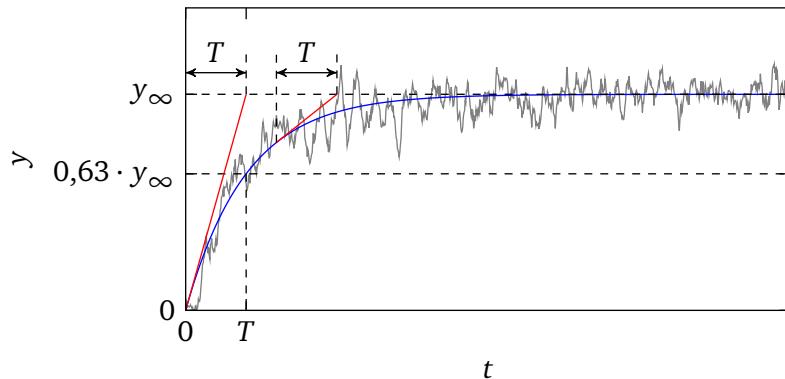


Abbildung 2.2: Sprungantwort PT1-Glied

Für die Ableitung der Sprungantwort gilt

$$\dot{y}(t) = \frac{K \cdot u_0}{T} \cdot e^{-\frac{1}{T} \cdot t} = \frac{y_\infty}{T} \cdot e^{-\frac{1}{T} \cdot t}$$

und es kann daher auch

$$T \cdot \dot{y}(t) = y_\infty - y(t)$$

geschrieben werden. Damit entspricht die Zeitkonstante T dem Zeitabstand eines Punktes auf der Sprungantwort zu dem Schnittpunkt der durch diesen Punkt gehenden Tangente mit dem Endwert y_∞ . Insbesondere kann damit T direkt als die Zeit abgelesen werden, an der die Tangente um Ursprung den Endwert y_∞ schneidet. Aufgrund von Rauschen ist es aber häufig kritisch, eine Tangente sauber anzulegen, so dass es günstiger ist, die Zeitkonstante an der Stelle $y(T) = y_\infty \cdot e^{-1} = 0,63 \cdot y_\infty$ abzulesen.

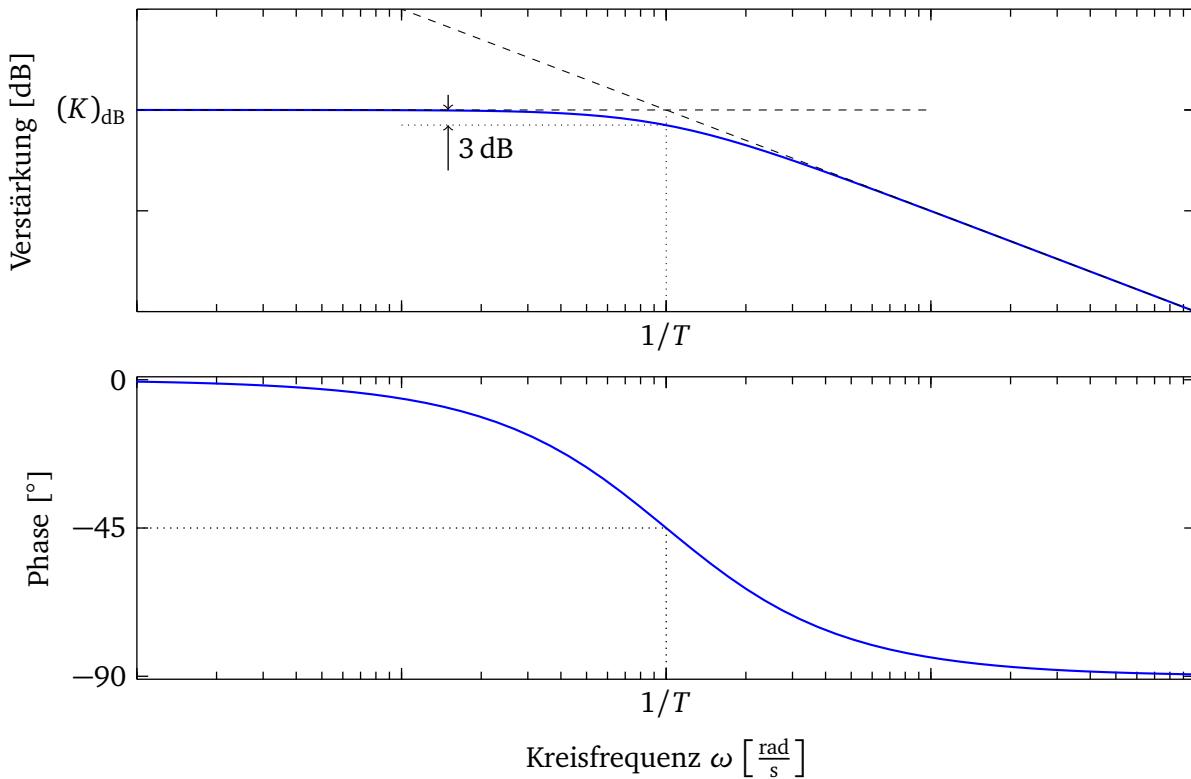


Abbildung 2.3: Frequenzgang PT₁-Glied

Bestimmung der Parameter aus dem Frequenzgang

In Abbildung 2.3 ist der Frequenzgang eines PT₁-Gliedes als Bodediagramm dargestellt. In diesem Diagramm sind auch die wesentlichen Kenngrößen eingezeichnet.

Die Verstärkung K kann für $\omega = 0$ aus dem Frequenzgang abgelesen werden. Der Kehrwert der Zeitkonstante T ist der Frequenzpunkt, an dem sich die beiden Asymptoten schneiden. Auch ist $1/T$ die Frequenz, bei der die Amplitude um $1/\sqrt{2}$ (-3 dB) gegenüber dem stationären Wert abgesunken ist, bzw. bei der die Phase -45° beträgt.

2.2 DT1-Glied

Beispiele

Als Beispiel für ein DT1-Glied kann wieder die RL-Reihenschaltung dienen, wenn nicht der Strom $i(t)$, sondern die Spulenspannung $u_L(t)$ als Ausgangsgröße gewählt wird (Abbildung 2.4a). Die Eingangsgröße ist weiterhin die Spannung $u(t)$. Dieses System wird durch die Differentialgleichung

$$\dot{u}_L(t) + \frac{R}{L}u_L(t) = \dot{u}(t)$$

beschrieben, woraus die Übertragungsfunktion

$$G(s) = \frac{s}{s + \frac{R}{L}}$$

folgt.

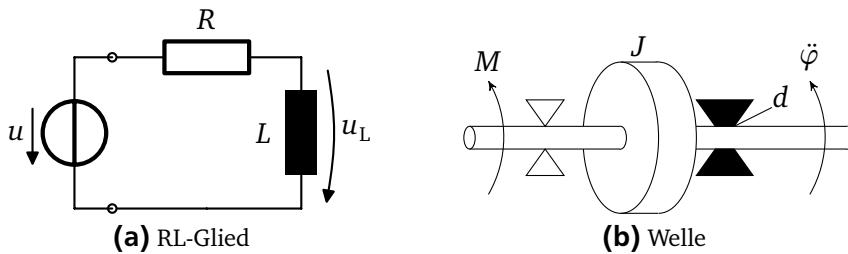


Abbildung 2.4: Beispiele für DT₁-Glieder

Betrachtet man bei der Welle nicht die Winkelgeschwindigkeit $\omega(t)$ sondern die Winkelbeschleunigung $\ddot{\varphi}(t)$ als Ausgangsgröße, so erhält man ebenfalls ein System mit DT₁-Verhalten. Die Differentialgleichung lautet dann

$$J \cdot \frac{d}{dt} \ddot{\varphi}(t) = -d \cdot \ddot{\varphi}(t) + M(t)$$

und die Übertragungsfunktion

$$G(s) = \frac{s}{Js + d}.$$

Übertragungsfunktion

Allgemein besitzt ein DT₁-Glied eine Übertragungsfunktion der Form

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

und ist durch die beiden Parameter K und T vollständig bestimmt.

Bestimmung der Parameter aus der Sprungantwort

Die Sprungantwort lautet

$$y(t) = K \cdot \frac{1}{T} \cdot e^{-\frac{t}{T}} \cdot u_0.$$

Damit kann der Parameter K aus der Sprunghöhe $y(0) = \frac{K}{T} \cdot u_0$ abgelesen werden, wenn T bekannt ist. Die Zeitkonstante T kann, wie in Abbildung 2.5a dargestellt, wieder über das Anlegen einer Tangente bestimmt werden. Bezuglich Rauschens kann es aber wieder etwas sicherer sein, die Zeitkonstante T an der Stelle abzulesen, bei der der Systemausgang noch gerade 36,8% des Startwertes y_0 beträgt ($y(T) = y_0 \cdot e^{-1}$). Aufgrund dessen, dass zur Bestimmung von y_0 nur ein einzelner Punkt, nämlich $t = 0$ zur Verfügung steht, kann dieses Vorgehen bei vorhandenem Rauschen schnell problematisch werden. Zudem könnten eigentlich zu vernachlässigende, schnellere Dynamikanteile diesen Sprung glätten und damit den Wert verfälschen. Im Gegensatz dazu kann der Parameter K bei einem PT₁-Glied sehr gut geschätzt werden, da dabei eine Mittelung über viele Messwerte vorgenommen werden kann.

Wenn möglich, bietet es sich hier daher an, dass System über eine Rampe $u(t) = K_R t$ anzuregen. Damit erhält man

$$Y'(s) = K \cdot \frac{s}{Ts + 1} \cdot \frac{K_R}{s^2} = \frac{K \cdot K_R}{Ts + 1} \cdot \frac{1}{s}.$$

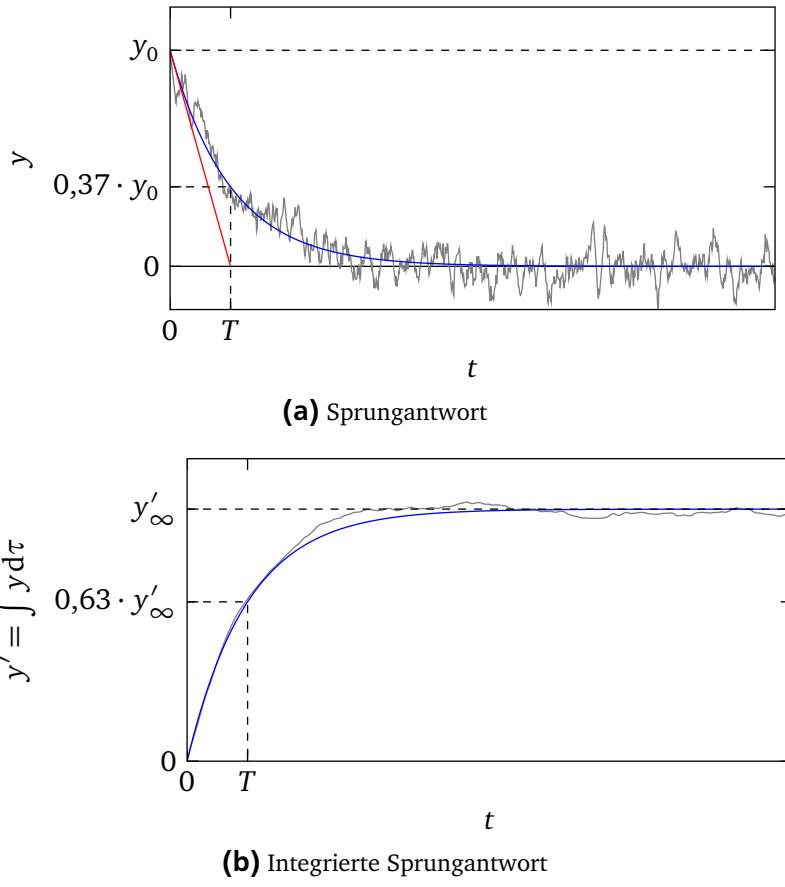


Abbildung 2.5: DT1-Glied

Dieses Signal entspricht der Antwort eines PT₁-Gliedes mit der stationären Verstärkung $K \cdot K_R$ auf einen Einheitssprung, und so kann man in diesem Fall das System als PT₁-Glied identifizieren.

Alternativ kann auch der Systemausgang integriert werden. Auch in diesem Fall kann ein (künstliches) PT₁-Glied

$$G'(s) = \frac{1}{s} \cdot G(s) = \frac{K}{Ts + 1}$$

identifiziert werden, dessen Parameter K und T dem des eigentlich betrachteten DT₁-Gliedes entsprechen. Die Integration hat auch den Vorteil, dass durch die Integration das Rauschen geglättet wird, was auch in Abbildung 2.5b deutlich wird.

Eine weitere Alternative besteht darin, zu versuchen eine andere Messgröße für die Kenngrößenermittlung zu wählen. Würde man in dem vorliegenden Beispiel nicht die Spannung über die Spule, sondern die Spannung über den Widerstand messen, so würde man wieder ein PT₁-Verhalten erhalten. In diesem Fall wäre es nicht einmal nötig, die Messung umzubauen, sondern man könnte einfach $u_R(t) = u(t) - u_L(t)$ verwenden.

Bestimmung der Parameter aus dem Frequenzgang

In Abbildung 2.6 ist das Bodediagramm eines DT₁-Gliedes dargestellt. Hier lässt sich K bei $\omega \rightarrow \infty$ einfach ablesen.¹ Die Zeitkonstante T ergibt sich wieder durch den Schnittpunkt der Asymptoten oder dem -45° -Phasenpunkt.

¹ Handelt es sich um einen messtechnisch ermittelten Frequenzgang, so nimmt der Anteil des Rauschens für hohe Frequenzen zu, so dass das Ablesen erschwert wird.

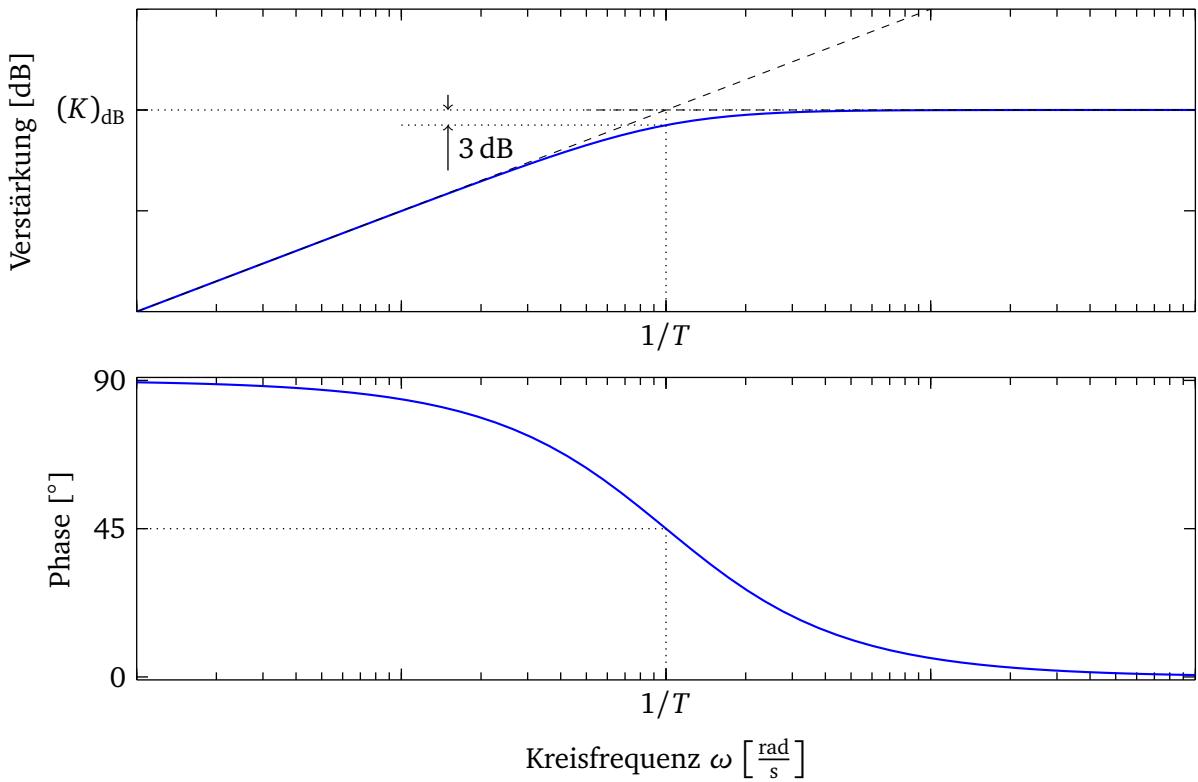


Abbildung 2.6: Frequenzgang DT1-Glied

2.3 IT1-Glied

Beispiel

Misst man beim Beispiel der Welle nicht die Winkelgeschwindigkeit oder -beschleunigung sondern den Winkel $\varphi(t)$ (Abbildung 2.7), erhält man

$$J \cdot \ddot{\varphi}(t) = -d\dot{\varphi}(t) + M(t)$$

bzw.

$$G(s) = \frac{1}{Js^2 + ds} = \frac{1}{s \cdot (Js + d)},$$

also ein IT_1 -Glied.

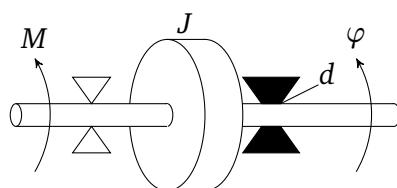


Abbildung 2.7: Beispiel für IT_1 -Glied: Welle

Übertragungsfunktion

Auch das IT_1 -Glied ist durch zwei Parameter, K und T , vollständig beschrieben. Die Übertragungsfunktion lautet

$$G(s) = \frac{K}{s \cdot (Ts + 1)}.$$

Bestimmung der Parameter aus der Sprungantwort

An dieser Stelle kann auch ohne größeren Mehraufwand direkt ein IT_n -Glied der Form

$$G(s) = \frac{K}{s \cdot \prod_{k=1}^n (T_k s + 1)}$$

betrachtet werden.

Zunächst wird über den Endwertsatz der Laplacetransformation der Wert der Steigung der Sprungantwort $y_s(t)$ für $t \rightarrow \infty$ bestimmt:

$$\lim_{t \rightarrow \infty} \frac{dy_s(t)}{dt} = \lim_{s \rightarrow 0} s \cdot \left(s \cdot G(s) \cdot \frac{u_0}{s} \right) = \lim_{s \rightarrow 0} \frac{K u_0}{\prod_{k=1}^n (T_k s + 1)} = K u_0.$$

Daraus folgt, dass der Parameter K gerade dem Endwert der Steigung der Sprungantwort geteilt durch die Höhe u_0 des Eingangssprunges entspricht.

Die Endsteigung entspricht also genau der Steigung, die ein Integrator der Form $G(s) = \frac{K}{s}$, also ohne die Verzögerungsglieder besäße. Damit ergibt sich für große t auch eine endliche Differenz

$$\Delta y_s(t) = y_s(t) - K \cdot \int_0^t u_0 \cdot d\tau$$

zwischen dem hier betrachteten IT_n -Glied und dem idealen Integrator. Im Laplacebereich wird dieses Signal durch

$$\Delta Y_s(s) = G(s) \cdot \frac{u_0}{s} - \frac{K}{s} \cdot \frac{u_0}{s}$$

beschrieben, und mit dem Endwertsatz erhält man zunächst

$$\lim_{t \rightarrow \infty} \Delta y_s(t) = \lim_{s \rightarrow 0} s \cdot \left(G(s) \cdot \frac{u_0}{s} - \frac{K}{s} \cdot \frac{u_0}{s} \right) = K u_0 \cdot \lim_{s \rightarrow 0} \frac{1 - \prod_{k=1}^n (T_k s + 1)}{s \cdot \prod_{k=1}^n (T_k s + 1)}.$$

Setzt man $s = 0$, so nehmen Zähler und Nenner gleichermaßen den Wert null an, so dass sich mit der Regel von l'Hôpital weiter²

$$\lim_{t \rightarrow \infty} \Delta y_s(t) = K u_0 \cdot \lim_{s \rightarrow 0} \frac{-\sum_{k=1}^n T_k \cdot \prod_{\substack{k'=1 \\ k' \neq k}}^n (T_{k'} s + 1)}{\prod_{k=1}^n (T_k s + 1) + s \cdot \sum_{k=1}^n T_k \cdot \prod_{\substack{k'=1 \\ k' \neq k}}^n (T_{k'} s + 1)} = -K u_0 \cdot \sum_{k=1}^n T_k$$

²

$$\begin{aligned} \frac{d}{ds} \prod_{k=1}^n (T_k s + 1) &= \frac{d}{ds} ((T_1 s + 1) \cdot (T_2 s + 1) \cdots (T_n s + 1)) \\ &= T_1 \cdot (T_2 s + 1) \cdots (T_n s + 1) + \cdots + T_n \cdot (T_1 s + 1) \cdots (T_{n-1} s + 1) = \sum_{k=1}^n T_k \cdot \prod_{\substack{k'=1 \\ k' \neq k}}^n (T_{k'} s + 1) \end{aligned}$$

ergibt. Führt man noch die Summenzeitkonstante $T_\Sigma = \sum_{k=1}^n T_k$ als Summe aller Zeitkonstanten ein, so kann auch kürzer

$$\lim_{t \rightarrow \infty} \Delta y_s(t) = -Ku_0 \cdot T_\Sigma$$

geschrieben werden. Damit lässt sich die Asymptote der Sprungantwort als

$$y_s(t) \xrightarrow{t \rightarrow \infty} Ku_0 \cdot t - Ku_0 \cdot T_\Sigma$$

angeben. Für $t = T_\Sigma$ ist der Wert der Asymptoten gerade null, d. h. die Summenzeitkonstante kann am Schnittpunkt der Asymptoten mit der Abszisse direkt abgelesen werden.

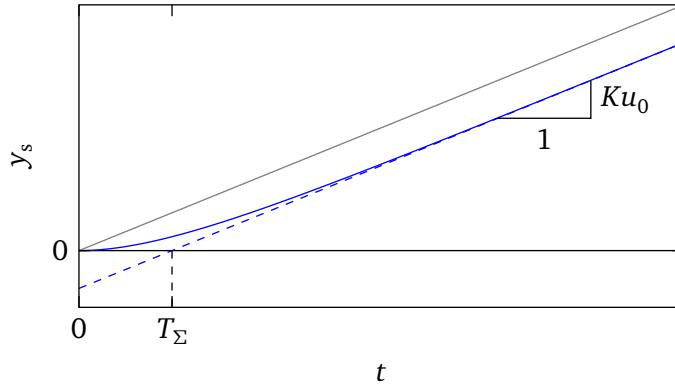


Abbildung 2.8: Sprungantwort eines IT_n -Gliedes

Alternativ zu diesem Vorgehen kann, wenn kaum Rauschen vorliegt, auch der Ausgang numerisch abgeleitet werden. Dann erhält man – hier wieder ein IT_1 -Glied mit nur einer Verzögerung vorausgesetzt – das künstliche System

$$G'(s) = s \cdot G(s) = \frac{K}{Ts + 1},$$

welches als PT_1 -Glied identifiziert werden kann. Gegebenenfalls liegen auch Freiheitsgrade in der Wahl der Ausgangsgröße vor, die es erlauben, das System als (echtes) PT_1 -System zu betrachten.

Bestimmung der Parameter aus dem Frequenzgang

Der Frequenzgang eines IT_1 -Gliedes ist in Abbildung 2.9 gezeigt. Die Zeitkonstante T kann wieder über den Asymptotenschnittpunkt oder dem -135° -Phasenpunkt abgelesen werden.

Der Ordinatenwert des Schnittpunkts der Asymptoten entspricht $K \cdot T$. Allgemein ist die Asymptote für $\omega \rightarrow 0$ über

$$|H(\omega)| \xrightarrow{\omega \ll \frac{1}{T}} \frac{K}{\omega}$$

gegeben. Damit kann anhand eines beliebig vorzugebenden ω der Wert für K bestimmt werden. In Abbildung 2.9 ist dies für $\omega = \frac{1}{T'}$ angedeutet. Für diesen Frequenzpunkt hat die Asymptote den Ordinatenwert $K \cdot T'$.

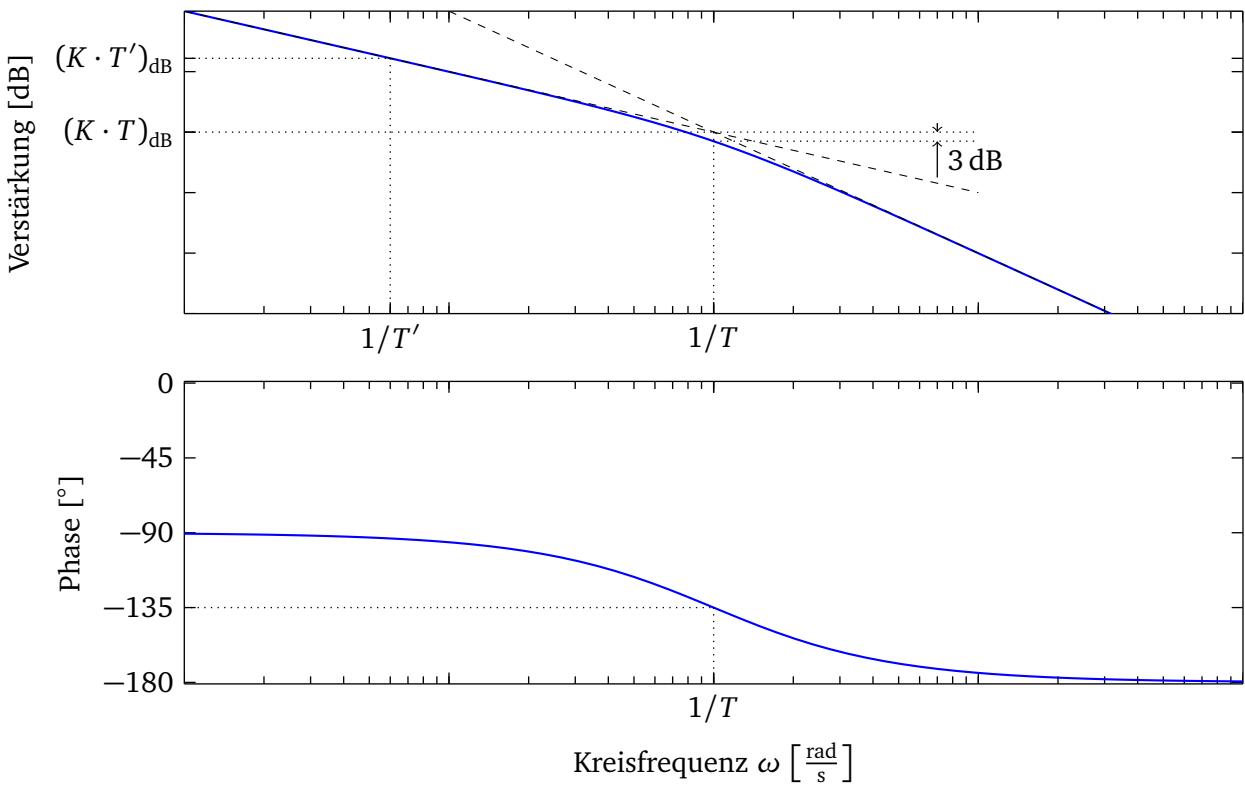


Abbildung 2.9: Frequenzgang IT1-Glied

2.4 PT2-Glied

Beispiel

Das Standardbeispiel für ein System mit PT_2 -Verhalten ist der in Abbildung 2.10 dargestellte Ein-Massen-Schwinger. Dieser wird durch die Differentialgleichung

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

beschrieben, was im Laplace-Bereich auf die Übertragungsfunktion

$$G(s) = \frac{1}{ms^2 + ds + k}$$

führt.

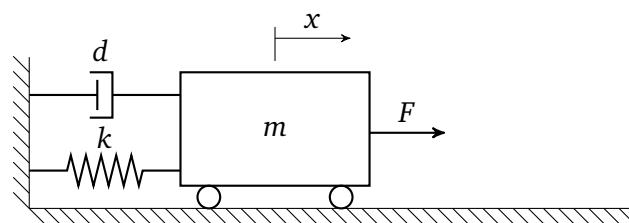


Abbildung 2.10: Beispiel für PT_2 -Glied

Übertragungsfunktion

Die Standardform eines PT₂-Gliedes lautet

$$G(s) = \frac{K}{s^2 \cdot \frac{1}{\omega_0^2} + s \cdot \frac{2D}{\omega_0} + 1} \quad (2.1)$$

$$= \frac{K \omega_0^2}{s^2 + s \cdot 2D\omega_0 + \omega_0^2}. \quad (2.2)$$

Die zwei Pole des PT₂-Gliedes liegen bei

$$\lambda_{1,2} = -D\omega_0 \pm \omega_0 \cdot \sqrt{D^2 - 1}.$$

Je nach Dämpfungsgrad liegen ein konjugiert-komplexes Polpaar oder zwei reelle Pole vor, und es ist sinnvoll, diese beiden Fälle zu unterscheiden.

2.4.1 Schwingfähiges PT2-Glied

Ist der Dämpfungsgrad D kleiner als eins, so ist der Radikant $D^2 - 1$ kleiner als Null und dessen Quadratwurzel damit imaginär. Die beiden Pole der Übertragungsfunktion (2.1) können dann in der Form

$$\lambda_{1,2} = -\delta \pm j\omega_e$$

mit

$$\delta = D \cdot \omega_0 \quad \text{und}$$

$$\omega_e = \omega_0 \cdot \sqrt{1 - D^2}$$

geschrieben werden.

Die homogene Zeitlösung ergibt sich aus einer Linearkombination der zu den beiden Polen λ_1 und λ_2 gehörenden Elementarlösungen,

$$y_h(t) = C_1 \cdot e^{\lambda_1 t} + C_2 \cdot e^{\lambda_2 t},$$

wobei die Konstanten C_1 und C_2 noch aus den Anfangsbedingungen bestimmt werden müssen. Da $\lambda_2 = \lambda_1^*$ gilt, muss (bei reellwertigen Anfangsbedingungen für $y(t)$) auch $C_2 = C_1^*$ gelten. Damit lässt sich die Lösung auch als

$$y_h(t) = e^{-\delta t} \cdot (C_1 e^{j\omega_e t} + C_1^* e^{-j\omega_e t})$$

und weiter als

$$y_h(t) = C \cdot e^{-\delta t} \cdot \sin(\omega_e t + \varphi)$$

schreiben.

Die homogene Lösung wird also durch eine abklingende harmonische Schwingung beschrieben, wobei der Realteil des konjugiert-komplexen Polpaares angibt, wie schnell diese Schwingung abklingt ($e^{-\delta t}$) und der Imaginärteil die „Frequenz“ ω_e der freien, abklingenden Schwingung beschreibt. Daher wird diese auch als Eigenfrequenz bezeichnet.

Neben der Eigenfrequenz ω_e und der Kennkreisfrequenz ω_0 ist noch die Resonanzfrequenz

$$\omega_r = \omega_0 \cdot \sqrt{1 - 2D^2}$$

eine wichtige Kenngröße für das PT₂-Glied. Diese ist die Frequenz, für der der Betrag des Frequenzgangs den maximalen Wert annimmt, $\omega_r = \arg \max_{\omega} |H(\omega)|$.

Diese drei besonderen Frequenzen sind natürlich keine unabhängigen Kenngrößen sondern über den Dämpfungsgrad miteinander gekoppelt.

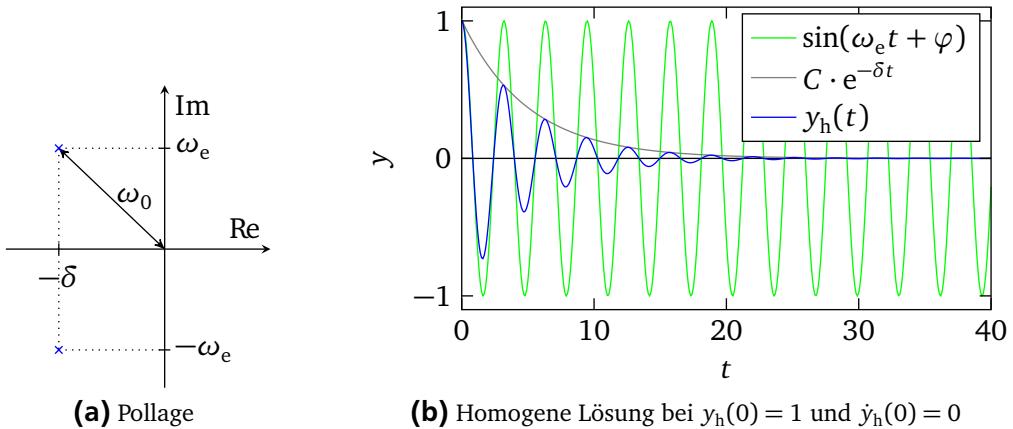


Abbildung 2.11: Zusammenhang zwischen Polen und Eigenverhalten beim PT_2 -Glied

Kennkreisfrequenz

Die Kennkreisfrequenz ω_0 gibt an, an welcher Stelle im Bodediagramm sich die Asymptoten schneiden (Knickfrequenz). Darüber hinaus ist es die Frequenz, bei der die Phase -90° beträgt. (Unabhängig vom Dämpfungsgrad!)

Eigen(kreis)frequenz

Die Eigenfrequenz ω_e ist die Frequenz der (gedämpften) Schwingung, die das System ohne äußeren Einflüsse ausführt. Das Eigenverhalten ist direkt über die Pole der Übertragungsfunktion gegeben, da die Pole den Exponenten der Exponentialfunktionen der Elementarlösungen entsprechen. D. h. die Eigenfrequenz kann direkt aus den Polen abgelesen werden.

Resonanz(kreis)frequenz

Die Resonanzfrequenz ω_r ist die Frequenz, bei der das System sinusförmige Eingangsgrößen maximal verstärkt. D. h. die Resonanzfrequenz ist mit der partikulären Lösung für sinusförmige Eingangsgrößen verbunden. Diese werden aber gerade durch den Frequenzgang dargestellt. Die Resonanzfrequenz kann direkt als die Frequenz abgelesen werden, für die der Frequenzgang betragsmäßig den größten Wert annimmt.

Es gilt

$$\omega_r \leq \omega_e \leq \omega_0$$

Für $D \geq \frac{1}{\sqrt{2}}$ gibt es keine Resonanz mehr, für $D \geq 1$ ist das System nicht mehr schwingungsfähig

Tabelle 2.1: Kenngrößen PT_2 -Glied

Stationäre Verstärkung	K
Kennkreisfrequenz	ω_0 „undamped natural frequency“
Eigenkreisfrequenz	ω_e „natural frequency“
Resonanzkreisfrequenz	ω_r
Dämpfungsgrad	D
Abklingkonstante	δ negativer Realteil der Pole

Identifikation der Kenngrößen

Die stationäre Verstärkung K kann bei der Sprunganregung wieder einfach über den Endwert $y_\infty = u_0 K$ abgelesen werden bzw. im Frequenzgang als Wert für $\omega \rightarrow 0$, $H(0) = K$. Damit sind noch zwei weitere Kenngrößen nötig, um das PT_2 -Glied vollständig zu beschreiben. Diese können aus der Menge der verbleibenden Kenngrößen Kennkreisfrequenz, Eigenfrequenz, Resonanzfrequenz, Dämpfungsgrad und Abklingfaktor gewählt werden. Alle weiteren Größen können dann bei Bedarf berechnet werden.

Ermittlung über Sprunganregung

Die Zeitlösung für den Ausgang $y(t)$ bei dem Eingang $u(t) = u_0 \cdot \sigma(t)$ lautet

$$y(t) = K \cdot u_0 \cdot \left(1 - \frac{1}{\sqrt{1-D^2}} \cdot e^{-\delta t} \cdot \sin(\omega_e t + \varphi) \right), \quad \tan \varphi = \frac{\sqrt{1-D^2}}{D}. \quad (2.3)$$

Das hier fast die homogene Lösung erscheint kann man sich darüber erklären, dass man anstelle des Sprungs auch das System bezüglich $\Delta y(t) = y(t) - y_\infty$ betrachten könnte. Hierbei stellt der Sprung eine Anfangsstörung dar.

Daher können anhand der Sprungantwort gut die Kenngrößen, die das Eigenverhalten beschreiben, d. h. Eigenfrequenz und Abklingkonstante, bestimmt werden.

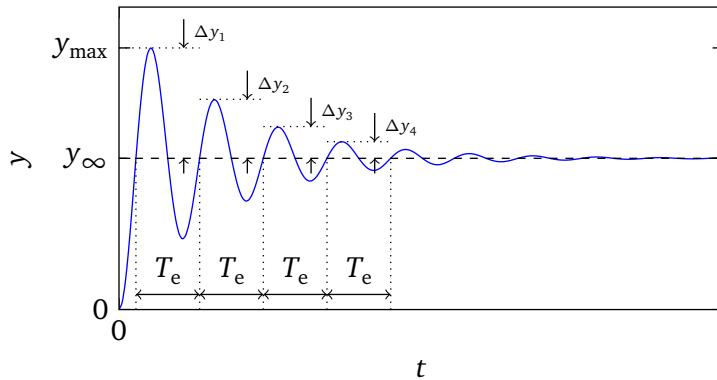


Abbildung 2.12: Sprungantwort schwingfähiges PT_2 -Glied

Der Verstärkungsfaktor K ergibt sich aus dem stationären Endwert über $y_\infty = K \cdot u_0$. Die Eigenfrequenz ω_e kann über die Schnittpunkte der Schwingung mit dem stationären Endwert y_∞ abgelesen werden.

Die Abklingkonstante δ kann über den maximalen Wert $\Delta y_{\max} = y_{\max} - y_\infty$ des Überschwingens bestimmt werden. Dazu wird die Sprungantwort (2.3) abgeleitet, bzw. die Impulsantwort

$$\dot{y}(t) = \frac{K u_0 \omega_0}{\sqrt{1-D^2}} \cdot e^{-\delta t} \cdot \sin(\omega_e \cdot t)$$

aus einer geeigneten Tabelle abgelesen. Damit ein Zeitpunkt t_{\max} ein (lokales) Extremum sein kann, muss $\dot{y}(t_{\max}) = 0$ und damit $\sin(\omega_e \cdot t_{\max}) = 0$ gelten. Daraus folgt $t_{\max} \in \{k \cdot \frac{\pi}{\omega_e} | k \in \mathbb{N}\}$. Für das erste (lokale und auch globale) Maximum gilt $t_{\max} = \pi / \omega_e$ und dies in Gl. (2.3) eingesetzt ergibt zunächst

$$y_{\max} = y(t_{\max}) = K \cdot u_0 \cdot \left(1 - \frac{1}{\sqrt{1-D^2}} \cdot e^{-\delta \cdot \frac{\pi}{\omega_e}} \cdot \sin(\pi + \varphi) \right), \quad \tan \varphi = \frac{\sqrt{1-D^2}}{D}.$$

Unter Beachtung, dass $\sin(\pi + \varphi) = -\sin(\varphi)$ und $\arctan x = \arcsin \frac{x}{\sqrt{1+x^2}}$ gilt, folgt daraus

$$y_{\max} = K \cdot u_0 \cdot \left(1 + e^{-\delta \cdot \frac{\pi}{\omega_e}} \right)$$

und weiter

$$\Delta y_{\max} = y_{\max} - y_{\infty} = y_{\infty} \cdot e^{-\delta \cdot \frac{\pi}{\omega_e}}.$$

Der Endwert y_{∞} kann einfach abgelesen werden, und so kann diese Gleichung einfach nach dem gesuchten δ aufgelöst werden,

$$\delta = \frac{\omega_e}{\pi} \cdot \ln \frac{y_{\infty}}{\Delta y_{\max}}.$$

Alternativ kann das logarithmische Dekrement verwendet werden. Bildet man den Quotienten von zwei Werten von $\Delta y(t)$, die ein ganzes Vielfaches n der Eigenschwingungsdauer $T_e = \frac{2\pi}{\omega_e}$ auseinander liegen, erhält man

$$\begin{aligned} \frac{\Delta y(t)}{\Delta y(t + n \cdot T_e)} &= \frac{\frac{K \cdot u_0}{\sqrt{1-D^2}} \cdot e^{-\delta t} \cdot \sin(\omega_e t + \varphi)}{\frac{K \cdot u_0}{\sqrt{1-D^2}} \cdot e^{-\delta \cdot (t+n \cdot T_e)} \cdot \sin(\omega_e (t + n \cdot T_e) + \varphi)} \\ &= \frac{e^{-\delta t}}{e^{-\delta \cdot (t+n \cdot T_e)}} = e^{\delta \cdot n T_e}. \end{aligned}$$

Dies ist gültig für alle t für die $\Delta y(t) \neq 0$ gilt und der Quotient damit definiert ist.

Bildet man den natürlichen Logarithmus des Quotienten erhält man

$$\ln \frac{\Delta y(t)}{\Delta y(t + n \cdot T_e)} = n \cdot (\delta \cdot T_e).$$

Den geklammerten Faktor $\delta \cdot T_e$, also der Wert des logarithmierten Quotienten für $n = 1$ wird als logarithmisches Dekrement Λ bezeichnet,

$$\Lambda = \delta \cdot T_e = \delta \cdot \frac{2\pi}{\omega_e} = \ln \frac{\Delta y(t)}{\Delta y(t + T_e)} = \frac{1}{n} \cdot \ln \frac{\Delta y(t)}{\Delta y(t + n \cdot T_e)}. \quad (2.4)$$

Sinnvollerweise wird man die Werte $\Delta y(t)$ und $\Delta y(t + n \cdot T_e)$ dort ablesen, wo diese möglichst groß sind und nicht gerade nahe an null liegen. Es sind immer Werte einzusetzen, die um ein Vielfaches der Schwingungsdauer T_e auseinander liegen. Hier bieten sich besonders die lokalen Maxima jeder „Periode“ an.

Wird das logarithmische Dekrement verwendet, kann das Dekrement für verschiedene lokale Maxima bestimmt und gemittelt werden. Ebenfalls ist eine grafische Mittelung (Regression) möglich. Formt man Gl. (2.4) nach

$$\ln |\Delta y(t + n \cdot T_e)| = -\Lambda \cdot n + \ln |\Delta y(t)|$$

um erkennt man, dass die logarithmierten Werte $\ln |\Delta y(t + n \cdot T_e)|$ für $n \in \mathbb{N}$ auf einer Geraden mit der Steigung $-\Lambda$ liegen.

Das logarithmische Dekrement bietet sich an, wenn das System sehr schwach gedämpft ist und viele deutliche Schwingungen zu erkennen sind, da dann eine einfache Mittelung möglich ist. Für stärker gedämpfte Systeme bietet sich das Vorgehen über das maximale Überschwingen an. Ist auch dieses nicht mehr klar zu erkennen, oder nur mit großen Ungenauigkeiten abzulesen, kann es auch für $D < 1$ schon problematisch werden, die Parameter aus der Sprungantwort abzulesen.

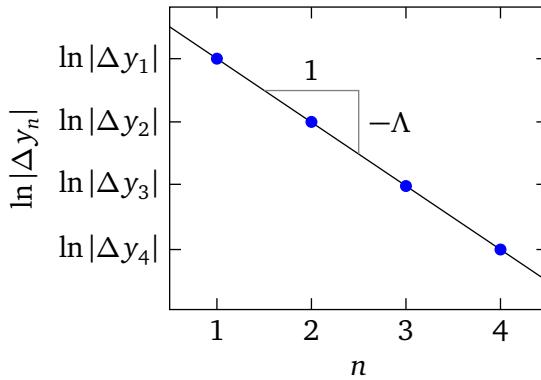


Abbildung 2.13: Zum logarithmischen Dekrement

Ermittlung über Frequenzgang

Die stationäre Verstärkung K lässt sich direkt aus dem Frequenzgang ablesen,

$$|H(\omega \rightarrow 0)| = K .$$

Die noch benötigten zwei Parameter können wieder aus der Menge der verbleibenden Kenngrößen gewählt werden, wobei natürlich diejenigen gewählt werden sollten, für die – für das vorliegende System bzw. am vorliegenden Frequenzgang – das sicherste Ablesen möglich ist.

Wenn die Resonanz ausgeprägt ist, kann man ω_r am Maximum des Frequenzganges ablesen. ω_0 ergibt sich am der Schnittpunkt der Asymptoten bzw. am Phasenwinkel von -90° .

Durch Einsetzen von $j\omega_0$ und $j\omega_r$ in die Übertragungsfunktion kann man leicht

$$|H(\omega_0)| = \frac{K}{2D}$$

bzw., nach etwas Umformen,

$$|H(\omega_r)| = \frac{K}{2D \cdot \sqrt{1-D^2}} , \quad 0 < D \leq \frac{1}{\sqrt{2}} ,$$

ablesen.

Schwach gedämpfte Systeme

Bei schwach gedämpften Systemen kann es für das Ablesen günstiger sein, den Frequenzgang nicht als übliches Bodediagramm, d. h. als logarithmisch aufgetragene Betrag-Phase-Darstellung, sondern getrennt nach Real- und Imaginärteil zu plotten, wie in Abbildung 2.15b dargestellt.³

Die Maxima des Realteils ergeben sich für $D < \frac{1}{2}$ allgemein zu

$$\begin{aligned} \omega_a &= \omega_0 \cdot \sqrt{1-2D} && \text{und} \\ \omega_b &= \omega_0 \cdot \sqrt{1+2D} . \end{aligned}$$

Damit lässt sich durch Einsetzen der Zusammenhang

$$D = \frac{1}{2} \cdot \frac{\omega_b^2 - \omega_a^2}{\omega_b^2 + \omega_a^2}$$

³ Sehr schwach gedämpfte schwingungsfähige Systeme treten vor allem bei mechanischen Strukturen auf. Ist eine solche zu identifizieren, lohnt sich auch ein Blick in die Literatur zur Schwingungsmesstechnik. So sind die in diesem Abschnitt vorgestellten Kenngrößen dem Skript „Schwingungsmesstechnik“ von Markert entnommen.

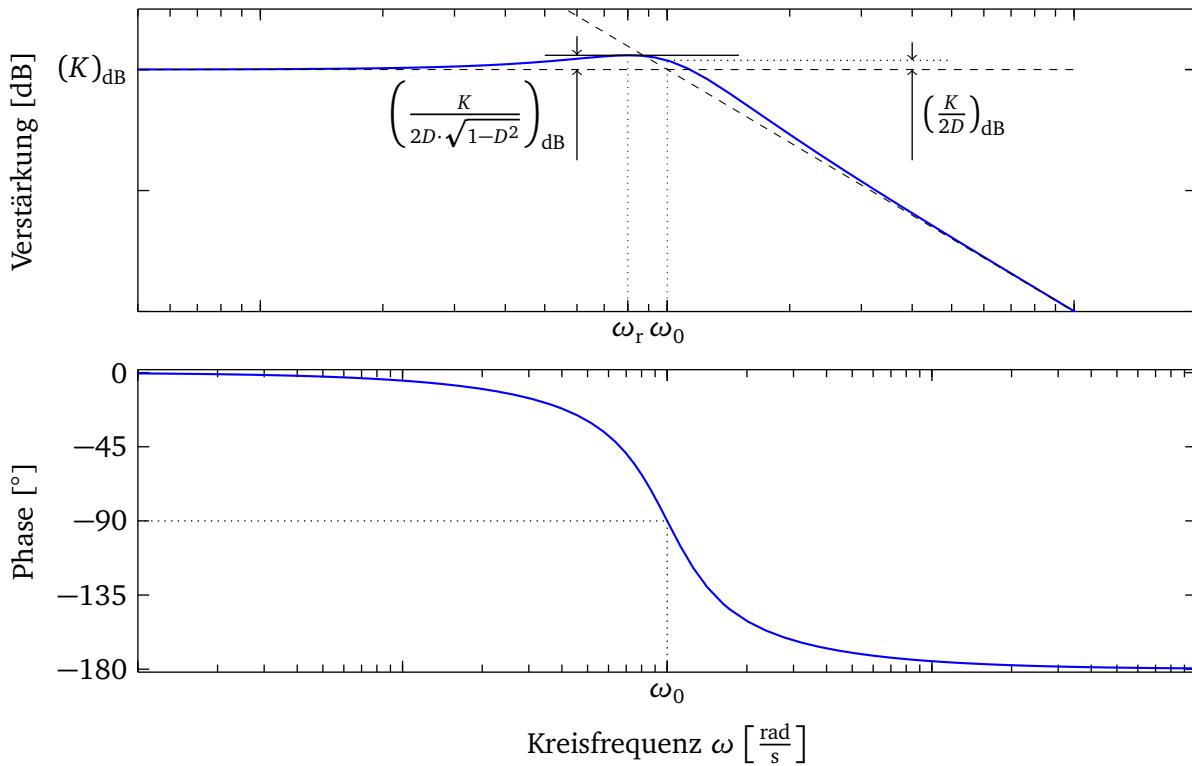


Abbildung 2.14: Frequenzgang schwingfähiges PT2-Glied

überprüfen. Der quadratische Mittelwert von ω_b , und ω_a entspricht gerade der Kennkreisfrequenz

$$\omega_0^2 = \frac{1}{2} \cdot (\omega_b^2 + \omega_a^2).$$

Die Kennkreisfrequenz entspricht ebenso dem Schnittpunkt des Realteils der Übertragungsfunktion mit der Abszisse,

$$\operatorname{Re}(H(\omega_0)) = 0,$$

was auch an der Ortskurve (Abbildung 2.15a) abgelesen werden kann.

Für sehr schwach gedämpfte Systeme, $D \ll 1$, entspricht die Resonanzfrequenz ungefähr der Kennkreisfrequenz und diese kann damit am Maximum des Imaginärteils abgelesen werden,

$$\omega_r \approx \omega_0 \approx \arg \max_{\omega} |\operatorname{Im}(H(\omega))|$$

mit

$$\max_{\omega} |\operatorname{Im}(H(\omega))| \approx |\operatorname{Im}(H(\omega_0))| = \frac{K}{2D}.$$

Der Vorteil gegenüber dem Ablesen der Resonanzfrequenz aus dem Betrag des Frequenzgangs, dessen Maximum für alle $D < 1$ exakt der Resonanzfrequenz entspricht, liegt darin, dass das Maximum im Imaginärteil etwas deutlicher, d. h. enger, ausgeprägt ist. Um dies zu verdeutlichen, ist im Plot des Imaginärteils in Abbildung 2.15b noch der negative Betrag des Frequenzganges dargestellt.

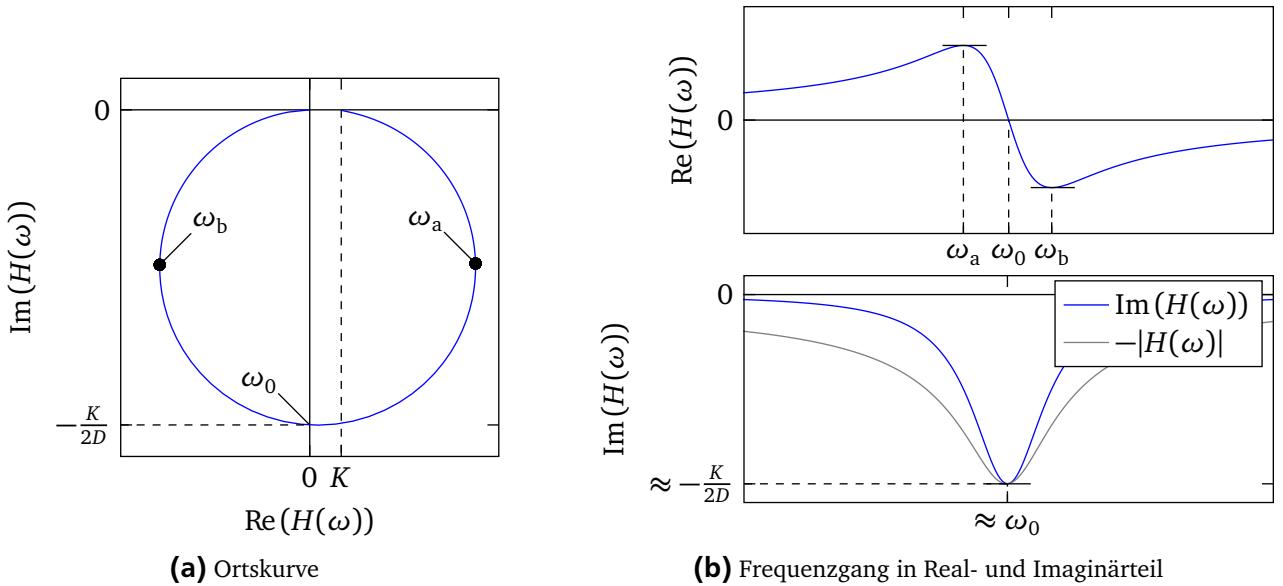


Abbildung 2.15: Ortskurve und Frequenzgang

2.4.2 Nichtschwingfähiges PT2-Glied

Erhöht man den Dämpfungsgrad D auf Werte über eins, so ist die Darstellung der Übertragungsfunktion in der Form (2.1) immer noch möglich, jedoch gilt dann für die Pole

$$\lambda_{1,2} = -D\omega_0 \pm \omega_0 \cdot \sqrt{D^2 - 1}.$$

D. h. diese sind beide reell. Die Kennkreisfrequenz ω_0 behält ihre Bedeutung in dem Sinne, dass diese immer noch jene Frequenz ist, bei dem die Phase des Frequenzgangs -90° beträgt, der Betrag des Frequenzganges $\frac{K}{2D}$ beträgt und sich die äußeren Asymptoten treffen. Eine Eigenfrequenz oder Resonanzfrequenz kann aber natürlich nicht mehr angegeben werden.

Hier ist es auch anschaulich die Übertragungsfunktion als

$$G(s) = \frac{K}{(T_1 s + 1) \cdot (T_2 s + 1)}$$

mit den reellen Polen

$$\lambda_{1,2} = -\frac{1}{T_1}, -\frac{1}{T_2}$$

darzustellen.

Die Anzahl der zu identifizierenden Parametern ist dabei unverändert drei.

Ermittlung über Frequenzgang

Im Frequenzgang kann, neben $K = H(\omega \rightarrow 0)$, wieder ω_0 und $|H(\omega_0)|$ zum Ablesen verwendet werden. Wenn die beiden Zeitkonstanten weit auseinanderliegen, können auch die Knickfrequenzen ω_{T_1} und ω_{T_2} abgelesen werden.

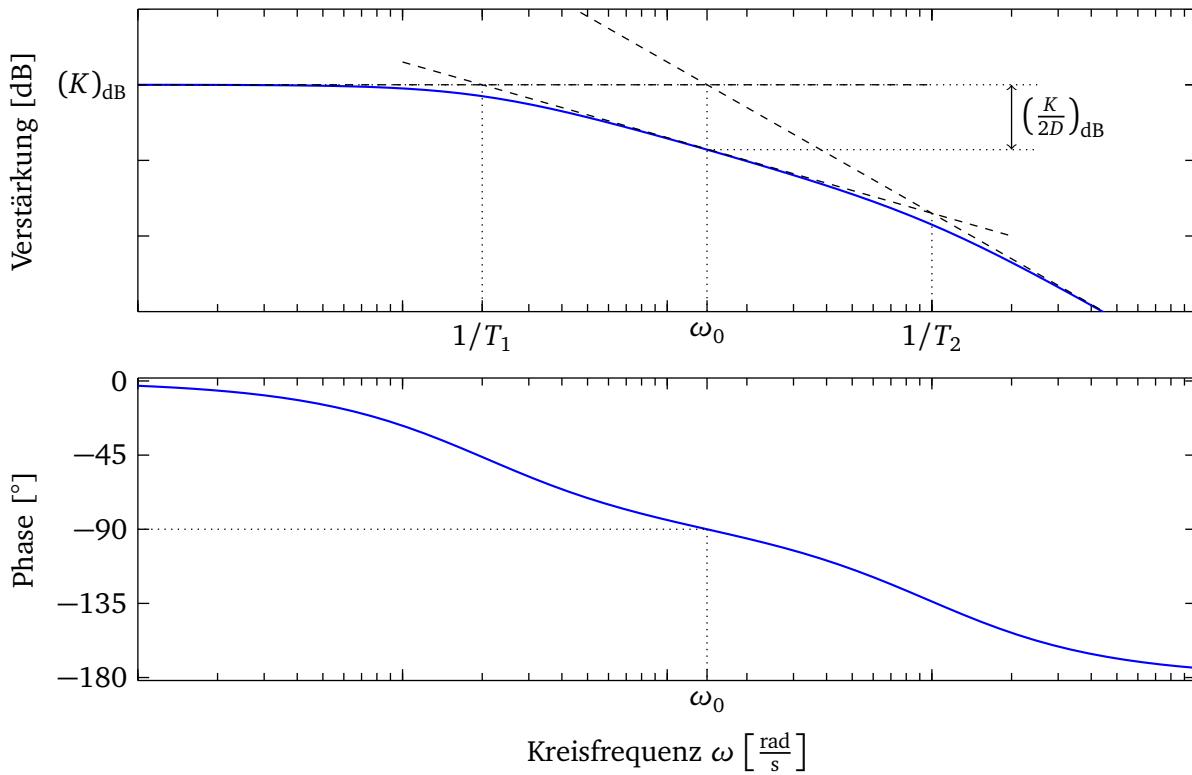


Abbildung 2.16: Frequenzgang nichtschwingfähiges PT2-Glied

2.5 PT_n-Glied

Die Übertragungsfunktion eines allgemeinen PT_n-Glieds lautet

$$G(s) = \frac{K}{\prod_{k=1}^n (T_k s + 1)}$$

und wird über $n+1$ Parameter, K sowie n Zeitkonstanten T_k , beschrieben. Um ein PT_n-Glied mit einfachen Verfahren zu identifizieren, muss diese Anzahl sinnvoll reduziert werden. Häufig kann angenommen werden, dass alle n Zeitkonstanten gleich einer Konstanten T sind. Dann lautet die Übertragungsfunktion

$$G(s) = \frac{K}{(Ts + 1)^n}.$$

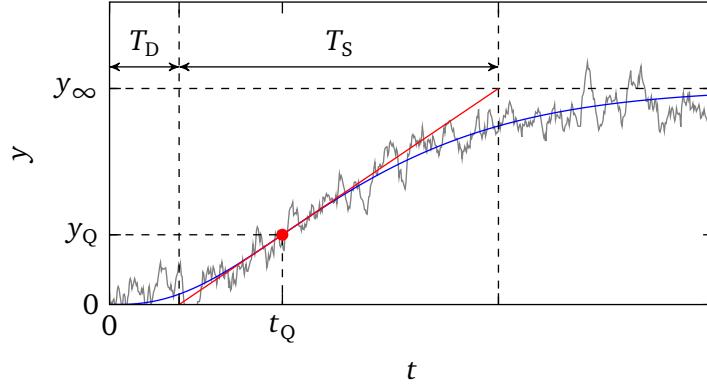
Häufig ist n jedoch nicht bekannt, so dass hier im Grunde drei Parameter zu bestimmen sind.

Bestimmung der Parameter aus der Sprungantwort

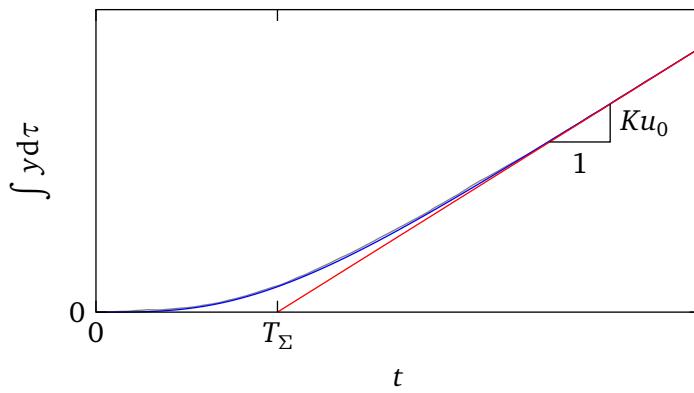
Die Zeitlösung der Sprungantwort eines PT_n-Systems mit gleicher Zeitkonstante T lautet

$$y(t) = Ku_0 \cdot \left(1 - e^{-\frac{t}{T}} \cdot \sum_{k=0}^{n-1} \frac{1}{k!} \cdot \left(\frac{t}{T}\right)^k \right).$$

K ergibt sich wieder aus dem Endwert $y_\infty = Ku_0$ der Sprungantwort. Um die Ordnung n und die Zeitkonstante T aus der Sprungantwort zu schätzen sind in der Literatur Ansätze zu finden, die auf Kenngrößen



(a) Sprungantwort



(b) Integrierte Sprungantwort

Abbildung 2.17: PTn-Glied

beruhen, die sich aus dem Wendepunkt (t_Q, y_Q) , der Verzögerungszeit T_D und der Übergangszeit T_S berechnen lassen. Diese Größen sind in Abbildung 2.17a an einem Beispiel eingetragen.

Für die genannten Größen gilt

$$\frac{t_Q}{T} = n - 1 ,$$

$$\frac{y_Q}{y_\infty} = 1 - e^{-(n-1)} \cdot \sum_{k=0}^{n-1} \frac{1}{k!} \cdot (n-1)^k ,$$

$$\frac{T_S}{T} = \frac{T \cdot (n-2)!}{(n-1)^{n-2}} \cdot e^{-(n-1)}$$

$$\text{und} \quad \frac{T_D}{T} = (n-1) - \frac{(n-2)!}{(n-1)^{n-2}} \cdot \left(e^{-(n-1)} - \sum_{k=0}^{n-1} \frac{(n-1)^k}{k!} \right) .$$

Ist der Wendepunkt und die Tangente gut ablesbar, so haben diese Kenngrößen den Vorteil, dass $\frac{y_Q}{y_\infty}$ und $\frac{T_D}{T_S}$ nicht von der Zeitkonstante T sondern nur der Ordnung n abhängen, so dass diese damit identifiziert werden kann.

Allerdings ist es sehr kritisch, diese Kenngrößen abzulesen, da zum Anlegen der Tangente in den Wendepunkt kaum Rauschen zugelassen werden kann. In Abbildung 2.17a ist in grau eine verrauschte Sprungantwort angegeben, und es wird deutlich, dass bei dieser der Wendepunkt nicht sicher identifiziert werden kann.

Da das betrachtete System nur drei Parameter besitzt, bietet sich auch die im nächsten Kapitel vorgestellten einfachen, nichtlinearen Optimierungsverfahren an. Stehen entsprechende Algorithmen nicht

zur Verfügung, liegen die Daten jedoch digital vor, dann kann auch eine nachträgliche Integration auf die Messdaten angewandt werden, also das (künstliche) System

$$G'(s) = \frac{K}{s \cdot \prod_{k=1}^n (Ts + 1)}$$

betrachtet werden. Dies wurde schon beim IT_1 -Glied behandelt, und die Summenzeitkonstante $T_\Sigma = nT$ kann wie dort besprochen bestimmt werden. Durch die Integration findet auch eine gewisse Glättung statt und darüber hinaus ist die Integrationszeitkonstante und damit die Endsteigung Ku_0 in diesem Fall auch bekannt, so dass man auch bei etwas Rauschen die Summenzeitkonstante noch gut ablesen kann, was in Abbildung 2.17b gezeigt ist. Die Ordnung n kann dann über ausprobieren und einem Vergleich mit der gemessenen Sprungantwort ermittelt werden.

Bestimmung der Parameter aus dem Frequenzgang

Die Zeitkonstante T ergibt sich aus dem Schnittpunkt der Asymptote. Die Ordnung n aus der Steigung der rechten Asymptoten ($-n \cdot 20 \text{ dB/Dekade}$) oder dem Phasenverlust ($-n \cdot 90^\circ$).

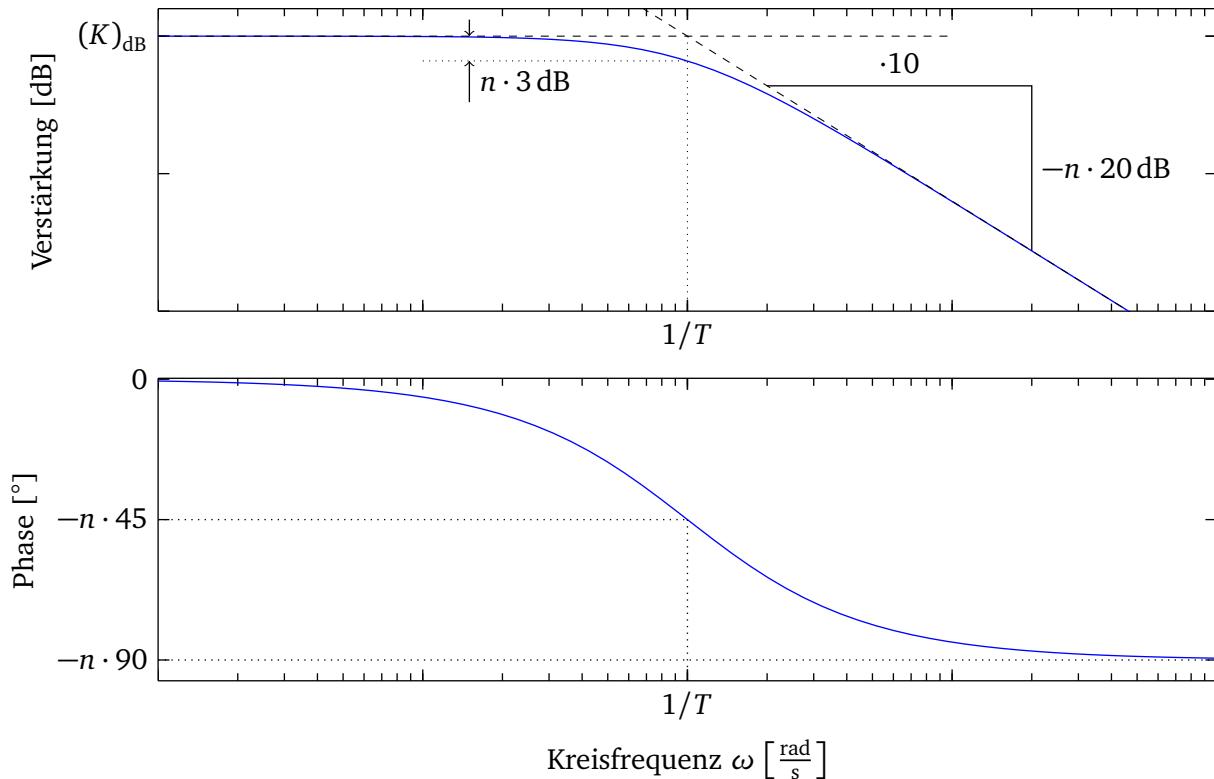


Abbildung 2.18: Frequenzgang PT_n -Glied (Gleiche Zeitkonstante T)

2.6 Totzeitglied

Die Übertragungsfunktion eines Totzeitgliedes lautet

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

und ist durch den einzigen Parameter T_t definiert.

Bestimmung der Parameter aus der Sprungantwort

Reine Totzeitglieder wären über die Sprungantwort

$$y(t) = \sigma(t - T_t) \cdot u_0$$

relativ einfach zu identifizieren, aber meist tritt eine Totzeit zusätzlich zu einem anderen Übertragungsverhalten auf.

Über die Sprungantwort im Zeitbereich ist die Totzeit nicht unbedingt gut zu identifizieren. Besteht das Gesamtsystem aus einer Totzeit und einem PT_1 -Verhalten, und ist die Sprungantwort nicht zu sehr verrauscht, dann ist die Totzeit noch gut zu erkennen. Bei stärkerem Rauschen kann es aber schon problematisch werden, die Totzeit sicher abzulesen. Liegt ein System aus Totzeit und einem Verzögerungsglied höherer Ordnung vor, dann ist es schon im unverrauschten Fall schwierig, die Totzeit abzulesen bzw. von den Verzögerungen zu trennen.

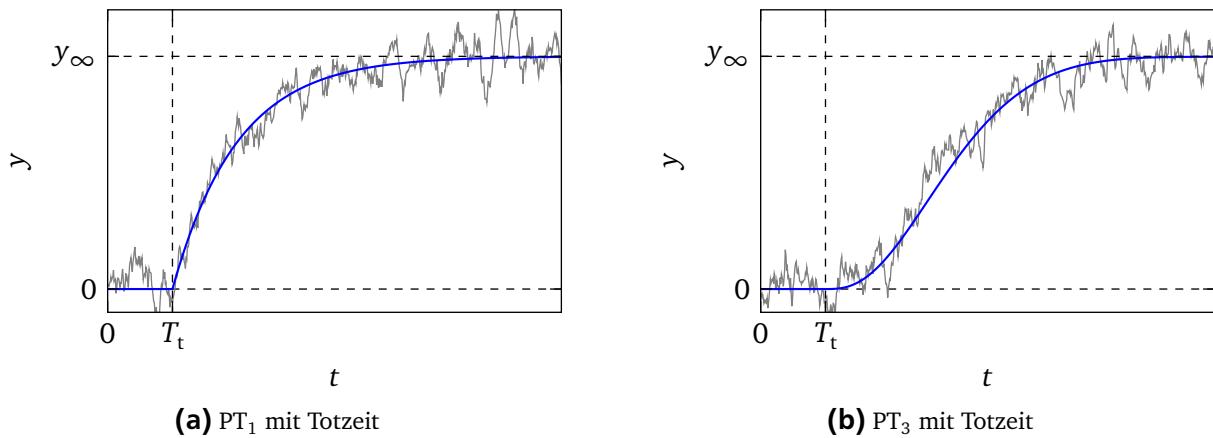


Abbildung 2.19: Sprungantwort von Systemen mit Totzeit

Bestimmung der Parameter aus dem Frequenzgang

Im Frequenzbereich hat man hingegen gute Chancen, eine Totzeit zu identifizieren. Diese lässt bekanntlich den Amplitudengang unverändert, führt aber zu einem linearen Abfall der Phase. In Abbildung 2.21a ist das Bodediagramm eines Totzeitgliedes mit $T_t = 0,2\text{s}$ dargestellt. Zur Verdeutlichung des linearen Phasenabfalls ist die Phase nochmals in Abbildung 2.21b über einer linearen Frequenzachse dargestellt.

Besteht das Gesamtsystem aus einer Totzeit $G_t(s)$ und einem weiteren, allgemeinen dynamischen System $G_d(s)$,

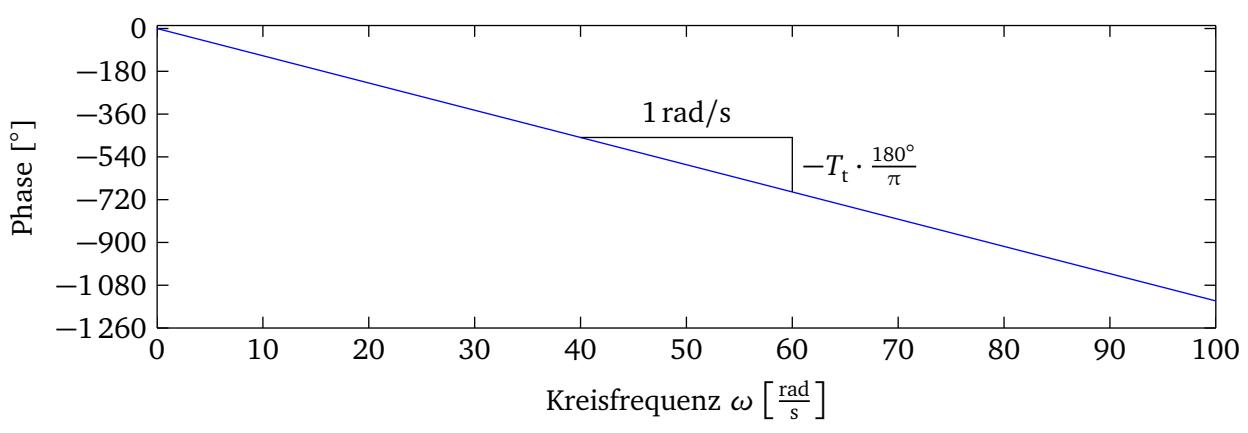
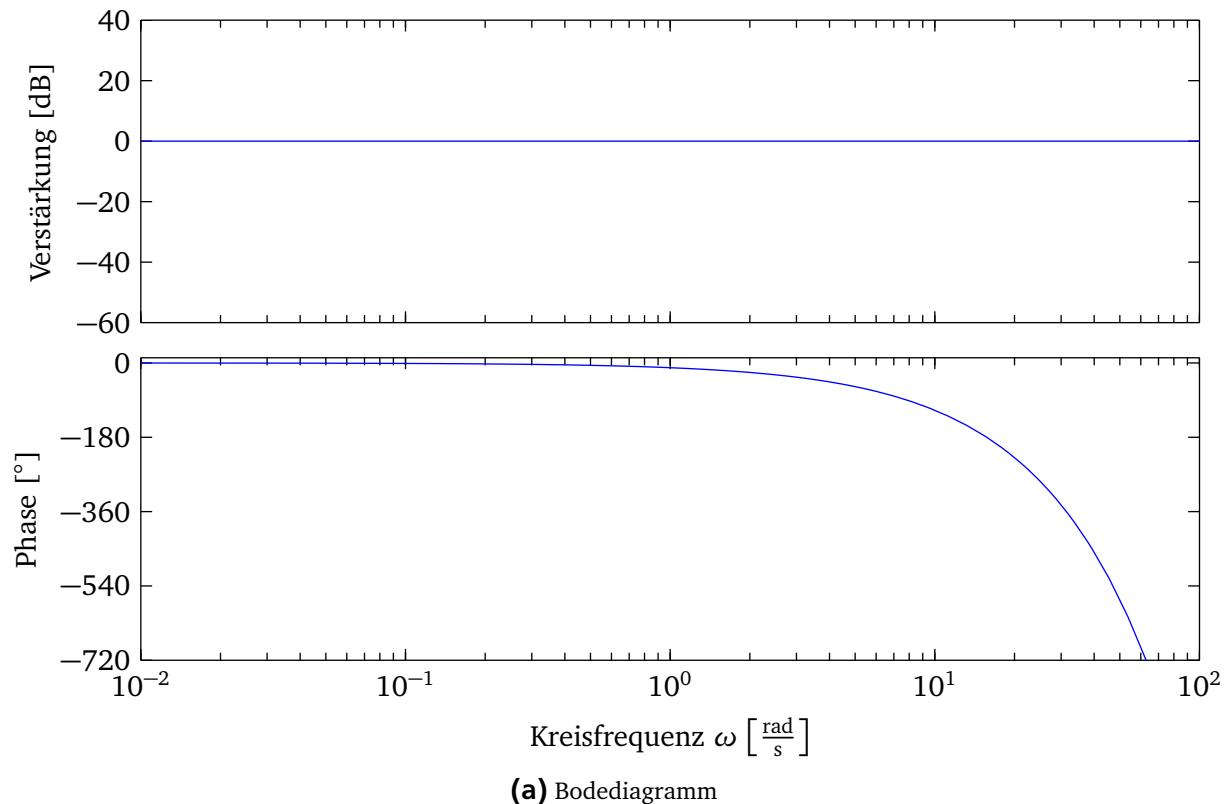
$$G_g(s) = G_t(s) \cdot G_d(s),$$

so gilt für den Frequenzgang $H_g(\omega)$ des Gesamtsystems, da $|G_t(j\omega)| = 1$ und $\arg G_t(j\omega) = -T_t \omega$,

$$\begin{aligned} |H_g(\omega)| &= |H_d(\omega)| \quad \text{und} \\ \arg H_g(\omega) &= -T_t \cdot \omega + \arg H_d(\omega). \end{aligned}$$

Ist die zu erwartende Phase an einem oder mehreren Frequenzpunkten bekannt, so kann dies ausgenutzt werden, den Phasengang entsprechend „zurückzudrehen“.

Bei PT_2 -Gliedern weiß man beispielsweise, dass bei der Kennkreisfrequenz ω_0 die Phase exakt -90° betragen muss. Dies ist in Abbildung 2.21a ausgenutzt.



(b) Darstellung der Phase über linearen Frequenzmaßstab

Abbildung 2.20: Frequenzgang Totzeitglied ($T_t = 0,2 \text{ s}$)

Auch kann man damit arbeiten, dass für alle Systeme ohne Totzeit der Phasengang einen konstanten Wert für $\omega \rightarrow \infty$ annehmen sollte, wenn alle Dynamiken erfasst sind. Fällt die Phase dennoch weiter ab, so kann die Totzeit aus dieser Steigung abgelesen werden. (Diese Steigung ist dabei bezüglich einer linearen Skalierung der Frequenzachse zu bestimmen, nicht einer logarithmischen.) Dies ist in Abbildung 2.21b dargestellt.

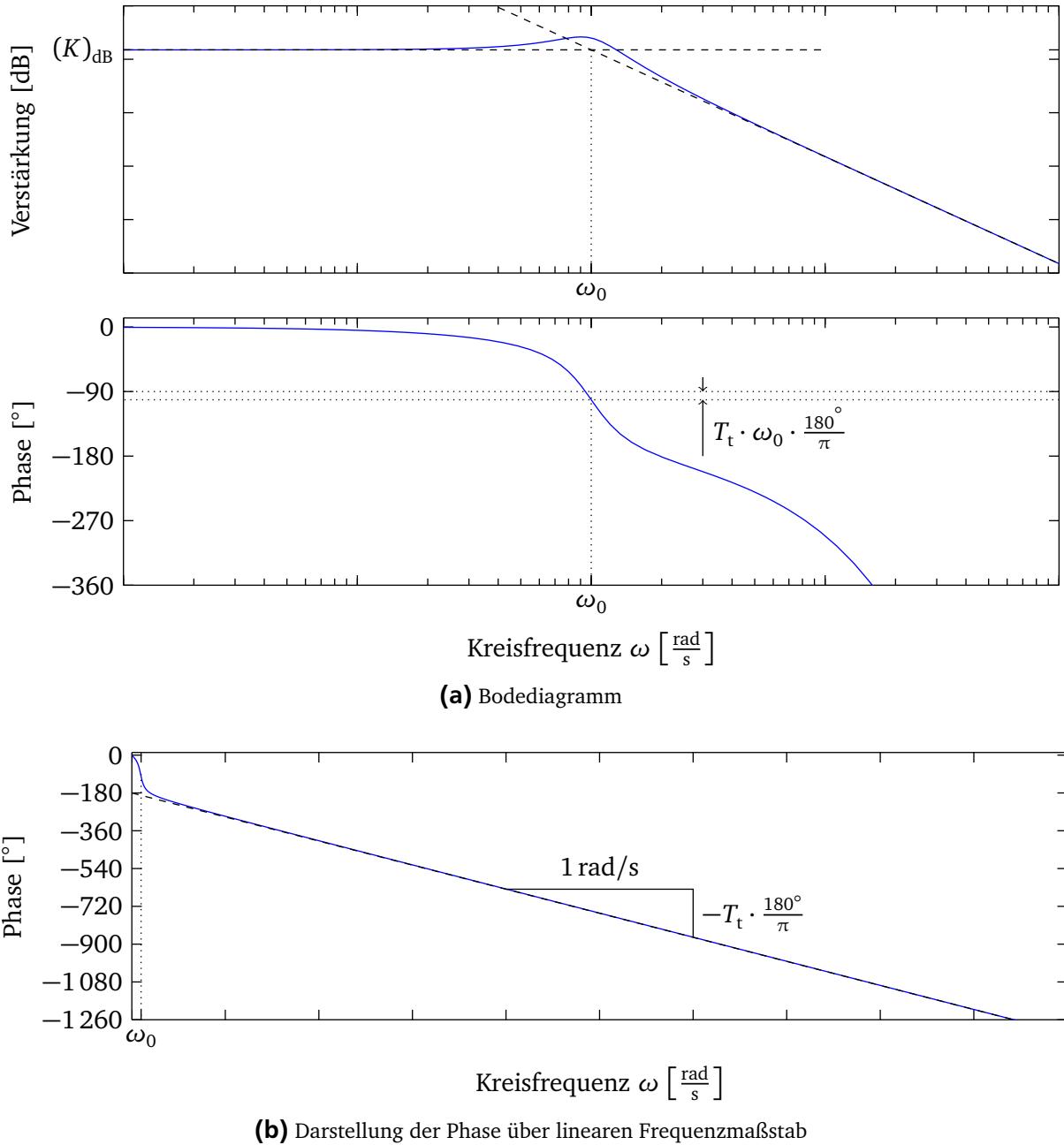


Abbildung 2.21: Frequenzgang PT_2 -Glied mit Totzeitglied

Zuletzt kann bei minimalphasigen Systemen der asymptotischen Steigung im Amplitudengang direkt ein asymptotischer Phasengang zugeordnet werden. Beträgt die Steigung des Amplitudengangs für hohe Frequenzen $-k \cdot 20 \text{ dB/Dekade}$, so muss die Phase asymptotisch $-k \cdot 90^\circ$ betragen. Dies ist im Beispiel von Abbildung 2.22 veranschaulicht. Dort sind im Amplitudengang die Knickfrequenzen eingezeichnet, und daraus entsprechend im Phasengang die Asymptoten. Bei $\omega = 100 \text{ rad/s}$ ist der Wert der Asymptoten -180° , und die tatsächliche Phase sollte auf jeden Fall über dieser Linie verlaufen. Tatsächlich liegt diese

aber $54,9^\circ$ unterhalb den -180° , und dies stellt damit eine untere Grenze für den Phasenverzug $\Delta\varphi_t$ dar, der durch die Totzeit verursacht wird, $\Delta\varphi_t > 54,9^\circ$. Damit lässt sich die Totzeit über

$$T_t = \frac{\Delta\varphi_t}{\omega} > \frac{54,9^\circ \cdot \frac{\pi \text{ rad}}{180^\circ}}{100 \frac{\text{rad}}{\text{s}}} = 9,6 \text{ ms}$$

nach unten abschätzen. (Tatsächlich wurde der Bode-Plot mit einer Totzeit von 10 ms erzeugt.)

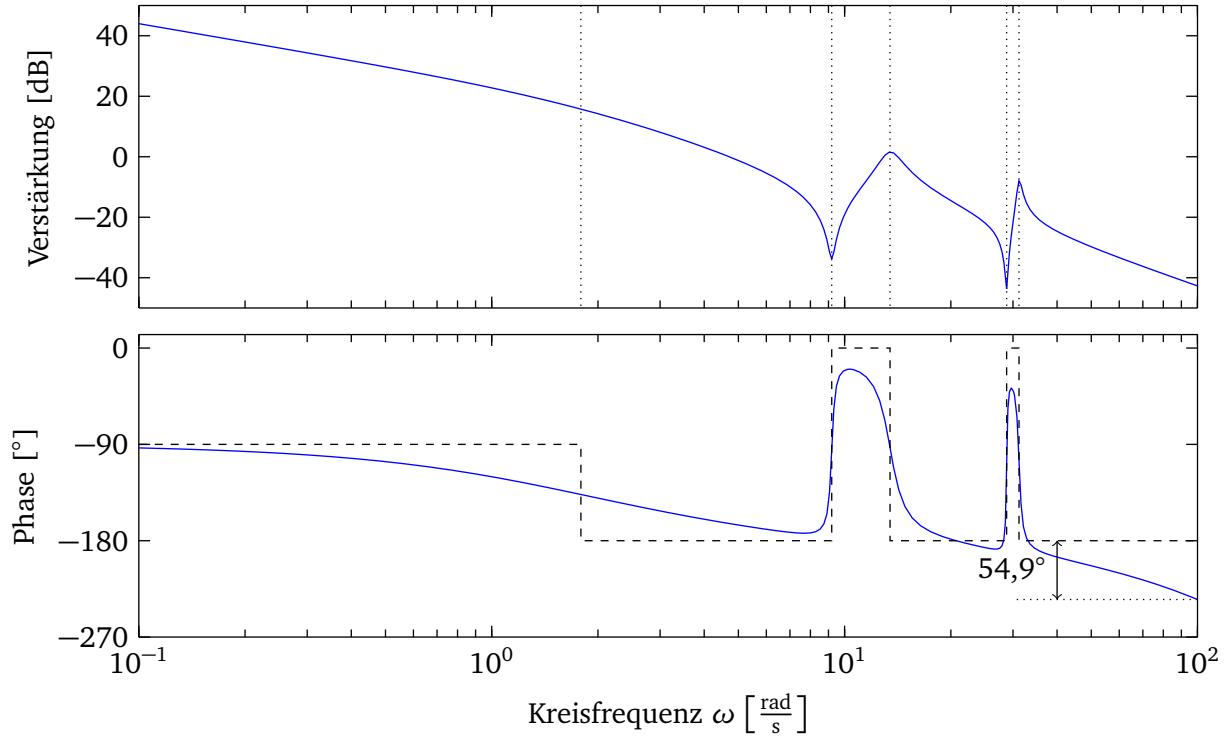


Abbildung 2.22: Frequenzgang eines Systems sechster Ordnung mit Totzeitglied

3 Minimierung des Ausgangsfehlers

In diesem Kapitel wird die Identifikation über die nichtlineare Optimierung des Ausgangsfehlers angeprochen. Dies ist vielleicht die naheliegendste Möglichkeit, insbesondere bei den heute vorhandenen Rechnern und Programmen, die Methoden zur nichtlinearen Optimierung bereitstellen und auch über genügend Rechenleistung verfügen.

Es wird sich aber auch zeigen, wo die Grenzen bzw. wo die Probleme damit liegen. Diese Art der Optimierung wird hierbei nur angerissen, da sich diese Vorlesung im Weiteren auf andere Methoden stützen wird.

3.1 Ansatz

Wenn eine Sprungantwort eines Systems vorliegt, von dem beispielsweise aus Vorüberlegungen ein PT₂-Verhalten zu erwarten ist, dann liegt es Nahe, ein Simulationsmodell eines PT₂-Gliedes mit der gleichen Eingangsgröße zu beaufschlagen, und die Parameter dann so einzustellen, dass die Abweichung

$$\tilde{e}(t) = \tilde{y}(t) - \hat{y}(t)$$

zwischen dem gemessenen Ausgang $\tilde{y}(t)$ und dem simulierten Ausgang $\hat{y}(t)$ in einer gewissen Weise minimal wird. Dieser Fehler wird auch als Ausgangsfehler bezeichnet. Verlangt man einen minimalen quadratischen Abstand, so entspräche dies dem Gütekriterium

$$J(\theta) = \int_0^{t_e} (\hat{y}(\tau, \theta) - \tilde{y}(\tau))^2 \cdot d\tau$$

bzw., wenn den abgetasteten Daten direkt Rechnung getragen werden soll,

$$J(\theta) = \sum_{k=0}^{N-1} (\hat{y}(t_k, \theta) - \tilde{y}(t_k))^2. \quad (3.1)$$

Der Schätzwert für die im Parametervektor θ zusammengefassten Daten ist dann jener Wert, bei dem das Gütekriterium den minimalen Wert annimmt,

$$\hat{\theta} = \arg \min_{\theta} J(\theta).$$

Das Gütekriterium kann auch im Frequenzbereich formuliert werden,

$$J(\theta) = \sum_{l=0}^{N-1} |\hat{H}(\omega_l, \theta) - \tilde{H}(\omega_l)|^2.$$

(Dies ist jetzt nicht zu der Formulierung im Zeitbereich äquivalent.)

Schon für einfachste Systeme ist dieses Optimierungsproblem nichtlinear in den Parametern, und es kann keine geschlossene Lösung angegeben werden. So lautet die Gütfunktion um ein PT₁-Glied anhand des gemessenen Ausgangsverlaufes $\tilde{y}(t_k)$ einer Sprungantwort ($u = u_0 \cdot \sigma(t)$) zu bestimmen

$$J([\hat{K}, \hat{T}]^T) = \sum_{k=0}^{N-1} \left(u_0 \cdot \hat{K} \cdot (1 - e^{-\frac{1}{T} \cdot t_k}) - \tilde{y}(t_k) \right)^2.$$

Prinzipiell kann man natürlich versuchen, das so entstehende Optimierungsproblem mit Methoden der nichtlinearen Optimierung zu lösen. Für Systeme niedriger Ordnung ist dies auch noch gut möglich. Im Folgenden werden anhand einiger Beispiele die Möglichkeiten aber – vor allem – auch die Probleme mit diesem Ansatz diskutiert.

3.2 Beispiele

3.2.1 PT₂-Glied

Ungestörtes System

In Abbildung 3.1 ist die ungestörte Sprungantwort eines PT₂-Gliedes ($K = 10$, $\omega_0 = 100$, $D = 0,1$) dargestellt.

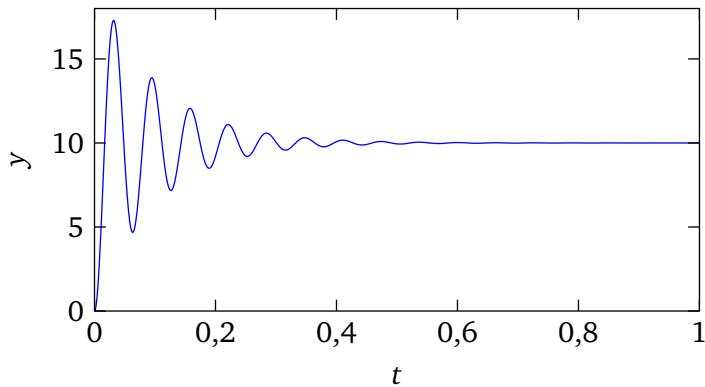


Abbildung 3.1: Sprungantwort eines PT₂-Gliedes mit $K = 10$, $\omega_0 = 100$, $D = 0,1$

Zur Optimierung ist die Sprungantwort mit $T_s = 0,001$ abgetastet, d. h. es liegen $N = 1000$ Messwerte vor.

Zu Vergleichszwecken ist die Gütfunktion (3.1) mit zwei Optimierungsmethoden minimiert. Zum einen das in der zum Standardumfang von Matlab gehörenden Funktion fminsearch implementierte Nelder-Meads-Simplex-Verfahren und zum anderen ein in fminunc verwendetes Quasi-Newton-Verfahren mit numerischer Bildung des Gradienten. Letztere Funktion ist Bestandteil der Optimization-Toolbox.

Für ein PT₂-Glied sind drei Parameter zu bestimmen. Verwendet man hier wieder die im vorherigen Kapitel eingeführte Kenngrößen K , ω_0 und D , so lautet der Parametervektor entsprechend

$$\boldsymbol{\theta} = [K \quad \omega_0 \quad D]^T.$$

Startet man die Optimierung anhand der unverrauschten Ausgangswerte mit den Anfangswerten

$$\boldsymbol{\theta}_{\text{init}} = [1 \quad 1 \quad 1]^T,$$

so ergibt sich mit dem Simplex-Verfahren die „optimale“ Lösung

$$\hat{\boldsymbol{\theta}}_{\text{Simplex}} = [\hat{K} = 10 \quad \hat{\omega}_0 = -100 \quad \hat{D} = -0,1]^T.$$

Es werden also die tatsächlichen Parameterwerte gefunden. Das ω_0 und D negativ geschätzt werden ist hier kein Fehler, sondern ein Freiheitsgrad der Darstellung 2.1. Mit dem verwendeten Quasi-Newton-Verfahren ergeben sich die Parameterwerte

$$\hat{\boldsymbol{\theta}}_{\text{QN}} = [\hat{K} = 10,1 \quad \hat{\omega}_0 = 30\,008 \quad \hat{D} = 115,1]^T.$$

Diese Parameter haben mit Ausnahme von \hat{K} offensichtlich wenig mit den tatsächlichen Parametern zu tun. Die Sprungantwort, die sich mit diesen Werten ergibt, ist in Abbildung 3.2 in rot dargestellt. Es zeigt

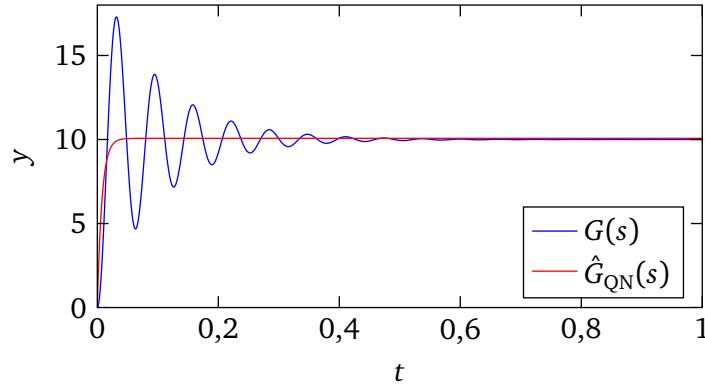


Abbildung 3.2: Sprungantwort eines PT_2 -Gliedes mit $K = 10$, $\omega_0 = 100$, $D = 0,1$

sich deutlich, dass die gefundenen Parameter auch nicht zu einer ähnlich aussehenden Sprungantwort führen.¹

Mit den Anfangswerten

$$\theta_{\text{init}} = [1 \ 10 \ 1]^T$$

konvergierten beide Optimierer gegen die tatsächliche Lösung.

Dieses Beispiel zeigt, dass auch bei sehr einfachen Systemen die Optimierung des Ausgangsfehlers schon auf einigermaßen passende Anfangswerte angewiesen sein kann. Startwerte könnten beispielsweise näherungsweise durch eine Kenngrößenermittlung wie in Kapitel 2 oder eine theoretische Modellbildung abgeschätzt werden.

Verrauschte Messdaten

In Abbildung 3.3 ist dieselbe Sprungantwort wie in Abbildung 3.1 gezeigt, diesmal jedoch mit einem Messrauschen überlagert.

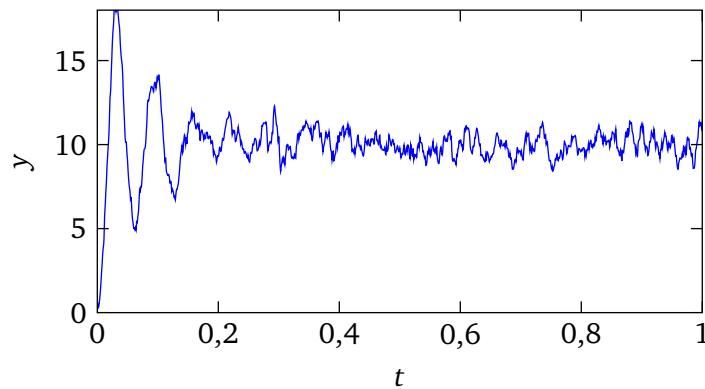


Abbildung 3.3: Verrauschte Sprungantwort eines PT_2 -Gliedes mit $K = 10$, $\omega_0 = 100$, $D = 0,1$

Bei der Optimierung zeigt sich, dass für das Quasi-Newton-Verfahren der Vektor der Anfangswerte noch etwas näher an den tatsächlichen Werten liegen muss,

$$\theta_{\text{init}} = [1 \ 50 \ 1]^T ,$$

¹ Dies soll hier nicht als Benchmark der verschiedenen Optimierer verstanden werden. Dies würde an dieser Stelle viel zu weit führen. Dies ist demnach hier als das Ergebnis eines einzelnen Falles anzusehen, und es soll nicht als Aussage über die prinzipielle Eignung des einen oder anderen Verfahrens verstanden werden. Es soll aber zeigen, dass verschiedene Optimierer für dasselbe nichtlineare Optimierungsproblem und denselben Startwerten durchaus deutlich verschiedene „Optimalwerte“ bestimmen können.

dann finden aber beide Verfahren die Lösung

$$\hat{\theta}_{\text{Simplex}} = \hat{\theta}_{\text{QN}} = [\hat{K} = 10,05 \quad \hat{\omega}_0 = 100,3 \quad \hat{D} = 0,1055]^T.$$

als Optimum. Der sich damit ergebende Verlauf ist in Abbildung 3.4 im Vergleich zu den verrauschten Daten sowie den unverrauschten Daten dargestellt.

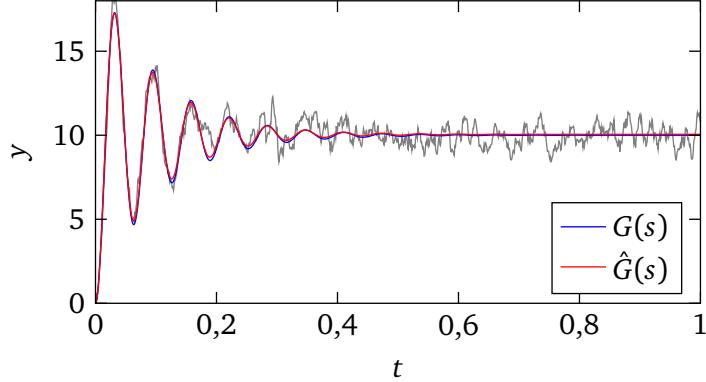


Abbildung 3.4: Sprungantwort eines PT_2 -Gliedes mit $K = 10$, $\omega_0 = 100$, $D = 0,1$

3.2.2 Drei-Massen-Schwinger

Als nächstes Beispiel wird der Drei-Massen-Schwinger mit dem Motormoment als Ein- und der Geschwindigkeit der dritten Trägheit als Ausgang betrachtet. Die hier als tatsächlich anzusehende Übertragungsfunktion ist mit Gl. (1.1) gegeben. Die Sprungantwort dieses Systems ist in Abbildung 3.5 dargestellt.

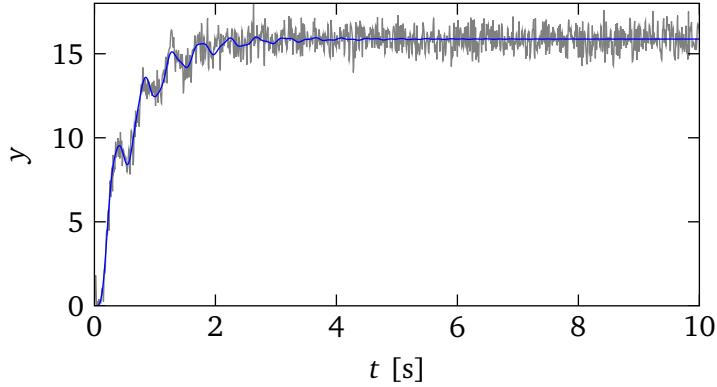


Abbildung 3.5: Sprungantwort des Drei-Massen-Schwingers

Hier könnte man aufgrund des Vorwissens über das System direkt ein System fünfter Ordnung ohne Nullstellen ansetzen,

$$G(s) = \frac{b_0}{s^5 + a_4 s^4 + a_3 s^3 + a_2 s^2 + a_1 s + a_0}.$$

Zur Optimierung müssen für die Koeffizienten Startwerte vorgegeben werden. Gäbe man einfach den Wert eins für alle Koeffizienten vor, dann würde sich das damit ergebende System instabil. Daher ist es günstiger Pole vorzugeben, und sich die Koeffizienten damit zu berechnen. Man könnte die Pole alle auf einen Standardwert wie z. B. -1 legen, aber es hat sich bei dem vorangegangenen Beispiel schon gezeigt,

dass es für die numerische Optimierung günstiger ist, Startwerte zu erhalten, die möglichst nahe an den Werten im Optimum liegen.

Hierzu ist so vorgegangen, dass zunächst ein PT_1 -Glied mit der Sprungantwort geschätzt wird. Das System zeigt ein mit Schwingungen überlagertes PT_1 -Verhalten, und so erhält man einen Anhaltswert von $0,56\text{ s}$ für die dominante Zeitkonstante und man schätzt die stationäre Verstärkung zu $15,86 \frac{\text{rad/s}}{\text{Nm}}$. Die Schwingungen besitzen eine ungefähre Periodendauer von $T_e = 0,5\text{ s}$, d. h. man erhält $\omega_e = 2\pi/T_e = 12,6 \frac{1}{\text{s}}$ als Anhaltswert für zumindest eine der Eigenfrequenzen. Die Schwingungen sind relativ schwach gedämpft, so dass der Realteil des dominanten Schwingungspaares zu $-0,25 \cdot \omega_e = -3,15 \frac{1}{\text{s}}$ abgeschätzt wird. Damit liegt eine grobe Einschätzung der Lage von drei der fünf Pole vor. Die verbleibenden beiden Pole werden gleich dem konjugiert-komplexen Polpaar gesetzt, und damit wird der Startvektor für die Optimierung anhand eines Systems mit den Polen $\{-1,781, -3,15 \pm j \cdot 12,6, -3,15 \pm j \cdot 12,6\}$ sowie der stationären Verstärkung $15,86$ bestimmt.

Für die Optimierung anhand der unverrauschten Messdaten finden damit beide hier verwendete Optimierer auch die Parameterwerte der gegebenen Übertragungsfunktion. Bei den verrauschten Messungen finden beide Optimierer das gleiche Optimum, wobei die dazugehörigen Pole sich etwas von den des tatsächlichen Systems unterscheiden. In diesem Fall ist es aber auch so, dass das Gütekriterium an der gefundenen Stelle einen geringeren Wert annimmt, als bei den tatsächlichen Parametern. Dies ist auf das Rauschen zurückzuführen.

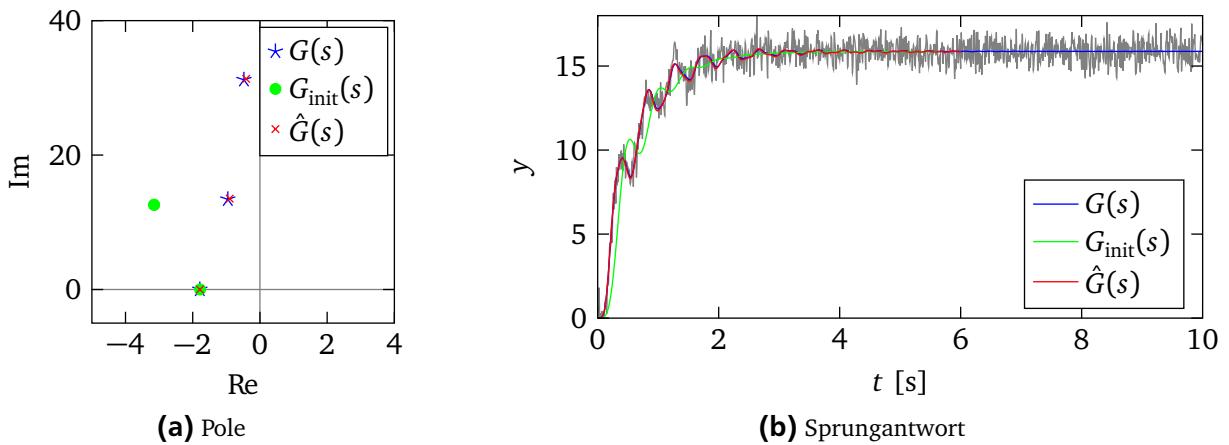


Abbildung 3.6: Ergebnis der Optimierung an den verrauschten Daten mit $n = 5$

Ansatz mit höherer Ordnung

Setzt man ein System siebter Ordnung ohne Nullstellen an,

$$G(s) = \frac{b_0}{s^7 + a_6s^6 + a_5s^5 + a_4s^4 + a_3s^3 + a_2s^2 + a_1s + a_0},$$

so liegt das tatsächliche System nicht mehr in der dadurch gegebenen Klasse. Jedoch könnte man ggf. erwarten, dass die beiden Pole, die dieses System zu viel besitzt, durch die Optimierung weit vom Ursprung entfernt geschoben werden, so dass diese keinen sichtbaren Einfluss auf die Sprungantwort mehr haben.

Zur Bestimmung des Anfangsvektors wurde einfach ein weiteres Polpaar bei $-3,15 \pm j \cdot 12,6$ vorgegeben. In Abbildung 3.7 sind die Ergebnisse bei der Optimierung mit den unverrauschten Daten gezeigt. Das Quasi-Newton-Verfahren belässt den zusätzlichen Pol in der Nähe und ist damit in ein lokales Minimum gelassen. Betrachtet man die Pollagen, so könnte man anhand derer nicht einfach sagen, dass das System eigentlich nur die Ordnung fünf besitzt. Das Simplexverfahren trifft die ursprünglichen Pole genau und

hat auch das überzählige Polpaar weit vom Ursprung weg geschoben, allerdings etwas in die rechte Halbebene.

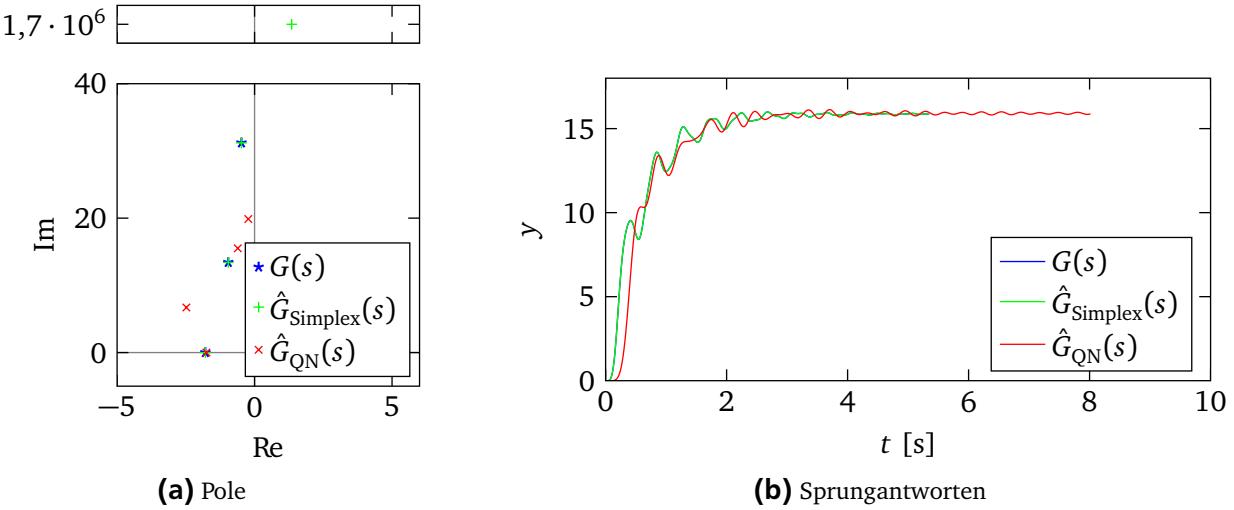


Abbildung 3.7: Ergebnis der Optimierung an den unverrauschten Daten mit $n = 7$ und $m = 0$

Allgemeiner Ansatz höherer Ordnung

Durch ein LZI-System mit n Polen und $m = n - 1$ Nullstellen,

$$G(s) = \frac{b_m s^m + b_{m-1} s^{m-1} + \dots + b_1 s + b_0}{s^n + a_{n-1} s^{n-1} + \dots + a_1 s + a_0}, \quad m \leq n,$$

kann durch eine geeignete Wahl der Koeffizienten a_i und b_i jedes durchgriffs- und totzeitfreie LZI-System bis zur Ordnung n dargestellt werden.

Wählt man im Fall des Drei-Massen-Schwingers den Ansatz $n = 7$ und $m = 6$,

$$G(s) = \frac{b_6 s^6 + b_5 s^5 + b_4 s^4 + b_3 s^3 + b_2 s^2 + b_1 s + b_0}{s^7 + a_6 s^6 + a_5 s^5 + a_4 s^4 + a_3 s^3 + a_2 s^2 + a_1 s + a_0},$$

so ergibt sich mit $a_1 = a_0 = 0$ und $b_3 = b_4 = b_5 = b_6 = 0$ sowie $b_1 = b_0 = 0$ wieder die Übertragungsfunktion eines Systems mit fünf Polen und keiner Nullstelle. Jedoch ist diese Wahl nicht eindeutig. Das $b_3 = b_4 = b_5 = b_6 = 0$ sein muss, ergibt sich aus der notwendigen Differenzenordnung von fünf. In der verbleibenden Übertragungsfunktion

$$G(s) = \frac{b_2 s^2 + b_1 s + b_0}{s^7 + a_6 s^6 + a_5 s^5 + a_4 s^4 + a_3 s^3 + a_2 s^2 + a_1 s + a_0}$$

müssen dann noch zwei Pol-Nullstellen-Kürzungen auftreten. Mit der Wahl $a_1 = a_0 = b_1 = b_0$ findet diese zwischen Pol- und Nullstellen in Null statt. Jedoch kann diese bei jedem beliebigen Wert auftreten. D.h. es bestehen von dieser Seite her Freiheitsgrade für die Optimierung.

Hierbei treten dann jedoch schon vierzehn Parameter auf, wobei die Lösung auch nicht eindeutig ist.

In Abbildung 3.8 ist das Ergebnis der Optimierung mit dem Quasi-Newton-Verfahren gezeigt. Es ein konjugiert-komplexes Polpaar ist gegen eine Nullstelle gekürzt (wenn auch in der rechten Halbebene), so dass zumindest die Systemordnung korrekt geschätzt wird. Das nicht-dominante Polpaar wird jedoch nicht getroffen, und die verbleibenden Nullstellen finden sich auch nicht weit vom Ursprung entfernt, sondern in der Nähe der Pole. Eine Nullstelle befindet sich auch in der rechten Halbebene, so dass das System als nicht-minimalphasig identifiziert wird.

Es zeigt sich damit, dass durch das Hinzufügen der Freiheitsgrade offenbar zusätzliche lokale Minima entstehen, die die Optimierung erschweren.

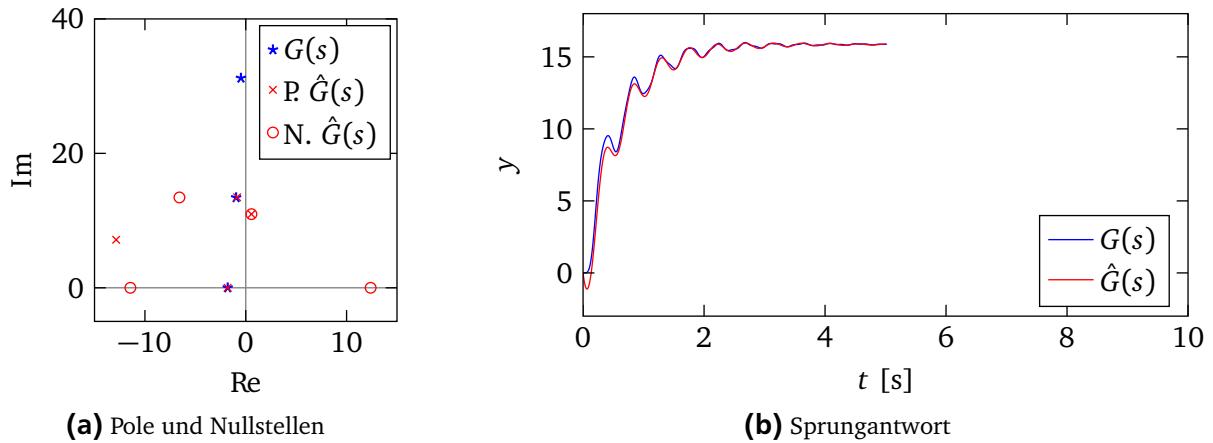


Abbildung 3.8: Ergebnis der Optimierung an den unverrauschten Daten mit $n = 7$ und $m = 6$

3.2.3 Doppelter Drei-Massen-Schwinger

Um noch ein Beispiel für ein System noch höherer Ordnung zu zeigen, wird einfach

$$G(s) = G_{\text{DMS}}(s) \cdot G_{\text{DMS}}(s)$$

verwendet, wobei $G_{\text{DMS}}(s)$ die Übertragungsfunktion des Drei-Massen-Schwingers vom Motormoment auf die Winkelgeschwindigkeit der dritten Trägheit ist.

Mit einem Ansatz $n = 10$ und $m = 0$ ergibt sich an den unverrauschten Daten mit dem Quasi-Newton-Verfahren die in Abbildung 3.9 gezeigten Lösungen. Dabei ist eine Lösung mit dem bisher verwendeten Pollagen (ebenfalls alle Pole verdoppelt) als Anfangswert bestimmt worden. Für die zweite Lösung sind die Startwerte der Pole mit 0,9 multipliziert worden.

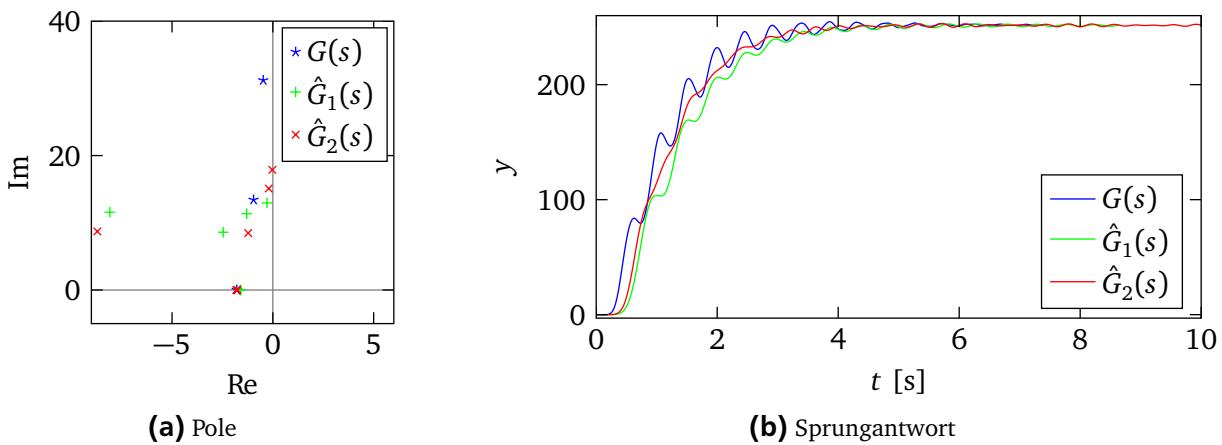


Abbildung 3.9: Ergebnis der Optimierung an den unverrauschten Daten mit $n = 10$ und zwei verschiedenen Anfangswerten

Es zeigt sich, dass die Optimierung immer kritischer wird. Bei dem einfachen Drei-Massen-Schwinger wurde das Minimum noch relativ sicher gefunden, hier hängt das gefundene Optimum schon sehr stark von den Anfangswerten ab, obwohl diese eigentlich vernünftig gewählt sind.

3.3 Fazit

Die Optimierung des Ausgangsfehlers führt auf ein nichtlineares Optimierungsproblem. Damit kann nicht garantiert werden², dass mit numerischen Methoden ein globales Optimum gefunden wird.

Zumindest für Systeme niedriger Ordnung stellt die Identifikation über die nichtlineare Optimierung aber eine Möglichkeit dar, deren Anwendung mit modernen Computerprogrammen auch wenig Zeit kostet, so dass es durchaus einen Versuch wert sein kann. Die Beispiele haben hier gezeigt, dass folgende Punkte beachtet werden sollten:

- Der Systemansatz sollte keine redundanten Parameter besitzen.
- Die Systemordnung sollte möglichst nicht überschätzt werden.
- Es sollte Vorwissen genutzt werden, um die Startwerte für die Parameter möglichst sinnvoll vorzugeben.
- Die Startwerte sollten variiert werden, um eine Abschätzung zu erhalten, wie „sicher“ die gefundenen Parameterwerte sind, bzw. um abzuschätzen, wie „wellig“ die Gütfunktion ist.
- Und vor allem sollten sowohl das Ergebnis als auch die ermittelten Parameterwerte kritisch betrachtet werden.

² Da im Allgemeinen auch nicht konvex

4 Grundlagen Stochastik

In diesem Abschnitt werden kurz die in diesem Skript verwendeten Notationen eingeführt und die wesentlichen, verwendeten Eigenschaften stochastischer Variablen und Prozesse zusammengestellt. Dabei wird auf eine mathematische Einführung in die Wahrscheinlichkeitsrechnung verzichtet. Eine solche ist z. B. in [Hänsler 2001] oder, mathematisch etwas strenger, in [Hoffmann, Marx und Vogt 2006, 631ff, 683ff] zu finden.

4.1 Zufallsvariablen

Eine Zufallsvariable, die hier in der Form

$$\{x\}_r$$

geschrieben wird, hat die Eigenschaft, dass sie bei jeder Wiederholung eines Experiments einen anderen Wert x annehmen kann, der nicht vorhersagbar ist. Der Wert x , den eine Zufallsvariable $\{x\}_r$ bei einer bestimmten Ausführung des Experiments annimmt, wird als Realisierung der Zufallsvariablen bezeichnet¹ und auch als

$$x = \{x\}_r^e$$

geschrieben², wenn in einem Ausdruck die Realisierungen von Zufallsvariablen deutlich gemacht werden sollen.

Die Wahrscheinlichkeit, dass eine Realisierung x bzw. $\{x\}_r^e$ einer Zufallsvariablen $\{x\}_r$ einen Wert kleiner oder gleich einer Konstanten a annimmt, wird hier als

$$P(\{x\}_r^e \leq a)$$

geschrieben.³ Falls, was hier meist der Fall ist, mit reellen (und nicht diskreten) Zufallsvariablen mit stetigen Verteilungsfunktionen gearbeitet wird, gilt $P(\{x\}_r^e \leq a) = P(\{x\}_r^e < a)$.

Wahrscheinlichkeitsverteilung

Die Verteilungsfunktion oder kurz Verteilung einer Zufallsvariablen $\{x\}_r$ wird mit $F_{\{x\}_r}(a)$ bezeichnet. Die Verteilung $F_{\{x\}_r}(a)$ gibt an, mit welcher Wahrscheinlichkeit eine Realisierung x der Zufallsvariablen $\{x\}_r$ einen Wert annimmt, der kleiner oder gleich a ist,

$$F_{\{x\}_r}(a) = P(\{x\}_r^e \leq a).$$

Im Folgenden wird das Argument häufig wie die (Realisierung der) Zufallsvariable benannt, also z. B. $F_{\{x\}_r}(x)$, obwohl damit nicht eine Realisierung von $\{x\}_r$, sondern ein bestimmter, vorgegebener Wert gemeint ist.

In Abbildung 4.1a sind die Verteilungsfunktionen für die Standardnormalverteilung und für eine Gleichverteilung auf dem Intervall $[-1, 3]$ dargestellt.

¹ Genauer ist die Realisierung der Wert, den die Zufallsvariable bei Eintreten eines bestimmten Ergebnis e annimmt. Und bei verschiedenen Ausführungen eines Experiments kann das gleiche Ergebnis auftreten.

² Zwischen dieser und der Notation von [Hänsler 2001] gilt damit: $\{x\}_r \Leftrightarrow x$ und $\{x\}_r^e \Leftrightarrow x(e)$.

³ Dies wird hier als etwas kompaktere Schreibweise für $P(e | \{x\}_r^e \leq a)$ verwendet.

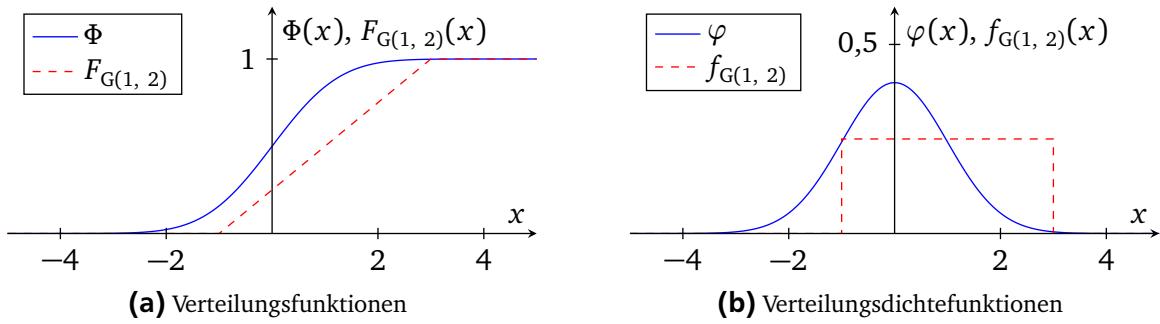


Abbildung 4.1: Standardnormalverteilung $\Phi \sim N(0, 1)$ und Gleichverteilung $G(1, 2)$

Die Verteilungsdichtefunktion ist die Ableitung der Verteilungsfunktion,

$$f_{\{x\}_r}(x) = \frac{dF_{\{x\}_r}(x)}{dx},$$

und damit gilt „rückwärts“

$$F_{\{x\}_r}(x) = \int_{-\infty}^x f_{\{x\}_r}(\xi) \cdot d\xi.$$

Da die Verteilung für $x \rightarrow \infty$ eins wird, muss auch $\int_{-\infty}^{\infty} f_{\{x\}_r}(\xi) \cdot d\xi = 1$ gelten.

In Abbildung 4.1b sind die Verteilungsdichtefunktionen zu den Verteilungen aus Abbildung 4.1a gezeigt.

Die Wahrscheinlichkeit, dass eine Realisierung x einer Zufallsvariable $\{x\}_r$ mit der Verteilungsdichte $f_{\{x\}_r}(x)$ einen Wert kleiner a annimmt, ist

$$P(\{x\}_r^e \leq a) = F_{\{x\}_r}(a) = \int_{-\infty}^a f_{\{x\}_r}(\xi) \cdot d\xi.$$

Dieses Integral ist in Abbildung 4.2a als rote Fläche dargestellt.

Die Wahrscheinlichkeit, dass eine Realisierung x einer Zufallsvariable $\{x\}_r$ mit der Verteilungsdichte $f_{\{x\}_r}(x)$ einen Wert größer a annimmt, ist

$$P(\{x\}_r^e > a) = 1 - F_{\{x\}_r}(a) = \int_a^{\infty} f_{\{x\}_r}(\xi) \cdot d\xi.$$

Dieses Integral ist in Abbildung 4.2b als rote Fläche dargestellt.

Die Wahrscheinlichkeit, dass eine Realisierung x einer Zufallsvariable $\{x\}_r$ mit der Verteilungsdichte $f_{\{x\}_r}(x)$ einen Wert im Intervall (a, b) annimmt, ist

$$P(a \leq \{x\}_r^e < b) = F_{\{x\}_r}(b) - F_{\{x\}_r}(a) = \int_a^b f_{\{x\}_r}(\xi) \cdot d\xi.$$

Dieses Integral ist in Abbildung 4.2c als rote Fläche dargestellt.

Besitzen zwei Zufallsvariablen $\{x\}_r$ und $\{y\}_r$ die gleiche Verteilung, so wird dies hier über

$$\{x\}_r \sim \{y\}_r$$

ausgedrückt. Dagegen soll die Schreibweise

$$\{x\}_r = \{y\}_r$$

bedeuten, dass für jedes Experiment bzw. jedes Ergebnis die Realisierungen $\{x\}_r^e$ und $\{y\}_r^e$ von $\{x\}_r$ bzw. $\{y\}_r$ den gleichen Wert annehmen, d. h. $\{x\}_r^e = \{y\}_r^e \forall e$. Damit impliziert $\{x\}_r = \{y\}_r$ auch $\{x\}_r \sim \{y\}_r$.

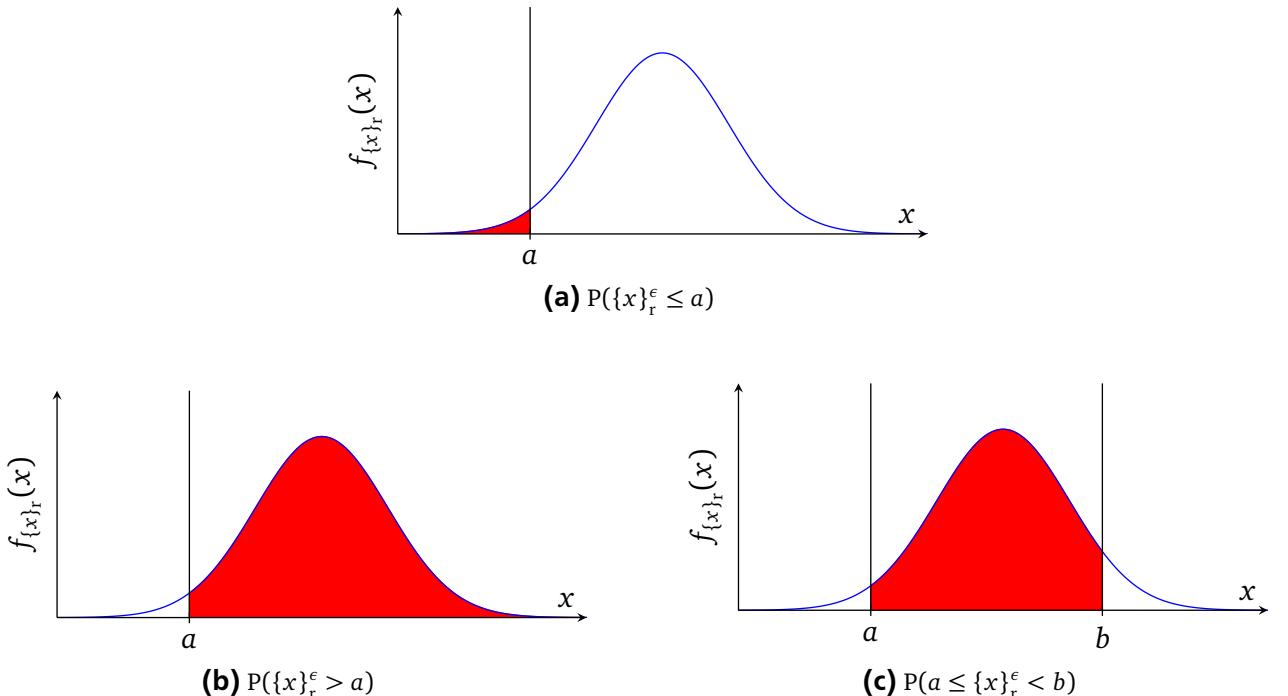


Abbildung 4.2: Darstellung der Wahrscheinlichkeiten dafür, dass $\{x\}_r^e$ in bestimmten Intervallen liegt

Gemeinsame Verteilungen

Betrachtet man zwei Zufallsvariablen $\{x\}_r$ und $\{y\}_r$, so kann man eine gemeinsame Verteilung

$$F_{\{x\}_r, \{y\}_r}(x, y) = P((\{x\}_r^e < x) \wedge (\{y\}_r^e < y))$$

angeben.⁴ Für die gemeinsame Verteilungsdichte zweier stetiger Zufallsvariablen gilt

$$f_{\{x\}_r, \{y\}_r}(x, y) = \frac{\partial^2 F_{\{x\}_r, \{y\}_r}(x, y)}{\partial x \partial y}$$

bzw.

$$F_{\{x\}_r, \{y\}_r}(x, y) = \int_{-\infty}^x \int_{-\infty}^y f_{\{x\}_r, \{y\}_r}(\xi, \eta) \cdot d\eta \cdot d\xi .$$

Integriert man über eine der Variablen, so erhält man die Randdichte der anderen Variablen,

$$f_{\{x\}_r}(x) = \int_{-\infty}^{\infty} f_{\{x\}_r, \{y\}_r}(x, y) \cdot dy .$$

Erwartungswert

Der Erwartungswert ist der Mittelwert aller Realisierungen und kann für eine einzelne stetige Zufallsvariable (die eine Verteilungsdichte besitzt) über

$$E\{x\}_r = \int_{-\infty}^{\infty} x \cdot f_{\{x\}_r}(x) \cdot dx$$

⁴ Eigentlich $P(\{\epsilon \mid \{x\}_r^e \leq x\} \cap \{\epsilon \mid \{y\}_r^e \leq y\})$.

berechnet werden. In dieser Arbeit wird der Erwartungswert auch mit

$$\mu_x = E\{x\}_r$$

abgekürzt.⁵

Der Erwartungswert der Zufallsvariablen $\{x\}_r$, dessen Verteilung über die gemeinsame Verteilungsdichte $f_{\{x\}_r, \{y\}_r}(x, y)$ gegeben ist, berechnet sich über⁶

$$E\{x\}_r = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x \cdot f_{\{x\}_r, \{y\}_r}(x, y) \cdot dy \cdot dx .$$

Eine wesentliche Eigenschaft der Erwartungswertbildung ist, dass diese linear ist, d. h. mit zwei beliebigen Zufallsvariablen $\{x\}_r$ und $\{y\}_r$ sowie der Konstanten a gilt

$$E\{ax\}_r = a \cdot E\{x\}_r$$

und

$$E\{x + y\}_r = E\{x\}_r + E\{y\}_r .$$

Hierbei ist hervorzuheben, dass diese Beziehungen immer gelten, unabhängig davon, welche Verteilungen $\{x\}_r$ und $\{y\}_r$ besitzen⁷, oder in welcher Beziehung diese zueinander stehen.

Als direkte Folge daraus ergibt sich, dass der Erwartungswert immer in Summen gezogen werden kann, d. h.

$$E\left\{\sum x_k\right\}_r = \sum E\{x_k\}_r ,$$

wobei bei unendlichen Summen gefordert sein kann, dass die Summen aller Realisierungen für sich konvergieren müssen.

Der Erwartungswert einer Zufallsvariablen, die sich aus der Anwendung einer Funktion $g(x)$ auf die Zufallsvariable $\{x\}_r$ ergibt, kann über

$$E\{g(x)\}_r = \int_{-\infty}^{\infty} g(x) \cdot f_{\{x\}_r}(x) \cdot dx$$

bestimmt werden.

Als Beispiel wird die im Intervall $[1, 3]$ gleichverteilte Zufallsvariable $\{x\}_r \sim G(2, 1)$ betrachtet. Deren Verteilungsdichte lautet

$$f_{\{x\}_r} = \begin{cases} \frac{1}{2} & 1 < x < 3 \\ 0 & \text{sonst.} \end{cases}$$

⁵ Eigentlich müsste es $\mu_{\{x\}_r}$ geschrieben werden, da der Erwartungswert eine Operation ist, die sinnvoll nur auf eine Zufallsvariable $\{x\}_r$ und nicht auf eine einzelne Realisierung $x = \{x\}_r^e$ angewandt werden kann. Bei der Verteilung und der Verteilungsdichte wird dies auch konsequent gemacht, bei den Größen wie Erwartungswert, Varianz, Korrelation etc. jedoch nicht, um die Indizes nicht zu überladen.

⁶ Für mehrdimensionale stetige Verteilungen treten entsprechend gemeinsame Verteilungen von mehr Zufallsvariablen auf.

⁷ Solange der Erwartungswert für $\{x\}_r$ und $\{y\}_r$ existiert.

Für den Erwartungswert erhält man damit erwartungsgemäß

$$E\{x\}_r = \int_{-\infty}^{\infty} x \cdot f_{\{x\}_r}(x) \cdot dx = \int_1^3 x \cdot \frac{1}{2} \cdot dx = 2.$$

Für den Erwartungswert der Zufallsvariablen $\{y\}_r$, die sich über $y = x^2$ aus der Zufallsvariablen $\{x\}_r$ ergibt, erhält man

$$E\{y\}_r = E\{x^2\}_r = \int_{-\infty}^{\infty} x^2 \cdot f_{\{x\}_r}(x) \cdot dx = \int_1^3 x^2 \cdot \frac{1}{2} \cdot dx = \frac{1}{2} \cdot \left[\frac{1}{3} \cdot x^3 \right]_1^3 = \frac{13}{3} = 4,333.$$

Ist $y = \frac{1}{x}$, dann gilt

$$E\{y\}_r = \int_{-\infty}^{\infty} \frac{1}{x} \cdot f_{\{x\}_r}(x) \cdot dx = \int_1^3 \frac{1}{x} \cdot \frac{1}{2} \cdot dx = \frac{1}{2} \cdot (\ln 3 - \ln 1) = 0,5493.$$

Varianz

Die Varianz einer Zufallsvariablen $\{x\}_r$ ist über

$$\text{Var}\{x\}_r = E\{(x - E\{x\}_r)^2\}_r$$

definiert. Es wird auch häufig die Abkürzung

$$\text{Var}\{x\}_r = \sigma_x^2$$

verwendet. Die Wurzel der Varianz wird als Standardabweichung

$$\sigma_x = \sqrt{\text{Var}\{x\}_r}$$

bezeichnet.

Betrachtet man zwei Zufallsvariablen $\{x\}_r$ und $\{y\}_r$, so kann auch die Kovarianz

$$\text{Cov}(\{x\}_r, \{y\}_r) = E\{(x - E\{x\}_r) \cdot (y - E\{y\}_r)\}_r$$

definiert werden. Aus dieser Definition folgt, dass die Reihenfolge unerheblich ist,

$$\text{Cov}(\{y\}_r, \{x\}_r) = \text{Cov}(\{x\}_r, \{y\}_r).$$

Für die Varianz der Zufallsvariablen $\{ax\}_r$ und $\{x + y\}_r$ ergibt sich damit⁸

$$\text{Var}\{ax\}_r = a^2 \cdot \text{Var}\{x\}_r$$

und

$$\text{Var}\{x + y\}_r = \text{Var}\{x\}_r + \text{Var}\{y\}_r + 2 \cdot \text{Cov}(\{x\}_r, \{y\}_r).$$

Ist eine Zufallsvariable $\{y\}_r$ durch

$$y = ax + \beta$$

mit $\text{Var}\{x\}_r = \sigma_x^2$ und festen Koeffizienten a und β gegeben, dann gilt $\text{Var}\{y\}_r = a^2 \sigma_x^2$ und

$$\begin{aligned} \text{Cov}(\{x\}_r, \{y\}_r) &= E\{(x - \mu_x) \cdot (y - \mu_y)\}_r = E\{(x - \mu_x) \cdot (ax - \beta - a\mu_x + \beta)\}_r \\ &= a \cdot E\{(x - \mu_x) \cdot (x - \mu_x)\}_r = a \cdot \text{Var}\{x\}_r. \end{aligned}$$

Für die Varianz von $\{x\}_r \sim G(2, 1)$ ergibt sich

$$\begin{aligned} \text{Var}\{x\}_r &= E\{(x - E\{x\}_r)^2\}_r = E\{(x - 2)^2\}_r = \int_{-\infty}^{\infty} (x - 2)^2 \cdot f_{\{x\}_r}(x) \cdot dx \\ &= \int_1^3 (x - 2)^2 \cdot \frac{1}{2} \cdot dx = \frac{1}{2} \cdot \left[\frac{1}{3} \cdot (x - 2)^3 \right]_1^3 = \frac{1}{3}. \end{aligned}$$

⁸ $\text{Var}\{ax\}_r = E\{(ax - a\mu_x)^2\}_r = a^2 \cdot E\{(x - \mu_x)^2\}_r = a^2 \cdot \text{Var}\{x\}_r$
 $\text{Var}\{x + y\}_r = E\{(x + y - \mu_x - \mu_y)^2\}_r = E\{(x - \mu_x)^2 + (y - \mu_y)^2 + 2 \cdot (x - \mu_x) \cdot (y - \mu_y)\}_r = \text{Var}\{x\}_r + \text{Var}\{y\}_r + 2 \cdot \text{Cov}(\{x\}_r, \{y\}_r)$

Mittlerer quadratischer Fehler

Für x kann auch immer $x + \mu_x - \mu_x$ geschrieben werden. Für den Erwartungswert einer quadrierten Zufallsvariablen $\{x\}_r$ ergibt sich so

$$E\{x^2\}_r = E\{(\mu_x + (x - \mu_x))^2\}_r.$$

Ausmultiplizieren führt dann auf

$$E\{x^2\}_r = E\{\mu_x^2 + 2 \cdot \mu_x \cdot (x - \mu_x) + (x - \mu_x)^2\}_r.$$

Der Erwartungs- bzw. Mittelwert μ_x ist keine Zufallsvariable, daher ist $E\{\mu_x^2\}_r = \mu_x^2 = (E\{x\}_r)^2$ und $E\{2 \cdot \mu_x \cdot (x - \mu_x)\}_r = 0$. $E\{(x - \mu_x)^2\}_r$ ist gerade die Varianz von $\{x\}_r$. Somit erhält man

$$E\{x^2\}_r = (E\{x\}_r)^2 + \text{Var}\{x\}_r.$$

Stellt $\{x\}_r$ einen Fehler oder eine sonstige Größe dar, die idealerweise den Wert null annimmt, dann stellt dessen quadratische Abweichung in vielen Fällen ein gutes Maß für diesen Fehler dar. Bei Stichproben wird der quadratische Fehler meist gemittelt, um den mittleren quadratischen Fehler anzugeben. Bei Zufallsgrößen entspricht die Erwartungswertbildung der Mittelung, so dass hier

$$\text{MSE}\{x\}_r = E\{x^2\}_r$$

festgelegt wird. Hierbei ist zu beachten, dass nicht zu jeder beliebigen Zufallsvariablen diese Festlegung interpretativ sinnvoll ist, sondern nur für Zufallsvariablen, die einen Fehler oder eine Abweichung darstellen.

Unkorrelierte und unabhängige Variablen

Kann man die Erwartungswertbildung und die Multiplikation zweier Zufallsvariablen $\{x\}_r$ und $\{y\}_r$ vertauschen,

$$E\{x \cdot y\}_r = E\{x\}_r \cdot E\{y\}_r,$$

so nennt man $\{x\}_r$ und $\{y\}_r$ unkorreliert.

Aus dieser Festlegung folgt, dass für unkorrelierte Zufallsvariablen $\{x\}_r$ und $\{y\}_r$ die Kovarianz verschwindet, $\text{Cov}(\{x\}_r, \{y\}_r) = 0$. Damit folgt weiter

$$\text{Var}\{x + y\}_r = \text{Var}\{x\}_r + \text{Var}\{y\}_r.$$

Der Korrelationskoeffizient ist über

$$\rho_{xy} = \frac{\text{Cov}(\{x\}_r, \{y\}_r)}{\sqrt{\text{Var}\{x\}_r \cdot \text{Var}\{y\}_r}}$$

gegeben und ist Null, wenn $\{x\}_r$ und $\{y\}_r$ unkorreliert sind. Sind beide Zufallsvariablen ideal (negativ) korreliert, so ist der Korrelationskoeffizient 1 (bzw. -1).

Betrachtet man wieder eine Zufallsvariablen $\{x\}_r$ und $y = \alpha x + \beta$, so ergibt sich

$$\rho_{xy} = \begin{cases} 1 & \alpha > 0 \\ -1 & \alpha < 0. \end{cases}$$

Zwei Zufallsvariablen werden statistisch unabhängig genannt, wenn

$$f_{\{x\}_r, \{y\}_r}(x, y) = f_{\{x\}_r}(x) \cdot f_{\{y\}_r}(y)$$

gilt. Unabhängigkeit ist eine stärkere Eigenschaft als Unkorreliertheit, d. h. zwei unabhängige Zufallsvariablen sind immer unkorreliert, die Umkehrung gilt aber nicht.

Es werden die beiden Zufallsvariablen $\{\tilde{x}_1\}_r$ und $\{\tilde{x}_2\}_r$ betrachtet, die sich aus einem konstanten Wert x sowie den unkorrelierten Zufallsvariablen $\{n_{x,1}\}_r$ bzw. $\{n_{x,2}\}_r$ zusammensetzen,

$$\tilde{x}_1 = x + n_{x,1},$$

$$\tilde{x}_2 = x + n_{x,2}.$$

Es gilt dann

$$\begin{aligned} E\{\tilde{x}_1 \cdot \tilde{x}_2\}_r &= E\{(x + n_1) \cdot (x + n_2)\}_r \\ &= x^2 + x \cdot E\{n_1\}_r + x \cdot E\{n_2\}_r + E\{n_1\}_r \cdot E\{n_2\}_r \\ &= E\{x + n_1\}_r \cdot E\{x + n_2\}_r \\ &= E\{\tilde{x}_1\}_r \cdot E\{\tilde{x}_2\}_r. \end{aligned}$$

Die beiden Zufallsvariablen $\{\tilde{x}_1\}_r$ und $\{\tilde{x}_2\}_r$ sind also ebenfalls unkorreliert. (Man kann sich aus diesem Beispiel auch merken, dass Gleichanteile nie zwei Zufallsvariablen korrelieren.⁹) Für die Kovarianz gilt damit

$$\text{Cov}(\{\tilde{x}_1\}_r, \{\tilde{x}_2\}_r) = 0.$$

Bedingte Verteilungen

Es seien die Zufallsvariablen $\{x\}_r$ und $\{y\}_r$ gegeben, die im Allgemeinen korreliert sind. Wenn die Verteilung von $\{x\}_r$ betrachtet wird, die unter der Bedingung vorliegt, dass die Realisierung der Zufallsvariable $\{y\}_r$ den (festen) Wert y_0 annimmt, so wird diese Verteilung ausführlich als

$$\{x|\{y\}_r^e = y_0\}_r$$

bezeichnet. Um eine etwas kürzere Notation zu erhalten und da Verwechslungen hier ausgeschlossen sind, wird im Weiteren kurz

$$\{x|y_0\}_r$$

geschrieben. Dies ist besonders zu beachten, wenn für die Zufallsvariable und Wert der speziellen Realisierung derselbe Name verwendet wird, d. h. $\{x|\{y\}_r^e = y\}_r = \{x|y\}_r$.

Für die abkürzenden Schreibweisen für Mittelwerte und Varianzen gilt dieselbe Notation entsprechend, d. h. $E\{x|y\}_r = \mu_{x|y}$ und $\text{Var}\{x|y\}_r = \sigma_{x|y}^2$.

Die bedingte Dichte der Verteilung $\{x|y_0\}_r$ ist über

$$f_{\{x|y_0\}_r} = \frac{f_{\{x\}_r, \{y\}_r}(x, y_0)}{f_{\{y\}_r}(y_0)}$$

gegeben. Sind $\{x\}_r$ und $\{y\}_r$ unabhängig, dann gilt

$$\{x|y\}_r = \{x\}_r.$$

⁹ Später wird noch von den Korrelationsfunktionen von Zufallsprozessen die Rede sein. Unkorreliertheit bedeutet dann nicht, dass die Korrelationsfunktion den Wert null annimmt, sondern dass die Kovarianzfunktion den Wert null annimmt. Addiert man auf die dabei betrachteten Signale jeweils ein beliebiges konstantes Signal, so verändert dies den Wert der Korrelationsfunktion, aber nicht den der Kovarianzfunktion.

Vektorielle Zufallsvariablen

Eine vektorielle Zufallsvariable ist zunächst eine einfache Zusammenfassung mehrerer Zufallsvariablen in einem Vektor. Damit ändert sich auch nichts an der Linearität des Erwartungswertes, d. h. für zwei vektorielle Zufallsvariablen $\{\mathbf{x}\}_r$ und $\{\mathbf{y}\}_r$ sowie der konstanten Matrix \mathbf{A} gilt

$$E\{\mathbf{Ax}\}_r = \mathbf{A} \cdot E\{\mathbf{x}\}_r$$

und

$$E\{\mathbf{x} + \mathbf{y}\}_r = E\{\mathbf{x}\}_r + E\{\mathbf{y}\}_r.$$

Für eine n -dimensionale vektorielle Zufallsvariable $\{\mathbf{x}\}_r$ ist die Kovarianzmatrix

$$\text{Cov}(\{\mathbf{x}\}_r, \{\mathbf{x}\}_r) = E\{(\mathbf{x} - E\{\mathbf{x}\}_r) \cdot (\mathbf{x} - E\{\mathbf{x}\}_r)^T\}_r = \begin{bmatrix} \text{Var}\{x_1\}_r & \cdots & \text{Cov}(\{x_1\}_r, \{x_n\}_r) \\ \vdots & \ddots & \vdots \\ \text{Cov}(\{x_n\}_r, \{x_1\}_r) & \cdots & \text{Var}\{x_n\}_r \end{bmatrix}$$

definiert. Für $\text{Cov}(\{\mathbf{x}\}_r, \{\mathbf{x}\}_r)$ kann auch kürzer $\text{Cov}\{\mathbf{x}\}_r$ oder C_{xx} geschrieben werden. Sind die einzelnen in \mathbf{x} zusammengefassten Zufallsvariablen unkorreliert, so ist $\text{Cov}\{\mathbf{x}\}_r$ eine Diagonalmatrix.

Allgemein kann die Kovarianzmatrix auch für zwei vektorielle Zufallsvariablen $\{\mathbf{x}\}_r$ und $\{\mathbf{y}\}_r$ angegeben werden,

$$\text{Cov}(\{\mathbf{x}\}_r, \{\mathbf{y}\}_r) = E\{(\mathbf{x} - E\{\mathbf{x}\}_r) \cdot (\mathbf{y} - E\{\mathbf{y}\}_r)^T\}_r = \begin{bmatrix} \text{Cov}(\{x_1\}_r, \{y_1\}_r) & \cdots & \text{Cov}(\{x_1\}_r, \{y_m\}_r) \\ \vdots & \ddots & \vdots \\ \text{Cov}(\{x_n\}_r, \{y_1\}_r) & \cdots & \text{Cov}(\{x_n\}_r, \{y_m\}_r) \end{bmatrix}.$$

Hierbei ist $\{\mathbf{y}\}_r$ eine m -dimensionale vektorielle Zufallsvariable. Aus dieser Definition folgen mit den konstanten Matrizen \mathbf{A} und \mathbf{B} die Zusammenhänge

$$\text{Cov}(\{\mathbf{y}\}_r, \{\mathbf{x}\}_r) = (\text{Cov}(\{\mathbf{x}\}_r, \{\mathbf{y}\}_r))^T,$$

$$\text{Cov}\{\mathbf{Ax}\}_r = \mathbf{A} \cdot \text{Cov}\{\mathbf{x}\}_r \cdot \mathbf{A}^T,$$

$$\text{Cov}(\{\mathbf{Ax}\}_r, \{\mathbf{By}\}_r) = \mathbf{A} \cdot \text{Cov}(\{\mathbf{x}\}_r, \{\mathbf{y}\}_r) \cdot \mathbf{B}^T \quad \text{und}$$

$$\text{Cov}\{\mathbf{x} + \mathbf{y}\}_r = \text{Cov}\{\mathbf{x}\}_r + \text{Cov}\{\mathbf{y}\}_r + \text{Cov}(\{\mathbf{x}\}_r, \{\mathbf{y}\}_r) + \text{Cov}(\{\mathbf{y}\}_r, \{\mathbf{x}\}_r)$$

Eine alternative, abkürzende Schreibweise für die Kovarianz ist $\text{Cov}(\{\mathbf{x}\}_r, \{\mathbf{y}\}_r) = C_{xy}$.

Für den Erwartungswert des äußeren Produktes zweier vektorieller Zufallsvariablen $\{\mathbf{x}\}_r$ und $\{\mathbf{y}\}_r$ gilt

$$\begin{aligned} E\{\mathbf{x} \cdot \mathbf{y}^T\}_r &= E\{((\mathbf{x} - \mu_x) + \mu_x) \cdot ((\mathbf{y}^T - \mu_y^T) + \mu_y^T)\}_r \\ &= E\{(\mathbf{x} - \mu_x) \cdot (\mathbf{y}^T - \mu_y^T)\}_r + E\{(\mathbf{x} - \mu_x)\}_r \cdot \mu_y^T + \mu_x \cdot E\{(\mathbf{y}^T - \mu_y^T)\}_r + \mu_x \cdot \mu_y^T \\ &= \text{Cov}(\{\mathbf{x}\}_r, \{\mathbf{y}\}_r) + E\{\mathbf{x}\}_r \cdot E\{\mathbf{y}^T\}_r. \end{aligned} \tag{4.1}$$

4.2 Ausgewählte Verteilungen

4.2.1 Normalverteilung

Die Verteilungsdichte $\varphi(x; \mu, \sigma^2) = f_{N(\mu, \sigma^2)}(x)$ der Normalverteilung $N(\mu, \sigma^2)$ ist als

$$\varphi(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

definiert. Damit ergibt sich für die Verteilungsfunktion $\Phi(x; \mu, \sigma^2) = F_{N(\mu, \sigma^2)}(x)$

$$\Phi(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma^2} \int_{-\infty}^x e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt.$$

Mit der Abkürzung

$$\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-\tau^2} d\tau$$

erhält man auch

$$\Phi(x; \mu, \sigma^2) = \frac{1}{2} \left(1 + \operatorname{erf} \left(\frac{x-\mu}{\sqrt{2}\sigma} \right) \right).$$

Die Normalverteilung $N(\mu, \sigma^2)$ wird über deren Erwartungswert μ und Varianz σ^2 vollständig beschrieben.

Standardnormalverteilung

Die Standardnormalverteilung ist eine Normalverteilung mit dem Mittelwert $\mu = 0$ und der Varianz $\sigma^2 = 1$ ($N(0, 1)$),

$$\varphi(x) = \varphi(x; 0, 1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}},$$

$$\Phi(x) = \Phi(x; 0, 1) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt.$$

Die Verteilungsdichtefunktion und die Verteilungsfunktion sind in Abbildung 4.3 dargestellt.

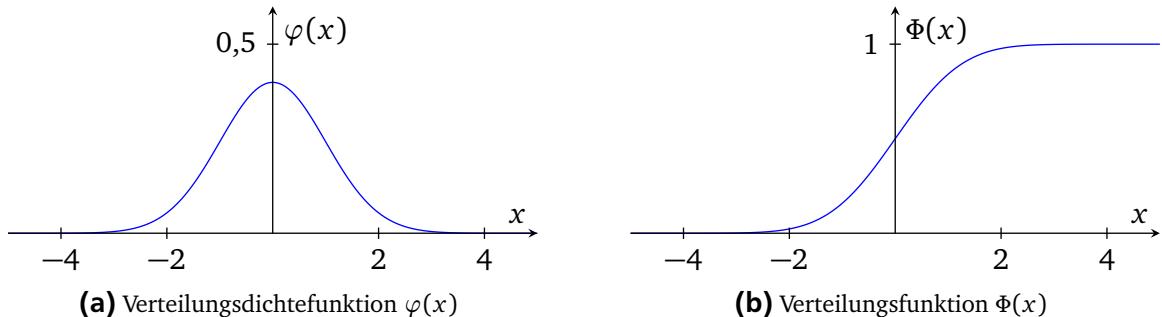


Abbildung 4.3: Standardnormalverteilung $N(0, 1)$

Normalverteilung

Die Verteilungsdichtefunktion und Verteilungsfunktion einer beliebigen Normalverteilung $N(\mu, \sigma)$ kann über

$$\varphi(x; \mu, \sigma^2) = \frac{1}{\sigma} \cdot \varphi \left(\frac{x-\mu}{\sigma} \right)$$

und

$$\Phi(x; \mu, \sigma^2) = \Phi \left(\frac{x-\mu}{\sigma} \right)$$

aus den jeweiligen Funktionen der Standardnormalverteilung berechnet werden.¹⁰

Für folgende speziellen Intervalle sind die Wahrscheinlichkeiten $P(|\{x\}_r - \mu| < c)$ hier angegeben:

68,27 % aller Werte liegen in einer $\pm\sigma$ -Umgebung von μ

95,45 % aller Werte liegen in einer $\pm 2\sigma$ -Umgebung von μ

99,73 % aller Werte liegen in einer $\pm 3\sigma$ -Umgebung von μ

Summen normalverteilter Zufallsvariablen

Eine wesentliche Eigenschaft der Normalverteilung ist, dass das Produkt einer normalverteilten Zufallsvariablen mit einer Konstanten sowie die Summe von zwei normalverteilten Zufallsvariablen jeweils wieder eine normalverteilte Zufallsvariable ergibt. Mit den oben angegebenen allgemeinen Beziehungen für den Erwartungswert und die Varianz ergibt sich Folgendes:

- Wenn $\{x\}_r \sim N(\mu_x, \sigma_x^2)$, dann gilt

$$\{a \cdot x\}_r \sim N(a \cdot \mu_x, a^2 \cdot \sigma_x^2). \quad (4.2)$$

- Wenn $\{x\}_r \sim N(\mu_x, \sigma_x^2)$ und $\{y\}_r \sim N(\mu_y, \sigma_y^2)$, dann gilt

$$\{x + y\}_r \sim N(\mu_x + \mu_y, \sigma_x^2 + 2\sigma_x\sigma_y\rho_{xy} + \sigma_y^2),$$

mit dem Korrelationskoeffizienten ρ_{xy} . Für unkorrelierte Zufallsvariablen gilt $\rho_{xy} = 0$ und damit

$$\{x + y\}_r \sim N(\mu_x + \mu_y, \sigma_x^2 + \sigma_y^2). \quad (4.3)$$

4.2.2 Gleichverteilung

Bei einer Gleich- oder Rechtecksverteilung $\{x\}_r \sim G(\mu, a)$ kann x nur Werte im Intervall $[\mu - a, \mu + a]$ annehmen, wobei die Wahrscheinlichkeitsdichte innerhalb dieses Intervalls konstant ist.

In dem in Abbildung 4.1 dargestellten Beispiel ist $\{x\}_r \sim G(1, 2)$, und damit hat die Wahrscheinlichkeitsdichte im Intervall $[-1, 3]$ den konstanten Wert $1/4$, und sie ist Null außerhalb dieses Intervalls.

Allgemein gilt für die Verteilungsfunktion

$$F_{G(\mu, a)}(x) = \begin{cases} 0 & \text{für } x < \mu - a \\ \frac{1}{2a} \cdot (x + a) & \text{für } \mu - a \leq x \leq \mu + a \\ 1 & \text{für } x > \mu + a \end{cases}$$

und für die Verteilungsdichte

$$f_{G(\mu, a)}(x) = \begin{cases} 0 & \text{für } x < \mu - a \vee x > \mu + a \\ \frac{1}{2a} & \text{für } \mu - a \leq x \leq \mu + a. \end{cases}$$

Damit ergibt sich auch der Mittelwert und die Varianz von $\{x\}_r \sim G(\mu, a)$ zu

$$E\{x\}_r = \mu \quad \text{und}$$

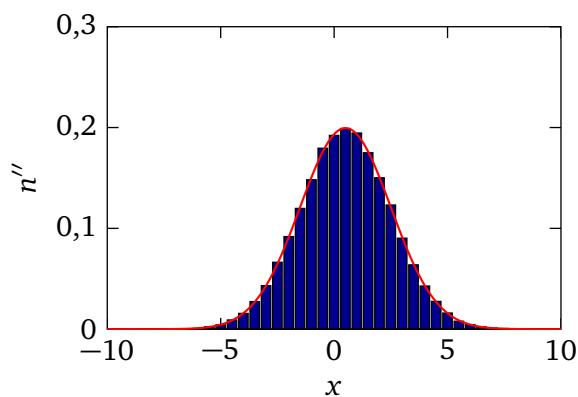
$$\text{Var}\{x\}_r = \int_{\mu-a}^{\mu+a} (x - \mu)^2 \cdot \frac{1}{2a} \cdot dx = \int_{-\mu}^{\mu} \tilde{x}^2 \cdot \frac{1}{2a} \cdot d\tilde{x} = \frac{1}{3}a^2.$$

¹⁰ $\Phi\left(\frac{x-\mu}{\sigma}\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x-\mu}{\sigma}} e^{-\frac{t^2}{2}} dt = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\frac{x-\mu}{\sigma\sigma}} e^{-\frac{(u-\mu)^2}{2\sigma^2}} du$ mit der Substitution $u = \sigma \cdot t + \mu$.

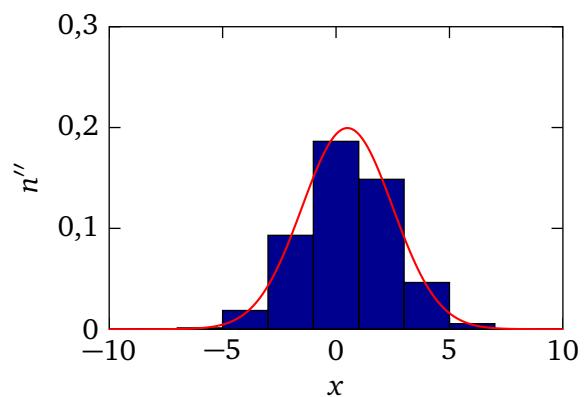
4.2.3 Simulation und Messung von Zufallsvariablen

Histogramm

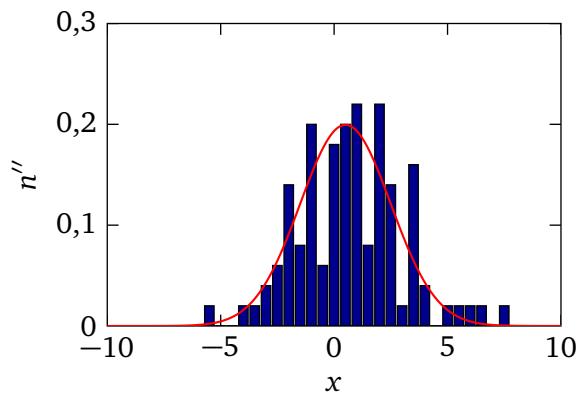
Werden viele Realisierungen einer Zufallsvariablen simulativ oder messtechnisch beobachtet, dann können die erhaltenen Werte als Histogramm dargestellt werden. Dabei wird die Menge an Werten, die die Zufallsvariable annehmen kann, in Intervalle unterteilt und für jedes der Intervalle die Anzahl der beobachteten Realisierungen gezählt. Diese werden dann als Höhe eines Rechtecks dargestellt, dessen Grundseite dem jeweiligen Intervall entspricht.



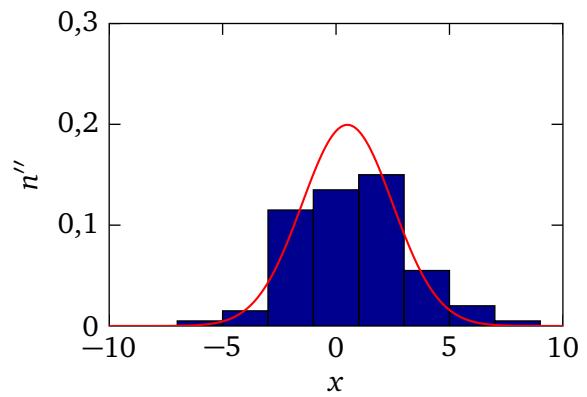
(a) $N = 100\,000, B = 41$



(b) $N = 100\,000, B = 11$



(c) $N = 100, B = 41$



(d) $N = 100, B = 11$

Abbildung 4.4: Histogramme (In Rot ist die Verteilungsdichte der Normalverteilung $N(0,5, 4)$ dargestellt.)

Es werden N Realisierungen berücksichtigt und die Ergebnismenge in B Intervalle mit den Grenzen x_b , $b = 0, \dots, B$ aufgeteilt. Die Intervalle müssen dabei nicht die vollständige Ergebnismenge abdecken, sondern nur jene Untermenge, die alle berücksichtigten Realisierungen umfasst. D. h. alle beobachteten Ergebnisse sollten im Intervall $[x_0, x_B]$ liegen.

Dann ist

$$n_b = \text{Anzahl der beobachteten Realisierungen } (x_k) \text{ für die gilt: } (x_b \leq x_k < x_{b+1}).$$

Diese Anzahl n_b kann direkt oder geteilt durch die Gesamtanzahl N als

$$n'_b = n_b/N$$

in einem Histogramm dargestellt werden. Will man diese experimentell ermittelte „Verteilungsdichte“ mit einer theoretisch erwarteten vergleichen, bietet es sich noch an, jeweils durch die Breite der Intervalle zu teilen,

$$n_b'' = \frac{n_b}{N} \cdot \frac{1}{x_{b+1} - x_b} .$$

Dadurch ist das durch die Rechtecke angenäherte Integral über alle Intervalle eins, womit das Histogramm als Verteilungsdichte skaliert ist und sich gut mit tatsächlichen Verteilungsdichten vergleichen lässt.

In Abbildung 4.4 sind Beispiele von Histogrammen gezeigt. Denen liegt jeweils eine (am Rechner simulierte) Normalverteilung mit dem Mittelwert 0,5 und der Varianz vier zu Grunde. In den ersten beiden Teilplots sind 100 000 Realisierungen simuliert und als Histogramm mit 41 bzw. 11 Intervallen dargestellt. Zusätzlich ist in rot die „wahre“ Verteilung angegeben. In den Teilplots (c) und (d) sind nur 100 Realisierungen simuliert. Es zeigt sich, dass die Anzahl B der Intervalle sinnvoll in Abhängigkeit der Anzahl N der betrachteten Grundmenge abgestimmt sein sollte. Für sehr große N wird ein großes B ermöglicht und man kann die Histogramme in erster Näherung als simulativ bzw. experimentell bestimmte Verteilungsdichten ansehen.

Einfacher ist die Darstellung der simulierten oder beobachteten Daten als Verteilungsfunktion, da diese Darstellung wenig kritisch gegenüber der gewählten Intervallanzahl B ist, jedoch lassen sich die Verteilungen optisch besser an der Verteilungsdichte erkennen bzw. beurteilen.

Reproduzierbare Ergebnisse

Die Betrachtung von Zufallsvariablen bringt es mit sich, dass die Ergebnisse nicht exakt vorhersagbar sind. Für Simulationen ist es aber zum Teil praktisch, diese genau reproduzieren zu können. Computerprogramme stellen daher die Möglichkeit zur Verfügung, die Generatoren der Zufallszahlen zu initialisieren.

In (dem Release 2013b von) Matlab beispielsweise kann dies über den Befehl `rng(seed)` gemacht werden. Die Variable `seed` steht für eine Zahl, die zur Initialisierung verwendet wird. Verwendet man denselben Wert für `seed`, dann werden nach Ausführen dieses Befehls immer wieder die gleichen Zufallszahlen erzeugt.

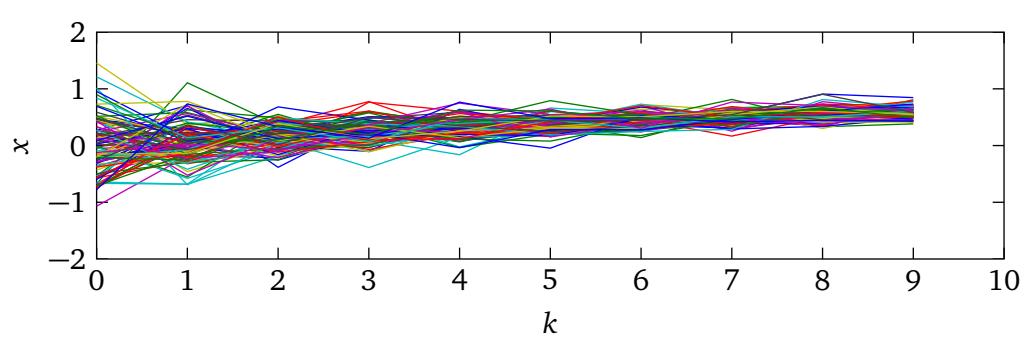
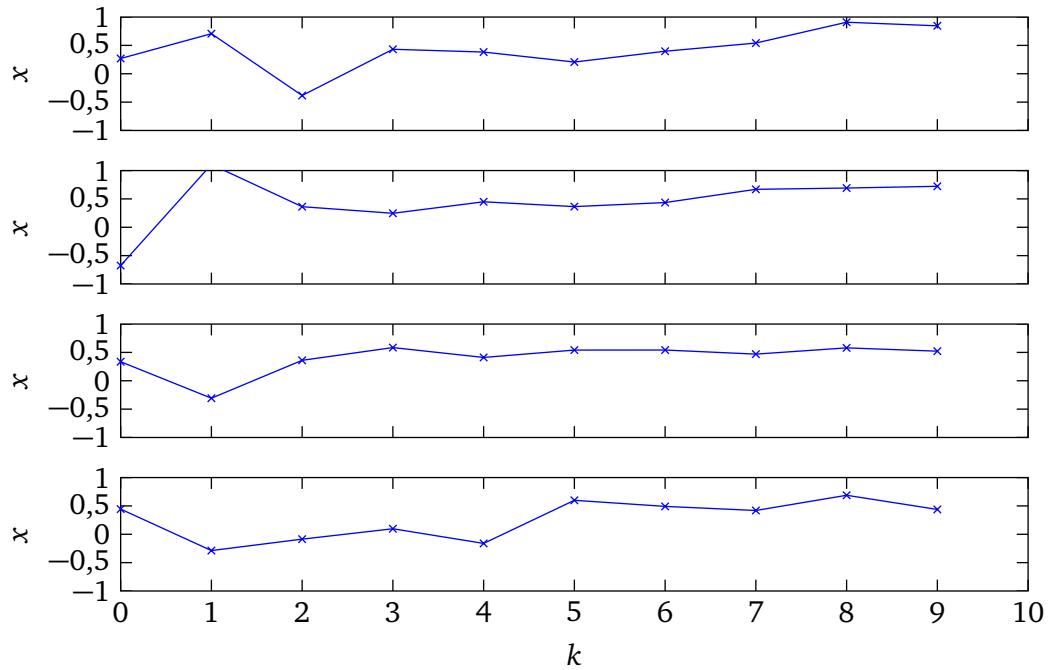
4.3 Zufallsprozesse

Bei diskreten Zufallsprozessen $\{x(t)\}_t$ oder $\{x_k\}_t$ gehört zu jeder Realisierung nicht ein einzelner Wert, sondern eine Funktion $x(t)$ bzw. eine Folge von Variablen x_k .

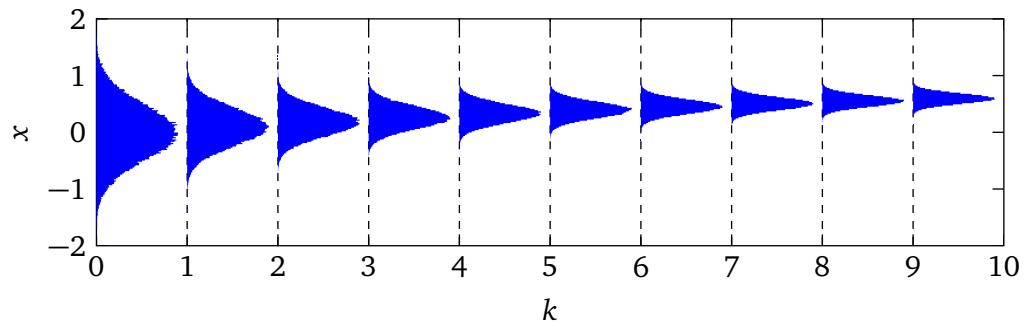
Für ein festes t bzw. k ergibt sich damit eine Zufallsvariable, für die auch die schon vorgestellten Größen wie z. B. der Erwartungswert und die Varianz definiert sind.

In Abbildung 4.5a sind beispielhaft vier Realisierungen eines (simulierten) diskreten Zufallsprozesses dargestellt. Die Beurteilung des Prozesses ist anhand dieser wenigen Realisierungen schwer möglich. Plottet man viele Realisierungen in ein Achsenystem wie in Abbildung 4.5b, dann lassen sich zumindest grob Tendenzen bezüglich des Verlaufs der Mittelwerte und der Streuung feststellen.

Günstiger ist aber die Darstellung der Verteilungen der einzelnen Zufallsvariablen $\{x_k\}_t$ zu jedem Zeitpunkt k . Dies geschieht in Abbildung 4.5c durch die Darstellung der Histogramme für die Zufallsvariablen zu den einzelnen Zeitpunkten k , wobei die Histogramme um 90° auf die Seite gedreht sind. Die Histogramme sind so skaliert, dass der Maximalwert zu jedem k 90 % des möglichen Platzes zum nächsten Zeitpunkt ausfüllt.



(b) Plot mehrerer Realisierungen ($N = 100$)



(c) Darstellung der Verteilungen der Zufallsvariablen $\{x_k\}_r$ über Histogramme ($N = 100\,000$)

Abbildung 4.5: Darstellung eines diskreten Zufallprozesses $\{x_k\}_r$

Diese Darstellung gibt einen sehr guten Eindruck von der Entwicklung der wesentlichen stochastischen Parameter der Zufallsvariablen $\{x_k\}_r$ eines Prozesses, dafür enthält sie keine Informationen über den Zusammenhang der Werte einer Realisierung zu verschiedenen Zeitpunkten.

Korreliertheit

Zwei Zufallsprozesse sind unkorreliert, wenn

$$E\{x_k \cdot y_l\}_r = E\{x_k\}_r \cdot E\{y_l\}_r \quad \forall k, l$$

gilt.

Des Weiteren wird hier auch von *einem* unkorrelierten Zufallsprozess gesprochen, wenn die zu zwei unterschiedlichen Zeitpunkten gehörenden Zufallsvariablen $\{x_k\}_r$ und $\{x_l\}_r$ eines Zufallsprozesses unkorreliert sind, d. h. wenn

$$E\{x_k \cdot x_l\}_r = E\{x_k\}_r \cdot E\{x_l\}_r \quad \text{für } k \neq l$$

gilt.

Weißes Rauschen

Als (diskretes) „Weißes Rauschen“ wird ein Prozess bezeichnet, dessen einzelne Verteilungen $\{x_k\}_r$ zu den Zeitpunkten k unabhängig und identisch verteilt sind.¹¹ Aus der Unabhängigkeit folgt

$$\text{Cov}(\{x_k\}_r, \{x_l\}_r) = \begin{cases} \text{Var}\{x_k\}_r & \text{für } k = l \\ 0 & \text{für } k \neq l, \end{cases}$$

wobei aufgrund der identischen Verteilung $\text{Var}\{x_k\}_r$ für alle k denselben Wert besitzt.

Sind darüber hinaus die einzelnen Variablen $\{x_k\}_r$ normalverteilt, so spricht man von einem „Gauß'schen Weissen Rauschen“.

Vektorielle Zufallsprozesse

Wie bei Zufallsvariablen kann auch ein Zufallsprozess aus vektoriellen Variablen bestehen. In diesem Kontext werden die Bezeichnungen

$$\text{Cov}(\{\mathbf{x}_k\}_r, \{\mathbf{x}_l\}_r) = C_{\mathbf{xx}, k, l} = E\{(\mathbf{x}_k - E\{\mathbf{x}_k\}_r) \cdot (\mathbf{x}_l - E\{\mathbf{x}_l\}_r)^T\}_r$$

und

$$\text{Cov}(\{\mathbf{x}_k\}_r, \{\mathbf{y}_l\}_r) = C_{\mathbf{xy}, k, l} = E\{(\mathbf{x}_k - E\{\mathbf{x}_k\}_r) \cdot (\mathbf{y}_l - E\{\mathbf{y}_l\}_r)^T\}_r$$

verwendet.

Später treten häufig nur noch Kovarianzmatrizen in Gleichungen auf, für die beide Zeitpunkte gleich sind, $l = k$. In diesen Fällen wird dann kürzer

$$\mathbf{P}_{\mathbf{x}, k} = C_{\mathbf{xx}, k, k}$$

geschrieben. Diese Matrix wird auch kurz als Kovarianzmatrix (zum Zeitpunkt k) von $\{\mathbf{x}_k\}_r$ bezeichnet. Diese wird mit einem eigenen Formelzeichen geschrieben, um Verwechslungen mit der einparametrischen Kovarianz $C_{\mathbf{xy}, k}$ zu vermeiden, die analog der einparametrischen Korrelation für stationäre Prozesse definiert werden kann.

¹¹ Im Englischen werden diese Eigenschaften als *i. i. d.* – independent identically distributed – bezeichnet. Das „identically“ darf nicht mit „gleich“ übersetzt werden, da dieser Begriff im Deutschen schon für eine spezielle Verteilungsform steht. Die (stetige) Gleich- oder Rechteckverteilung heißt im Englischen (*continuous*) *uniform distribution* oder *rectangular distribution*.

Stationäre Prozesse

Ein Zufallsprozess heißt stationär, wenn seine statistischen Eigenschaften invariant gegenüber Verschiebungen der Zeit sind. Zwei Zufallsprozesse heißen verbunden stationär, wenn beide stationär und ihre gemeinsamen statistischen Eigenschaften invariant gegenüber Verschiebungen der Zeit sind. [Hänsler 2001, S. 66]

Gilt die Invarianz nur für die Erwartungswerte erster und zweiter Ordnung, d. h. für Mittelwert und Korrelationsfunktionen, dann spricht man von schwacher Stationarität.

So gilt z. B. für den Mittelwert eines stationären Prozesses

$$\mu_x = E\{x_k\}_r ,$$

d. h. dieser hängt nicht von k (oder t , bei zeitkontinuierlichen Prozessen) ab.

Ergodische Prozesse

Ein stationärer Zufallsprozess heißt ergodisch, wenn die Zeitmittelwerte jeder beliebigen Musterfunktion mit Wahrscheinlichkeit Eins mit den entsprechenden Scharmittelwerten übereinstimmen. [Hänsler 2001, S. 69]

Entsprechend der schwachen Stationarität spricht man von „schwach ergodisch“, wenn dies nur für die Erwartungswerte erster und zweiter Ordnung zutrifft.

Für den Mittelwert eines ergodischen, zeitdiskreten Prozesses $\{x_k\}_r$ gilt also

$$\mu_x = E\{x_k\}_r = \lim_{k \rightarrow \infty} \frac{1}{2k+1} \cdot \sum_{k'=-k}^k x_{k'} .$$

Die zeitkontinuierliche Entsprechung ist

$$\mu_x = E\{x(t)\}_r = \lim_{T \rightarrow \infty} \frac{1}{2T} \cdot \int_{-T}^T x(t) \cdot dt .$$

Die Ergodizität nachzuweisen ist aufwendiger und in der Regel nicht durchzuführen. Daher wird dies häufig als Annahme formuliert. Voraussetzung dafür, dass ein Prozess ergodisch ist, ist dessen Stationarität. Jedoch ist diese keine hinreichende Bedingung.

4.3.1 Korrelation und Leistungsdichtespektren

Korrelation

Die Autokorrelationsfunktion eines Prozesses $\{x(t)\}_r$ ist als

$$R_{xx}(t_1, t_2) = E\{x(t_1)x(t_2)\}_r$$

und die Kreuzkorrelationsfunktion zweier Prozesse $\{x(t)\}_r$ und $\{y(t)\}_r$ als

$$R_{xy}(t_1, t_2) = E\{x(t_1)y(t_2)\}_r$$

definiert. Im Fall diskreter Zufallsprozesse wird hier

$$R_{xx,k_1,k_2} = E\{x_{k_1}x_{k_2}\}_r$$

bzw.

$$R_{xy,k_1,k_2} = E\{x_{k_1}y_{k_2}\}_r$$

geschrieben.

Die folgenden anhand der Kreuzkorrelation für kontinuierliche Signale diskutierten Festlegungen treffen sinngemäß auch auf die Autokorrelation sowie für deren zeitdiskreten Entsprechungen zu.

Sind die Prozesse (verbunden) stationär, so hängt der Erwartungswert $E\{x(t_1)y(t_2)\}_r$ nur von der Zeitdifferenz $\tau = t_2 - t_1$ ab. In diesem Fall gilt

$$R_{xy}(\tau) = E\{x(t)y(t + \tau)\}_r .$$

Sind die betrachteten Prozesse darüberhinaus ergodisch, dann kann auch die Definition

$$R_{xy}^E(\tau) = \frac{1}{2T} \cdot \lim_{T \rightarrow \infty} \int_{-T}^T x(t)y(t + \tau)dt$$

verwendet werden. Für ergodische Prozesse gilt $R_{xy}^E(\tau) = R_{xy}(\tau)$.¹²

Will man den Begriff der Korrelation auch auf deterministische Signale erweitern, so ist dies nur sinnvoll möglich, wenn auch nicht-stationäre (also nicht-konstante) Signale zugelassen werden. Die zweiparametrische Korrelation $R_{xy}(t_1, t_2)$ kann direkt auch für deterministische Signale verwendet werden. Jedoch zeigt sich, dass auch eine der einparametrischen Korrelationsfunktion ähnlichen Definition sinnvoll ist. Für periodische Signale kann man sich dabei an der Definition der ergodischen Korrelation orientieren und legt

$$R_{xy}^P(\tau) = \frac{1}{NT} \int_{t_0}^{t_0+NT} x(t)y(t + \tau)dt$$

fest, wenn beide Signale T -periodisch sind.¹³ Für periodische Signale gilt $R_{xy}^P(\tau) = R_{xy}^E(\tau)$, aber nicht $R_{xy}^P(\tau) = R_{xy}(\tau)$, da $R_{xy}(\tau)$ für nicht-stationäre Signale nicht sinnvoll angegeben werden kann.

Für zeitbegrenzte Signale $x(t)$ und $y(t)$ – bzw. allgemeiner für Signale, bei denen das uneigentliche Integral $\int_{-\infty}^{\infty} x(t)y(t + \tau)dt$ existiert – ist diese Definition (mit $T \rightarrow \infty$) nicht sinnvoll, da diese damit immer den Wert null besitzen würden. Daher wird für diese Art von Signalen die Impuls-Korrelationsfunktion

$$R_{xy}^I(\tau) = \int_{-\infty}^{\infty} x(t)y(t + \tau)dt$$

eingeführt.

¹² Die Tatsache, dass für ergodische Prozesse $R_{xy}^E(\tau) = R_{xy}(\tau)$ gilt, macht die besondere Kennzeichnung $R_{xy}^E(\tau)$ nicht hinfällig. Würde man auf das „E“ verzichten, so müsste man im Weiteren immer darauf achten, dass $x(t)$ und $y(t)$ auch ergodische, verbunden stationäre Zufallsprozesse sind. So hingegen kann $R_{xy}^E(\tau)$ von beliebigen Prozessen bestimmt werden, wobei die wesentliche Interpretation als Korrelationsfunktion nur für ergodische Prozesse gilt.

¹³ Im Unterschied zu $R_{xy}^E(\tau)$ wird hier durch das Integral ein Durchschnittswert bestimmt, während das Integral bei $R_{xy}^E(\tau)$ den Erwartungswert ersetzt.

Leistungsdichtesprektren

Autoleistungsdichtespektrum

Das Autoleistungsdichtesprektum ist als die Fouriertransformierte der Autokorrelation definiert. Für das zeitkontinuierliche Autoleistungsdichtespektrum gilt damit

$$S_{xx}(\omega) = \int_{-\infty}^{\infty} R_{xx}(\tau) e^{-j\omega\tau} d\tau.$$

Im Zeitdiskreten tritt an die Stelle der Fouriertransformierten die zeitdiskrete Fouriertransformation (DTFT – discrete time fourier transform), so dass

$$S_{xx}(\Omega) = \sum_{k=-\infty}^{\infty} R_{xx,k} \cdot e^{-j\Omega k}$$

gilt.

Kreuzleistungsdichtespektrum

Analog ist die Kreuzleitungsdichte als Fouriertransformierte der Kreuzkorrelationsfunktion definiert. Im Zeitkontinuierlichen gilt also

$$S_{xy}(\omega) = \int_{-\infty}^{\infty} R_{xy}(\tau) e^{-j\omega\tau} d\tau.$$

Im Zeitdiskreten ergibt sich

$$S_{xy}(\Omega) = \sum_{k=-\infty}^{\infty} R_{xy,k} \cdot e^{-j\Omega k}.$$

Analog kann man auch bezüglich der Variationen der Korrelationsfunktionen entsprechend „Leistungsdichtespektren“ definieren. Dies ist auch insoweit sinnvoll, dass diese den gleichen Transformationen bezüglich LZI-Systemen gehorchen.

Eigenschaften für reelle, stationäre Zufallsprozesse

Eigenschaften der Korrelationsfunktionen

Die Autokorrelationsfunktion eines reellen, stationären Zufallsprozess ist ebenfalls reell und nach

$$R_{xx}(-\tau) = E\{x(t) \cdot x(t - \tau)\}_r = E\{x(t - \tau) \cdot x(t)\}_r = E\{x(\tilde{t}) \cdot x(\tilde{t} + \tau)\}_r = R_{xx}(\tau)$$

gerade.

Für die Kreuzkorrelation ergibt der analoge Weg

$$R_{xy}(-\tau) = E\{x(t) \cdot y(t - \tau)\}_r = E\{y(t - \tau) \cdot x(t)\}_r = E\{y(\tilde{t}) \cdot x(\tilde{t} + \tau)\}_r = R_{yx}(\tau),$$

diese ist also nicht gerade.

Für die Autokorrelationsfunktion gilt, dass diese bei $\tau = 0$ ihren (betragsgrößten) Wert annimmt, d. h.

$$R_{xx}(0) \geq |R_{xx}(\tau)| . \quad (4.4)$$

Dies kann wie folgt gezeigt werden ([Hänsler 2001, S. 75]). Ausgangspunkt ist der Erwartungswert

$$E\{(a \cdot x(t) - x(t + \tau))^2\}_r = a^2 \cdot R_{xx}(0) - 2a \cdot R_{xx}(\tau) + R_{xx}(0) \geq 0 \quad (4.5)$$

der sicher nicht negativ ist. Multipliziert man die Ungleichung mit $R_{xx}(0)$, so ändert dies nichts an der Relation, da $R_{xx}(0) \geq 0$ ist,

$$a^2 \cdot (R_{xx}(0))^2 - 2a \cdot R_{xx}(\tau) \cdot R_{xx}(0) + (R_{xx}(0))^2 \geq 0 .$$

Dies kann in die Form

$$(a \cdot R_{xx}(0) - R_{xx}(\tau))^2 - (R_{xx}(\tau))^2 + (R_{xx}(0))^2 \geq 0$$

gebracht werden, woraus

$$(R_{xx}(0))^2 \geq (R_{xx}(\tau))^2 - (a \cdot R_{xx}(0) - R_{xx}(\tau))^2$$

folgt. Da die Ungleichung (4.5) für alle a gilt, muss auch die letzte Ungleichung für alle a gelten. Da $(R_{xx}(\tau))^2$ nach oben begrenzt werden soll, ist a so zu wählen, dass $a \cdot R_{xx}(0) - R_{xx}(\tau) = 0$ ist. (Alle anderen Werte für a würden ein größeres zulässiges $(R_{xx}(\tau))^2$ ergeben.) Es folgt

$$(R_{xx}(0))^2 \geq (R_{xx}(\tau))^2$$

und damit letztlich Gl. (4.4).

Eigenschaften der Leistungsdichtespektren

Für das Autoleistungsdichtespektrum $S_{xx}(-\omega)$ gilt

$$\begin{aligned} S_{xx}(-\omega) &= \int_{-\infty}^{\infty} R_{xx}(t) \cdot e^{j\omega t} \cdot dt = \int_{-\infty}^{-\infty} R_{xx}(-\tau) \cdot e^{-j\omega\tau} \cdot (-1) \cdot d\tau \\ &= \int_{-\infty}^{\infty} R_{xx}(\tau) \cdot e^{-j\omega\tau} \cdot d\tau = S_{xx}(\omega) . \end{aligned}$$

Das Autoleistungsdichtespektrum ist demnach eine gerade Funktion. Eine analoge Rechnung ergibt

$$S_{xy}(-\omega) = S_{yx}(\omega)$$

für das Kreuzleistungsdichtespektrum. Es ist hierbei die Reihenfolge der Indizes zu beachten.

Die Fouriertransformierte einer geraden, reellen Funktion ist reell. Demnach ist auch das Autoleistungsdichtespektrum eine reelle Funktion. Ein Rechenweg zum Nachweis ist

$$\begin{aligned} S_{xx}(\omega) &= \int_{-\infty}^{\infty} R_{xx}(t) \cdot e^{-j\omega t} dt = \int_{-\infty}^0 R_{xx}(t) \cdot e^{-j\omega t} dt + \int_0^{\infty} R_{xx}(t) \cdot e^{-j\omega t} dt \\ &= \int_0^{\infty} R_{xx}(-\tau) \cdot e^{j\omega\tau} d\tau + \int_0^{\infty} R_{xx}(t) \cdot e^{-j\omega t} dt \\ &= \int_0^{\infty} R_{xx}(t) \cdot (e^{j\omega t} + e^{-j\omega t}) \cdot dt \\ &= 2 \cdot \int_0^{\infty} R_{xx}(t) \cdot \cos(\omega t) \cdot dt . \end{aligned}$$

Die Eigenschaft „reelle Funktion“ kann kompakt auch als

$$S_{xx}(\omega) = S_{xx}^*(\omega)$$

geschrieben werden.

Die Kreuzleistungsdichte ist nicht gerade, sondern „nur“ reell. Damit ist der Realteil der Fouriertransformierten gerade und der Imaginärteil ungerade, d. h. es gilt

$$S_{xy}(\omega) = S_{xy}^*(-\omega).$$

Dies kann man schnell überprüfen, indem man in beiden Fällen das Fourierintegral aufschreibt. Es ergibt sich

$$S_{xy}(\omega) = \int_{-\infty}^{\infty} R_{xy}(t) \cdot e^{-j\omega t} \cdot dt = \int_{-\infty}^{\infty} R_{xy}(t) \cdot \cos(\omega t) \cdot dt - j \cdot \int_{-\infty}^{\infty} R_{xy}(t) \cdot \sin(\omega t) \cdot dt$$

und

$$S_{xy}(-\omega) = \int_{-\infty}^{\infty} R_{xy}(t) \cdot e^{j\omega t} \cdot dt = \int_{-\infty}^{\infty} R_{xy}(t) \cdot \cos(\omega t) \cdot dt + j \cdot \int_{-\infty}^{\infty} R_{xy}(t) \cdot \sin(\omega t) \cdot dt.$$

Für das Autoleistungsdichtespektrum kann noch bewiesen werden, dass es eine nicht-negative Funktion ist, d. h.

$$S_{xx}(\omega) \geq 0.$$

4.4 Übertragung stochastischer Eigenschaften bei linearen Systemen

4.4.1 Übertragung durch Gewichtsfunktion und Frequenzgang

Zeitdiskret

Bei gegebener Gewichtsfunktion $g_{k'}$ kann der Ausgang y_k eines zeitdiskreten LTI-Systems durch die Faltung der Gewichtsfunktion mit der Eingangsfolge u_k berechnet werden,

$$y_k = \sum_{k'=-\infty}^{\infty} g_{k'} \cdot u_{k-k'}.$$

Handelt es sich bei u_k nicht um eine deterministische Folge, sondern um eine Realisierung eines Zufallsprozesses, so ändert dies nicht an dem angegebenen Zusammenhang. In diesem Fall ist y_k dann ebenso eine Realisierung eines Zufallsprozesses. In dem Fall, dass es sich bei $\{u_k\}_r$ um einen Zufallsprozess handelt, ist man in der Regel nicht an dem Verlauf einzelner Realisierungen für $\{y_k\}_r$ interessiert, sondern an Aussagen über die stochastischen Eigenschaften des Zufallsprozesses $\{y_k\}_r$ in Abhängigkeit der stochastischen Eigenschaften des Zufallsprozesses $\{u_k\}_r$.

Dabei wird sehr häufig davon ausgegangen, dass $\{u_k\}_r$ ein stationärer Prozess ist.

Mittelwert

Der Mittelwert der Ausgangsgröße $\{y_k\}_r$ ist durch deren Erwartungswert

$$\mu_{y,k} = E\{y_k\}_r = E \left\{ \sum_{k'=-\infty}^{\infty} g_{k'} \cdot u_{k-k'} \right\}_r$$

gegeben. Es folgt

$$\mu_{y,k} = \sum_{k'=-\infty}^{\infty} g_{k'} \cdot E\{u_{k-k'}\}_r = \sum_{k'=-\infty}^{\infty} g_{k'} \cdot \mu_{u,k-k'} .$$

Bezüglich der Mittelwerte verhält sich das System demnach genauso wie im deterministischen Fall.

Ist $\{u_k\}_r$ ein schwach stationärer Prozess, d. h. der Mittelwert $E\{u_k\}_r = \mu_{u,k} = \mu_u$ ist für alle k gleich, so kann dies zu

$$\mu_{y,k} = \mu_u \cdot \sum_{k'=-\infty}^{\infty} g_{k'}$$

vereinfacht werden. Die Summe entspricht gerade dem diskreten Frequenzgang¹⁴ bei $e^{j0} = 1$, also

$$\mu_{y,k} = \mu_u \cdot H(e^{j0} = 1) .$$

$H(e^{j0} = 1)$ ist die stationäre Verstärkung des Systems.

Korrelationsfunktionen

Die Kreuzkorrelationsfunktion des Ein- und Ausgangsprozesses ist als

$$R_{uy,k_1,k_2} = E\{u_{k_1} \cdot y_{k_2}\}_r$$

definiert. Einsetzen der Faltungssumme für y_{k_2} führt dann auf

$$R_{uy,k_1,k_2} = E \left\{ u_{k_1} \cdot \sum_{k'=-\infty}^{\infty} g_{k'} \cdot u_{k_2-k'} \right\}_r = E \left\{ \sum_{k'=-\infty}^{\infty} g_{k'} \cdot u_{k_1} \cdot u_{k_2-k'} \right\}_r = \sum_{k'=-\infty}^{\infty} g_{k'} \cdot E\{u_{k_1} \cdot u_{k_2-k'}\}_r .$$

Der verbleibende Erwartungswert entspricht gerade der Autokorrelationsfunktion der Eingangsfolge, und damit erhält man

$$R_{uy,k_1,k_2} = \sum_{k'=-\infty}^{\infty} g_{k'} \cdot R_{uu,k_1,k_2-k'} .$$

Bei schwach stationären Prozessen, d. h. $R_{uu,k_1,k_2} = R_{uu,k_2-k_1}$, gilt damit

$$R_{uy,k} = \sum_{k'=-\infty}^{\infty} g_{k'} \cdot R_{uu,k-k'} .$$

¹⁴ Der diskrete Frequenzgang ergibt sich, sofern die Summe konvergiert, aus der DTFT der diskreten Gewichtsfunktion,

$$H(e^{j\Omega}) = \sum_{k=-\infty}^{\infty} g_k \cdot e^{-j\Omega k} .$$

Wird die Reihenfolge von Ein- und Ausgang bei der Korrelationsfunktion vertauscht, so ergibt sich mit

$$R_{yu,k} = \sum_{k'=-\infty}^{\infty} g_{-k'} \cdot R_{uu,k-k'}$$

ein ähnlicher Ausdruck. Lediglich die Gewichtsfunktion ist gespiegelt.

Für die Autokorrelationsfunktion des Ausgangs erhält man

$$R_{yy,k} = \sum_{k'=-\infty}^{\infty} \sum_{k''=-\infty}^{\infty} g_{-k'} g_{k''} \cdot R_{uu,k-k'-k''} .$$

Leistungsdichtespektren

Die Leistungsdichtespektren sind als die Fouriertransformierten (im Zeitdiskreten die DTFT) der Korrelationsfunktionen definiert. Da auch für die DTFT die entsprechende Faltungssätze gelten, d. h. die Faltung im Zeitbereich entspricht der Multiplikation im Frequenzbereich, können diese hier einfach angegeben werden.

Es gilt

$$\begin{aligned} S_{uy}(\Omega) &= H(e^{j\Omega}) \cdot S_{uu}(\Omega), \\ S_{yu}(\Omega) &= H^*(e^{j\Omega}) \cdot S_{uu}(\Omega) \quad \text{und} \\ S_{yy}(\Omega) &= H(e^{j\Omega}) \cdot H^*(e^{j\Omega}) \cdot S_{uu}(\Omega). \end{aligned}$$

Zeitkontinuierlich

An die Stelle der Faltungssumme tritt bei zeitkontinuierlichen Systemen das Faltungsintegral

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

Mittelwert

Für den Mittelwert gilt (wobei hier und im Weiteren immer davon ausgegangen wird, dass die auftretenden Integrale für alle Realisierungen des Zufallsprozesses existieren)

$$\mu_y(t) = E\{y(t)\}_r = \int_{-\infty}^{\infty} g(\tau) \cdot E\{u(t - \tau)\}_r \cdot d\tau = \int_{-\infty}^{\infty} g(\tau) \cdot \mu_u(t - \tau) \cdot d\tau .$$

Bei schwach stationären Prozessen hängt der Mittelwert nicht von der Zeit ab, und es ergibt sich das zum Zeitdiskreten analoge Ergebnis

$$\mu_y(t) = H(0) \cdot \mu_u(t) .$$

$H(0)$ ist auch hier die stationäre Verstärkung des betrachteten Systems.

Korrelationsfunktion

Für die Korrelationsfunktionen ergeben sich die den Faltungssummen im Zeitdiskreten entsprechenden Faltungsintegrale

$$R_{uy}(\tau) = \int_{-\infty}^{\infty} g(\nu) \cdot R_{uu}(\tau - \nu) \cdot d\nu,$$

$$R_{yu}(\tau) = \int_{-\infty}^{\infty} g(-\nu) \cdot R_{uu}(\tau - \nu) \cdot d\nu$$

und

$$R_{yy}(\tau) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(-\nu)g(\nu') \cdot R_{uu}(\tau - \nu - \nu') \cdot d\nu' d\nu.$$

Leistungsdichtespektren

Ebenfalls analog ergeben sich für die Leistungsdichtespektren

$$S_{uy}(\omega) = H(\omega) \cdot S_{uu}(\omega),$$

$$S_{yu}(\omega) = H^*(\omega) \cdot S_{uu}(\omega)$$

und

$$S_{yy}(\omega) = H(\omega) \cdot H^*(\omega) \cdot S_{uu}(\omega).$$

4.4.2 Zustandsraumdarstellung

Zeitdiskret

Es wird die Fragestellung betrachtet, wie sich der Mittelwert und die Varianz einer Eingangs- oder Störgröße auf die Zustände oder den Ausgang eines Systems überträgt. Dabei wird dem Vorgehen aus [Brammer und Siffling 1975, 220ff] gefolgt.

Es wird ein zeitvariantes, diskretes Zustandsraummodell (\mathbf{A}_k , \mathbf{B}_k , \mathbf{C}_k , \mathbf{D}_k) betrachtet. Die Eingangsgröße \mathbf{u}_k ist durch einen, zunächst allgemeinen, Zufallsprozess gegeben, dessen Mittelwerte $\mu_{u,k} = E\{\mathbf{u}_k\}_r$ und Kovarianzmatrizen $C_{uu,k,m} = E\{(\mathbf{u}_k - \mu_{u,k}) \cdot (\mathbf{u}_m - \mu_{u,m})^T\}_r$ bekannt seien.

Mittelwerte

Bezüglich der Mittelwerte verhält sich das System genauso wie gegenüber deterministischen Signalen, d. h.

$$\begin{aligned}\mu_{x,k+1} &= \mathbf{A}_k \cdot \mu_{x,k} + \mathbf{B}_k \cdot \mu_{u,k} \\ \mu_{y,k} &= \mathbf{C}_k \cdot \mu_{x,k} + \mathbf{D}_k \cdot \mu_{u,k},\end{aligned}$$

was direkt aus der Linearität des Erwartungswertes folgt.

Kovarianzen

Die Kovarianzmatrix des Zustandsvektors zum Zeitpunkt $k + 1$ ist als

$$C_{xx,k+1,k+1} = E\{(\mathbf{x}_{k+1} - \mu_{x,k+1}) \cdot (\mathbf{x}_{k+1} - \mu_{x,k+1})^T\}_r$$

definiert. Ausmultiplizieren ergibt zunächst

$$\begin{aligned} C_{xx,k+1,k+1} &= E\{\mathbf{x}_{k+1}\mathbf{x}_{k+1}^T\}_r - E\{\mathbf{x}_{k+1}\}_r \cdot \mu_{x,k+1}^T - \mu_{x,k+1} \cdot E\{\mathbf{x}_{k+1}^T\}_r + \mu_{x,k+1} \cdot \mu_{x,k+1}^T \\ &= E\{\mathbf{x}_{k+1}\mathbf{x}_{k+1}^T\}_r - \mu_{x,k+1} \cdot \mu_{x,k+1}^T. \end{aligned} \quad (4.6)$$

Im Erwartungswert werden die Zustandsvektoren durch die Zustandsgleichung ersetzt, so dass man

$$\begin{aligned} E\{\mathbf{x}_{k+1}\mathbf{x}_{k+1}^T\}_r &= E\{(\mathbf{A}_k\mathbf{x}_k + \mathbf{B}_k\mathbf{u}_k) \cdot (\mathbf{A}_k\mathbf{x}_k + \mathbf{B}_k\mathbf{u}_k)^T\}_r \\ &= \mathbf{A}_k \cdot E\{\mathbf{x}_k\mathbf{x}_k^T\}_r \cdot \mathbf{A}_k^T + \mathbf{A}_k \cdot E\{\mathbf{x}_k\mathbf{u}_k^T\}_r \cdot \mathbf{B}_k^T + \mathbf{B}_k \cdot E\{\mathbf{u}_k\mathbf{x}_k^T\}_r \cdot \mathbf{A}_k^T + \mathbf{B}_k \cdot E\{\mathbf{u}_k\mathbf{u}_k^T\}_r \cdot \mathbf{B}_k^T \end{aligned}$$

erhält. Wird Gl. (4.1) auf Seite 56 auf die Erwartungswerte der rechten Seite angewendet, so folgt

$$\begin{aligned} E\{\mathbf{x}_{k+1}\mathbf{x}_{k+1}^T\}_r &= \mathbf{A}_k \cdot C_{xx,k,k} \cdot \mathbf{A}_k^T + \mathbf{A}_k \cdot C_{xu,k,k} \cdot \mathbf{B}_k^T + \mathbf{B}_k \cdot C_{ux,k,k} \cdot \mathbf{A}_k^T + \mathbf{B}_k \cdot C_{uu,k,k} \cdot \mathbf{B}_k^T \\ &\quad + \mathbf{A}_k \cdot \mu_{x,k} \mu_{x,k}^T \cdot \mathbf{A}_k^T + \mathbf{A}_k \cdot \mu_{x,k} \mu_{u,k}^T \cdot \mathbf{B}_k^T + \mathbf{B}_k \cdot \mu_{u,k} \mu_{x,k}^T \cdot \mathbf{A}_k^T + \mathbf{B}_k \cdot \mu_{u,k} \mu_{u,k}^T \cdot \mathbf{B}_k^T \\ &= \mathbf{A}_k \cdot C_{xx,k,k} \cdot \mathbf{A}_k^T + \mathbf{A}_k \cdot C_{xu,k,k} \cdot \mathbf{B}_k^T + \mathbf{B}_k \cdot C_{ux,k,k} \cdot \mathbf{A}_k^T + \mathbf{B}_k \cdot C_{uu,k,k} \cdot \mathbf{B}_k^T \\ &\quad + (\mathbf{A}_k \cdot \mu_{x,k} + \mathbf{B}_k \cdot \mu_{u,k}) \cdot (\mathbf{A}_k \cdot \mu_{x,k} + \mathbf{B}_k \cdot \mu_{u,k})^T \\ &= \mathbf{A}_k \cdot C_{xx,k,k} \cdot \mathbf{A}_k^T + \mathbf{A}_k \cdot C_{xu,k,k} \cdot \mathbf{B}_k^T + \mathbf{B}_k \cdot C_{ux,k,k} \cdot \mathbf{A}_k^T + \mathbf{B}_k \cdot C_{uu,k,k} \cdot \mathbf{B}_k^T + \mu_{x,k+1} \mu_{x,k+1}^T. \end{aligned}$$

Wird $E\{\mathbf{x}_{k+1}\mathbf{x}_{k+1}^T\}_r$ in dieser Form in Gl. (4.6) eingesetzt, so ergibt sich die erste Zeile von

$$\begin{aligned} C_{xx,k+1,k+1} &= \mathbf{A}_k \cdot C_{xx,k,k} \cdot \mathbf{A}_k^T + \mathbf{A}_k \cdot C_{xu,k,k} \cdot \mathbf{B}_k^T + \mathbf{B}_k \cdot C_{ux,k,k} \cdot \mathbf{A}_k^T + \mathbf{B}_k \cdot C_{uu,k,k} \cdot \mathbf{B}_k^T \\ C_{yy,k,k} &= \mathbf{C}_k \cdot C_{xx,k,k} \cdot \mathbf{C}_k^T + \mathbf{C}_k \cdot C_{xu,k,k} \cdot \mathbf{D}_k^T + \mathbf{D}_k \cdot C_{ux,k,k} \cdot \mathbf{C}_k^T + \mathbf{D}_k \cdot C_{uu,k,k} \cdot \mathbf{D}_k^T. \end{aligned}$$

Die zweite Zeile erhält man durch eine analoge Rechnung bezüglich des Ausgangsvektors. Zur Berechnung sind neben der Matrix $C_{xx,k,k}$ des vorangehenden Schritts auch die Kenntnis der Kovarianzmatrizen $C_{uu,k,k}$ und $C_{xu,k,k}$ nötig.¹⁵ Wird der Eingang von außen vorgegeben und nicht (teilweise) über eine Regelung bestimmt, so hängt $C_{uu,k,k}$ nur von den stochastischen Eigenschaften des Eingangsprozesses $\{\mathbf{u}_k\}_r$ ab. Wird eine Regelung unter Störeinfluss $\{\mathbf{n}_k\}_r$ betrachtet, so könnte $C_{uu,k,k}$ in Abhängigkeit von $C_{xx,k,k}$ oder $C_{yy,k,k}$ und $C_{nn,k,k}$ angegeben werden. Allerdings ist es in diesem Fall einfacher, zunächst die Zustandsraumdarstellung des geregelten Systems aufzustellen und dann die hier aufgeführten Gleichungen für dieses neue System mit dem Eingang \mathbf{n}_k anzuwenden.

$C_{xu,k,k}$ kann über

$$C_{xu,k,k} = \mathbf{A}_k \cdot C_{xu,k-1,k} + \mathbf{B}_k \cdot C_{uu,k-1,k}$$

bestimmt werden, wobei dazu $C_{xu,k-1,k}$ und $C_{uu,k-1,k}$ bekannt sein müssen. Dabei ergibt sich $C_{uu,k-1,k}$ wieder aus dem Eingangsprozess $\{\mathbf{u}_k\}_r$. Für $C_{xu,k-1,k}$ ergibt sich

$$C_{xu,k-1,k} = \mathbf{A}_k \cdot C_{xu,k-2,k} + \mathbf{B}_k \cdot C_{uu,k-2,k},$$

¹⁵ $C_{ux,k,k} = C_{xu,k,k}^T$

was zu der iterativen Berechnungsvorschrift

$$C_{xu,k',k} = A_k \cdot C_{xu,k'-1,k} + B_k \cdot C_{uu,k'-1,k}$$

verallgemeinert werden kann.

Es ist $C_{xu,k,k} = 0$ wenn $C_{uu,k-k',k} = 0$, $k' = 0, \dots, k-1$, und $C_{xu,0,k} = 0$ gilt. Die erste Bedingung sagt aus, dass der Eingangsprozess unkorreliert ist. Die zweite Bedingung schließt auch eine Korrelation des Eingangsprozesses mit dem Anfangszustand $\{x_0\}_r$ aus. Sind diese Bedingungen erfüllt (was z. B. eine Rückführung ausschließt), vereinfachen sich die Gleichungen für die Kovarianzmatrizen zu

$$\begin{aligned} C_{xx,k+1,k+1} &= A_k \cdot C_{xx,k,k} \cdot A_k^T + B_k \cdot C_{uu,k,k} \cdot B_k^T \\ C_{yy,k,k} &= C_k \cdot C_{xx,k,k} \cdot C_k^T + D_k \cdot C_{uu,k,k} \cdot D_k^T. \end{aligned}$$

Mit der oben eingeführten Notation der Kovarianzmatrizen kann dies auch etwas kompakter als

$$\begin{aligned} P_{x,k+1} &= A_k \cdot P_{x,k} \cdot A_k^T + B_k \cdot P_{u,k} \cdot B_k^T \\ P_{y,k} &= C_k \cdot P_{x,k} \cdot C_k^T + D_k \cdot P_{u,k} \cdot D_k^T. \end{aligned}$$

geschrieben werden.

Zeitkontinuierlich

Auch im Zeitkontinuierlichen verhalten sich die Mittelwerte genauso wie deterministische Signale, d. h. es gilt

$$\begin{aligned} \dot{\mu}_x(t) &= A(t) \cdot \mu_x(t) + B(t) \cdot \mu_u(t) \\ \mu_y(t) &= C(t) \cdot \mu_x(t) + D(t) \cdot \mu_u(t). \end{aligned}$$

Für die Kovarianzmatrix der Zustandsgröße $\{x(t)\}_r$ gilt die Differentialgleichung

$$\dot{P}_x(t) = A(t) \cdot P_x(t) + P_x(t) \cdot A^T(t) + B(t) \cdot Q_u(t) \cdot B^T(t),$$

wobei dazu davon ausgegangen wurde, dass der Eingangsprozess $\{u(t)\}_r$ durch ein weißes, normalverteiltes Rauschen mit der Kovarianzmatrix $C_{uu}(t_1, t_2) = Q_u \cdot \delta(t_1 - t_2)$ beschrieben wird. [Brammer und Siffling 1975]

4.5 Stochastische Konvergenz

Häufig ist es von Interesse, wie sich die Eigenschaften eines stochastischen Prozesses – z. B. dem Schätzwert eines Schätzers – ändern, wenn k gegen unendlich geht.

Die folgenden Definitionen geben verschiedene Kriterien an, wie sich ein Zufallsprozesses $\{x_k\}_r$ für $k \rightarrow \infty$ einer Zufallsvariablen $\{x\}_r$ annähert. Die „Referenzvariable“ $\{x\}_r$ kann dabei auch ein fester Wert sein, muss also keine Zufallsvariable sein. Tatsächlich wird das im Rahmen dieser Vorlesung meistens der Fall sein. Dabei ist der betrachtete Prozess dann der Schätzwert $\{\hat{\theta}_k\}_r$, dessen Verhalten für $k \rightarrow \infty$ Messungen anhand des tatsächlichen Parameterwertes θ bewertet werden soll.

4.5.1 Konvergenzbegriffe

Starke Konvergenz

Die Folge (der Zufallsprozess) $\{x_k\}_r$ konvergiert stark (fast sicher, mit Wahrscheinlichkeit Eins) gegen die Zufallsvariable $\{x\}_r$, wenn

$$P(\lim_{k \rightarrow \infty} (\{x_k\}_r^\epsilon - \{x\}_r^\epsilon) = 0) = 1$$

gilt. Hierfür wird auch die abkürzende Schreibweise

$$\{x_k\}_r \xrightarrow{\text{f.s.}} \{x\}_r$$

verwendet.

Schwache Konvergenz

Die Folge (der Zufallsprozess) $\{x_k\}_r$ konvergiert schwach (stochastisch, in Wahrscheinlichkeit) gegen die Zufallsvariable $\{x\}_r$, wenn

$$\lim_{k \rightarrow \infty} P(|\{x_k\}_r^\epsilon - \{x\}_r^\epsilon| \geq \epsilon) = 0 \quad \forall \epsilon > 0$$

gilt. Dies wird auch

$$\operatorname{plim}_{k \rightarrow \infty} \{x_k\}_r = \{x\}_r$$

geschrieben.

Konvergenz im quadratischen Mittel

Die Folge (der Zufallsprozess) $\{x_k\}_r$ konvergiert im quadratischen Mittel gegen die Zufallsvariable $\{x\}_r$, wenn

$$\lim_{k \rightarrow \infty} E(|\{x_k\}_r - \{x\}_r|^2) = 0$$

gilt, wobei vorausgesetzt wird, dass endliche zweiten Momente existieren. Abkürzend wird dies

$$\operatorname{l.i.m.}_{k \rightarrow \infty} \{x_k\}_r = \{x\}_r$$

geschrieben.

Konvergenz in Verteilung

Konvergieren lediglich die Verteilungsfunktionen $F_{\{x_k\}_r}(x)$ gegen eine Verteilungsfunktion $F_{\{x\}_r}(x)$, so spricht man von der Konvergenz in Verteilung. Diese wird symbolisch durch ein großgeschriebenes Lim gekennzeichnet, Schreibweise:

$$\operatorname{Lim}_{k \rightarrow \infty} \{x_k\}_r \sim \{x\}_r,$$

Diskussion

Wenn es darum geht, die Eigenschaft eines Schätzers zu beurteilen, spielt es letztlich keine Rolle, ob man dazu die starke Konvergenz, die schwache Konvergenz oder die Konvergenz im quadratischen Mittel heranzieht. Alle diese Konvergenzen sagen aus, dass für $k \rightarrow \infty$ der Schätzer gegen den tatsächlichen Parameterwert geht. Eine Aussage über die Geschwindigkeit mit der die Schätzer konvergieren wird mit keinem der Kriterien getroffen.

Für den rechnerischen Nachweis der Konvergenz eines Schätzers bietet sich die Konvergenz im quadratischen Mittel an, da man den dort auftretenden Erwartungswert (wirklich) einfach aufschreiben kann, und dann nach etwas ermüdender Rechnung die entstehende Gleichung meist in eine Form bringen kann, in der der Grenzwert explizit angegeben werden kann.

Die Bedeutung der schwachen Konvergenz liegt im *Satz von Slutsky* begründet.

Satz 1 Satz von Slutsky, [Rinne 2008, S. 411]

Ist $z = g(x, y)$ eine an der Stelle (ξ, η) stetige Funktion und sind $\{x_k\}_r$ und $\{y_k\}_r$ zwei Folgen von Zufallsvariablen (d. h. Zufallsprozesse) mit $\text{plim}\{x_k\}_r = \xi$ und $\text{plim}\{y_k\}_r = \eta$, dann ist $\{z_k\}_r = g(\{x_k\}_r, \{y_k\}_r)$ mit $\text{plim}\{z_k\}_r = g(\xi, \eta) = g(\text{plim}\{x_k\}_r, \text{plim}\{y_k\}_r)$.

Für die Anwendbarkeit ist noch wichtig zu wissen, dass die Konvergenz im quadratischen Mittel immer die schwache Konvergenz impliziert, d. h. beinhaltet. Dies ist in Abbildung 4.6 veranschaulicht. Damit kann die Konvergenz im quadratischen Mittel rechnerisch nachgewiesen werden, und dann anhand des Satzes von Slutsky die Eigenschaften von Funktionen, die beispielsweise mit dem Schätzwert arbeiten, beschrieben werden. (Wobei diese Eigenschaften dann nur bezüglich der schwachen Konvergenz gelten, was jedoch wie schon erwähnt für die praktische Beurteilung irrelevant ist.)

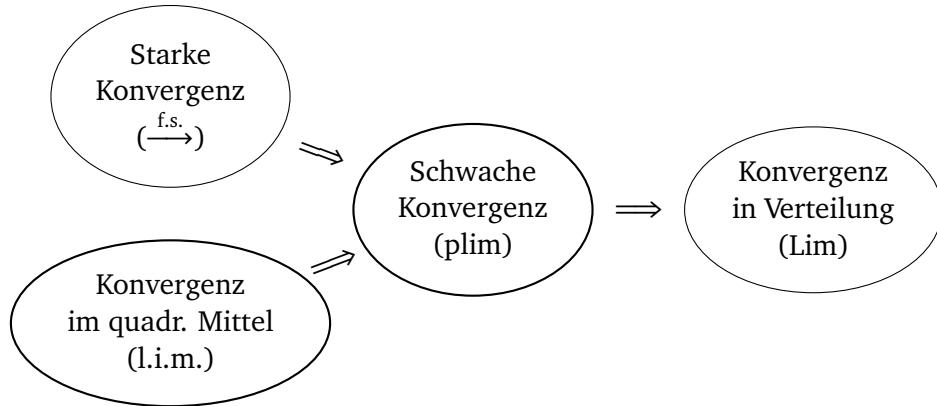


Abbildung 4.6: Implikationen der verschiedenen Konvergenzbegriffe, nach [Rinne 2008, S. 412]

Die starke Konvergenz und die Konvergenz in Verteilung sind hier nur der Vollständigkeit halber angegeben und spielen im Rahmen der Vorlesung keine Rolle.

4.5.2 Eigenschaften von Schätzern

Biasfreiheit

Unter Bias wird die Abweichung des Erwartungswerts des Schätzwertes $\{\hat{\theta}\}_r$ vom tatsächlichen Wert θ verstanden,

$$b = E\{\hat{\theta}_k\}_r - \theta .$$

Gilt

$$E\{\hat{\theta}_k\}_r - \theta = 0 ,$$

dann wird der Schätzer betrachtete Schätzer für θ_k als „biasfrei“ oder „erwartungstreu“ bezeichnet. Biasfreiheit gilt für jedes beliebige k und nicht nur für $k \rightarrow \infty$.

Gilt nur

$$\lim_{k \rightarrow \infty} E\{\hat{\theta}_k\}_r = \theta ,$$

so spricht man von asymptotischer Biasfreiheit. „Biasfreiheit“ ist damit eine stärkere Eigenschaft als „asymptotische Biasfreiheit“.

Konsistenz

Ein Schätzer wird als konsistent bezeichnet, wenn der Schätzwert für $k \rightarrow \infty$ schwach gegen den wahren Wert konvergiert,

$$\plim_{k \rightarrow \infty} \{\theta_k\}_r = \theta .$$

Ein Schätzer wird als „konsistent im quadratischen Mittel“ bezeichnet, wenn

$$\text{l.i.m.}_{k \rightarrow \infty} \{\theta_k\}_r = \theta$$

gilt. Ein im quadratischen Mittel konsistenter Schätzer ist (zumindest) asymptotisch biasfrei und die Varianz geht für $k \rightarrow \infty$ gegen Null.

In Abbildung 4.7 sind die Eigenschaften (asymptotische) Biasfreiheit und Konsistenz an Beispielen verdeutlicht. In allen der vier gezeigten Fällen wird ein Schätzer für den Parameter θ betrachtet. Im ersten Fall, Abbildung 4.7a, ist der Schätzer biasfrei, aber nicht konsistent. D. h. der Erwartungs- bzw. Mittelwert des Zufallsprozesses $\{\hat{\theta}_k\}_r$ entspricht für jedes k dem wahren Parameterwert θ , jedoch bleibt die Varianz ab $k = 2$ gleich und nimmt nicht weiter ab. Daher konvergiert der Schätzwert für $k \rightarrow \infty$ nicht.

Dagegen nimmt die Varianz in Abbildung 4.7b stetig ab. Darüber hinaus ist auch dieser Schätzer biasfrei, womit er konsistent ist.

Die Schätzwerte $\{\hat{\theta}_k\}_r$ in Abbildung 4.7c besitzen einen Bias, der aber für $k \rightarrow \infty$ verschwindet. Der Schätzer ist damit asymptotisch biasfrei. Da er auch konvergiert, ist er darüber hinaus auch konsistent.

In Abbildung 4.7d ist ein Beispiel für einen Schätzer gezeigt, der zwar konvergiert, jedoch nicht gegen den wahren Parameterwert. Damit ist dieser Schätzer weder (asymptotisch) biasfrei noch konsistent.

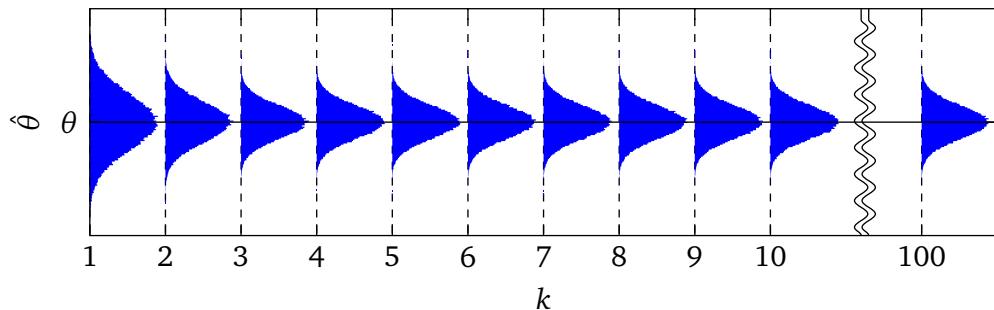
4.6 Anwendungen

4.6.1 Eigenschaften von Schätzern

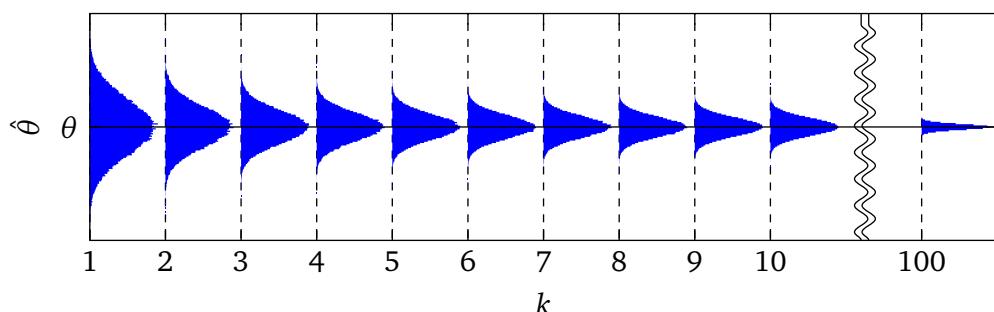
Zur Übung der Anwendung der bisherigen Eigenschaften und zur Diskussion der (wünschenswerten) Eigenschaften von Schätzern werden in diesem Abschnitt einfache Schätzer betrachtet.

Es wird mit dem Schätzer

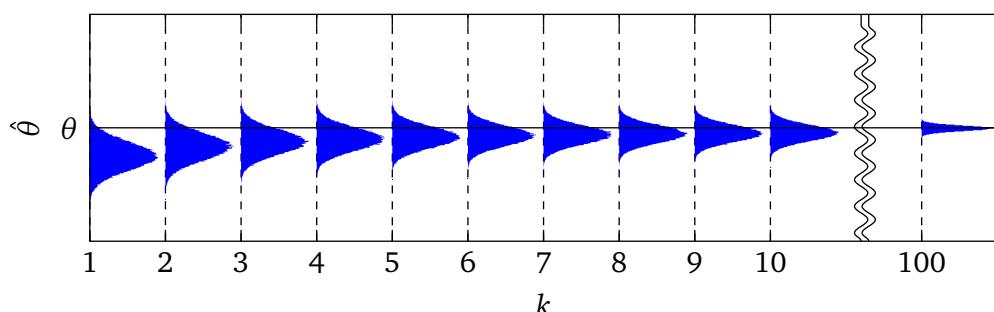
$$\hat{x}_k = \frac{1}{k} \cdot \sum_{k'=0}^{k-1} \tilde{x}_{k'}$$



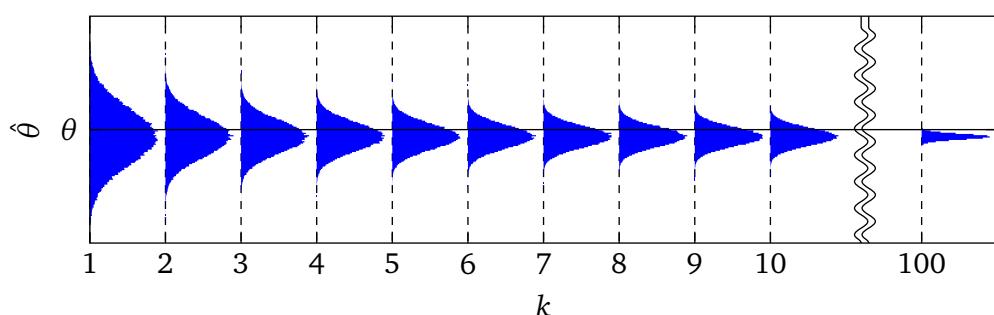
(a) Biasfrei, aber nicht konsistent



(b) Biasfrei und konsistent



(c) Asymptotisch biasfrei und konsistent



(d) Konvergierend, aber weder asymptotisch biasfrei noch konsistent

Abbildung 4.7: Schätzprozesse

begonnen, der aus k unabhängig verrauschten Messwerten $\{\tilde{x}_{k'}\}_r$ mit $E\{\tilde{x}_{k'}\}_r = x$ und $\text{Var}\{\tilde{x}_{k'}\}_r = \sigma_x^2$ den wahren Parameterwert x schätzen soll.¹⁶

Der Erwartungswert dieses Schätzers ist

$$E\{\hat{x}_k\}_r = E\left\{\frac{1}{k} \cdot \sum_{k'=0}^{k-1} \tilde{x}_{k'}\right\}_r = \frac{1}{k} \cdot \sum_{k'=0}^{k-1} E\{\tilde{x}_{k'}\}_r = \frac{1}{k} \cdot k \cdot x = x$$

Der Erwartungswert dieses Schätzers entspricht demnach für jede Anzahl k an Messungen dem tatsächlichen Wert. Diese Eigenschaft wird als Biasfreiheit bezeichnet.

Aufgrund dessen, dass die Messwerte unkorreliert sind, gilt

$$\text{Var}\{\hat{x}_k\}_r = \sum_{k'=0}^{k-1} \text{Var}\left\{\frac{1}{k} \cdot \tilde{x}_{k'}\right\}_r = \sum_{k'=0}^{k-1} \frac{1}{k^2} \cdot \text{Var}\{\tilde{x}_{k'}\}_r = \frac{1}{k} \cdot \sigma_x^2.$$

Für $k \rightarrow \infty$ geht die Varianz des Schätzwertes gegen null. Zusammen mit der Eigenschaft der Biasfreiheit folgt daraus die Konsistenz des Schätzers.

Ein alternativer Weg die Varianz zu bestimmen beginnt mit dem Aufschreiben der Definition der Varianz,

$$\text{Var}\{\hat{x}_k\}_r = E\{(\hat{x}_k - E\{\hat{x}_k\}_r)^2\}_r = E\{(\hat{x}_k - x)^2\}_r.$$

Ausmultiplizieren der Klammer führt zunächst auf

$$\begin{aligned} \text{Var}\{\hat{x}_k\}_r &= E\{\hat{x}_k^2\}_r - 2 \cdot E\{\hat{x}_k \cdot x\}_r + x^2 = E\{\hat{x}_k^2\}_r - 2 \cdot E\{\hat{x}_k\}_r \cdot x + x^2 = E\{\hat{x}_k^2\}_r - 2 \cdot \hat{x} \cdot x + x^2 \\ &= E\{\hat{x}_k^2\}_r - x^2. \end{aligned}$$

Im verbleibenden Erwartungswert wird \hat{x}_k durch die Schätzgleichung ersetzt¹⁷,

$$E\{\hat{x}_k^2\}_r = \frac{1}{k^2} \cdot E\left\{\sum_{k'} \tilde{x}_{k'} \cdot \sum_{k''} \tilde{x}_{k''}\right\}_r.$$

Für den hier auftretenden Erwartungswert gilt

$$\begin{aligned} E\left\{\left(\sum_{k'} \tilde{x}_{k'}\right) \cdot \left(\sum_{k''} \tilde{x}_{k''}\right)\right\}_r &= E\left\{\sum_{k'} \left(\tilde{x}_{k'} \cdot \sum_{k''} \tilde{x}_{k''}\right)\right\}_r = E\left\{\sum_{k'} \left(\tilde{x}_{k'} \cdot \left(\tilde{x}_{k'} + \sum_{k'' \neq k'} \tilde{x}_{k''}\right)\right)\right\}_r \\ &= E\left\{\sum_{k'} \tilde{x}_{k'} \tilde{x}_{k'}\right\}_r + E\left\{\sum_{k'} \tilde{x}_{k'} \cdot \sum_{k'' \neq k'} \tilde{x}_{k''}\right\}_r \\ &= \sum_{k'} E\{\tilde{x}_{k'}^2\}_r + \sum_{k'} \sum_{k'' \neq k'} E\{\tilde{x}_{k'}\}_r \cdot E\{\tilde{x}_{k''}\}_r \end{aligned}$$

und weiter, mit $E\{\tilde{x}_{k'}^2\}_r = x^2 + \sigma_x^2$,

$$E\left\{\left(\sum_{k'} \tilde{x}_{k'}\right) \cdot \left(\sum_{k''} \tilde{x}_{k''}\right)\right\}_r = k \cdot x^2 + k \cdot \sigma_x^2 + k \cdot (k-1) \cdot x^2 = k \cdot \sigma_x^2 + k^2 \cdot x^2.$$

Die in die Ausgangsgleichung eingesetzt ergibt

$$\text{Var}\{\hat{x}_k\}_r = \frac{1}{k^2} \cdot (k \cdot \sigma_x^2 + k^2 \cdot x^2) - x^2 = \frac{1}{k} \cdot \sigma_x^2.$$

¹⁶ Für die Schätzwerte gibt der Index k die Anzahl der gemittelten Messwerte an. D.h. \hat{x}_1 berücksichtigt nur \tilde{x}_0 und \hat{x}_4 die Werte $\tilde{x}_0, \dots, \tilde{x}_3$. Fasst man die k als Zeitpunkte auf, wird dadurch eine Verzögerung von einem Schritt eingeführt. Jedoch ist diese Darstellung für die Rechnungen etwas günstiger und klarer.

¹⁷ Sind keine Summationsgrenzen angegeben, ist immer 0 als untere und $k-1$ als obere Grenze gemeint.

Schätzer für die Varianz

Die Varianz der Messwerte lässt sich über einen Schätzer der Form

$$\hat{\sigma}_{x,k}^2 = \frac{1}{m} \cdot \sum_{k'} (\tilde{x}_{k'} - \hat{x}_k)^2$$

schätzen, wobei m noch festgelegt wird.

Für den Erwartungswert ergibt sich zunächst, wobei Ergebnisse von oben verwendet werden,

$$\begin{aligned} E\{\hat{\sigma}_{x,k}^2\}_r &= \frac{1}{m} \cdot \left(\sum_{k'} E\{\tilde{x}_{k'}^2\}_r - 2 \cdot \sum_{k'} E\{\tilde{x}_{k'} \hat{x}_k\}_r + \sum_{k'} E\{\hat{x}_k^2\}_r \right) \\ &= \frac{1}{m} \cdot \left(k \cdot (x^2 + \sigma_x^2) - 2 \cdot \sum_{k'} E\{\tilde{x}_{k'} \hat{x}_k\}_r + k \cdot \frac{1}{k^2} \cdot (k^2 \cdot x^2 + k \cdot \sigma_x^2) \right) \\ &= \frac{1}{m} \cdot \left(2k \cdot x^2 + (k+1) \cdot \sigma_x^2 - 2 \cdot \sum_{k'} E\{\tilde{x}_{k'} \hat{x}_k\}_r \right). \end{aligned}$$

Für den verbleibenden Erwartungswert gilt

$$\begin{aligned} E\{\tilde{x}_{k'} \hat{x}_k\}_r &= \frac{1}{k} \cdot E\{\tilde{x}_{k'} \cdot \sum_{k''} \tilde{x}_{k''}\}_r = \frac{1}{k} \cdot \left(E\{\tilde{x}_{k'}^2\}_r + \sum_{\substack{k'' \\ k'' \neq k'}} E\{\tilde{x}_{k'}\}_r \cdot E\{\tilde{x}_{k''}\}_r \right) \\ &= \frac{1}{k} \cdot (x^2 + \sigma_x^2 + (k-1) \cdot x^2) = \frac{1}{k} \cdot (kx^2 + \sigma_x^2) \end{aligned}$$

und damit ergibt sich

$$E\{\hat{\sigma}_{x,k}^2\}_r = \frac{1}{m} \cdot (2k \cdot x^2 + (k+1) \cdot \sigma_x^2 - 2 \cdot (kx^2 + \sigma_x^2)) = \frac{k-1}{m} \cdot \sigma_x^2.$$

Für die naheliegende Wahl $m = k$ ist der Schätzer

$$\hat{\sigma}_{x,k}^2 = \frac{1}{k} \cdot \sum_{k'} (\tilde{x}_{k'} - \hat{x}_k)^2$$

demnach nicht biasfrei, da $E\{\hat{\sigma}_{x,k}^2\}_r = \frac{k-1}{k} \cdot \sigma_x^2 \neq \sigma_x^2$ ist. Da aber zumindest $\lim_{k \rightarrow \infty} \frac{k-1}{k} = 1$ gilt, ist er asymptotisch biasfrei.

Eine biasfreie Schätzung erhält man mit $m = k - 1$,

$$\hat{\sigma}_{x,k}^2 = \frac{1}{k-1} \cdot \sum_{k'} (\tilde{x}_{k'} - \hat{x}_k)^2.$$

Da bei beiden Schätzern die Varianz $\text{Var}\{\hat{\sigma}_{x,k}^2\}_r$ des Schätzwertes $\hat{\sigma}_{x,k}^2$ für $k \rightarrow \infty$ gegen Null geht (hier jetzt nicht gezeigt), sind beide Schätzer konsistent. Dabei ist die Varianz für endliche k beim biasbehafteten Schätzer (etwas) geringer.¹⁸

¹⁸ Dies ist keine Besonderheit. Aus jedem biasfreien Schätzer der Form $\hat{x}_k = \frac{1}{m(k)} \cdot \sum(\cdot)$ lässt sich ein immer noch asymptotisch biasfreier Schätzer mit geringerer Varianz erzeugen, indem einfach zu dem Nenner eine positive Konstante addiert wird, also z. B. $\hat{x}_k = \frac{1}{m(k)+1} \cdot \sum(\cdot)$. Ein solches „Schätzerpaar“ ist in den Abbildungen 4.7b und (c) dargestellt. Wenn Überschätzungen aus irgendeinem Grund möglichst vermieden werden sollen, dann kann die Verwendung der biasbehafteten Variante günstig sein, und auch der $\text{MSE}\{\hat{x}_k - x\}_r$ kann geringer sein. Letzteres hängt von der Größe des zu schätzenden Wertes und der Varianz des Rauschens ab. Ohne aber den Nachweis zu Erbringen, dass der MSE für den biasbehafteten Schätzer geringer ist, ist der Schätzer $\hat{\sigma}_{x,k}^2 = \frac{1}{k-1} \cdot \sum_{k'} (\tilde{x}_{k'} - \hat{x}_k)^2$ als „der“ Schätzer für die Varianz anzusehen.

Schätzer für den Mittelwert

Der (wahre) Mittelwert von k Werten $x_{k'}$ ist

$$\bar{x}_k = \frac{1}{k} \cdot \sum_{k'=0}^{k-1} x_{k'}.$$

Der Schätzer lautet hier wieder

$$\hat{\bar{x}}_k = \frac{1}{k} \cdot \sum_{k'=0}^{k-1} \tilde{x}_{k'}.$$

Der Unterschied zu dem vorher behandelten Fall ist der, dass die unverrauschten Werte $x_{k'}$ hier nicht gleich sind, sondern von k' abhängen, d. h. $E\{\tilde{x}_{k'}\}_r = x_{k'}$.

Der Erwartungswert ist hier

$$E\{\hat{\bar{x}}_k\}_r = \frac{1}{k} \cdot \sum_{k'=0}^{k-1} E\{\tilde{x}_{k'}\}_r = \frac{1}{k} \cdot \sum_{k'=0}^{k-1} x_{k'} = \bar{x}_k.$$

Ebenso ergibt sich für die Varianz wieder

$$\text{Var}\{\hat{\bar{x}}_k\}_r = \frac{1}{k} \cdot \sigma_x^2.$$

Der Schätzer ist also biasfrei und konsistent.

Schätzer für die mittleren quadrierten Werte

Als nächstes werden Schätzer für die mittleren quadrierten Werte

$$\bar{r}_{xx,k} = \frac{1}{k} \cdot \sum_{k'} x_{k'}^2 \quad \left(= \frac{1}{k} \cdot \sum_{k'} r_{xx,k'}^2 \right) \quad (4.7)$$

und

$$\bar{r}_{xy,k} = \frac{1}{k} \cdot \sum_{k'} x_{k'} y_{k'} \quad (4.8)$$

betrachtet.

Ein naheliegender Schätzer für $\bar{r}_{xx,k}$ ist

$$\hat{\bar{r}}_{xx,k} = \frac{1}{k} \cdot \sum_{k'} \tilde{x}_{k'}^2. \quad (4.9)$$

Mit

$$\tilde{x}_{k'} = x_{k'} + n_{x,k'}$$

ergibt sich für den Erwartungswert

$$\begin{aligned} E\{\hat{\bar{r}}_{xx,k}\}_r &= E\left\{ \frac{1}{k} \cdot \sum_{k'} (x_{k'}^2 + 2x_{k'} n_{x,k'} + n_{x,k'}^2) \right\}_r \\ &= \frac{1}{k} \cdot \sum_{k'} x_{k'}^2 + \frac{1}{k} \cdot \sum_{k'} 2x_{k'} E\{n_{x,k'}\}_r + \frac{1}{k} \cdot \sum_{k'} E\{n_{x,k'}^2\}_r \\ &= \bar{r}_{xx,k} + \sigma_x^2. \end{aligned}$$

Der Schätzer (4.9) ist also nicht biasfrei. Er ist nicht einmal asymptotisch biasfrei, da der Term σ_x^2 , der den Bias beschreibt, nicht mit wachsendem k verschwindet.

Für die Varianz gilt

$$\text{Var}\{\hat{r}_{\text{xx},k}\}_r = E\{(\hat{r}_{\text{xx},k} - E\{\hat{r}_{\text{xx},k}\}_r)^2\}_r = E\{(\hat{r}_{\text{xx},k} - \bar{r}_{\text{xx},k} - \sigma_x^2)^2\}_r .$$

Weitere Rechnungen, die in Abschnitt A.1 gezeigt sind, führen auf

$$\text{Var}\{\hat{r}_{\text{xx},k}\}_r = \frac{1}{k} \cdot (4 \cdot \sigma_x^2 \cdot \bar{r}_{\text{xx},k} + 4 \cdot \mu_{\text{xxx}} \cdot \bar{x}_k + \mu_{\text{xxxx}} - (\sigma_x^2)^2) ,$$

wobei μ_{xxx} und μ_{xxxx} das dritte bzw. vierte stochastische Moment des Messrauschens bezeichnet, die als konstant und endlich anzunehmen sind. Damit gilt $\text{Var}\{\hat{r}_{\text{xx},k}\}_r \rightarrow 0$ für $k \rightarrow \infty$, d. h. der Schätzer konvergiert im quadratischen Mittel. Allerdings nicht gegen die tatsächliche Folge $\bar{r}_{\text{xx},k}$, so dass der Schätzer nicht konsistent ist.

$\bar{r}_{\text{xy},k}$ kann über

$$\hat{r}_{\text{xy},k} = \frac{1}{k} \cdot \sum_{k'} \tilde{x}_{k'} \tilde{y}_{k'} \quad (4.10)$$

geschätzt werden. Die Überprüfung des Erwartungswertes

$$\begin{aligned} E\{\hat{r}_{\text{xy},k}\}_r &= E\left\{ \frac{1}{k} \cdot \sum_{k'} (x_{k'} y_{k'} + x_{k'} n_{y,k'} + y_{k'} n_{x,k'} + n_{x,k'} n_{y,k'}) \right\}_r \\ &= \frac{1}{k} \cdot \sum_{k'} x_{k'} y_{k'} + \frac{1}{k} \cdot \sum_{k'} x_{k'} E\{n_{y,k'}\}_r + \frac{1}{k} \cdot \sum_{k'} y_{k'} E\{n_{x,k'}\}_r + \frac{1}{k} \cdot \sum_{k'} E\{n_{x,k'}\}_r \cdot E\{n_{y,k'}\}_r \\ &= \bar{r}_{\text{xy},k} \end{aligned}$$

zeigt, dass der Schätzer (4.10) biasfrei ist. Da auch die Varianz für $k \rightarrow \infty$ gegen null geht (hier jetzt nicht gezeigt), ist dieser Schätzer auch konsistent.

4.6.2 Optimales gewichtetes Mittel

In diesem Abschnitt wird die Aufgabenstellung betrachtet, zwei Zufallszahlen bzw. zwei Vektoren von Zufallszahlen so zu mitteln, dass die Varianz des Mittelwertes minimal ist.

Mittelung skalarer Zufallsvariablen

Gegeben sind jeweils eine Realisierung der Zufallsvariablen $\{x\}_r$ und $\{y\}_r$. Der Erwartungswert dieser beiden Zufallsvariablen ist gleich und beträgt μ . Damit auch der Mittelwert den Erwartungswert μ besitzt, erfolgt die Mittelung in der Form

$$z = (1 - \gamma)x + \gamma y . \quad (4.11)$$

Mit dem Parameter γ kann der Mittelwert zugunsten von x oder y verschoben werden. Dabei ist γ zwischen 0 und 1 zu wählen.

Unkorreliert

Die Varianz des Mittelwertes lautet

$$\sigma_z^2 = (1 - \gamma)^2 \sigma_x^2 + \gamma^2 \sigma_y^2 \quad (4.12)$$

Notwendige Bedingung für ein (lokales) Extremum ist, dass die erste Ableitung

$$\frac{d\sigma_z^2}{d\gamma} = -2 \cdot (1 - \gamma) \cdot \sigma_x^2 + 2\gamma\sigma_y^2$$

null ist. Hieraus ergibt sich als einziges Extremum

$$\gamma_{\text{opt}} = \frac{\sigma_x^2}{\sigma_x^2 + \sigma_y^2}.$$

Da Gl. (4.12) nach oben unbeschränkt ist, muss es sich dabei auch um ein Minimum handeln.

Umformen ergibt¹⁹

$$\sigma_{z,\text{opt}}^2 = \frac{\sigma_x^2 \sigma_y^2}{\sigma_x^2 + \sigma_y^2}.$$

Ist $\sigma_x^2 \ll \sigma_y^2$, d.h. ist x viel sicherer als y bekannt, so gilt $\gamma_{\text{opt}} \rightarrow 0$ und es wird im Wesentlichen der Wert x verwendet. Ist dagegen y sicherer bekannt, $\sigma_x^2 \gg \sigma_y^2$, so geht $\gamma_{\text{opt}} \rightarrow 1$ und es wird der Wert y verwendet.

Ist $\sigma_x^2 = \sigma_y^2$, dann ist $\gamma_{\text{opt}} = 0,5$ und es erfolgt eine gleichmäßige Mittelung. Die Varianz in diesem Fall ist $\sigma_{z,\text{opt}}^2 = 0,5 \cdot \sigma_x^2 = 0,5 \cdot \sigma_y^2$.

Mehrere Variablen

Dieses Ergebnis kann noch verallgemeinert werden. So werden k unkorrelierte Zufallsvariablen $\{x_{k'}\}_r$ dem gleichen Erwartungswert x und den Varianzen $\text{Var}\{x_{k'}\}_r = \sigma_{k'}^2$ nach

$$x_m = \frac{\sum_{k'} \frac{1}{\sigma_{k'}^2} \cdot x_{k'}}{\sum_{k'} \frac{1}{\sigma_{k'}^2}}$$

optimal gemittelt. D.h. dass die Varianz $\text{Var}\{x_m\}_r$ des Mittelwertes minimal ist. Die Varianz des Mittelwertes beträgt dann

$$\text{Var}\{x_m\}_r = \left(\sum_{k'} \frac{1}{\sigma_{k'}^2} \right)^{-1}.$$

Dies ist hier nicht gezeigt, da es später noch einfacher gezeigt werden kann.

¹⁹ $\sigma_{z,\text{opt}}^2 = \left(1 - \frac{\sigma_x^2}{\sigma_x^2 + \sigma_y^2}\right)^2 \cdot \sigma_x^2 + \left(\frac{\sigma_x^2}{\sigma_x^2 + \sigma_y^2}\right)^2 \cdot \sigma_y^2 = \frac{\sigma_x^2 \sigma_y^4 + \sigma_x^4 \sigma_y^2}{(\sigma_x^2 + \sigma_y^2)^2} = \frac{(\sigma_x^2 + \sigma_y^2) \cdot \sigma_x^2 \sigma_y^2}{(\sigma_x^2 + \sigma_y^2)^2}$

Korreliert

Sind $\{x\}_r$ und $\{y\}_r$ korreliert, dann beträgt die Varianz von $\{z\}_r$ aufgrund von Gl. (4.11) allgemein

$$\sigma_z^2 = (1 - \gamma)^2 \sigma_x^2 + \gamma^2 \sigma_y^2 + 2 \cdot (1 - \gamma) \cdot \gamma \cdot \sigma_x \sigma_y \rho_{xy}.$$

Die ersten beiden Summanden entsprechen dem Ausdruck der Varianz für den Mittelwert im unkorrelierten Fall. Der dritte Summand ist positiv, wenn die Korrelation positiv ist ($\rho_{xy} > 0$) und negativ bei negativer Korrelation ($\rho_{xy} < 0$). Hieran kann man schon erkennen, dass für positive Korrelation die Varianz des Mittelwertes höher sein wird als im unkorrelierten Fall. Bei negativer Korrelation sind geringere Varianzen des Mittelwertes erreichbar.

Über

$$\frac{d\sigma_z^2}{d\gamma} = -2 \cdot (1 - \gamma) \cdot \sigma_x^2 + 2\gamma\sigma_y^2 + 2 \cdot (-\gamma + (1 - \gamma)) \cdot \sigma_x \sigma_y \rho_{xy} \stackrel{!}{=} 0$$

ergibt sich

$$\gamma_{\text{opt}} = \frac{\sigma_x^2 - \sigma_x \sigma_y \rho_{xy}}{\sigma_x^2 + \sigma_y^2 - 2 \cdot \sigma_x \sigma_y \rho_{xy}}.$$

Hierbei kann sich unter Umständen für γ_{opt} ein Wert außerhalb des zulässigen Intervalls [0, 1] ergeben. Dann liegt das Minimum auf dem Rand des Definitionsbereiches [0, 1].

Mittelung vektorieller Zufallsvariablen

Im Fall vektorieller Zufallsvariablen $\{x\}_r$ und $\{y\}_r$ wird vom gleichen Ansatz ausgegangen, d. h. die Mittelung erfolgt über eine Gewichtungsmatrix Γ in der Form

$$\mathbf{z} = (\mathbf{I} - \Gamma^T) \cdot \mathbf{x} + \Gamma^T \mathbf{y}.$$

Ausschreiben der Definition der Kovarianzmatrix bzw. das Anwenden der oben formulierten Rechenregeln führt auf die Kovarianzmatrix

$$\begin{aligned} \mathbf{P}_z &= (\mathbf{I} - \Gamma^T) \cdot \mathbf{P}_x \cdot (\mathbf{I} - \Gamma) + \Gamma^T \mathbf{P}_y \Gamma + (\mathbf{I} - \Gamma^T) \cdot \mathbf{P}_{xy} \cdot \Gamma + \Gamma^T \mathbf{P}_{yx} \cdot (\mathbf{I} - \Gamma) \\ &= \mathbf{P}_x - \Gamma^T \cdot (\mathbf{P}_x - \mathbf{P}_{yx}) - (\mathbf{P}_x - \mathbf{P}_{xy}) \cdot \Gamma + \Gamma^T \cdot (\mathbf{P}_x + \mathbf{P}_y - \mathbf{P}_{xy} - \mathbf{P}_{yx}) \cdot \Gamma \end{aligned}$$

von $\{\mathbf{z}\}_r$.

Ziel ist es, die Varianzen der einzelnen in $\{\mathbf{z}\}_r$ enthaltenen Zufallsvariablen zu minimieren. Diese Varianzen stehen auf der Diagonalen der Kovarianzmatrix, und können einzeln in der Form

$$\sigma_i^2 = \mathbf{e}_i^T \mathbf{P}_z \mathbf{e}_i$$

angegeben werden. Für die Varianz der i -ten Komponente von $\{\mathbf{z}\}_r$ folgt damit

$$\sigma_i^2 = \mathbf{e}_i^T \mathbf{P}_x \mathbf{e}_i - \mathbf{e}_i^T \Gamma^T \cdot (\mathbf{P}_x - \mathbf{P}_{yx}) \mathbf{e}_i - \mathbf{e}_i^T (\mathbf{P}_x - \mathbf{P}_{xy}) \cdot \Gamma \mathbf{e}_i + \mathbf{e}_i^T \Gamma^T \cdot (\mathbf{P}_x + \mathbf{P}_y - \mathbf{P}_{xy} - \mathbf{P}_{yx}) \cdot \Gamma \mathbf{e}_i$$

und weiter

$$\begin{aligned} \sigma_i^2 &= \mathbf{e}_i^T \mathbf{P}_x \mathbf{e}_i - \mathbf{e}_i^T \Gamma^T \cdot (\mathbf{P}_x - \mathbf{P}_{xy}^T) \cdot \mathbf{e}_i - \mathbf{e}_i^T \Gamma^T \cdot (\mathbf{P}_x - \mathbf{P}_{xy}^T) \cdot \mathbf{e}_i + \mathbf{e}_i^T \Gamma^T \cdot (\mathbf{P}_x + \mathbf{P}_y - \mathbf{P}_{xy} - \mathbf{P}_{xy}^T) \cdot \Gamma \mathbf{e}_i \\ &= \mathbf{e}_i^T \mathbf{P}_x \mathbf{e}_i - 2 \cdot \mathbf{e}_i^T \Gamma^T \cdot (\mathbf{P}_x - \mathbf{P}_{xy}^T) \mathbf{e}_i + \mathbf{e}_i^T \Gamma^T \cdot (\mathbf{P}_x + \mathbf{P}_y - \mathbf{P}_{xy} - \mathbf{P}_{xy}^T) \cdot \Gamma \mathbf{e}_i. \end{aligned}$$

Die letzte Zeile kann mit der Abkürzung

$$\boldsymbol{\gamma}_i = \boldsymbol{\Gamma} \mathbf{e}_i$$

auch als

$$\sigma_i^2 = \mathbf{e}_i^T \mathbf{P}_x \mathbf{e}_i - 2 \cdot \boldsymbol{\gamma}_i^T \cdot (\mathbf{P}_x - \mathbf{P}_{xy}^T) \cdot \mathbf{e}_i + \boldsymbol{\gamma}_i^T \cdot (\mathbf{P}_x + \mathbf{P}_y - \mathbf{P}_{xy} - \mathbf{P}_{xy}^T) \cdot \boldsymbol{\gamma}_i$$

geschrieben werden. Die Ableitung dieses Ausdrucks lautet²⁰

$$\frac{d\sigma_i^2}{d\boldsymbol{\gamma}_i} = -2 \cdot \mathbf{e}_i^T \cdot (\mathbf{P}_x - \mathbf{P}_{xy}) + 2 \cdot \boldsymbol{\gamma}_i^T \cdot (\mathbf{P}_x + \mathbf{P}_y - \mathbf{P}_{xy} - \mathbf{P}_{xy}^T)$$

und mit $\left. \frac{d\sigma_i^2}{d\boldsymbol{\gamma}_i} \right|_{\boldsymbol{\gamma}_{i,\text{opt}}} = 0$ folgt

$$\boldsymbol{\gamma}_{i,\text{opt}}^T = \mathbf{e}_i^T \cdot (\mathbf{P}_x - \mathbf{P}_{xy}) \cdot (\mathbf{P}_x + \mathbf{P}_y - \mathbf{P}_{xy} - \mathbf{P}_{xy}^T)^{-1}$$

für den optimalen Wert für $\boldsymbol{\gamma}_i$. Zuletzt können die zu den jeweiligen Einträgen gehörenden Vektoren $\boldsymbol{\gamma}_i$ wieder zu einer Matrix

$$\boldsymbol{\Gamma}_{\text{opt}}^T = (\mathbf{P}_x - \mathbf{P}_{xy}) \cdot (\mathbf{P}_x + \mathbf{P}_y - \mathbf{P}_{xy} - \mathbf{P}_{xy}^T)^{-1}$$

zusammengesetzt werden.

Im Sonderfall unkorrelierter Zufallsvektoren folgt aus $\mathbf{P}_{xy} = \mathbf{0}$

$$\boldsymbol{\Gamma}_{\text{opt}}^T = \mathbf{P}_x \cdot (\mathbf{P}_x + \mathbf{P}_y)^{-1}.$$

Für diesen Sonderfall ist $\boldsymbol{\Gamma}_{\text{opt}}^T \cdot \mathbf{P}_x$ symmetrisch, und es ergibt sich für die Kovarianzmatrix des gemittelten Vektors $\{\mathbf{z}\}_r$

$$\begin{aligned} \mathbf{P}_{z,\text{opt}} &= \mathbf{P}_x - \boldsymbol{\Gamma}_{\text{opt}}^T \cdot \mathbf{P}_x - \mathbf{P}_x \cdot \boldsymbol{\Gamma}_{\text{opt}} + \boldsymbol{\Gamma}_{\text{opt}}^T \cdot (\mathbf{P}_x + \mathbf{P}_y) \cdot \boldsymbol{\Gamma}_{\text{opt}} \\ &= \mathbf{P}_x - 2 \cdot \boldsymbol{\Gamma}_{\text{opt}}^T \cdot \mathbf{P}_x + \boldsymbol{\Gamma}_{\text{opt}}^T \cdot (\mathbf{P}_x + \mathbf{P}_y) \cdot \boldsymbol{\Gamma}_{\text{opt}} \\ &= \mathbf{P}_x - \boldsymbol{\Gamma}_{\text{opt}}^T \cdot \mathbf{P}_x \\ &= (\mathbf{I} - \boldsymbol{\Gamma}_{\text{opt}}^T) \cdot \mathbf{P}_x \\ &= (\mathbf{I} - \mathbf{P}_x \cdot (\mathbf{P}_x + \mathbf{P}_y)^{-1}) \cdot \mathbf{P}_x \\ &= \mathbf{P}_y \cdot (\mathbf{P}_x + \mathbf{P}_y)^{-1} \cdot \mathbf{P}_x. \end{aligned}$$

Im letzten Schritt wurde die Matrix $(\mathbf{P}_x + \mathbf{P}_y)^{-1}$ nach rechts aus der äußeren Klammer ausgeklammert.

Es ist zu beachten, dass die Mittelung der einzelnen Einträge der Vektoren durch jeweils eigene Zeilen von $\boldsymbol{\Gamma}^T$ erfolgt. D.h. die Mittelung des ersten Eintrags von x und y beeinflusst nicht die Mittelung der zweiten oder anderer Einträge, und sie wird auch nicht von den Mittelungen anderer Einträge beeinflusst. Damit konnte gefordert werden, dass die Varianz jedes Eintrags von $\{\mathbf{z}\}_r$ minimal ist, und nicht nur die Summe der Varianzen!

²⁰ $\frac{dx^T b}{dx} = b^T$, $\frac{dx^T A x}{dx} = 2x^T A$ wenn A symmetrisch

4.6.3 Quotientenbildung von Zufallsprozessen

Häufig liegt ein betrachteter Zusammenhang der Form

$$z = \frac{x}{y}$$

vor. Der Wert von z soll dabei aus Messgrößen $\{\tilde{x}_k\}_r$ und $\{\tilde{y}_k\}_r$ geschätzt werden. Als mögliche Schätzer bieten sich dazu zunächst

$$\hat{z}_k = \frac{1}{k} \cdot \sum_{k'} \frac{\tilde{x}_{k'}}{\tilde{y}_{k'}} \quad (4.13)$$

oder

$$\hat{z}_k = \frac{\frac{1}{k} \cdot \sum_{k'} \tilde{x}_{k'}}{\frac{1}{k} \cdot \sum_{k'} \tilde{y}_{k'}} = \frac{\hat{x}_k}{\hat{y}_k} \quad (4.14)$$

an.

Eigenschaften des Kehrwerts einer Zufallsvariablen

Als ersten Schritt wird

$$z = \frac{1}{y}$$

betrachtet, wobei $\{y\}_r$ eine Zufallsvariable ist. In Abbildung 4.8 sind die Verteilungsdichten von $\{z\}_r$ für verschiedene normalverteilte $\{y\}_r$ dargestellt.²¹

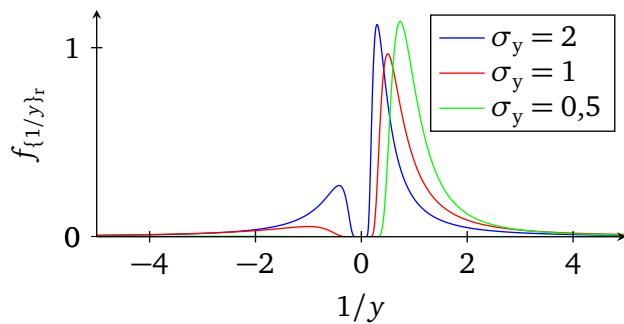


Abbildung 4.8: Verteilungsdichte der Zufallsvariablen $z = \frac{1}{y}$ mit $\{y\}_r \sim N(1, \sigma_y^2)$

Allgemein gilt für eine Funktion $g(y)$ einer Zufallsvariablen $\{y\}_r$, dass der Erwartungswert über

$$E\{g(y)\}_r = \int_{-\infty}^{\infty} g(y) \cdot f_{\{y\}_r}(y) \cdot dy$$

berechnet werden kann.

Geht man von einer normalverteilten Zufallsvariablen aus, d. h. $\{y\}_r \sim N(\mu_y, \sigma_y^2)$, dann ergibt sich damit

$$E\{z\}_r = \int_{-\infty}^{\infty} \frac{1}{y} \cdot \varphi(y, \mu_y, \sigma_y^2) \cdot dy$$

²¹ Für eineindeutige Funktionen $z = g(y)$ gilt $f_{\{z\}_r}(z) = \frac{1}{|g'(g^{-1}(z))|} \cdot f_{\{y\}_r}(g^{-1}(z))$ mit $g'(y) = \frac{dg(y)}{dy}$, [Hänsler 2001, S. 137].

für den Erwartungswert für $\{z\}_r$. Einsetzen der Verteilungsdichte der Normalverteilung führt zunächst auf

$$E\{z\}_r = \int_{-\infty}^{\infty} \frac{1}{y} \cdot \frac{1}{\sqrt{2\pi\sigma_y^2}} \cdot e^{-\frac{(y-\mu_y)^2}{2\sigma_y^2}} \cdot dy = \frac{1}{\sqrt{2\pi\sigma_y^2}} \cdot \int_{-\infty}^{\infty} \frac{1}{y} \cdot e^{-\frac{(y-\mu_y)^2}{2\sigma_y^2}} \cdot dy .$$

Das Integral kann in vier Teilintegrale aufgeteilt werden,

$$\begin{aligned} \int_{-\infty}^{\infty} \frac{1}{y} \cdot e^{-\frac{(y-\mu_y)^2}{2\sigma_y^2}} \cdot dy &= \int_{-\infty}^{-a} \frac{1}{y} \cdot e^{-\frac{(y-\mu_y)^2}{2\sigma_y^2}} \cdot dy + \int_{-a}^0 \frac{1}{y} \cdot e^{-\frac{(y-\mu_y)^2}{2\sigma_y^2}} \cdot dy \\ &\quad + \int_0^b \frac{1}{y} \cdot e^{-\frac{(y-\mu_y)^2}{2\sigma_y^2}} \cdot dy + \int_b^{\infty} \frac{1}{y} \cdot e^{-\frac{(y-\mu_y)^2}{2\sigma_y^2}} \cdot dy , \end{aligned}$$

mit $a > 0$ und $b > 0$. Die beiden „äußersten“ uneigentlichen Integrale ergeben endliche Werte, da $\int_a^{\infty} y^{-1} dy$, $a > 0$, endlich ist. Das zweite Integral kann nach oben durch

$$\int_{-a}^0 \frac{1}{y} \cdot e^{-\frac{(y-\mu_y)^2}{2\sigma_y^2}} \cdot dy < \min_{y \in [-a, 0]} \left\{ e^{-\frac{(y-\mu_y)^2}{2\sigma_y^2}} \right\} \cdot \int_{-a}^0 \frac{1}{y} \cdot dy \rightarrow -\infty$$

abgeschätzt werden. Dieses besitzt keinen endlichen Wert. Auf gleiche Weise ergibt sich

$$\int_0^b \frac{1}{y} \cdot e^{-\frac{(y-\mu_y)^2}{2\sigma_y^2}} \cdot dy \rightarrow \infty .$$

D. h. die Zufallsvariable $\{1/y\}_r$ besitzt keinen Erwartungswert, und damit auch keine Momente höherer Ordnung, also keine Varianz.²²

Dass der Erwartungswert der Zufallsvariablen $\{1/y\}_r$ nicht existiert, ist an die Verteilung von $\{y\}_r$ gebunden. Ist die Verteilungsdichte für Realisierungen um den Wert Null null, d. h. wird niemals durch Null oder eine Zahl geteilt, die beliebig nahe an Null herankommt, dann existiert auch der Erwartungswert.

In Abbildung 4.9a sind die Verteilungsdichten der auf das Intervall $[0, 1, 1, 9]$ gestützten Normalverteilungen mit den Standardabweichungen 2, 1 und 0,5 und dem gemeinsamen Mittelwert 1 dargestellt. Eine solche Stützung entspricht im Grunde dem Herausfiltern von offensichtlichen Ausreißern oder unrealistischen Werten bei einer Messung. (Jedoch kann diese nicht wie hier exakt symmetrisch bezogen auf den Mittelwert erfolgen, da dieser ja nicht bekannt ist.) Besitzt $\{y\}_r$ eine solche Verteilung, dann können nicht nur die Verteilungsdichtefunktionen von $\{1/y\}_r$, sondern auch die Erwartungswerte von $\{1/y\}_r$ berechnet werden. Die Verteilungsdichten sind in Abbildung 4.9b dargestellt. Die Erwartungs- bzw. Mittelwerte sind in den selben Plot als senkrechte, gestrichelte Linien eingezeichnet. Ebenfalls ist der Wert $1/y = 1$ markiert, der dem „wahren“ Wert entspricht.

Es ist deutlich, dass ein Bias vorliegt. Dieser entsteht dadurch, dass $1/y$ eine nichtlineare Funktion ist. Je geringer die Varianz von $\{y\}_r$ ist, desto geringer ist auch der Bias. Der genaue Wert hängt hier auch noch von dem Intervall ab, auf das die Verteilungsdichten gestützt sind.

²² Dies ist auch dann der Fall, wenn der Mittelwert von $\{y\}_r$ null ist. Zwar ergibt das Gesamtintegral aufgrund der Symmetrie dann Null, jedoch fordert die genaue Definition des Erwartungswertes, die hier jetzt nicht vorgestellt wurde, dass das Integral in den beiden Intervallen $(-\infty, 0]$ und $[0, \infty)$ endlich sein muss.

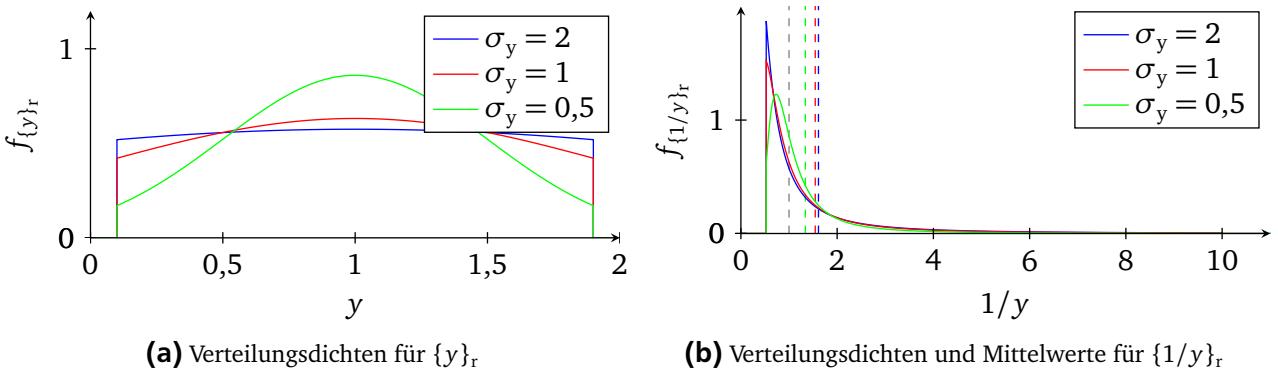


Abbildung 4.9

Eigenschaften des Quotienten zweier Zufallsvariablen

Im Fall

$$z = \frac{x}{y}$$

mit den Zufallsvariablen $\{x\}_r$ und $\{y\}_r$, „erbt“ die sich ergebende Zufallsvariable $\{z\}_r$ die Probleme der Verteilung $1/y$. Die genaue Verteilung hängt in diesem Fall aber natürlich auch noch von der Verteilung von $\{x\}_r$ sowie der Korrelation zwischen $\{x\}_r$ und $\{y\}_r$ ab.²³

Wenn die Standardabweichungen klein gegenüber den Mittelwerten sind, insbesondere für die Zufallsvariable im Nenner, dann kann man den Quotienten x/y näherungsweise durch dessen Taylorreihenentwicklung um die Erwartungswerte von $\{x\}_r$ und $\{y\}_r$ bis zum quadratischen Glied darstellen,

$$\begin{aligned} \frac{x}{y} &\approx \frac{\mu_x}{\mu_y} + \left. \frac{\partial \frac{x}{y}}{\partial x} \right|_{\mu_x, \mu_y} \cdot (x - \mu_x) + \left. \frac{\partial \frac{x}{y}}{\partial y} \right|_{\mu_x, \mu_y} \cdot (y - \mu_y) \\ &\quad + \frac{1}{2} \cdot \left. \frac{\partial^2 \frac{x}{y}}{\partial x^2} \right|_{\mu_x, \mu_y} \cdot (x - \mu_x)^2 + \frac{1}{2} \cdot \left. \frac{\partial^2 \frac{x}{y}}{\partial y^2} \right|_{\mu_x, \mu_y} \cdot (y - \mu_y)^2 + \left. \frac{\partial^2 \frac{x}{y}}{\partial x \partial y} \right|_{\mu_x, \mu_y} \cdot (x - \mu_x) \cdot (y - \mu_y). \end{aligned}$$

Damit erhält man

$$E\left\{\frac{x}{y}\right\}_r \approx \frac{\mu_x}{\mu_y} \cdot \left(1 + \left(\frac{\sigma_y}{\mu_y} \right)^2 - \rho_{xy} \frac{\sigma_x \sigma_y}{\mu_x \mu_y} \right)$$

als Näherungswert für den Erwartungswert. Die Varianz kann über

$$\text{Var}\left\{\frac{x}{y}\right\}_r \approx \left(\frac{\mu_x}{\mu_y} \right)^2 \cdot \left(\left(\frac{\sigma_x}{\mu_x} \right)^2 + \left(\frac{\sigma_y}{\mu_y} \right)^2 - 2 \rho_{xy} \frac{\sigma_x \sigma_y}{\mu_x \mu_y} \right)$$

angenähert werden.

Man kann daran erkennen, dass die Variation des Nenners einen Bias des Quotienten erzeugt. Auf die Varianz haben beide Zufallszahlen gleichermaßen Einfluss. Dies gilt aber wie schon gesagt nur, wenn der Nenner keine Werte Nahe null annimmt, da ansonsten die quadratische Näherung nicht gilt. Ist die Standardabweichung des Nenners bezogen auf seinen Mittelwert groß, dann dominiert dieser die stochastischen Eigenschaften.

²³ Für den Sonderfall $z = x/y$ mit standardnormalverteilten Zufallsvariablen $\{x\}_r$ und $\{y\}_r$, die darüber hinaus nicht korreliert sind, heißt die sich ergebende Verteilung von $\{z\}_r$ Standard-Cauchy-Verteilung.

Mittelung der Quotienten

Nach diesen Vorüberlegungen wird der erste Schätzer (4.13),

$$\hat{z}_k = \frac{1}{k} \cdot \sum_{k'} \frac{\tilde{x}_{k'}}{\tilde{y}_{k'}} ,$$

betrachtet. Und es lässt sich direkt die Problematik dieses Schätzers erkennen. Wenn die Varianz des Nenners gegenüber seinem Mittelwert groß ist, dann sind die einzelnen Schätzwerte sehr schlecht. Es treten extreme Ausreißer auf, die sich kaum durch eine Mittelung eliminieren lassen.

Wenn die Zufallsvariablen $\{y_k\}_r$ Verteilungen besitzen, bei der die einzelnen Quotienten einen Erwartungswert besitzen, dann ist dieser Schätzer im Allgemeinen biasbehaftet. Dieser Bias verschwindet auch nicht asymptotisch, da – bei gleicher Verteilung der Zufallsvariablen $\{x_k\}_r$ und $\{y_k\}_r$ für alle k – jeder der Quotienten $\frac{\tilde{x}_{k'}}{\tilde{y}_{k'}}$ den gleichen Bias besitzt.

Mittelung von Zähler und Nenner einzeln

Bei dem Schätzer (4.14),

$$\hat{z}_k = \frac{\frac{1}{k} \cdot \sum_{k'} \tilde{x}_{k'}}{\frac{1}{k} \cdot \sum_{k'} \tilde{y}_{k'}} = \frac{\hat{x}_k}{\hat{y}_k} , \quad (4.15)$$

dagegen werden erst die Messwerte des Zählers und des Nenners getrennt gemittelt, und dann der Quotient gebildet. Durch diese Mittelung kann die Varianz des Nenners $\{\hat{y}_k\}_r$ unter kritische Werte gedrückt werden, so dass die Quotientenbildung gut möglich ist.

Nimmt man normalverteilte Zufallsvariablen an, so existiert der Erwartungswert zwar theoretisch auch hier nicht, jedoch kann durch die vorhergehende Mittelung für $\{\hat{y}_k\}_r$ auch gut von einer gestützten Normalverteilungen ausgegangen werden. Aber auch ohne diese Annahme kann gezeigt werden, dass der Schätzer konsistent ist.

Wird normalverteiltes Rauschen angenommen, so ist auch der Schätzwert des Nenners normalverteilt und damit kann auch für \hat{z}_k kein Erwartungswert angegeben werden. Jedoch gilt

$$\text{plim}_{k \rightarrow \infty} \{\hat{x}_k\}_r = x$$

und

$$\text{plim}_{k \rightarrow \infty} \{\hat{y}_k\}_r = y$$

und damit folgt aus dem Satz von Slutsky

$$\text{plim}_{k \rightarrow \infty} \{\hat{z}_k\}_r = \frac{x}{y} = z .$$

Der Schätzer ist also konsistent.

Auslöschung

Die Mittelung von Zähler und Nenner einzeln ist nach dem bisher gesagten klar zu bevorzugen.

Durch die Mittelung können sich aber Nutzanteile der Signale teilweise wieder auslöschen, nicht aber das Rauschen. Damit sinkt das Signal-Rausch-Verhältnis unbeabsichtigt. Dies ist beispielhaft anhand Abbildung 4.10 diskutiert.

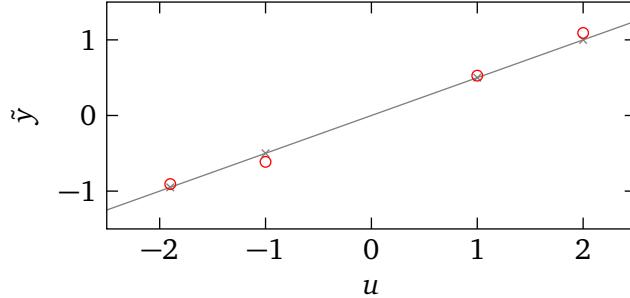


Abbildung 4.10: Beispiel zur Auslöschung von Nutzdaten (grau: Exakter Zusammenhang und unverrauschte Messungen)

Es liegt ein Zusammenhang $y = \theta \cdot u$ mit $\theta = 0,5$ vor. Der Parameter θ soll anhand gemessener Daten geschätzt werden. Es liegen vier Messwerte (u_k, \tilde{y}_k) vor,

$$(2, 1,0917) \quad (1, 0,5269) \quad (-1, -0,6129) \quad (-1,9, -0,9069).$$

Für die gemessenen Werte \tilde{y}_k gilt $\tilde{y}_k = y_k + n_{y,k}$, wobei $\{n_{y,k}\}_r$ ein mittelwertfreier und unkorrelierter Rauschprozess mit der Varianz $\text{Var}\{n_{y,k}\}_r = 0,05^2$ darstellt.

Bestimmt man den Schätzwert nach

$$\hat{\theta} = \frac{\tilde{y}_0 + \tilde{y}_1 + \tilde{y}_2 + \tilde{y}_3}{u_0 + u_1 + u_2 + u_3}$$

erhält man

$$\hat{\theta} = 0,9874,$$

ein offensichtlich sehr schlechter Schätzwert.

Betrachtet man nur den Nutzsignalanteil der Messgröße, so erhält man natürlich den exakten Parameterwert (solange die Summe der u_k' nicht Null ergibt)

$$\theta = \frac{y_0 + y_1 + y_2 + y_3}{u_0 + u_1 + u_2 + u_3} = \frac{0,05}{0,1}.$$

Dadurch, dass sich bei der Mittelung aber die positiven und negativen Anteile beinahe vollständig aufheben, gewinnt das Rauschen deutlich an Einfluss. Den Schätzwert kann man in einen Teil, der die Nutzsignale y_k' beinhaltet, und einen Teil, der den Rauscheinfluss beschreibt, aufteilen,

$$\hat{\theta} = \frac{\sum y_{k'}}{\sum u_{k'}} + \frac{\sum n_{y,k'}}{\sum u_{k'}} = \theta + \frac{\sum n_{y,k'}}{\sum u_{k'}}.$$

Da hier u_k unverrauscht und $\{n_{y,k}\}_r$ mittelwertfrei ist, ist $\{\hat{\theta}\}_r$ biasfrei. Aber für die Varianz gilt

$$\text{Var}\{\hat{\theta}\}_r = \frac{4}{(\sum u_{k'})^2} \cdot \text{Var}\{n_{y,k}\}_r = \frac{4}{0,01} \cdot 0,05^2 = 1.$$

Die Standardabweichung von $\{\hat{\theta}\}_r$ ist hier mit 1 schon höher als der eigentliche Parameterwert!

Verwendet man nur die beiden Messungen mit positiven Werten für die Schätzung, so erhält man

$$\hat{\theta} = \frac{\tilde{y}_0 + \tilde{y}_1}{u_0 + u_1} = 0,5395.$$

Für die Varianz gilt hier

$$\text{Var}\{\hat{\theta}\}_r = \frac{2}{(\sum_{k'=0}^1 u_{k'})^2} \cdot \text{Var}\{n_{y,k}\}_r = \frac{2}{3} \cdot 0,05^2 = 0,001667,$$

woraus sich für die Standardabweichung 0,041 ergibt. Diese ist deutlich kleiner, obwohl nur zwei der vier Messungen berücksichtigt sind.

Kann nicht sichergestellt werden, dass eine solche Auslöschung ausgeschlossen wird, was durchaus der Fall sein kann, wie später noch diskutiert wird, dann ist auch die Mittelung von Zähler und Nenner einzeln keine ideale Lösung. Auf die Mittelung der Quotienten überzugehen sollte aber auch nur in den Fällen in Erwägung gezogen werden, wenn das Rauschen des Nenners sehr klein oder gar nicht vorhanden ist.

Schätzung über quadrierte Werte (Korrelation)

Eine weitere Möglichkeit besteht darin, von

$$z = \frac{xx}{xy}$$

auszugehen. Hierbei wurde einfach der Quotient x/y mit x erweitert, was dessen Wert nicht verändert.

Analog zu Gl. (4.14) kann $\{z\}_r$ hier über

$$\hat{z}_k = \frac{\frac{1}{k} \cdot \sum_{k'} \tilde{x}_{k'} \tilde{x}_{k'}}{\frac{1}{k} \cdot \sum_{k'} \tilde{x}_{k'} \tilde{y}_{k'}} = \frac{\hat{\bar{r}}_{xx,k}}{\hat{\bar{r}}_{xy,k}}$$

geschätzt werden.

Es wurde schon gezeigt, dass

$$E\{\hat{\bar{r}}_{xx,k}\}_r = \bar{r}_{xx,k} + \sigma_x^2$$

und

$$E\{\hat{\bar{r}}_{xy,k}\}_r = \bar{r}_{xy,k}$$

gilt, und es wurde erwähnt, dass die Schätzer für $E\{\hat{\bar{r}}_{xx,k}\}_r$ und $E\{\hat{\bar{r}}_{xy,k}\}_r$ im quadratischen Mittel konvergieren. Daraus kann

$$\text{plim}_{k \rightarrow \infty} \{\hat{z}_k\}_r = \frac{\text{plim}_{k \rightarrow \infty} \{\hat{\bar{r}}_{xx,k}\}_r}{\text{plim}_{k \rightarrow \infty} \{\hat{\bar{r}}_{xy,k}\}_r} = \frac{\bar{r}_{xx,\infty} + \sigma_x^2}{\bar{r}_{xy,\infty}} = \frac{\bar{r}_{xx,\infty}}{\bar{r}_{xy,\infty}} \cdot \left(1 + \frac{\sigma_x^2}{\bar{r}_{xx,\infty}}\right) = z \cdot \left(1 + \frac{\sigma_x^2}{\bar{r}_{xx,\infty}}\right)$$

angegeben werden. Die beiden Größen $\bar{r}_{xx,\infty}$ und $\bar{r}_{xy,\infty}$ entsprechen den tatsächlichen mittleren quadrierten Werten für $k \rightarrow \infty$, und es wird davon ausgegangen, dass diese nicht null sind.

Analog kann der ursprüngliche Zusammenhang mit y erweitert werden,

$$z = \frac{xy}{yy},$$

was auf den Schätzer

$$\hat{z}_k = \frac{\frac{1}{k} \cdot \sum_{k'} \tilde{x}_{k'} \tilde{y}_{k'}}{\frac{1}{k} \cdot \sum_{k'} \tilde{y}_{k'} \tilde{y}_{k'}} = \frac{\hat{\bar{r}}_{xy,k}}{\hat{\bar{r}}_{yy,k}}$$

führt, der gegen

$$\text{plim}_{k \rightarrow \infty} \{\hat{z}_k\}_r = \frac{\bar{r}_{xy,\infty}}{\bar{r}_{yy,\infty} + \sigma_y^2} = z \cdot \left(\frac{1}{1 + \frac{\sigma_y^2}{\bar{r}_{yy,\infty}}} \right)$$

konvergiert.

Bei den beiden Schätzern dieses Abschnitts tritt zwar das Problem der Auslöschung nicht auf, jedoch sind beide nicht biasfrei, nicht einmal asymptotisch biasfrei. Es sollte der Schätzer gewählt werden, für den der aufgrund der gegebenen Rauscheigenschaften (σ_x^2, σ_y^2) und Nutzsignaleigenschaften ($\bar{r}_{xx,k}, \bar{r}_{yy,k}$) der geringere Bias zu erwarten ist.

5 Grundlagen Signalverarbeitung

Unter den Methoden die später vorgestellt werden, arbeiten viele im Frequenzbereich. Die den Methoden zugrundeliegenden Ansätze gehen dabei (häufig) von fouriertransformierten Daten aus. So lautet der einfachste Frequenzgangschätzer

$$\hat{H}(\omega) = \frac{\mathcal{F}(\tilde{y}(t))}{\mathcal{F}(\tilde{u}(t))},$$

d. h. es wird der Quotient aus den Fouriertransformierten des Aus- und Eingangssignales bestimmt.

Die Messdaten liegen zunächst im Zeitbereich vor. Da bei der Datenaufnahme diese immer abgetastet werden (die Realisierung der vorgestellten Methoden mit einem Analogrechner wird, auch wenn teilweise prinzipiell möglich, in dieser Vorlesung nicht betrachtet) und die Messdauer immer begrenzt ist, wird die Transformation über die DFT berechnet werden. Es stellt sich damit die Frage, ob bzw. unter welchen Umständen die Fouriertransformierte durch die DFT ersetzt werden kann.

Dies motiviert dieses Kapitel. Es werden die verschiedenen Transformationen, die kontinuierliche oder abgetastete Zeitdaten in den Frequenzbereich transformieren, betrachtet. Diese sind die Fourierreiheentwicklung für periodische Signale, die Fouriertransformation, die zeitdiskrete Fouriertransformation sowie die diskrete Fouriertransformation.¹ Mathematisch handelt es sich dabei um unabhängige, jeweils für sich definierte Transformationen. Unter gewissen Voraussetzungen an die Signale im Zeitbereich zeigen sich jedoch interpretative Zusammenhänge. (Genau genommen erlauben diese Interpretationen es erst, wie hier von *dem* Frequenzbereich zu sprechen.)

Es wird dabei davon ausgegangen, dass im Zeitbereich immer reelwertige Größen vorliegen.

5.1 Fourierreihe (FR)

Die Fourierreihe für T -periodische Signale $y(t)$ ist definiert als

$$y(t) = \sum_{l=-\infty}^{\infty} \check{y}_l e^{jl\omega_T t} \quad (5.1)$$

$$= \check{y}_0 + \sum_{l=1}^{\infty} (\check{y}_l e^{jl\omega_T t} + \check{y}_{-l} e^{-jl\omega_T t}) \quad (5.2)$$

mit $\omega_T = \frac{2\pi}{T}$ und den Fourier-Koeffizienten

$$\check{y}_l = \mathfrak{F}(y(t))_l = \frac{1}{T} \int_0^T y(t) \cdot e^{-jl\omega_T t} dt \quad (5.3)$$

bzw.

$$\check{y}_l = \mathfrak{F}(y(t))_l = \frac{1}{T} \int_{\tau}^{\tau+T} y(t) \cdot e^{-jl\omega_T t} dt,$$

¹ Daneben wird kurz noch die LaplaceTransformation betrachtet.

da der Startpunkt der Integration aufgrund der T -Periodizität von $y(t)$ und $e^{-jl\omega_T t}$ beliebig ist. Mit $e^{-jl\omega_T t} = \cos(l\omega_T t) - j \cdot \sin(l\omega_T t)$ folgt unter der gemachten Annahme eines reelwertigen $y(t)$, dass die negativen Fourierkoeffizienten gerade das konjugiert-komplexe des jeweils entsprechenden positiven Fourierkoeffizienten sind,

$$\check{y}_{-l} = \check{y}_l^* .$$

Damit kann Gl. (5.2) zu

$$y(t) = \check{y}_0 + \sum_{l=1}^{\infty} 2 \cdot (\operatorname{Re}(\check{y}_l) \cdot \cos(l\omega_T t) - \operatorname{Im}(\check{y}_l) \cdot \sin(l\omega_T t))$$

umgeformt werden. Für jedes l kann die Summe aus Kosinus- und Sinusschwingung mit dem Zusammenhang $B \cdot \cos(\Omega t) + C \cdot \sin(\Omega t) = A \cdot \sin(\Omega t + \varphi)$ mit $A = \sqrt{B^2 + C^2}$ und $\tan \varphi = \frac{B}{C}$ durch eine einzige Winkelfunktion beschrieben werden, und Gl. (5.1) so alternativ auch als

$$y(t) = \sum_{l=0}^{\infty} A_l \cdot \sin(\omega_l t + \varphi_l)$$

oder

$$y(t) = \sum_{l=0}^{\infty} A_l \cdot \cos(\omega_l t + \phi_l)$$

mit

$$\omega_l = l \cdot \omega_T ,$$

$$A_l = \begin{cases} \check{y}_0 & l = 0 \\ 2 \cdot |\check{y}_l| & l > 0 \end{cases}$$

und

$$\varphi_l = \begin{cases} 0 & l = 0 \\ \operatorname{atan2}(\operatorname{Re}(\check{y}_l), -\operatorname{Im}(\check{y}_l)) & l > 0 \end{cases}$$

bzw.

$$\phi_l = \begin{cases} 0 & l = 0 \\ \operatorname{atan2}(\operatorname{Im}(\check{y}_l), \operatorname{Re}(\check{y}_l)) & l > 0 \end{cases}$$

geschrieben werden. D. h. baut man die Zeitlösung aus Kosinusfunktionen auf, dann entsprechen deren Phasen den Argumenten der positiven (d. h. $l > 0$) Fourierkoeffizienten. Baut man die Zeitlösung aus Sinusfunktionen auf, dann entsprechen die Phasen den Argumenten der um 90° gedrehten, d. h. mit $1j$ multiplizierten \check{y}_l ,

$$\varphi_l = \begin{cases} 0 & l = 0 \\ \arg(1j \cdot \check{y}_l) & l > 0 \end{cases}$$

$$\phi_l = \begin{cases} 0 & l = 0 \\ \arg(\check{y}_l) & l > 0 . \end{cases}$$

In dieser Darstellung wird deutlich, dass mit Ausnahme von $l = 0$ die Fourierkoeffizienten die *halbe* Amplitude angeben. Man könnte diese auch entsprechend anders definieren, jedoch passt die Definition wie hier später besser zur DFT.

Interpretation

Die Fourierreihenzerlegung besagt, dass jedes periodische Signal mathematisch auch als Summe harmonischer Schwingungen geschrieben werden kann. Man kann damit sagen, dass das Signal die entsprechenden Frequenzen „enthält“.

Beispiel 1: Monofrequentes Signal

Es wird als erstes Beispiel das Signal

$$y(t) = \cos(\omega_T t) + 2 \cdot \sin(\omega_T t)$$

betrachtet. Setzt man dieses in die Gleichung (5.3) zur Bestimmung der Fourierkoeffizienten ein und ersetzt man dort noch $e^{-jl\omega_T t}$ durch die Eulerformel, so folgt aus den Orthogonalitätseigenschaften der Winkelfunktionen direkt, dass

$$\check{y}_l = 0 \quad \forall l \notin \{-1, 1\}$$

gilt. Für $l = 1$ ergibt sich dagegen

$$\check{y}_1 = \frac{1}{2} - j \cdot 1 .$$

(Und für $l = -1$ entsprechend $\check{y}_{-1} = \frac{1}{2} + j \cdot 1$.) Damit ist auch $y(t)$ eine harmonische Schwingung mit der Amplitude $A_1 = 2 \cdot \sqrt{5/4} = \sqrt{5}$ und der Phase

$$\varphi = \arg(j \cdot \check{y}_1) = \arg(1 + j \cdot 0,5) = \text{atan}(0,5) = 26,57^\circ$$

bzw.

$$\phi = \arg(\check{y}_1) = \text{atan}(-1, 0,5) = \text{atan}(-2) = -63,43^\circ .$$

D. h.

$$y(t) = \sqrt{5} \cdot \sin(\omega_T t + 26,57^\circ) = \sqrt{5} \cdot \cos(\omega_T t - 63,43^\circ) .$$

Für ein allgemeines monofrequentes Signal

$$y(t) = a \cdot \cos(\omega_T t) + b \cdot \sin(\omega_T t)$$

ergibt sich

$$\check{y}_l = \begin{cases} \frac{1}{2} \cdot (a - j \cdot b) & l = 1 \\ \frac{1}{2} \cdot (a + j \cdot b) & l = -1 \\ 0 & \text{sonst.} \end{cases}$$

Beispiel 2: Rechteckschwingung

Als zweites Beispiel wird die in Abbildung 5.1a dargestellte Rechteckschwingung

$$y(t) = \begin{cases} h & 0 \leq \text{mod}(t, T) < T/2 \\ -h & T/2 \leq \text{mod}(t, T) < T \end{cases}$$

betrachtet.² Für die Fourierkoeffizienten, die auch in Abbildung 5.1b dargestellt sind, ergibt sich

$$\check{y}_l = \begin{cases} j \cdot \frac{-2 \cdot h}{\pi \cdot l} & l \text{ ungerade} \\ 0 & l \text{ gerade} \end{cases}$$

und damit kann $y(t)$ auch als

$$y(t) = \frac{4h}{\pi} \cdot \sum_{l'=1}^{\infty} \frac{1}{(2l'-1)} \cdot \sin((2l'-1) \cdot \omega_T t)$$

geschrieben werden. Anders als im ersten Beispiel enthält das Signal $y(t)$ mehr als eine Frequenz. Genauer enthält es unendlich viele Frequenzen, da es kein L mit $\check{y}_l = 0 \forall l > L$ gibt. Die höheren Frequenzanteile haben jedoch einen geringeren Einfluss, da der Betrag der Fourierkoeffizienten \check{y}_l für $l \rightarrow \infty$ gegen null geht, und damit auch die Amplituden der Sinusschwingungen der Reihe für hohe Frequenzen gegen null geht. Der Rechteckimpuls ist damit auch ein Beispiel für ein nicht-bandbegrenztes Signal.

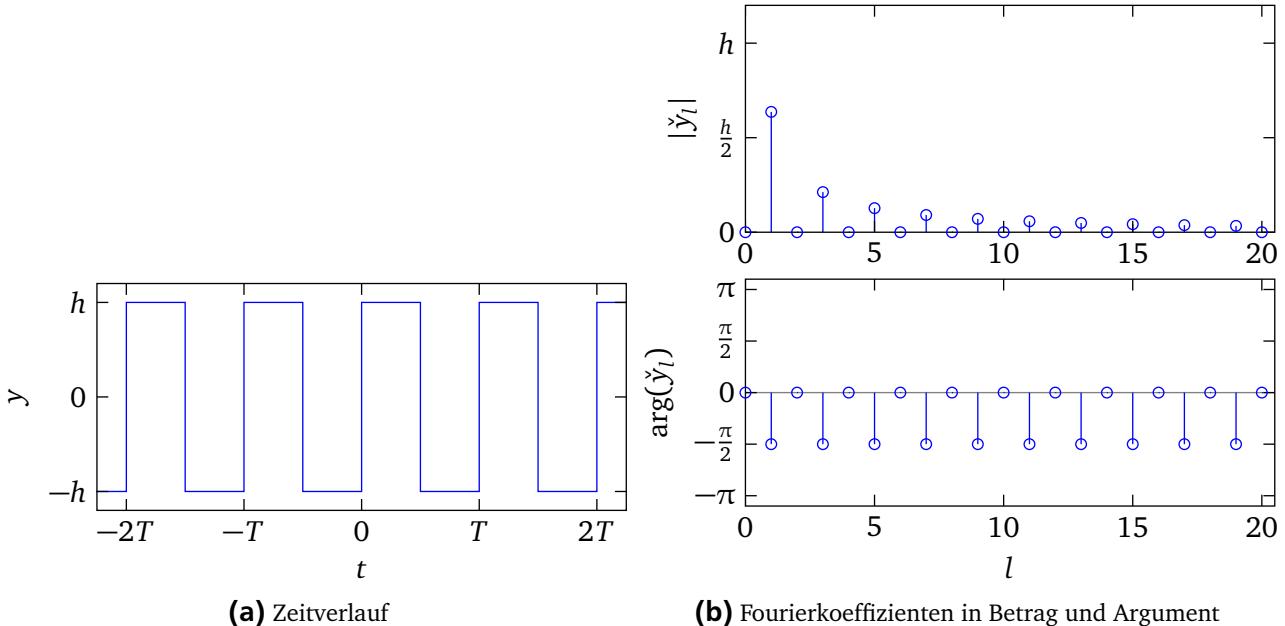


Abbildung 5.1: Rechteckschwingung

5.2 Fouriertransformation (FT)

Die Fouriertransformation für ein Signal $y(t)$ ist als

$$Y(\omega) = \mathcal{F}(y(t))|_{\omega} = \int_{-\infty}^{\infty} y(t) \cdot e^{-j\omega t} dt \quad (5.4)$$

² Mit $\text{mod}(a, b)$ ist der auf reelle Zahlen verallgemeinerte Modulo der Division mit Rest gemeint. D. h. $\text{mod}(a, b) = r$ wobei $r \in [0, b)$ die Gleichung $\frac{a}{b} = n \cdot b + r$ mit $n \in \mathbb{N}$ erfüllt. Es ist beispielsweise $\text{mod}(4,5, 2) = 0,5$ und $\text{mod}(4,5, 2,1) = 0,3$.

definiert³, vorausgesetzt dass uneigentliche Integral existiert. Die inverse Transformation lautet dann

$$y(t) = \mathcal{F}^{-1}(Y(\omega))|_t = \frac{1}{2\pi} \cdot \int_{-\infty}^{\infty} Y(\omega) \cdot e^{j\omega t} d\omega. \quad (5.5)$$

Analog zur den Koeffizienten der Fourierreihe folgt aus der Reellwertigkeit von $y(t)$ hier

$$Y(-\omega) = Y^*(\omega).$$

Berechnung

Für zeitbegrenzte Signale, d. h. Signale für die ein Intervall $[t_1, t_2]$ angegeben werden kann für das $y(t) = 0 \forall t \notin [t_1, t_2]$ gilt, reduziert sich das Integral auf das entsprechende Intervall,

$$Y(\omega) = \int_{t_1}^{t_2} y(t) \cdot e^{-j\omega t} dt.$$

Dieses kann dann analytisch oder numerisch gelöst werden.

Für den in Abbildung 5.2a dargestellten Rechteckimpuls lautet die Fouriertransformierte

$$Y(\omega) = h \cdot T \cdot \frac{\sin\left(\frac{\omega T}{2}\right)}{\frac{\omega T}{2}} \cdot e^{-j\frac{\omega T}{2}} = h \cdot T \cdot \text{sinc}\left(\frac{1}{\pi} \cdot \frac{\omega T}{2}\right) \cdot e^{-j\frac{\omega T}{2}}.$$

Diese ist in Abbildung 5.2b ähnlich einem Bodediagramm in Betrag und Argument dargestellt. In Abbildung 5.2c ist sie in Real- und Imaginärteil dargestellt.

Für den in Abbildung 5.3a gezeigten Dreiecksimpuls ist die Fouriertransformierte in Abbildung 5.3b und (c) dargestellt.

Interpretation

Auch bei der Fouriertransformation wird, wie die inverse Fouriertransformation zeigt, dass Signal $y(t)$ als Summe von harmonischen Schwingungen aufgebaut, da ein Integral als Grenzwert einer Reihe (unendliche Summe) aufgefasst werden kann. Jedoch werden die Koeffizienten dabei noch mit $d\omega$ multipliziert, so dass $Y(\omega)$ eine Amplitudendichte darstellt.

Damit stellt sich die Frage, ob und wie beispielsweise eine Sinusschwingung, für die das Integral $\int_{-\infty}^{\infty} y(t) \cdot e^{-j\omega t} dt$ nicht existiert, im Fourierbereich angegeben werden kann. Anschaulich kann dies darüber begründet werden, dass dieses Signal überhaupt keine Frequenzanteile mit Ausnahme der Frequenz ω_T enthält. Die ganze „Leistung“ des Signales liegt also an einer einzelnen, diskreten Frequenz vor. D. h. die Amplitudendichte ist null für alle Frequenzen ungleich $\pm\omega_T$. Für die Frequenz dagegen ist die Dichte unendlich hoch.

Nicht fourierintegrierbare Signale

Für periodische Signale gibt es zwar die Fourierreihe als geeignetes Mittel, diese im Frequenzbereich darzustellen. Somit ist man nicht zwingend darauf angewiesen, auch für diese Signale eine Fouriertransformierte zu definieren. Jedoch bietet es im weiteren Verlauf, beispielsweise bei der Diskussion der Wirkung von Fensterungen, Vorteile, auch mit periodischen Zeitsignalen im Bildbereich der Fouriertransformation zu arbeiten.

Wie schon erwähnt, muss für eine Sinusschwingung die Amplitudendichten für alle Frequenzen bis auf die Frequenz der Schwingung null sein. Bei der einen Schwingungsfrequenz dagegen ist diese „unendlich“ hoch, wofür es einer geeigneten mathematischen Modellierung bedarf.

³ Häufig ist die Fouriertransformierte auch mit dem Faktor $1/\sqrt{2\pi}$ definiert, $\mathcal{F}(y(t))|_\omega = \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{\infty} y(t) \cdot e^{-j\omega t} dt$. Dann lautet die Rücktransformation entsprechend $\mathcal{F}^{-1}(Y(\omega))|_t = \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{\infty} Y(\omega) \cdot e^{j\omega t} d\omega$.

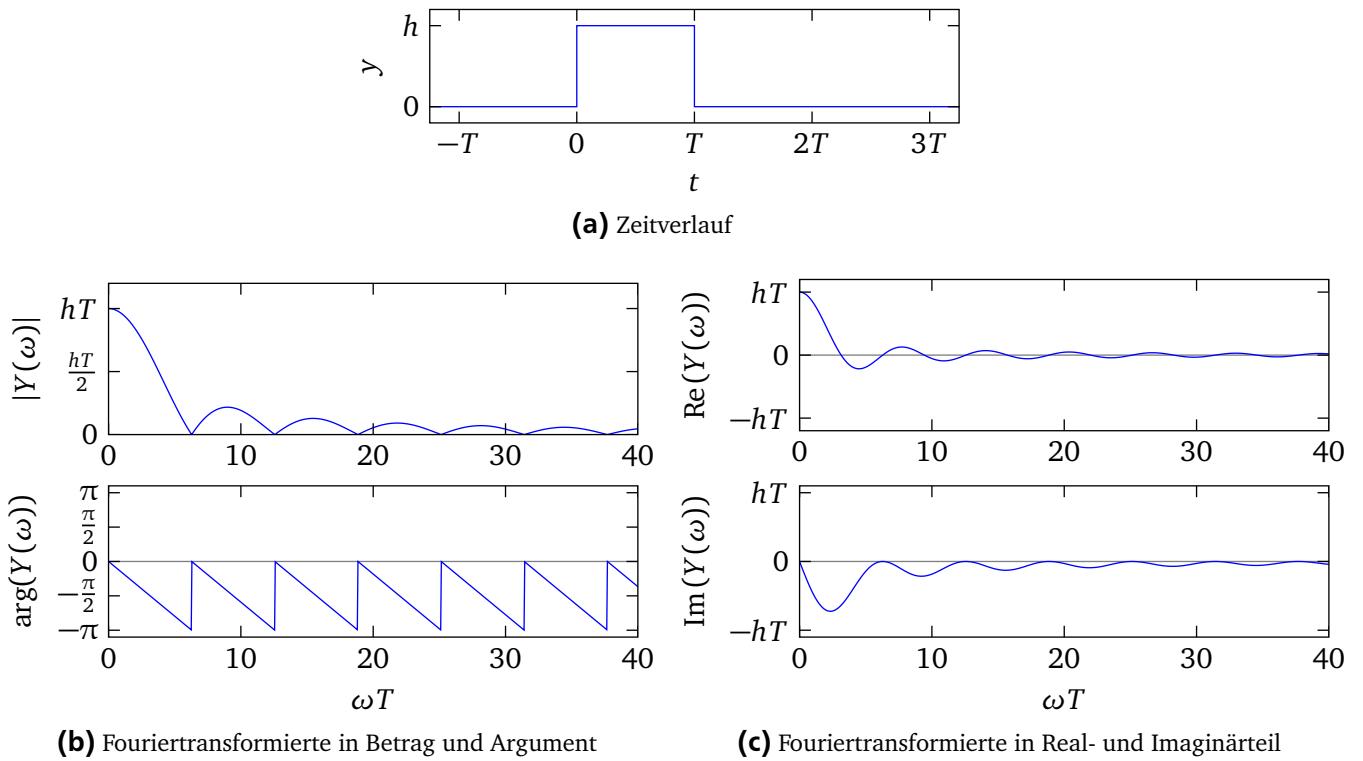


Abbildung 5.2: Rechteckimpuls

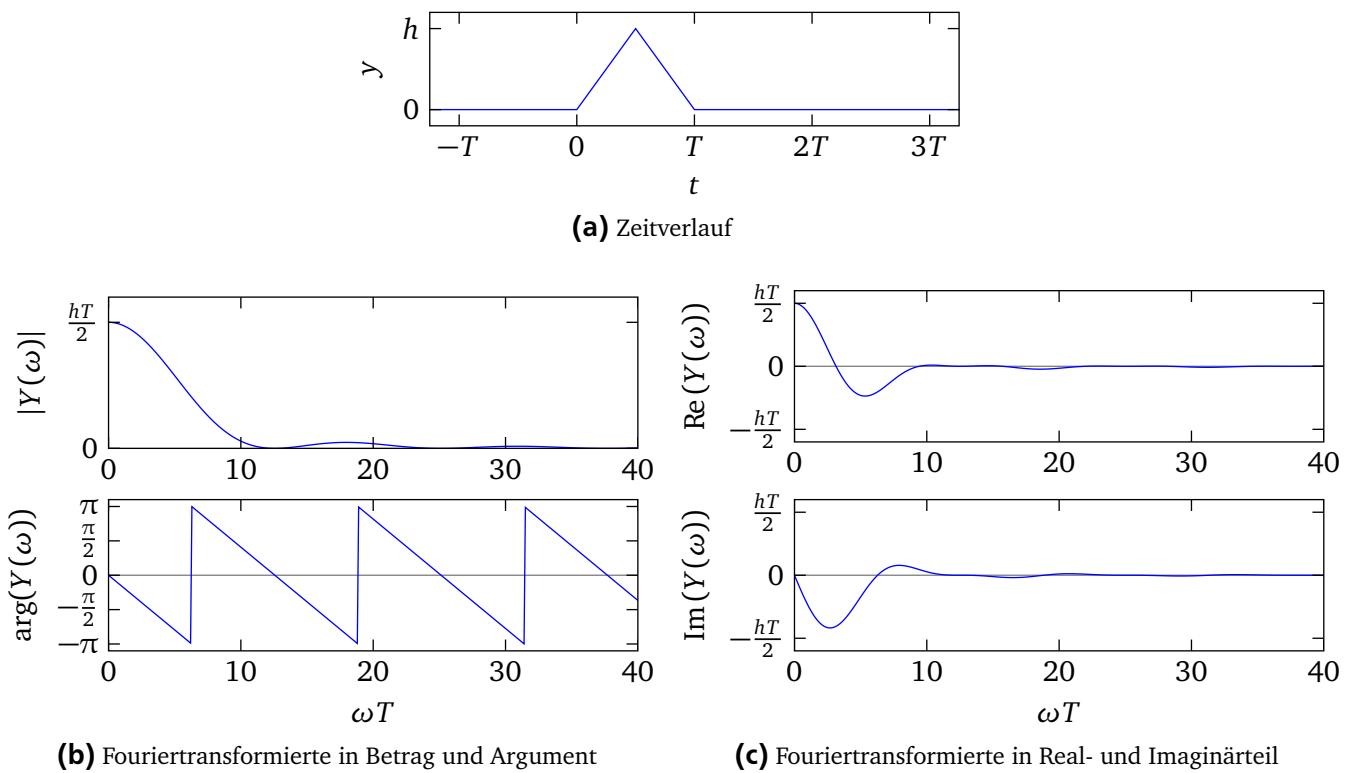


Abbildung 5.3: Dreieckimpuls

Dirac-Impuls

Definiert man Rechteckimpulse nach

$$\text{rect}(t, \tau, \varepsilon) = \begin{cases} \frac{1}{2\varepsilon} & \tau - \varepsilon \leq t \leq \tau + \varepsilon \\ 0 & \text{sonst,} \end{cases}$$

so ist deren Fläche

$$\int_{-\infty}^{\infty} \text{rect}(t, \tau, \varepsilon) dt$$

immer gleich eins und für $\varepsilon \rightarrow 0$ können so beliebig enge Impulse beschrieben werden, wie in Abbildung 5.4a angedeutet.

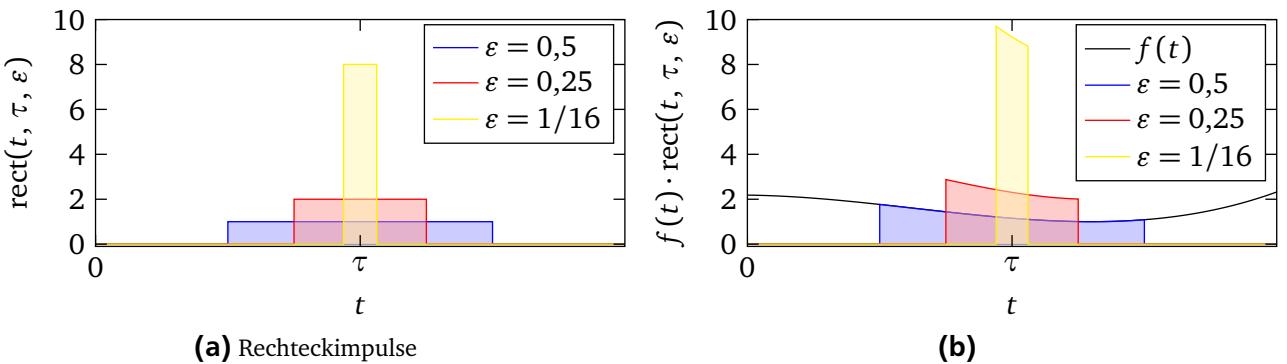


Abbildung 5.4: Zum Dirac-Impuls

Bildet man das Integral

$$I = \int_{-\infty}^{\infty} f(t) \cdot \text{rect}(t, \tau, \varepsilon) dt ,$$

so erhält man für I den Durchschnittswert der Funktion $f(t)$ auf dem Intervall $[\tau - \varepsilon, \tau + \varepsilon]$ auf dem der Rechteckimpuls ungleich null ist, bzw., nach dem Mittelwertsatz der Integralrechnung, $I = f(t')$, $t' \in [\tau - \varepsilon, \tau + \varepsilon]$.

Um so einen idealen Impuls zu erhalten, muss ε nicht nur beliebig nahe an Null herankommen, sondern genau null werden. Damit verlässt man jedoch das Gebiet der bekannten Funktionen und muss für eine mathematische Beschreibung auf Distributionen ausweichen. Dies führt für die Zwecke dieser Vorlesung allerdings zu weit.

Bezeichnet man einen bei τ auftretenden Impuls als $\delta(t - \tau)$, so wird in Anlehnung an das Verhalten des Rechteckimpulses für eine gegen Null gehende Impulsbreite ($\varepsilon \rightarrow 0$) die Notation

$$\int_{\tau-\vartheta}^{\tau+\vartheta} \delta(t - \tau) dt = 1 , \quad \vartheta > 0 ,$$

und

$$\int_{\tau-\vartheta}^{\tau+\vartheta} f(t) \delta(t - \tau) dt = f(\tau) , \quad \vartheta > 0 ,$$

festgelegt. Letzteres nennt man auch Ausblendeigenschaft. Liegt der Punkt τ außerhalb des Integrationsintervalls, so ist der Wert des Integrals als Null anzusehen. Für den Fall, dass eine der Grenzen genau auf den Punkt τ liegt, können ebenso passend Regeln festgelegt werden. Dabei bezeichnen die Integralzeichen nicht mehr die klassische Integration, sondern stellen zusammen mit den $\delta(t - \tau)$ eine eigene Notation da.

Zuletzt sagt die Festlegung

$$\int_a^b (f(t) + \sum_i g_i(t)\delta(t - \tau_i)) \cdot dt = \int_a^b f(t) \cdot dt + \sum_i \int_a^b g_i(t)\delta(t - \tau_i) \cdot dt$$

aus, dass sich das „überladene“ Integral für normale Funktionsanteile wie ein herkömmliches Integral verhält, und für die δ -Impulse kann einzeln von der Ausblendeigenschaft Gebrauch gemacht werden.

Lässt man in der Rücktransformation zur Fouriertransformation dann dieses „Integral“ zu, dann lassen sich zwar anhand dieser einfachen Festlegungen nicht direkt die Fouriertransformierten der interessierenden Signale bestimmen, jedoch können Fouriertransformierte, die δ -Impulse enthalten definiert und dann durch die Rücktransformation geprüft werden, zu welcher Zeitfunktion diese gehören.

Fouriertransformation für periodische Signale

Mit der gegebenen Eigenschaft des δ -Impulses kann für die konstante Funktion $y(t) = 1$ die Fouriertransformierte

$$\mathcal{F}(1)|_{\omega} = 2\pi \cdot \delta(\omega)$$

angegeben werden. Um zu prüfen, ob diese Wahl sinnvoll ist, wird diese in die inverse Fouriertransformation eingesetzt, und es muss sich wieder die Konstante 1 ergeben. Dies ist auch der Fall, wie

$$y(t) = \frac{1}{2\pi} \cdot \int_{-\infty}^{\infty} 2\pi \delta(\omega) e^{j\omega t} d\omega = e^{j \cdot 0 \cdot t} = 1$$

zeigt.

Für ein allgemeines T -periodisches Signal lautet die Fouriertransformierte

$$\mathcal{F}(y(t))|_{\omega} = \sum_{l=-\infty}^{\infty} 2\pi \cdot \check{y}_l \cdot \delta(\omega - l \cdot \omega_T)$$

wobei $\omega_T = 2\pi/T$ ist und die \check{y}_l die Fourierkoeffizienten der Fourierreihe sind. Das dies tatsächlich die Fouriertransformierte ist, kann ebenfalls durch die Anwendung der inversen Fouriertransformation geprüft werden,

$$\begin{aligned} y(t) &= \frac{1}{2\pi} \cdot \int_{-\infty}^{\infty} \sum_{l=-\infty}^{\infty} 2\pi \cdot \check{y}_l \delta(\omega - l \cdot \omega_T) \cdot e^{j\omega t} d\omega \\ &= \frac{1}{2\pi} \cdot 2\pi \cdot \sum_{l=-\infty}^{\infty} \check{y}_l \cdot \int_{-\infty}^{\infty} \delta(\omega - l \cdot \omega_T) \cdot e^{j\omega t} d\omega \\ &= \sum_{l=-\infty}^{\infty} \check{y}_l e^{jl \cdot \omega_T t}. \end{aligned}$$

In der letzten Zeile steht die Fourierreihenentwicklung von $y(t)$, die eine äquivalente Darstellung für den Zeitverlauf $y(t)$ ist.

D.h., dass, wenn der δ -Impuls zugelassen wird, die Fouriertransformierte auch für periodische Signale angegeben werden kann. Diese entspricht dann den mit 2π und den Fourierkoeffizienten gewichteten δ -Impulsen an den zu den jeweiligen Koeffizienten gehörenden Frequenzen.

Fouriertransformation für nicht integrierbare, nicht periodische Signale

Zur Identifikation wird in der Regel entweder das System im eingeschwungenen Zustand bei periodischer Anregung, d. h. mit periodischen Signalen, oder bei nicht periodischen Signalen, die aufgrund der endlichen Messdauer immer zeitbegrenzt sind, betrachtet. So stellt sich die Frage nach der Fouriertransformierten nicht periodischer, nicht fourierintegrierbarer Signale eigentlich nicht.

Der Vollständigkeit halber und zur Diskussion der Anregung über einen Sprung wird hier noch kurz auf die entsprechende Fouriertransformierte eingegangen.

Zunächst wird dazu bestimmt, welches die zur Fouriertransformierten

$$\mathcal{F}(y(t))|_{\omega} = \frac{1}{j\omega}$$

gehörende Zeitfunktion $y(t)$ ist.

Diese ergibt sich aus der inversen Fouriertransformation zunächst als

$$y(t) = \frac{1}{2\pi} \cdot \int_{-\infty}^{\infty} \frac{1}{j\omega} \cdot e^{j\omega t} d\omega = \frac{1}{2\pi} \cdot \int_{-\infty}^0 \frac{1}{j\omega} \cdot e^{j\omega t} d\omega + \frac{1}{2\pi} \cdot \int_0^{\infty} \frac{1}{j\omega} \cdot e^{j\omega t} d\omega .$$

Nach Vertauschen der Grenzen des ersten Integrals und der Substitution $\omega' = -\omega$ können die beiden Integrale zusammengefasst werden,

$$y(t) = \frac{1}{2\pi} \cdot \int_0^{\infty} \left(\frac{-1}{j\omega} \cdot e^{-j\omega t} + \frac{1}{j\omega} \cdot e^{j\omega t} \right) d\omega .$$

Die Anwendung der Eulerformel führt schließlich auf

$$y(t) = \frac{1}{\pi} \cdot \int_0^{\infty} \frac{\sin(\omega t)}{\omega} d\omega .$$

Der hier auftretende Integrand ist nicht elementar integrierbar, das uneigentliche Integral existiert jedoch. Der Wert dieses uneigentlichen Integrals kann in entsprechenden Tabellenwerken nachgeschlagen werden.⁴

Damit ergibt sich letztlich

$$y(t) = \begin{cases} -\frac{1}{2} & t < 0 \\ \frac{1}{2} & t > 0 \end{cases}$$

als die zur Fouriertransformierten $\frac{1}{j\omega}$ gehörenden Zeitfunktion. Dies entspricht gerade der halben Sigmoidfunktion, und damit gilt

$$\mathcal{F}(\text{sign}(t))|_{\omega} = \frac{2}{j\omega} .$$

Schließlich kann damit auch die Fouriertransformierte des Einheitssprungs $\sigma(t)$ angegeben werden. Dazu wird dieser als $\sigma(t) = \frac{1}{2} \cdot \text{sign}(t) + \frac{1}{2}$ geschrieben, und es folgt sofort

$$\mathcal{F}(\sigma(t))|_{\omega} = \frac{1}{j\omega} + \pi \cdot \delta(\omega) .$$

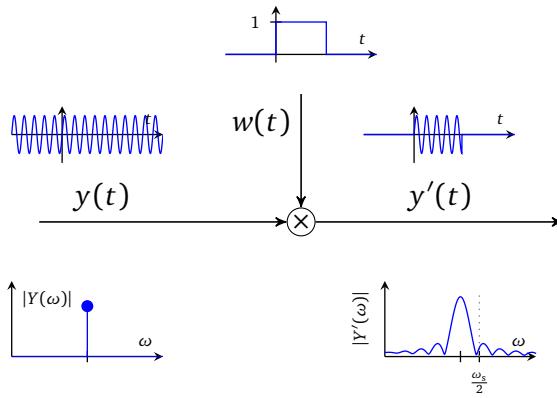


Abbildung 5.5: Fensterung

Fensterung

Die Fensterung eines Signals $y(t)$ stellt die Multiplikation dieses Signals mit einer Fensterfunktion $w(t)$ im Zeitbereich dar, wie in Abbildung 5.5 anhand einer Sinusschwingung und eines Rechteckfensters dargestellt.

Im Frequenzbereich entspricht dies einer Faltung der Fouriertransformierten $\mathcal{F}(y(t))|_{\omega} = Y(\omega)$ und $\mathcal{F}(w(t))|_{\omega} = W(\omega)$,

$$\mathcal{F}(y(t) \cdot w(t)) = \frac{1}{2\pi} \cdot \mathcal{F}(y(t)) * \mathcal{F}(w(t)) = \frac{1}{2\pi} \cdot \int_{-\infty}^{\infty} Y(\nu) \cdot W(\omega - \nu) \cdot d\nu. \quad (5.6)$$

Dies kann man sich am besten veranschaulichen, wenn man die Fensterung eines Sinussignals der Frequenz ω_1 mit der halben komplexen Amplitude \check{y}_1 , dessen Fouriertransformierte

$$Y(\omega) = 2\pi \cdot \check{y}_1^* \cdot \delta(\omega + \omega_1) + 2\pi \cdot \check{y}_1 \cdot \delta(\omega - \omega_1)$$

lautet, betrachtet. Dies in Gl. (5.6) eingesetzt führt auf

$$\begin{aligned} \mathcal{F}(y(t) \cdot w(t)) &= \frac{1}{2\pi} \cdot (2\pi \cdot \check{y}_1^* \cdot W(\omega + \omega_1) + 2\pi \cdot \check{y}_1 \cdot W(\omega - \omega_1)) \\ &= \check{y}_1^* \cdot W(\omega + \omega_1) + \check{y}_1 \cdot W(\omega - \omega_1). \end{aligned}$$

Das Spektrum des gefensterten Signals ergibt sich also aus der Addition des um ω_1 nach links und rechts verschobenen und mit \check{y}_1^* bzw. \check{y}_1 multiplizierten Spektrum des Fensters.

Für ein Rechteckfenster der Breite T mit

$$\mathcal{F}(w(t))|_{\omega} = T \cdot \frac{\sin\left(\frac{\omega \cdot T}{2}\right)}{\frac{\omega \cdot T}{2}} \cdot e^{-j\frac{\omega \cdot T}{2}}$$

gilt damit beispielsweise

$$\mathcal{F}(y(t) \cdot w(t))|_{\omega} = \check{y}_1^* \cdot T \cdot \frac{\sin\left(\frac{(\omega + \omega_1) \cdot T}{2}\right)}{\frac{(\omega + \omega_1) \cdot T}{2}} \cdot e^{-j\frac{(\omega + \omega_1) \cdot T}{2}} + \check{y}_1 \cdot T \cdot \frac{\sin\left(\frac{(\omega - \omega_1) \cdot T}{2}\right)}{\frac{(\omega - \omega_1) \cdot T}{2}} \cdot e^{-j\frac{(\omega - \omega_1) \cdot T}{2}}.$$

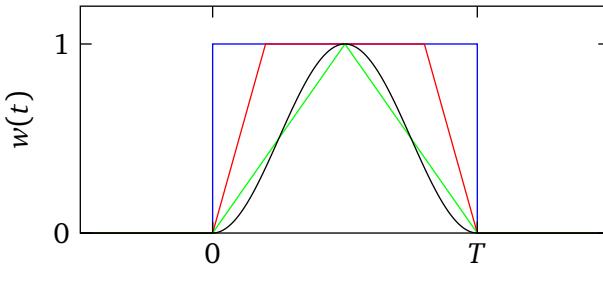
⁴ $\int_0^\infty \frac{\sin(\omega t)}{\omega} d\omega = \begin{cases} -\frac{\pi}{2} & t < 0 \\ \frac{\pi}{2} & t > 0 \end{cases}$. Alternativ kann es durch eine Substitution $x = \pm\omega t$ noch auf den Integralsinus $\int_0^\infty \frac{\sin(x)}{x} dx$ zurückgeführt werden, dessen Wert $\frac{\pi}{2}$ noch häufiger zu finden sein sollte. Bei [Föllinger 1977, 202ff] hingegen wird das ursprüngliche Integral direkt über den Residuensatz ausgewertet.

Der Betrag des sich damit ergebenden Spektrums ist in Abbildung 5.5 unter dem Signalweg skizziert.

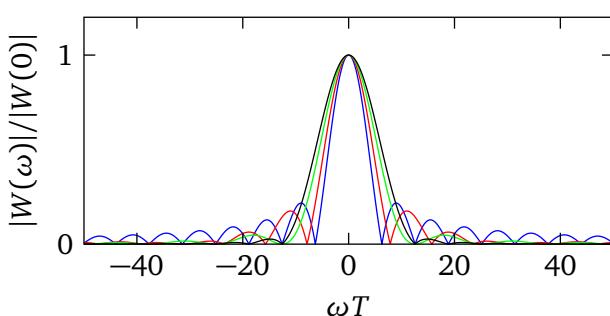
Wesentlich bei der Beurteilung eines Fensters ist dessen Fouriertransformierte. Diese ist für das Rechteckfenster zusammen mit der Fouriertransformierten eines Trapezfensters, eines Dreieckfensters und eines Von-Hann-Fensters

$$w_{\text{Hn}}(t) = \begin{cases} \frac{1}{2} \cdot \left(1 - \cos\left(\frac{2\pi}{T} \cdot t\right)\right) & 0 \leq t < T \\ 0 & \text{sonst,} \end{cases}$$

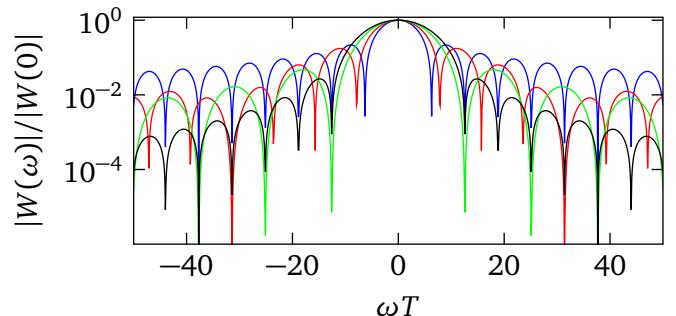
jeweils gleicher Breite, in Abbildung 5.6 gezeigt. Dabei ist jeweils noch auf den Wert des Spektrums bei $\omega = 0$ normiert.



(a) Zeitverläufe



(b) Spektren, Ordinate linear aufgetragen



(c) Spektren, Ordinate logarithmisch aufgetragen

Abbildung 5.6: Fensterfunktionen

Es zeigt sich, dass das Rechteckfenster das engste Hauptmaximum besitzt. Dafür sind die Nebenmaxima am deutlichsten ausgeprägt. Je flacher die Flanken der Fensterfunktion sind, desto geringer werden die Nebenmaxima und desto breiter wird das Hauptmaximum. Das Von-Hann-Fenster weist die geringsten Nebenmaxima auf, dafür besitzt es auch das breiteste Hauptmaximum.

Die Frequenz in Abbildung 5.6 ist auf $1/T$ bezogen angegeben. Betrachtet man das Rechteckfenster, so beträgt die Breite des Hauptmaximum $\Delta\omega T = 4\pi$, wenn hier die Breite an den beiden Punkten bestimmt wird, an denen der Betrag der Fouriertransformierten zum erstenmal Null wird. (Diese liegen bei $\omega T = \pm 2\pi$.) In Frequenzen ausgedrückt beträgt die Breite damit $\Delta\omega = \frac{4\pi}{T}$. Verdoppelt man also die Fensterbreite, dann halbieren sich die Breiten der Hauptmaxima. Die Höhen der Nebenmaxima bleiben durch die Normierung auf $W(0)$ dabei gleich, allerdings rücken diese enger an das Hauptmaximum heran.

In Abbildung 5.7 ist die Wirkung zweier Rechteckfenster unterschiedlicher Länge sowie des Von-Hann-Fensters anhand eines Testsignals, welches aus der Überlagerung zweier Sinusschwingungen unterschiedlicher Frequenz besteht, verglichen.

In Abbildung 5.7a sind die Zeitverläufe der gefensterten Signale gezeigt, in 5.7b die dazugehörigen Fouriertransformierten, welche wieder auf $W(0)$ normiert sind. Hier findet man die oben getroffenen Aussagen bestätigt.

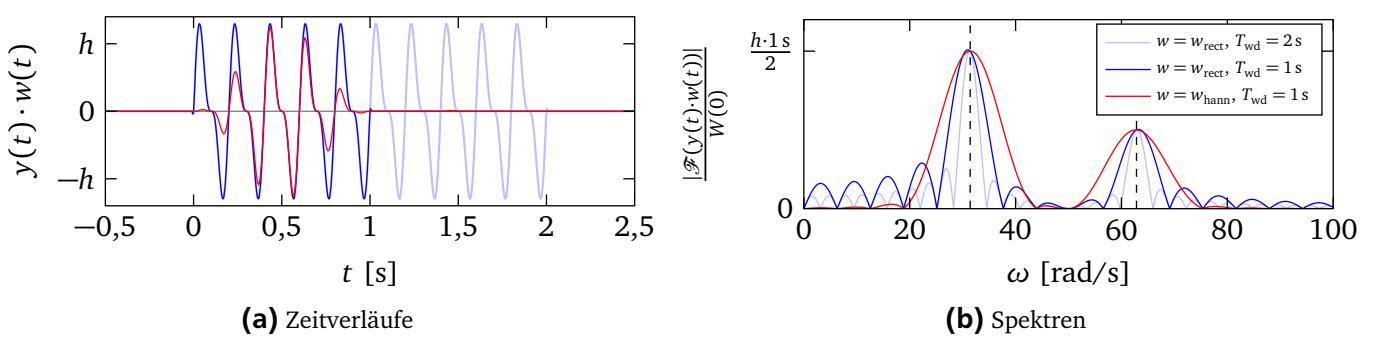


Abbildung 5.7: Wirkung der Fensterung

Prinzipiell sollte bei der Identifikation aber eine Fensterung möglichst vermieden werden, da diese das Signal verfälscht.

Fouriertransformation über ganzzahlige Periode bei periodischen Funktionen

Wendet man die Fouriertransformation (5.4) auf ein T -periodisches Signal $y(t)$, welches mit einem Rechteckfenster, dessen Breite nT ein ganzes Vielfaches n dieser Periodendauer beträgt, gefenstert ist an, so ergibt sich

$$Y'(\omega) = \mathcal{F}(y(t) \cdot w_{nT}(t))|_{\omega} = \int_0^{nT} y(t) \cdot e^{-j\omega t} dt = n \cdot \int_0^T y(t) \cdot e^{-j\omega t} dt.$$

Vergleicht man dies mit den Werten der Fourierkoeffizienten (5.3)

$$\check{y}_l = \mathfrak{F}(y(t))_l = \frac{1}{T} \int_0^T y(t) \cdot e^{-jl\omega_T t} dt$$

die sich für das Signal $y(t)$ ergeben, so erkennt man den Zusammenhang

$$\check{y}_l = \frac{1}{nT} \cdot Y'(l\omega_T) \quad (5.7)$$

zwischen den Fourierkoeffizienten \check{y}_l des ungefensterten, periodischen Signals und den Werten der Fouriertransformation $Y'(\omega)$ des gefensterten Signals zu den speziellen Frequenzpunkten $\omega = l\omega_T$.

5.3 Laplacetransformation (LT)

Die Laplace-Transformation eines Signales $y(t)$ mit $y(t) = 0$ für $t < 0$ ist über

$$Y(s) = \mathcal{L}(y(t))|_s = \int_0^{\infty} y(t) \cdot e^{-st} dt$$

gegeben.

Es ist im Allgemeinen nicht möglich die Fouriertransformation von Signalen auf dem Weg über die Laplace-Transformation bestimmen (die ja für alle technischen Signale angegeben werden kann), indem zunächst die Laplace-Transformation bestimmt wird, und dann $s = j\omega$ gesetzt wird. Dies zeigt schon das Beispiel der Einheitssprungfunktion. Deren Laplacetransformierte lautet

$$Y(s) = \frac{1}{s} .$$

Setzt man einfach $s = j\omega$, so ergibt sich

$$Y'(\omega) = \frac{1}{j\omega}.$$

Dies ist aber die Fouriertransformierte der halben Signumfunktion, wie im vorherigen Abschnitt besprochen wurde. Die Fouriertransformierte der Einheitssprungfunktion lautet

$$Y(\omega) = \frac{1}{j\omega} + \pi \cdot \delta(\omega).$$

Wenn dagegen die Fouriertransformierte eines Signals bestimmt werden soll, für dass das Fourierintegral „klassisch“ existiert, dann kann ohne weiteres die Korrespondenztabelle der Laplacetransformation genutzt und dort $s = j\omega$ gesetzt werden.

5.4 Zeitdiskrete Fouriertransformation (DTFT)

Mit der bezogenen Frequenz

$$\Omega = \omega \cdot T_s \quad (5.8)$$

wird

$$Y_D(\Omega) = \text{DTFT}(y_k)|_{\Omega} = \sum_{k=-\infty}^{\infty} y_k \cdot e^{-j\Omega k} \quad (5.9)$$

als die zeitdiskrete Fouriertransformation definiert. Die inverse Transformation dazu lautet

$$y_k = \text{DTFT}^{-1}(Y_D(\Omega))_k = \frac{1}{2\pi} \cdot \int_{-\pi}^{\pi} Y_D(\Omega) \cdot e^{j\Omega k} d\Omega. \quad (5.10)$$

Das Spektrum $Y_D(\Omega)$ ist 2π -periodisch, was aus Gl. (5.9) ersichtlich ist. Zur Rücktransformation (5.10) wird nur eine Periode berücksichtigt. Für reelle Signale y_k gilt auch hier wieder $Y_D(\Omega) = Y_D^*(-\Omega)$.

Beispiel

Als Beispiel wird zunächst der schon bei der Fouriertransformation verwendete Dreieckimpuls betrachtet, der hier mit einer Abtastzeit von $T_s = \frac{1}{3} \cdot T$, wobei T wieder die Impulsbreite ist, abgetastet wird. Der kontinuierliche sowie der abgetastete Dreieckimpuls ist in Abbildung 5.8a dargestellt.

Aufgrund dessen, dass die Signal für $k < 0$ und $k > 3$ identisch null ist kann die diskrete Fouriertransformation hier einfach berechnet werden. Das Ergebnis ist in Abbildung 5.8b in Betrag-Argument-Darstellung sowie in (c) aufgeteilt in Real- und Imaginärteil dargestellt. Da die diskrete Fouriertransformierte 2π -periodisch ist, ist auch nur ein entsprechender Ausschnitt gezeigt.

Interpretation

Ausgangspunkt für die folgende, [Kammeyer und Kroschel 2006, 20ff] entnommene Betrachtung ist die Gleichung (5.4) der inversen (kontinuierlichen) Fouriertransformation, die für den Zeitpunkt $t = kT_s$ ausgewertet wird,

$$y_k = y(kT_s) = \frac{1}{2\pi} \cdot \int_{-\infty}^{\infty} Y(\omega) \cdot e^{j\omega kT_s} \cdot d\omega.$$

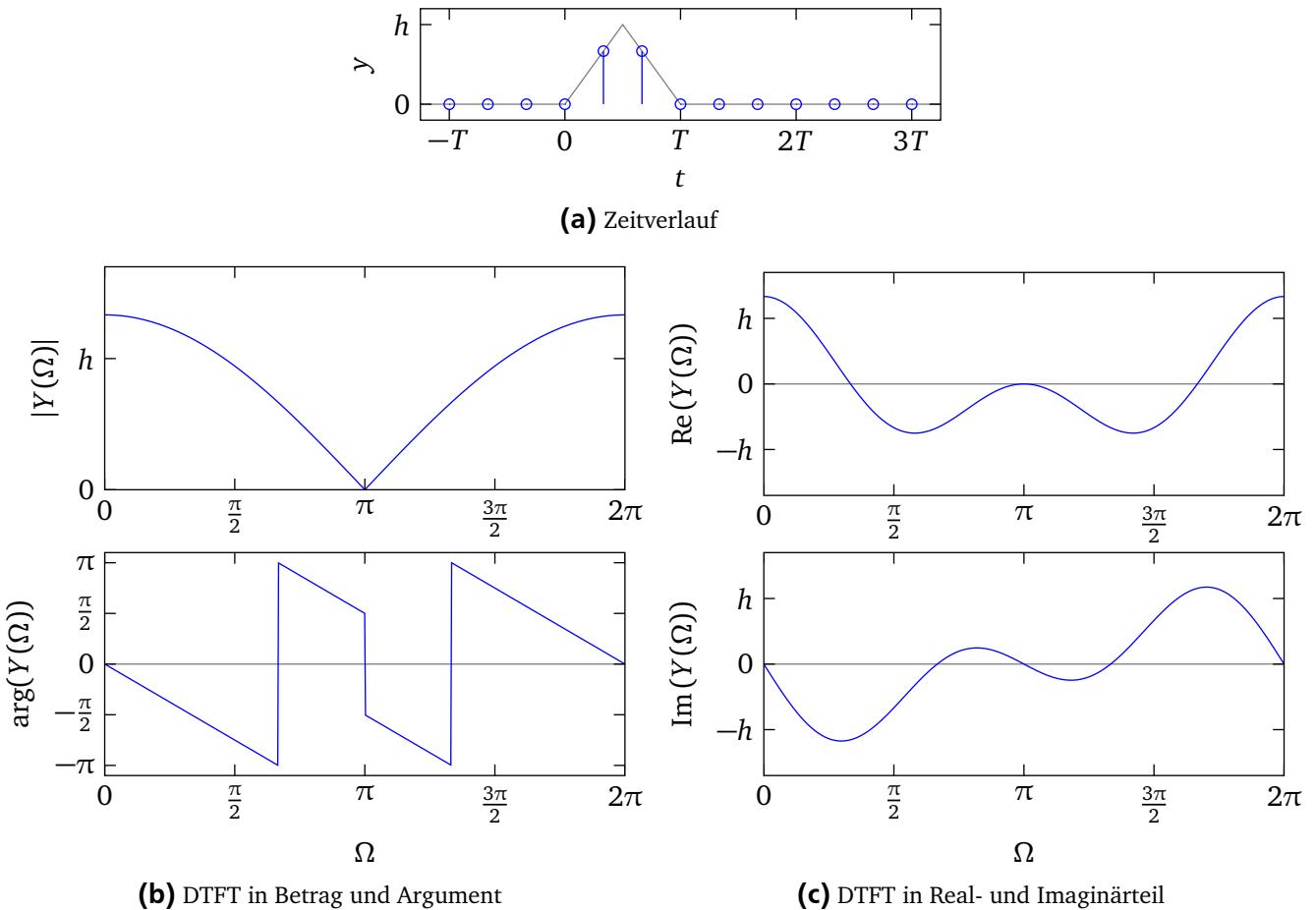


Abbildung 5.8: Abgetasteter Dreieckimpuls, $T_s = \frac{1}{3} \cdot T$

Das Integrationsgebiet wird jetzt in Intervalle der Breite $2\pi/T_s$ aufgeteilt,

$$y_k = \frac{1}{2\pi} \sum_{m=-\infty}^{\infty} \int_{\frac{2\pi m - \pi}{T_s}}^{\frac{2\pi m + \pi}{T_s}} Y(\omega) \cdot e^{j\omega k T_s} \cdot d\omega,$$

und weiter erhält man mit der Substitution $\omega' = \omega - \frac{2\pi m}{T_s}$

$$y_k = \frac{1}{2\pi} \sum_{m=-\infty}^{\infty} \int_{-\frac{\pi}{T_s}}^{\frac{\pi}{T_s}} Y\left(\omega' + \frac{2\pi m}{T_s}\right) \cdot e^{j\omega' k T_s} \cdot e^{j2\pi \frac{m}{T_s} k T_s} \cdot d\omega'.$$

Da immer $e^{j2\pi mk} = 1$ für ganzzahlige m und k gilt und die Summe – da die Integralgrenzen nun nicht mehr von m abhängen – in das Integral gezogen werden kann, erhält man weiter

$$y_k = \frac{1}{2\pi} \int_{-\frac{\pi}{T_s}}^{\frac{\pi}{T_s}} \sum_{m=-\infty}^{\infty} Y\left(\omega' + \frac{2\pi m}{T_s}\right) \cdot e^{j\omega' k T_s} \cdot d\omega'.$$

Mit der weiteren Substitution $\Omega = \omega' T_s$ ergibt sich damit

$$y_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} \underbrace{\frac{1}{T_s} \cdot \sum_{m=-\infty}^{\infty} Y\left(\frac{\Omega + 2\pi m}{T_s}\right)}_{Y_D(\Omega)} \cdot e^{j\Omega k} \cdot d\Omega.$$

Vergleicht man diese Gleichung mit Gl. (5.10), dann kann man

$$Y_D(\Omega) = \frac{1}{T_s} \cdot \sum_{m=-\infty}^{\infty} Y\left(\frac{\Omega + 2\pi m}{T_s}\right)$$

ablesen. Dieser Zusammenhang kann auch mit der unnormierten Frequenz ω geschrieben werden,

$$Y_D(\omega \cdot T_s) = \frac{1}{T_s} \cdot \sum_{m=-\infty}^{\infty} Y\left(\omega + m \cdot \frac{2\pi}{T_s}\right) = \frac{1}{T_s} \cdot \sum_{m=-\infty}^{\infty} Y(\omega + m \cdot \omega_s), \quad (5.11)$$

wobei im letzten Schritt die Abkürzung $\omega_s = \frac{2\pi}{T_s}$ für die Abtastfrequenz in rad/s verwendet wurde.

Die zeitdiskrete Fouriertransformation ergibt sich also aus der Addition der um ω_s verschobenen Fouriertransformation und einer Skalierung.

Dies ist an dem Beispiel des Dreieckimpulses in Abbildung 5.9 dargestellt, wobei in (a) der Real- und – analog – in (b) der Imaginärteil gezeigt ist. Im jeweils obersten Teilplot ist die Fouriertransformierte des Dreieckimpulses, die schon in Abbildung 5.3c gezeigt wurde, gezeichnet. Dabei sind diesmal auch die Werte für negative ω gezeigt. Mit der Abtastzeit von $T_s = \frac{1}{3}T$ ergibt sich $\omega_s = \frac{6\pi}{T}$.

Die zeitdiskrete Fouriertransformation der Abtastfolge berechnet sich also als die Summe des um ganze Vielfache von $\omega_s = \frac{6\pi}{T}$ verschobenen Spektrums des kontinuierlichen Dreieckimpulses. Da die Abszisse der Graphen in ωT angegeben ist, ergibt sich hier eine Verschiebung von $\omega_s T = 6\pi \approx 18,85$. Die einzelnen, noch nicht summierten verschobenen „Originalspektren“ sind im jeweils zweiten Plot in Abbildung 5.9a bzw. (b) gezeigt. (Genau genommen sind die Kurven für $m = -3, \dots, 3$ geplottet, wobei diejenigen für $m = \pm 3$ kaum mehr auszumachen sind.)

Im jeweils untersten Plot sind in hellgrün die Summe der verschobenen Originalspektren, d. h. die mit T_s skalierte DTFT dargestellt. Zum Vergleich ist in schwarz nochmals die Fouriertransformierte des Dreieckimpulses geplottet. Der rot gezeichnete Teil der DTFT ist genau derjenige, der auch schon in Abbildung 5.8c gezeigt ist.

Aliasing

In Abbildung 5.9 ist für das Originalspektrum der Bereich von $\pm \frac{\omega_s}{2}$ grau hinterlegt. Diese Markierung ist – bei entsprechender Farbgebung – auch für die verschobenen Spektren übernommen worden. Es lässt sich daran gut erkennen, dass alle Anteile des ursprünglichen Signals, die außerhalb dieser Markierung liegen, bei der Addition zu einer „Verfälschung“ des Spektrums in diesem Bereich führen. Oder anders gesagt, wenn das ursprüngliche Signal außerhalb des Bereiches exakt null wäre, dann würde das Spektrum der DTFT im Intervall $\Omega \in [-\pi, \pi]$ bzw. $\omega \in [-\omega_s/2, \omega_s/2]$ bis auf die Skalierung exakt dem Spektrum des ursprünglichen, kontinuierlichen Signales entsprechen.

D. h., wenn $Y(\omega) = 0$ für $\omega \notin [-\omega_s/2, \omega_s/2]$ gilt, dann gilt

$$\mathcal{F}(y(t))|_{\omega} = T_s \cdot \text{DTFT}(y_k)|_{\Omega=\omega T_s} \quad \forall \omega \in [-\omega_s/2, \omega_s/2].$$

Verringert man die Abtastzeit T_s , so erhöhen sich mit ω_s auch die Intervalle, um die die ursprünglichen Spektren verschoben werden. Bei einem Signal wie dem Dreieckimpuls, dessen Amplitudendichte mit wachsendem ω relativ schnell abfällt, reduziert sich damit auch die Verfälschung des Spektrums beim Abtasten. Als zweiter Effekt vergrößert sich dabei auch der Frequenzbereich, der nicht einfach durch die periodische Fortsetzung entsteht.

Dies ist in Abbildung 5.10 gezeigt, wo neben dem Spektrum des zeitkontinuierlichen Signals noch die jeweils mit T_s multiplizierten Spektren der mit $T_s = \frac{T}{3}$ und $T_s = \frac{T}{6}$ abgetasteten Signale eingezeichnet sind. Das Spektrum für $T_s = \frac{T}{3}$ ist dabei schon aus den vorherigen Diskussionen bekannt.

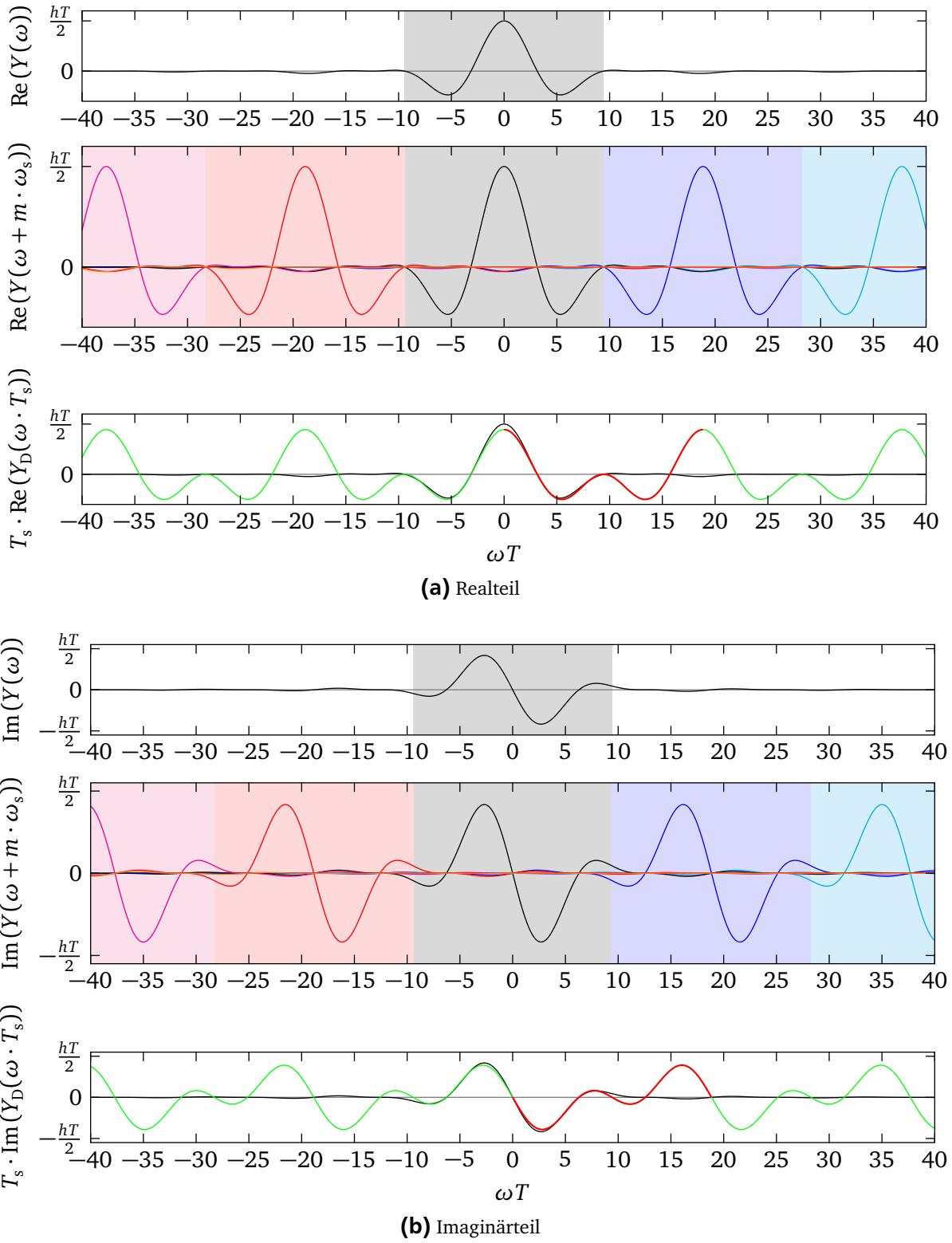


Abbildung 5.9: Zur Interpretation der DTFT

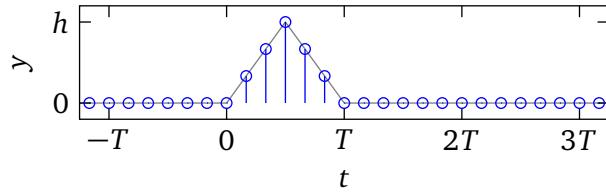
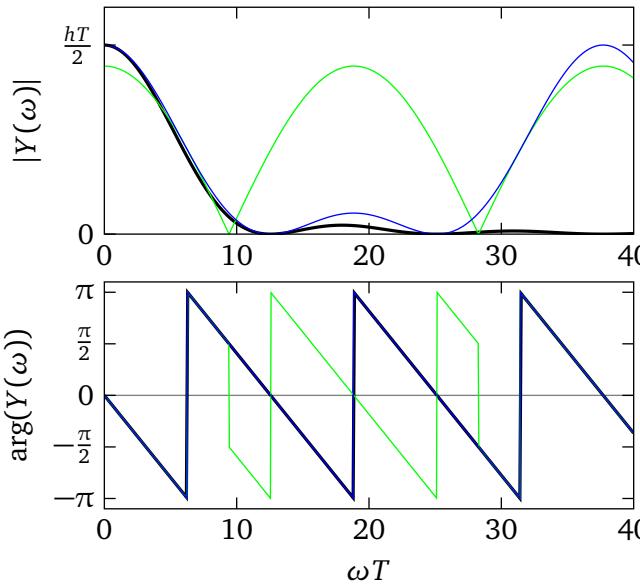
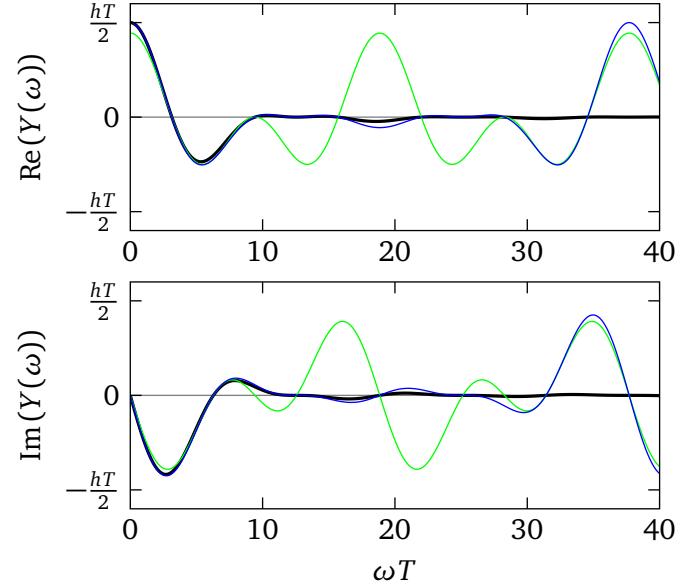


Abbildung 5.10: Abgetasteter Dreieckimpuls, $T_s = \frac{T}{6}$



(a) FT bzw. DTFT in Betrag und Argument



(b) FT bzw. DTFT in Real- und Imaginärteil

Abbildung 5.11: Abgetasteter Dreieckimpuls

schwarz: $Y(\omega) = \mathcal{F}(y(t))|_{\omega}$
grün: $Y(\omega) = T_s \cdot \text{DTFT}(y(kT_s))|_{\Omega=\omega T_s}, T_s = \frac{T}{3}$
blau: $Y(\omega) = T_s \cdot \text{DTFT}(y(kT_s))|_{\Omega=\omega T_s}, T_s = \frac{T}{6}$

Beispiel periodisches Signal

Ausgangspunkt ist das zeitkontinuierliche Signal

$$y(t) = h \cdot \cos\left(\frac{2\pi}{T} \cdot t\right)$$

welches im obersten Plot in Abbildung 5.12a in blau dargestellt ist. In der δ -Notation lautet die Fouriertransformierte

$$Y(\omega) = 2\pi h \cdot \delta\left(\omega + \frac{2\pi}{T}\right) + 2\pi h \cdot \delta\left(\omega - \frac{2\pi}{T}\right)$$

und ist im mittleren Plot dargestellt.

Dieses Signal $y(t)$ wird nun mit der Abtastzeit

$$T_s = 0,85 \cdot T$$

diskretisiert. Damit ergibt sich die mit den schwarzen Punkten dargestellte Abtastfolge $y_k = y(kT_s)$.

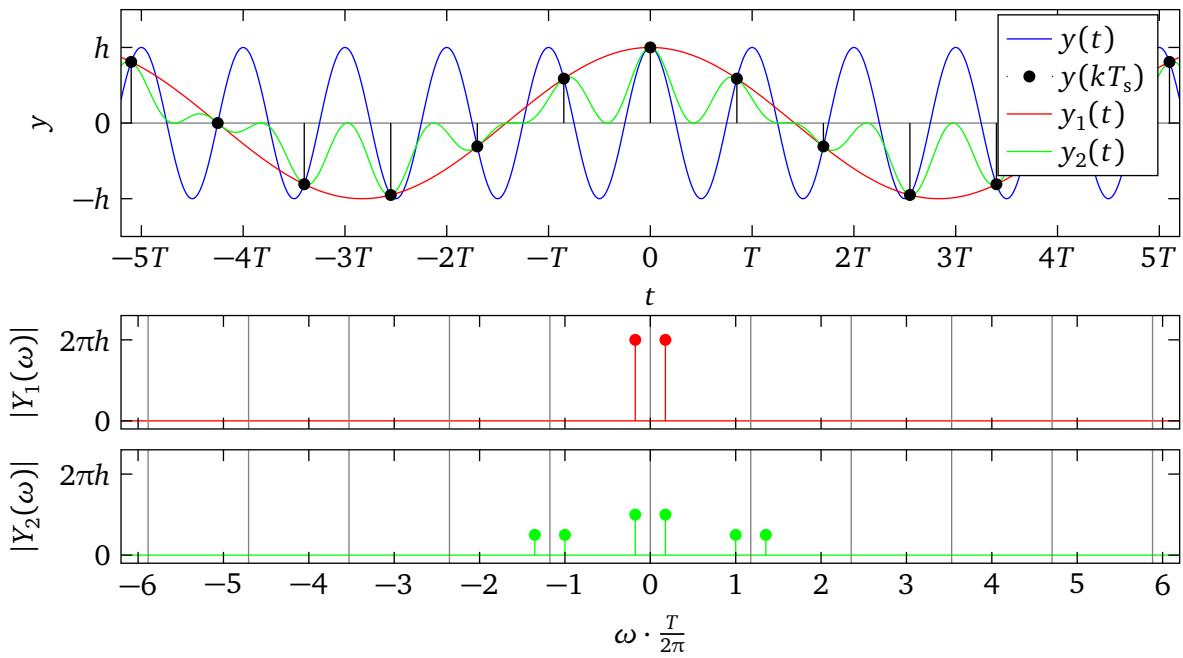
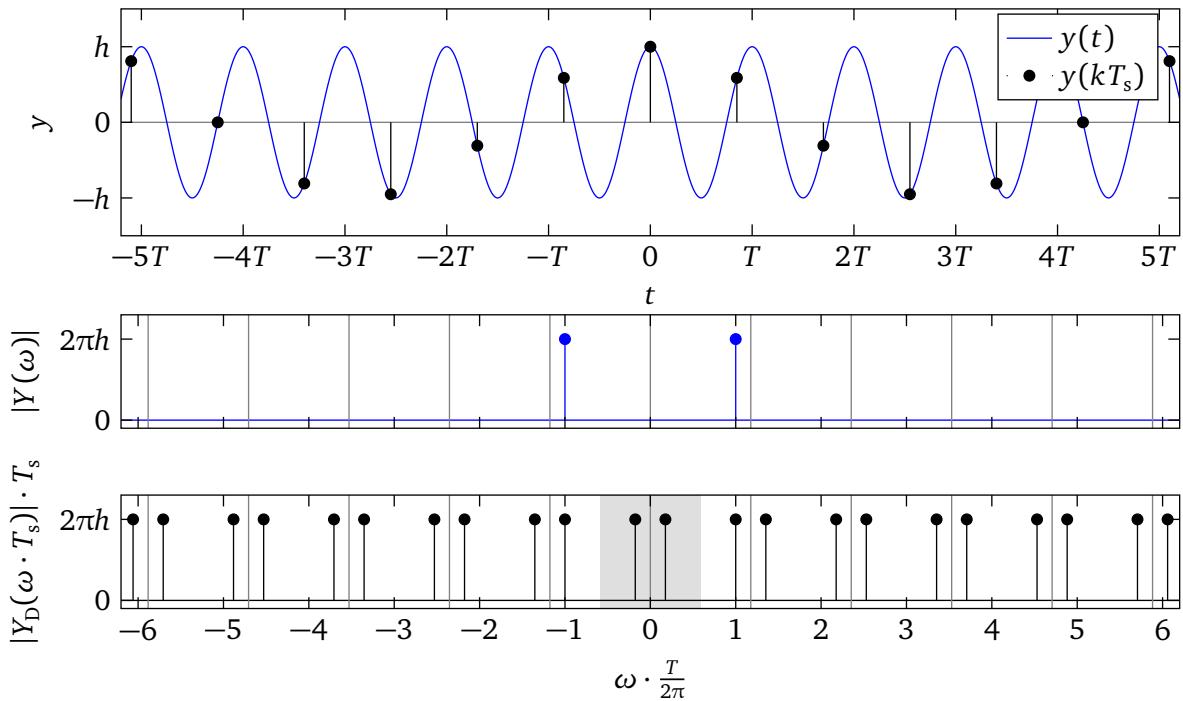


Abbildung 5.12: Aliasing (Abtastung mit $T_s = 0,85 \cdot T$)

Um das Spektrum der DTFT dieser Abtastfolge zu erhalten, wird Gl. (5.11) verwendet, nach der sich das DTFT-Spektrum durch das Verschieben des ursprünglichen Spektrums um ganze Vielfache der Abtastfrequenz und aufaddieren ergibt. Die Abtastfrequenz lautet hier

$$\omega_s = \frac{2\pi}{0,85 \cdot T} = 1,177 \cdot \frac{2\pi}{T}$$

und deren Vielfache sind mit senkrechten grauen Strichen in den relevanten Plots markiert.

Mathematisch lautet das DTFT-Spektrum

$$Y_D(\omega \cdot T_s) = \frac{1}{T_s} \cdot \sum_{m=-\infty}^{\infty} \left(2\pi h \cdot \delta \left(\omega + m\omega_s + \frac{2\pi}{T} \right) + 2\pi h \cdot \delta \left(\omega + m\omega_s - \frac{2\pi}{T} \right) \right)$$

welches auch im untersten Plot in Abbildung 5.12a dargestellt ist. Dabei ist diese nicht über der normierten Frequenz $\Omega = \omega T_s$ sondern über ω dargestellt, damit die Plots der Spektren die gleiche Achsen Skalierung besitzen. Die Impulse jeweils knapp rechts der markierten Abtastfrequenzen stammen von dem ursprünglichen linken Impuls, die auf der linken Seite von dem rechten Impuls.

Es ist gut zu erkennen, dass die ursprüngliche Frequenz von $\frac{2\pi}{T}$ nun (unter anderem) schon bei $0,177 \cdot \frac{2\pi}{T}$ im Spektrum erscheint. Dies wird als Aliasing bezeichnet.

Damit wird die Rekonstruktion des kontinuierlichen Zeitsignals problematisch. In der Regel wird das Zeitsignal rekonstruiert, indem die inverse Fouriertransformation (FT, nicht DTFT) auf das auf das Intervall $\Omega \in [-\pi, \pi]$ bzw. $\omega \in [-\omega_s/2, \omega_s/2]$ beschränkte, mit T_s skalierte DTFT angewendet wird. Dieses durch die Beschränkung entstehende Spektrum ist im mittleren Plot in Abbildung 5.12b dargestellt. Mathematisch kann dieses durch

$$Y_1(\omega) = 2\pi h \cdot \delta \left(\omega + 0,177 \cdot \frac{2\pi}{T} \right) + 2\pi h \cdot \delta \left(\omega - 0,177 \cdot \frac{2\pi}{T} \right)$$

ausgedrückt werden. Wichtiger soll es hier jedoch sein, sich in der grafischen Darstellung davon zu überzeugen, dass die Abtastung mit T_s hier zu dem selben periodischen DTFT-Spektrum führt.

Bildet man die inverse Fouriertransformation von $Y_1(\omega)$ erhält man⁵

$$y_1(t) = h \cdot \cos \left(0,177 \cdot \frac{2\pi}{T} \cdot t \right).$$

Diese Schwingung ist auch im oberen Plot von Abbildung 5.12b, welche auch den Verlauf des ursprünglichen $y(t)$ sowie die Abtastwerte $y(kT_s)$ enthält, in rot eingezeichnet. Es ist klar zu erkennen, dass die beiden Funktionen $y(t)$ und $y_1(t)$ an den Abtastpunkten übereinstimmen.

D. h. anstelle einer Schwingung mit der Frequenz $\frac{2\pi}{T}$ ergibt sich eine Schwingung mit deutlich niedrigerer Frequenz. Kennt man nur die DTFT ohne weitere Informationen, dann lässt sich das Signal zwischen den Abtastpunkten nicht mehr rekonstruieren. Das Spektrum eines weiteren Signals $y_2(t)$, welches ebenfalls

⁵ Hier wurde eine Kosinusschwingung gewählt, damit der Imaginärteil der Fouriertransformierten immer null und der Realteil positiv ist. Damit entspricht der Absolutwert direkt der Fouriertransformierten, und man hat es hier etwas einfacher. Würde man eine Sinusschwingung oder eine phasenverschobene Kosinusschwingung nehmen, dann müsste man die Phasenlagen entsprechend berücksichtigen.

die gleiche Abtastfolge und damit die gleiche DTFT erzeugt, ist im untersten Plot von Abbildung 5.12b dargestellt.⁶ Die dazugehörige Funktion

$$y_2(t) = \frac{h}{2} \cdot \cos\left(0,177 \cdot \frac{2\pi}{T} \cdot t\right) + \frac{h}{4} \cdot \cos\left(\frac{2\pi}{T} \cdot t\right) + \frac{h}{4} \cdot \cos\left(1,353 \cdot \frac{2\pi}{T} \cdot t\right)$$

setzt sich aus drei harmonischen Schwingungen zusammen, deren Amplituden in Summe gerade h ergeben. Es ist in Abbildung 5.12b auch zu erkennen, dass die Werte zu den Abtastpunkten exakt mit denen der anderen Möglichkeiten übereinstimmen.

Die Tatsache, dass das Spektrum periodisch ist, bedeutet auch nicht, dass das abgetastete Signal Anteile bei wie hier z. B. $0,177 \cdot \frac{2\pi}{T}$ und $\frac{2\pi}{T}$ und ... enthält. Dann wäre die Signalenergie unendlich hoch. Die Frequenzen $0,177 \cdot \frac{2\pi}{T}$ und $\frac{2\pi}{T}$ bezeichnen bei einer Abtastzeit von $T_s = 0,85 T$ schlicht dasselbe.

Fragt man sich, welches kontinuierliches Signal nach Abtastung zu dem Spektrum der DTFT geführt hat, dann ist die Periodizität als „oder“ anzusehen, was das Beispiel hier illustriert hat.

Shannon-Theorem

Um ein Signal rekonstruierbar abzutasten⁷, muss dieses bandbegrenzt sein, d. h. es muss für $Y(\omega) = 0 \forall \omega > \omega_{\max}$ gelten. Die Abtastzeit ist dann so zu wählen, dass $\omega_s > 2\omega_{\max}$ gilt. Beziehungsweise ist dafür zu sorgen, dass $Y(\omega) = 0 \forall \omega \geq \frac{1}{2} \cdot \omega_s$ ist. (Die erste Bedingung gibt die notwendige Abtastfrequenz bei gegebenen Signal, die zweite stellt Bedingungen an das Signal bei gegebener Abtastfrequenz.) Die halbe Abtastfrequenz wird auch als Nyquistfrequenz

$$\omega_{ny} = \frac{1}{2} \cdot \omega_s$$

bezeichnet.

Um das Signal entsprechend (näherungsweise, exakt lässt es sich bei zeitbegrenzten Signalen ohnehin nie erreichen) zu konditionieren können *analoge* Tiefpassfilter verwendet werden. Steht kein analoger Tiefpass zur Verfügung, dann kann auch mit entsprechend höherer Frequenz abgetastet werden, dann ein digitaler Tiefpassfilter verwendet und dann unterabgetastet werden. Diese beiden Möglichkeiten eines Anti-Aliasing-Filter sind in Abbildung 5.13, zusammen mit den Spektren der Ein- und Ausgangs- sowie Zwischensignale dargestellt.

Als letztes sei noch angemerkt, dass man auch Signale mit höherer Frequenz als ω_{ny} sinnvoll abtasten kann, so lange sichergestellt ist, dass es rekonstruierbar ist. So könnte in dem Beispiel hier das ursprüngliche Signal auch rekonstruiert werden, wenn man sicherstellen würde (z. B. durch ein entsprechendes Bandpassfilter), dass das Signal nur für $|\Omega| \in [\pi, 2\pi]$ bzw. $|\omega| \in [\omega_s/2, \omega_s]$ Anteile enthält.

Aliasing bei zeitbegrenzten Signalen

Abschließend ist in Abbildung 5.14 noch ein Beispiel für Aliasing bei einem Dreieckimpuls als Beispiel für ein zeitbegrenztes Signal gezeigt. Da das Signal zeitbegrenzt ist, erhält man eine Fouriertransformierte

⁶

$$\begin{aligned} Y_2(\omega) = & 2\pi \cdot \frac{h}{2} \cdot \delta\left(\omega + 0,177 \cdot \frac{2\pi}{T}\right) + 2\pi \cdot \frac{h}{2} \cdot \delta\left(\omega - 0,177 \cdot \frac{2\pi}{T}\right) \\ & + 2\pi \cdot \frac{h}{4} \cdot \delta\left(\omega + \frac{2\pi}{T}\right) + 2\pi \cdot \frac{h}{4} \cdot \delta\left(\omega - \frac{2\pi}{T}\right) \\ & + 2\pi \cdot \frac{h}{4} \cdot \delta\left(\omega + 1,353 \cdot \frac{2\pi}{T}\right) + 2\pi \cdot \frac{h}{4} \cdot \delta\left(\omega - 1,353 \cdot \frac{2\pi}{T}\right) \end{aligned}$$

⁷ Was auch eine notwendige Voraussetzung ist, sinnvoll im Frequenzbereich zu arbeiten. Man könnte auch sagen „so abzutasten, dass keine Überlappung der Spektren auftritt“, jedoch ist „rekonstruierbar“ etwas griffiger.

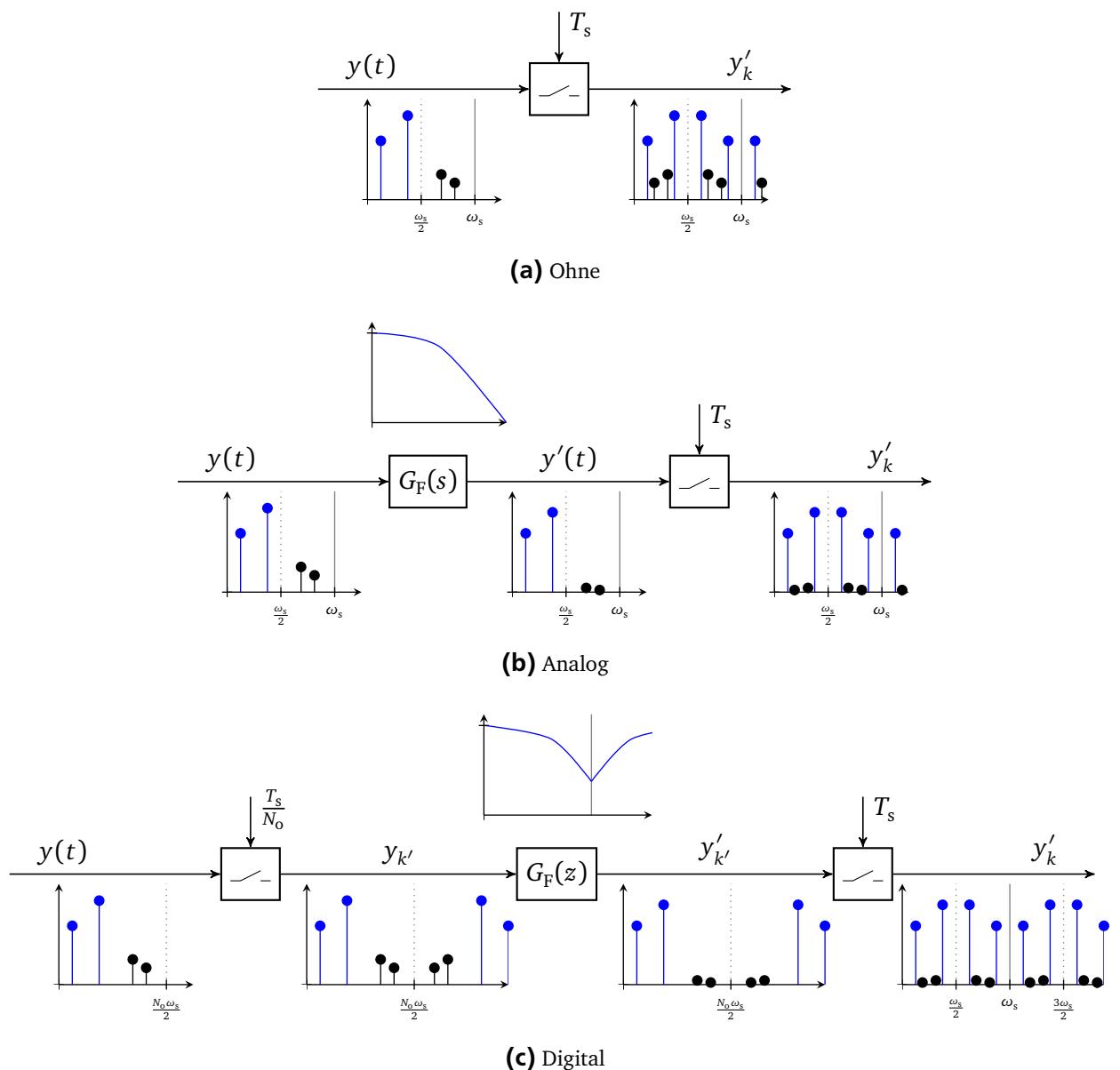


Abbildung 5.13: Abtastung mit Anti-Aliasingfilter

ohne Impulse, die als erste Fouriertransformierte unter dem Zeitplot dargestellt ist. In grau ist die DTFT, die anhand der gezeigten Abtastpunkte bestimmt ist, dargestellt.

Nimmt man zur Rekonstruktion wieder das Intervall $[-\Omega, \Omega]$, was der blauen Kurve $Y_1(\omega)$ entspricht, so erhält man den ebenfalls in blau dargestellten Zeitverlauf, der dem ursprünglichen Impuls relativ gut folgt, die Ecken jedoch glättet.

In rot und in grün sind noch zwei Beispiele für Signale angegeben, die dieselben Abtastwerte liefern, und damit auch dieselbe DTFT besitzen. Deren FT-Spektrum ist in den untersten beiden Plots dargestellt.

Zusammenhang zwischen Original- und Aliasingfrequenz

Zeigen sich im Spektrum unerwartete Störungen, so gilt es herauszufinden, welche Frequenz diese Störung ursprünglich hatte. D.h. gegeben ist eine Abtastfrequenz ω_s sowie eine Frequenz ω_n mit $0 \leq \omega_n < \frac{\omega_s}{2}$. Wenn Aliasing nicht ausgeschlossen werden kann, welche Frequenzen können ω_n im zeitdiskreten Signal hervorufen?

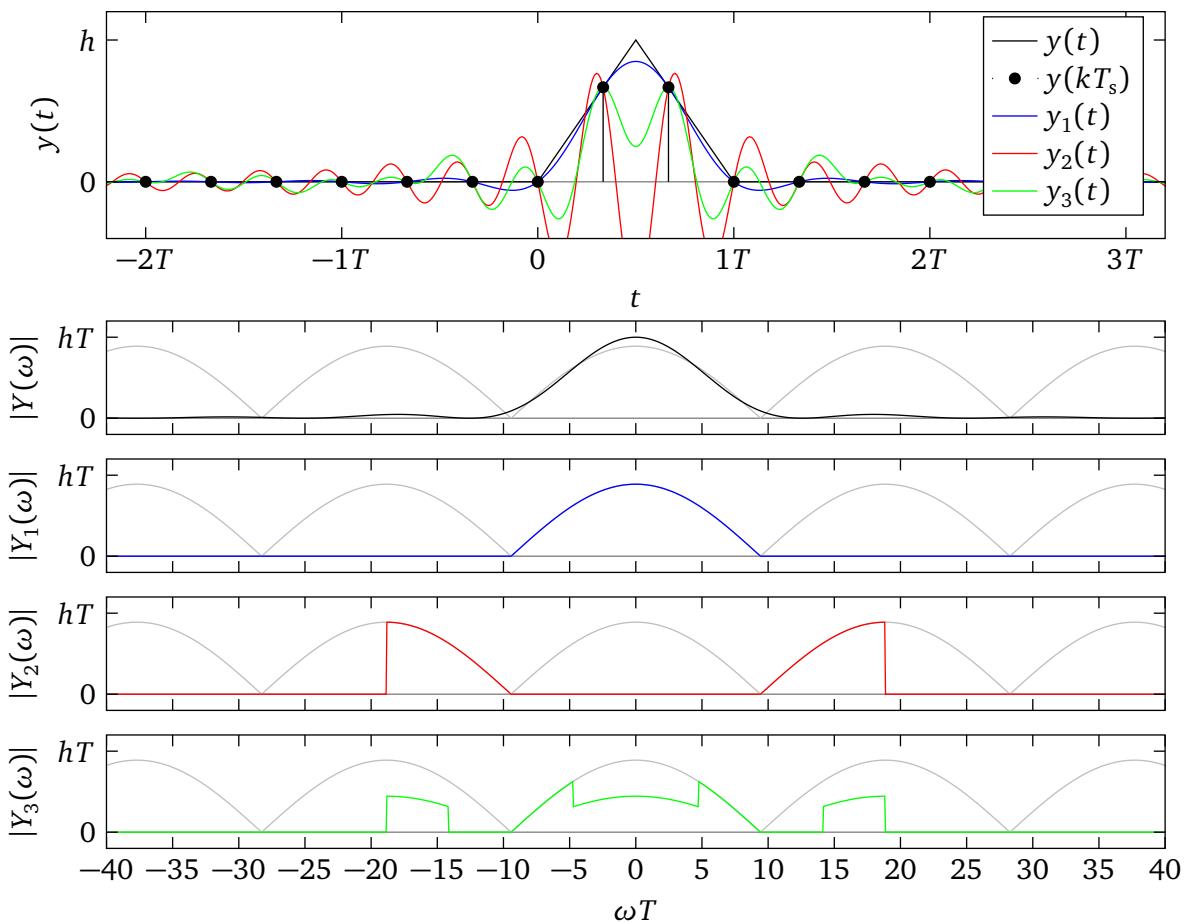


Abbildung 5.14: Aliasing bei zeitbegrenzten Signalen

Ein Frequenzanteil bei ω_{orig} erscheint bei einer Abtastfrequenz von ω_s bei

$$\omega_{\text{alias}} = \begin{cases} \text{mod}(\omega_{\text{orig}}, \omega_s) & \text{wenn } \text{mod}(\omega_{\text{orig}}, \omega_s) \leq \frac{\omega_s}{2} \\ \omega_s - \text{mod}(\omega_{\text{orig}}, \omega_s) & \text{wenn } \text{mod}(\omega_{\text{orig}}, \omega_s) > \frac{\omega_s}{2} \end{cases}$$

im Intervall von $[0, \omega_s/2]$. Grafisch ist dieser Zusammenhang in Abbildung 5.15 dargestellt.

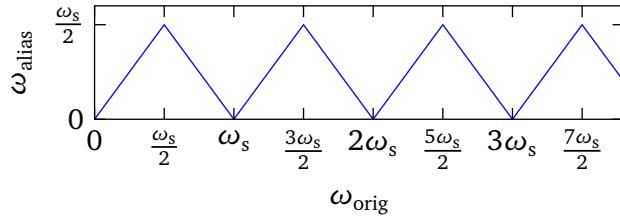


Abbildung 5.15: Aliasfrequenz über Originalfrequenz

Umgekehrt kann bei gegebener Frequenz $\omega_{\text{alias}} \in [0, \omega_s/2]$ im zeitdiskreten Signal für die Originalfrequenz(en) die Menge

$$\omega_{\text{orig}} \in \{m \cdot \omega_s + \omega_{\text{alias}} \mid m = 0, 1, \dots\} \cup \{m \cdot \omega_s - \omega_{\text{alias}} \mid m = 1, 2, \dots\}$$

angegeben werden.

Angenommen eine Abtastung erfolgt mit $f_s = 80 \text{ Hz}$. Bei welcher Frequenz f_{alias} würde man ggf. Störungen durch die Netzfrequenz von 50 Hz sehen? Da $\text{mod}(50 \text{ Hz}, 80 \text{ Hz}) = 50 \text{ Hz}$ und $50 \text{ Hz} > \frac{1}{2} \cdot 80 \text{ Hz}$ ergibt sich $f_{\text{alias}} = 80 \text{ Hz} - 50 \text{ Hz} = 30 \text{ Hz}$.

5.5 Diskrete Fouriertransformation (DFT)

Die diskrete Fouriertransformation ist als

$$Y_l = \text{DFT}(y_k)_l = \sum_{k=0}^{N-1} y_k e^{-j \frac{2\pi}{N} \cdot l \cdot k} \quad (5.12)$$

definiert. Diese transformiert eine Folge y_k , $k = 0, \dots, N-1$, in eine N -periodische Folge Y_l . Die inverse diskrete Fouriertransformation lautet

$$y_k = \text{DFT}^{-1}(Y_l)_k = \frac{1}{N} \cdot \sum_{l=0}^{N-1} Y_l \cdot e^{j \frac{2\pi}{N} \cdot l \cdot k}, \quad 0 \leq k < N, \quad (5.13)$$

wobei die Einschränkung auf den Bereich $0 \leq k < N$ eigentlich nicht gemacht wird, aber für den Gebrauch in dieser Vorlesung sinnvoll erscheint.

Prinzipiell kann die inverse Transformation auch dazu genutzt werden, Werte für y_k für $k < 0$ und $k \geq N$ zu bestimmen. Die so entstehende Folge y_k ist dann ebenfalls (wie das Spektrum) N -periodisch. D. h. in diesem Fall müsste die Bildung der DFT so interpretiert werden, dass man ausgehend von einem abgetasteten Signal y_k dieses zunächst auf N Abtastwerte beschränkt, diesen Teil des Signals dann periodisch fortsetzt und darauf die Gleichung der DFT anwendet, also letztlich das Spektrum bestimmt. Dieses im Zeitbereich periodisch fortgesetzte Signal entspricht aber nur in den wenigsten Fällen dem Signal, von welchem das Spektrum bestimmt werden sollte. (Nämlich im Fall von „korrekt“ abgetasteten periodischen Funktionen.) Und auch Maßnahmen wie das des Zeropadding führen dazu, dass man ein Spektrum von einem Signal bestimmt, welches „optisch“ nicht mehr viel mit dem eigentlichen Signal zu tun hat.

Daher wird hier die DFT aus der DTFT heraus motiviert.

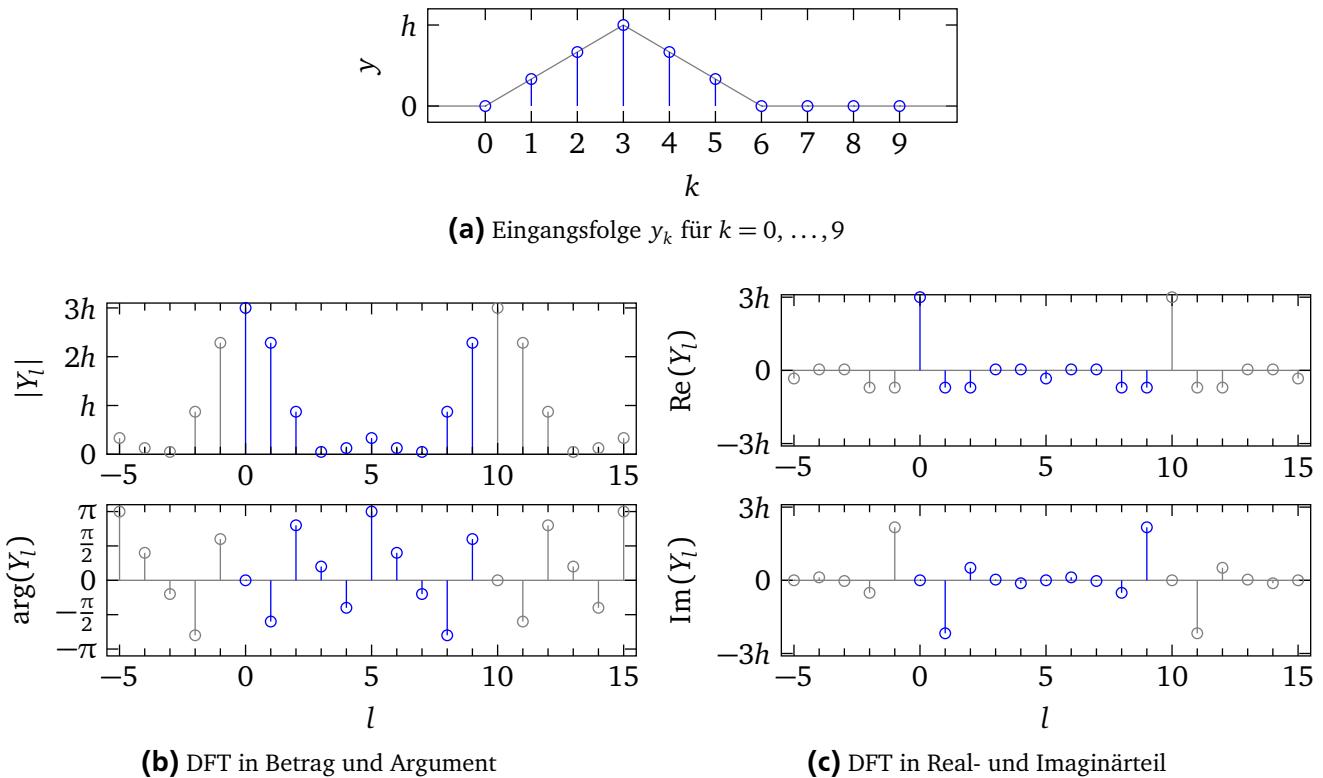


Abbildung 5.16: Dreieckimpuls, $N = 10$

Beispiel

In Abbildung 5.16 ist die DFT einer Eingangsfolge dargestellt, die aus der Abtastung eines Dreieckimpulses entstanden ist. Abbildung 5.16a zeigt die Werte der Folge, die zur Berechnung der DFT verwendet sind. Die sich damit ergebenden Werte für Y_l sind in Abbildung 5.16b und (c) gezeigt. Dabei sind die Werte für $l = 0, \dots, N - 1$, die auch die Werte sind, die in der Regel durch entsprechende Berechnungsalgorithmen zurückgegeben werden, in blau dargestellt.

Auch bei der DFT gilt für reelle Signale y_k , dass $Y_l = Y_{-l}^*$ ist. Aus der N -Periodizität folgt damit auch $Y_l = Y_{N-l}^*$.

Damit enthalten die Werte für $l = 0, \dots, \frac{N}{2}$ (N gerade) bzw. $l = 0, \dots, \frac{N-1}{2}$ (N ungerade) alle Informationen über die DFT. Dies ist auch von daher sinnvoll, dass bei einer reellen Folge von N Werten N Freiheitsgrade existieren. Daher sollte auch deren DFT N Freiheitsgrade besitzen, was durch die beschriebene Symmetrie der Fall ist. (Y_0 ist immer reell und für gerade N ist $Y_{N/2}$ ebenfalls immer reell.)

Interpretation

Vergleicht man die Gleichung (5.12) der DFT,

$$Y_l = \operatorname{DFT}(y_k)_l = \sum_{k=0}^{N-1} y_k e^{-j \cdot \frac{2\pi}{N} \cdot l \cdot k},$$

mit der Gleichung (5.9) der DTFT,

$$Y_D(\Omega) = \operatorname{DTFT}(y_k)|_{\Omega} = \sum_{k=-\infty}^{\infty} y_k \cdot e^{-j\Omega k},$$

so erkennt man schon eine gewisse Ähnlichkeit. Definiert man

$$y'_k = \begin{cases} y_k & 0 \leq k \leq N - 1 \\ 0 & \text{sonst,} \end{cases}$$

was man auch als Multiplikation von y_k mit einem Rechteckfenster w_k der Länge N auffassen kann, erhält man

$$Y'_l = \text{DFT}(y'_k)_l = \sum_{k=0}^{N-1} y'_k e^{-j \cdot \frac{2\pi}{N} \cdot l \cdot k}$$

und

$$Y'_D(\Omega) = \text{DTFT}(y'_k)|_{\Omega} = \sum_{k=0}^{N-1} y'_k \cdot e^{-j\Omega k}.$$

Berechnet man die DTFT nun für die speziellen Werte $\Omega = l \cdot \frac{2\pi}{N}$, so erhält man durch Einsetzen

$$Y'_D\left(l \cdot \frac{2\pi}{N}\right) = \text{DTFT}(y'_k)|_{\Omega=l \cdot \frac{2\pi}{N}} = \sum_{k=0}^{N-1} y'_k \cdot e^{-jl \cdot \frac{2\pi}{N} k} = \text{DFT}(y'_k)_l.$$

D. h. die DFT kann als eine DTFT angesehen werden, der eine Fensterung mit einem Rechteckfenster mit einer Breite von N Abtastwerten vorausgeht und die nur an speziellen Frequenzstützstellen ausgewertet wird. Dies ist in Abbildung 5.17 als Blockschaltbild dargestellt.

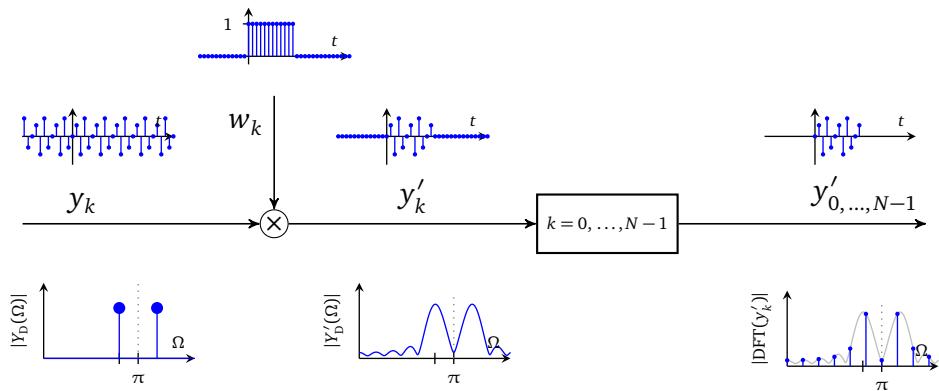


Abbildung 5.17: Zur Interpretation der DFT

(Die oberen Graphen zeigen die Zeitverläufe, die unteren die Beträge der Spektren der Signale.)

Ausgangspunkt in Abbildung 5.17 ist die unendliche Abtastfolge einer Sinusschwingung. Das Spektrum dieses Signals kann in der δ -Notation durch Impulse dargestellt werden, die sich periodisch wiederholen. Die Beschränkung auf N Abtastwerte kann nun in die Schritte

- Fensterung und
- Beschränkung der Abtastpunkte des gefensterten Signals

präzisiert und aufgeteilt werden.

Die Fensterung kann als die Multiplikation der Signalfolge y_k mit der Rechteckfensterfolge w_k angesehen werden und entspricht im Frequenzbereich einer (periodischen) Faltung. Das sich ergebende Signal y'_k samt dem Betrag des Spektrums $Y'_D(\Omega)$ ist ebenfalls in Abbildung 5.17 dargestellt.

Durch die Beschränkung auf N Abtastwerte entsteht im Weiteren kein Fehler im Spektrum mehr, da nur Abtastpunkte mit dem Wert null ignoriert werden, die ohnehin keinen Einfluss mehr auf das Spektrum haben. Jedoch gibt die DFT die Werte des DTFT-Spektrums nur an diskreten, durch N bestimmte Frequenzstützstellen aus.

Der Abstand dieser Frequenzstützstellen, bzw. die Frequenzauflösung beträgt

$$\begin{aligned}\Delta\omega_{\text{DFT}} &= \omega_{l+1} - \omega_l = \frac{1}{T_s} \cdot (\Omega_{l+1} - \Omega_l) = \frac{1}{T_s} \cdot \left((l+1) \cdot \frac{2\pi}{N} - l \cdot \frac{2\pi}{N} \right) \\ &= \frac{2\pi}{T_s N}\end{aligned}$$

bzw.

$$\Delta f_{\text{DFT}} = \frac{1}{T_s N}$$

in Hertz. Die Frequenzauflösung (in Hertz) entspricht also gerade dem Kehrwert der Messdauer $T_N = T_s N$.

Zeropadding

Anhand der Interpretation nach Abbildung 5.17 kann man auch gut das Zeropadding erklären. Hierbei wird der Folge y'_k noch eine (im Grunde beliebige) Anzahl N_Z an Nullen angehängt. Diese ändern nichts am Spektrum, bringen den DFT-Algorithmus aber dazu, den Wert des Spektrums für mehr Frequenzstützstellen zu berechnen. So kann das Spektrum numerisch feiner aufgelöst bestimmt werden. (Siehe auch Abbildung 5.18.)

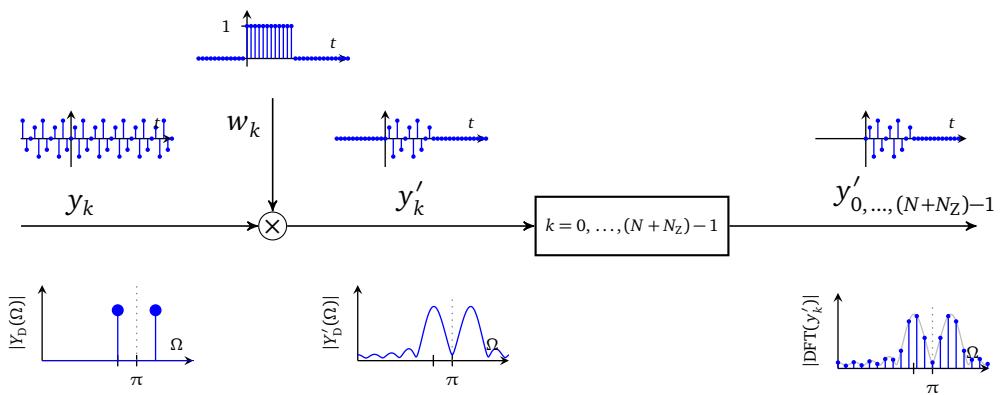


Abbildung 5.18: DFT mit Zeropadding

Wenn man den Wert zu einer ganz bestimmten Frequenz wissen will, dann sollte man Gl. (5.9), also die DTFT, für das gefensterte Signal direkt auswerten.

Wirkung bei kontinuierlichen Signalen

Geht man von einem kontinuierlichen Signal aus, dann stellt sich die Frage, was das über Abtastung und DFT bestimmte Spektrum mit dem ursprünglichen Spektrum des kontinuierlichen Signals zu tun hat. Man kann sich dazu die Wirkung der Abtastung und DFT über die Einzelschritte

- Abtastung mit der Abtastzeit T_s ,
- Fensterung mit Rechteckfenster und

- Beschränkung auf N Abtastwerte

erklären. Dabei können die Schritte „Fensterung“ und „Abtastung“ auch vertauscht werden. Die dazugehörigen Blockschaltbilder sind in Abbildung 5.19a bzw. (b) dargestellt.

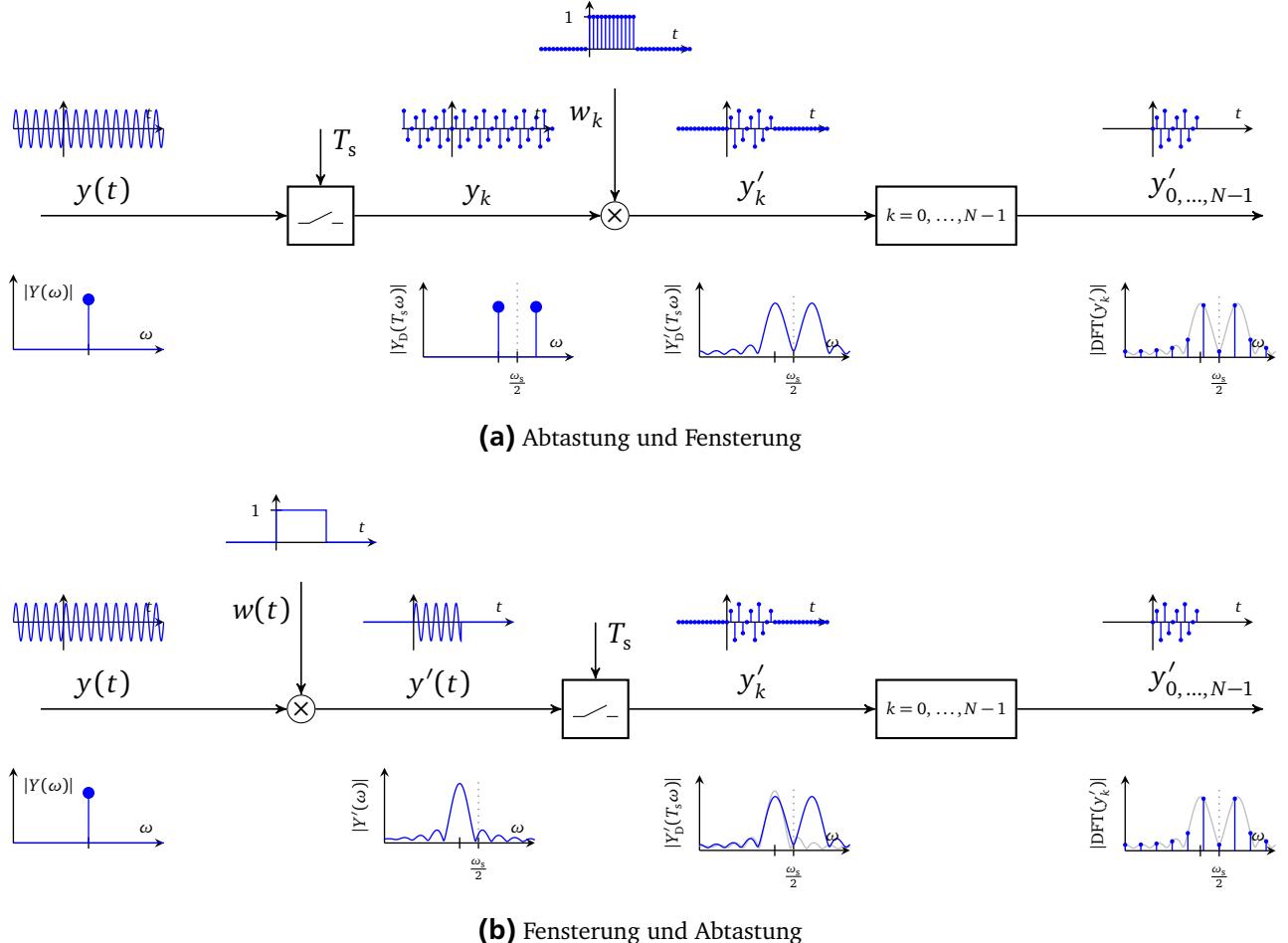


Abbildung 5.19: Zur Interpretation der DFT

(Die oberen Graphen zeigen die Zeitverläufe, die unteren die Beträge der Spektren der Signale. Bei den Spektren ist das Spektrum des jeweils vorhergehenden Schrittes geeignet skaliert in grau dargestellt.)

Ausgangspunkt in diesem Beispiel ist eine unendlich andauernde Sinusschwingung $y(t)$. Dessen Spektrum kann in der δ -Notation durch ein Impulspaar bei der positiven und negativen Schwingungsfrequenz dargestellt werden. (In der Abbildung sind bei den Spektren aber nur positive Werte der Frequenzachse gezeigt.)

Nach Abbildung 5.19a wird dieses zunächst mit der Abtastzeit T_s abgetastet. Hierbei könnte, was in diesem Beispiel aber nicht der Fall ist, Aliasing auftreten. Damit entsprechen die folgenden Schritte dann denen aus Abbildung 5.17.

Man kann aber auch die Einzelschritte in der Reihenfolge nach Abbildung 5.19b anordnen. Das Ergebnis ist natürlich dasselbe, aber zumindest dem Autor fällt es leichter, sich die Fensterung und damit die Faltung bei kontinuierlichen Signalen bzw. deren Spektren vorzustellen, weshalb diese Variante hier ebenfalls kurz besprochen wird.

Da praktisch nur eine endliche Anzahl an Werten abgetastet und verarbeitet werden kann, wird das Signal immer mit einem Rechteckfenster⁸ multipliziert. Entsprechend erhält man das Spektrum $Y'(\omega)$ des gefensterten Signals $y'(t)$ durch die Faltung des Originalspektrums $Y(\omega)$ und der Fouriertransformierten eines Rechteckfensters.

Durch die folgende Abtastung ergibt sich das Spektrum durch die zeitdiskrete Fouriertransformation und kann aus der Überlagerung der entsprechend verschobenen $Y'(\omega)$ konstruiert werden.

Letztlich erfolgt noch die Beschränkung auf die N Abtastwerte bei $k = 0, \dots, N - 1$. Hierbei ist wesentlich, dass sich am Spektrum nichts mehr ändert, sondern lediglich eine Beschränkung auf bestimmte Stützstellen l der im vorhergehenden Schritt entstandene DTFT erfolgt.

Im Vergleich mit Abbildung 5.19a zeigt sich hier auch, dass die Verfälschung durch das Aliasing auch dann auftritt, wenn das eigentliche Signal $y(t)$ nicht zum Aliasing führen würde. In diesem Beispiel ist dies nämlich der Fall. Das Aliasing kommt hier nur durch die Fensterung zu Stande. Die Sinusschwingung wird eigentlich schnell genug abgetastet.

Parameter der DFT

Aus den oben angegebenen Gleichungen lassen sich folgende Aussagen ableiten:

- Ist die Anzahl N der speicherbaren Samples (Speichertiefe) gegeben, so führt eine Verringerung der Abtastzeit T_s zu einer Verringerung der Frequenzauflösung! Soll die Auflösung maximiert werden, dann muss die Abtastzeit unter Beachtung des Aliasing so hoch wie möglich gewählt werden.⁹
- Ist die Messdauer $T_s \cdot N$ gegeben, so führt eine Verringerung der Abtastzeit zu keiner Änderung der Frequenzauflösung.
- Ist die Abtastzeit T_s gegeben, so führt eine Erhöhung der Anzahl der Abtastpunkte N zu einer höheren Frequenzauflösung.
- Die Frequenzauflösung (in Hertz) entspricht dem Kehrwert der Messdauer $T_s N$.

Beispiel Dreiecksimpuls (Zeitbegrenztes Signal)

Als Beispiel der Wirkung der Parameter der DFT wird in Abbildung 5.20 die DFT eines Dreieckimpulses bestimmt, wobei die Parameter der DFT variiert werden. Prinzipiell gilt, dass wenn der komplette Impuls abgetastet wird, das Fenster wirkungslos ist, und man an dieser Stelle keinen Fehler macht. Das Spektrum des Dreieckimpulses enthält auch unendlich hohe Frequenzanteile, so dass ein Aliasing nicht vollständig auszuschließen ist. Jedoch führt eine ausreichend kleine Abtastzeit T_s dazu, dass dieses vernachlässigt werden kann.

Die Verläufe der DTFT und DFT sind mit der Abtastzeit T_s multipliziert, um sie mit der ebenfalls angegebenen Fouriertransformierten des kontinuierlichen Dreieckimpuls vergleichen zu können. Die Verläufe der DTFT und der DFT sind nur bis zur Nyquistfrequenz gezeigt, da darüber hinaus der Vergleich sinnlos ist. Die Verläufe der DFT und DTFT stimmen an den Punkten, für die die DFT bestimmt ist, genau überein, da beide Transformationen das gleiche, zeitbegrenzte Signal verwenden. Daher müssen diese im Weiteren nicht getrennt diskutiert werden.

In Abbildung 5.20a sind 24 Abtastwerte bei einer Abtastzeit von $\frac{T}{12}$ berücksichtigt. Das Spektrum der DFT deckt sich für kleine Frequenzen sehr gut mit dem Referenzspektrum der FT, im mittleren und höheren Frequenzbereich wird ein geringes Aliasing sichtbar.

⁸ Soll ein anderes Fenster, z. B. ein Dreieckfenster, verwendet werden, dann kann man diese bewusste Fensterung auch direkt im Signal $y(t)$ berücksichtigen. Dann wäre die „Zwangsfensterung“ durch die begrenzte Abtastung ohne Wirkung.

⁹ Soll die DFT erst später berechnet werden und lässt sich das Signal noch anpassen, so ist dies keine Einschränkung, da das Signal dann beliebig mit Nullen aufgefüllt werden kann.

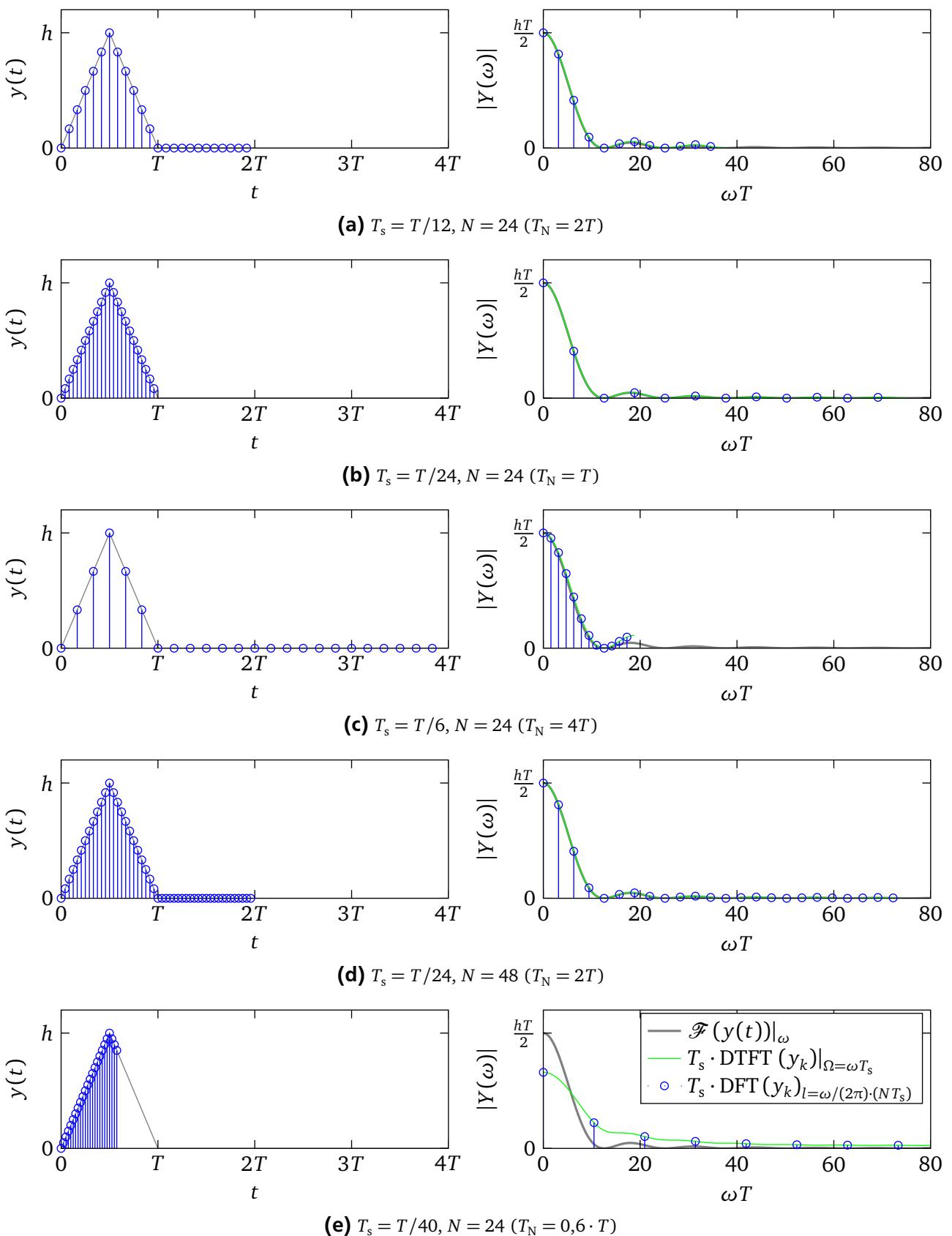


Abbildung 5.20: Vergleich der DFT eines abgetasteten Dreieckimpulses für verschiedene Parameterwerte

Wird die Abtastzeit halbiert, die Anzahl der Punkte aber beibehalten, dann halbiert sich die Messdauer, was eine Halbierung der Frequenzauflösung zur Folge hat, siehe Abbildung 5.20b. Dafür werden Werte des Spektrums auch für höheren Frequenzen berechnet, da sich die Nyquistfrequenz verdoppelt. Damit reduziert sich das Aliasing im Vergleich zum vorherigen Fall nochmals deutlich. (Was aber in der gedruckten Fassung nur sehr schwer zu erkennen sein dürfte.)

Verdoppelt man, ausgehend vom ersten Beispiel, die Abtastzeit bei gleicher Anzahl an Messpunkten, so verdoppelt sich mit der Verdoppelung der Messdauer auch die Frequenzauflösung, Abbildung 5.20c. Dafür halbiert sich die Nyquistfrequenz, womit die Verfälschung des Ursprungsspektrums durch Aliasing deutlich zunimmt.

In Abbildung 5.20d ist die Abtastzeit halbiert, dafür die Anzahl der Messpunkte verdoppelt (wieder bezogen auf das erste Beispiel). Damit bleibt die Frequenzauflösung gleich, bei verdoppelter Nyquistfrequenz. Es werden also auch höhere Frequenzanteile bestimmt und das Aliasing im mittleren Frequenzbereich nimmt ab.

Zuletzt ist in Abbildung 5.20e die Abtastzeit deutlich verringert, ohne dies durch eine höhere Anzahl an Messpunkten auszugleichen. Dies hat hier zu Folge, dass der Dreieckimpuls nicht vollständig erfasst wird. Das Spektrum der DTF (und DTFT, das diese hier auch vom gefensterten Signal bestimmt ist) weichen damit deutlich von der FT des Dreieckimpulses ab. Wenn man möchte, kann man dies als „Leakage“ bezeichnen. Man kann aber auch einfach sagen, dass man das Signal und damit auch dessen Spektrum kaputtmacht, wenn man es mittendrin abschneidet.

Verbindung zwischen DFT und der Fourier-Reihe bei periodischen Signalen

Falls die DFT von einem periodischen Signal bestimmt wird, kein Aliasing auftritt und die Messdauer NT_s einem ganzen Vielfachen n der Periodendauer T entspricht, dann kann ein Zusammenhang zwischen den Werten der DFT und den Fourier-Koeffizienten aufgezeigt werden.

Dies kann auf zwei Wegen geschehen. Zum einen auf direktem Weg, was in Anhang B.2 gezeigt ist. Zum anderen über den mit Gl. (5.7),

$$\check{y}_l = \frac{1}{nT} \cdot Y'(l\omega_T), \quad \omega_T = \frac{2\pi}{T},$$

gegebenen Zwischenschritt über die Fouriertransformierte eines periodischen Signals, welches mit einem Rechteckfenster multipliziert ist, dessen Breite nT ein ganzes Vielfaches n der Periodendauer T ist.

Mit den gemachten Voraussetzungen gilt

$$\text{DTFT}(y'_k)|_{\Omega=\omega_{Ts}} = \frac{1}{Ts} \cdot \mathcal{F}(y'(t))|_{\omega} = \frac{1}{Ts} \cdot Y'(\omega)$$

und

$$Y_{l'} = \text{DFT}(y'_k)|_{l'} = \text{DTFT}(y'_k)|_{\Omega=l'\cdot\frac{2\pi}{N}},$$

wobei hier der Index l' verwendet ist, um ihn von dem Index l der Fourierkoeffizienten \check{y}_l zu unterscheiden. Die Frequenz zum Index l' der DFT entspricht $\frac{\Omega}{Ts} = \frac{l'\cdot 2\pi}{NT_s}$, und diese soll gleich $l\omega_T$ sein. Daraus folgt

$$l' = l\omega_T \cdot \frac{NT_s}{2\pi} = \frac{NT_s}{T} \cdot l = \frac{nT}{T} \cdot l = nl,$$

und aus den vorhergehenden drei Gleichungen ergibt sich

$$\check{y}_l = \frac{1}{nT} \cdot T_s \cdot Y_{n \cdot l}$$

und letztlich, wenn die Messdauer nT durch NT_s ausgedrückt wird,

$$\check{y}_l = \frac{1}{N} \cdot Y_{n \cdot l} . \quad (5.14)$$

In diesem Fall muss der Wert der DFT nur durch die Anzahl der Abtastpunkte geteilt werden.

Hierbei wurde (implizit) angenommen, dass der erste Abtastwert y'_0 für die DFT mit dem Zeitpunkt $t = 0$ des Signals übereinstimmt. Gehört der erste Abtastwert zu $t = t_0$, so muss die Phase noch entsprechend korrigiert werden,

$$\check{y}_l = \frac{1}{N} \cdot Y_{n \cdot l} \cdot e^{-j\omega_T l t_0} . \quad (5.15)$$

Der Faktor $e^{-j\omega_T l t_0}$ entspricht bei der Fouriertransformation einer Zeitverschiebung um t_0 und wird bei den Skalierungen nicht verändert.

Beispiel Sinus (Periodisches Signal)

In Abbildung 5.21 ist die DFT von (Ausschnitten) einer Sinusschwingung gezeigt. Dabei werden die grundsätzlichen Zusammenhänge zwischen Abtastzeit, Anzahl Abtastpunkte und Frequenzauflösung nicht nochmals betrachtet, sondern speziell Punkte diskutiert, die den periodischen Charakter des Ursprungssignals betreffen.

In Abbildung 5.21a sind exakt drei Perioden der Schwingung abgetastet. Damit ergibt sich ein „schönes“ Spektrum der DFT, bei dem nur bei der Frequenz des Sinus ein Wert ungleich null erscheint. Dieser kann mit einer geeigneten Skalierung auch direkt als Fourierreihenentwicklung des periodischen, nicht gefensterten Signals interpretiert werden. Die Nullstellen des gefalteten Spektrums fallen hier genau auf die Frequenzpunkte, für die die DFT Werte berechnet.

Nimmt man einen Abtastpunkt mehr hinzu, so ist nur noch der Wert bei $l = 0$ bzw. $\omega = 0$ gleich null, Abbildung 5.21b. Dadurch, dass ein Abtastwert mit dem Wert null hinzugefügt wurde, ändert sich die DTFT, die aus diesen Punkten bestimmt wird, nicht. Der Frequenzabstand, in dem die DFT Werte liefert, wird aber etwas kleiner, so dass nun, ausgehend von $\omega = 0$, die Frequenzpunkte, an denen die DFT einen Wert liefert, nicht mehr mit den Punkten zusammenfällt, an denen das gefaltete Spektrum null ist. Wenn man Vielfache ganzer Perioden eines periodischen Signals abtasten möchte, dann bedeutet das nicht, dass mit dem letzten Abtastpunkt eine neue Periode beginnt, sondern gerade das der Abtastpunkt, mit dem eine neue Periode beginnen würde, nicht mehr hinzugenommen werden darf.

In Abbildung 5.21c sind zu den 24 abgetasteten Werten noch zwölf Nullen angehängt (zero-padding). Dies ändert wieder nichts an der DTFT, jedoch erhöht sich damit die Frequenzauflösung der DFT. Dies ist hier jedoch nicht unbedingt gewollt, da die eigentliche Frequenz nicht getroffen wird und Nebenfrequenzen im Spektrum erscheinen, die im ursprünglichen periodischen Signal nicht vorhanden sind.

Teilweise wird geraten, die Anzahl der Abtastpunkte für die DFT so festzulegen, dass die periodische Fortsetzung des Signals keine Sprünge aufweist. Dies ist bei periodischen Signalen aber kein sicheres Kriterium, wie in Abbildung 5.21d gezeigt ist. Hier führt dies nämlich zu schlechtesten Wahl, wenn das Treffen der Maximums des Spektrums als Gütemaß herangezogen wird. Auch allgemein ist so eine Herangehensweise fraglich. Sprünge am Rand der Abtastwerte bzw. in der periodischen Fortsetzung deuten zwar darauf hin, dass das Spektrum entsprechend „wildere“ Frequanzanteile hat. Ob diese aber nun „falsch“ sind, kommt natürlich auf das ursprüngliche Signal an. Wenn auf ein System – warum auch immer – ein abgeschnittener Sinus gegeben wird, dann muss auch das Spektrum eines abgeschnittenen Sinus bestimmt werden.

In Abbildung 5.21e sind genau 50 Perioden abgetastet, und entsprechend liegt nur ein DFT-Wert ungleich null vor, nämlich bei der Frequenz der Schwingung. In Abbildung 5.21f sind 50,5 Perioden abgetastet,

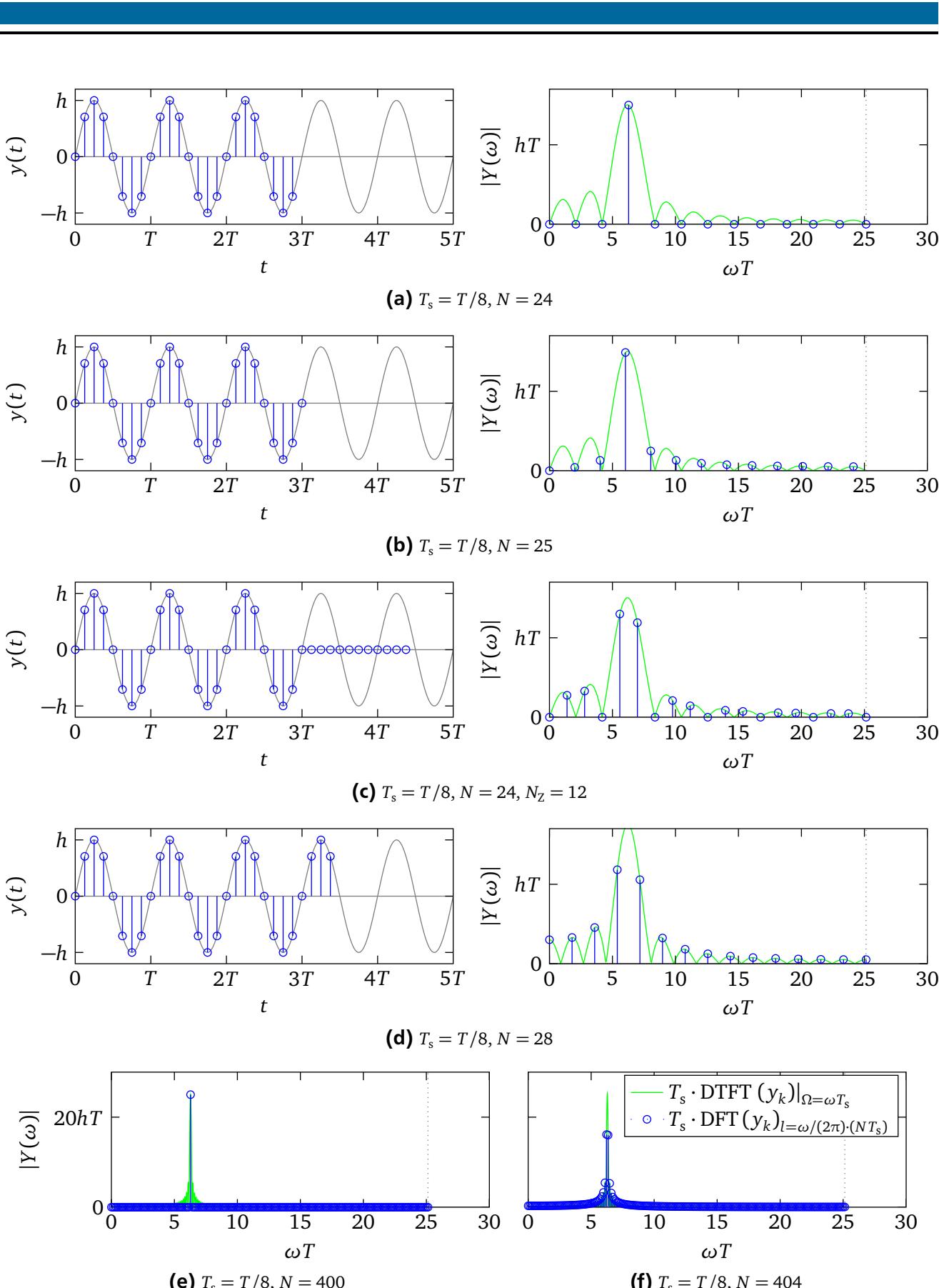


Abbildung 5.21: Vergleich der DFT einer abgetasteten Sinusschwingung für verschiedene Parameterwerte

was wieder zu einem gestörten Spektrum führt. Zwar wird die Breite des Einflussbereiches im Vergleich zur Abtastung von 3,5 Perioden geringer, die Amplitude der Schwingung wird jedoch genauso unterschätzt wie bei 28 Abtastwerten. Auch hierfür verwendet man den Begriff „Leakage-Effekt“. Im Kontext von periodischen Signalen ist diese Bezeichnung auch anschaulich, da das Spektrum von der eigentlichen Frequenz zu den Seiten hin „leckt“.

Damit lässt sich festhalten, dass man, auch wenn sehr viele Perioden gemessen werden, darauf achten sollte die Anzahl der Abtastpunkte so festzulegen, dass möglichst eine ganzzahlige Anzahl an Perioden verwendet wird. Dann erscheinen die harmonischen Anteile nur bei den richtigen Frequenzen,

Fazit zur DFT

- Zeitbegrenzte Signale
 - Ist das Signal zeitbeschränkt und das Intervall, über welches die DFT gebildet wird umfasst den ganzen relevanten Bereich, dann macht man durch die Fensterung keinen Fehler.
 - Sind alle Frequenzanteile des Signals (und des Fensters) über der Nyquistfrequenz vernachlässigbar, dann macht man durch die Abtastung keinen relevanten Fehler.
 - Erfüllt das Signal (mit der verwendeten Abtastfrequenz) beide Bedingungen, dann stellen die Werte der DFT sehr gute Näherungen für die Werte der Fouriertransformation dar.
- Periodische Signale
 - Wenn über ein ganzes Vielfaches der Periodendauer abgetastet wird und kein Aliasing auftritt, dann können die Werte der DFT direkt in Koeffizienten der Fourierreihenentwicklung umgerechnet werden.
- Alle hier nicht aufgezählten Fälle sind möglichst zu vermeiden.

Dabei können Messdaten auch beide Fälle enthalten. So kann eine zeitbegrenzte Systemantwort auf einen Gleitsinus mit periodischen Störungen überlagert sein.

Dies ist zusammenfassend mit den wichtigsten formelmäßigen Zusammenhängen in Abbildung 5.22 dargestellt.

5.6 Berechnung der Transformationen

5.6.1 Fourierreihe und Fouriertransformation

Die Integrale zur Berechnung der Fourierkoeffizienten nach Gl. (5.3),

$$\check{y}_l = \mathfrak{F}(y(t))_l = \frac{1}{T} \int_0^T y(t) \cdot e^{-jl\omega_T t} dt ,$$

bzw. zur Berechnung der Fouriertransformation für auf $[t_a, t_b]$ zeitbegrenzte Signale nach Gl. (5.4),

$$Y(\omega) = \mathcal{F}(y(t))|_{\omega} = \int_{t_a}^{t_b} y(t) \cdot e^{-j\omega t} dt ,$$

können analytisch oder numerisch gelöst werden. (Analytisch können auch nichtzeitbegrenzte Signale behandelt werden.) Darüberhinaus können diese aufgrund der in diesem Kapitel besprochenen Zusammenhänge auch in abgetaster Form in sehr guter Näherung durch eine DFT bestimmt werden.

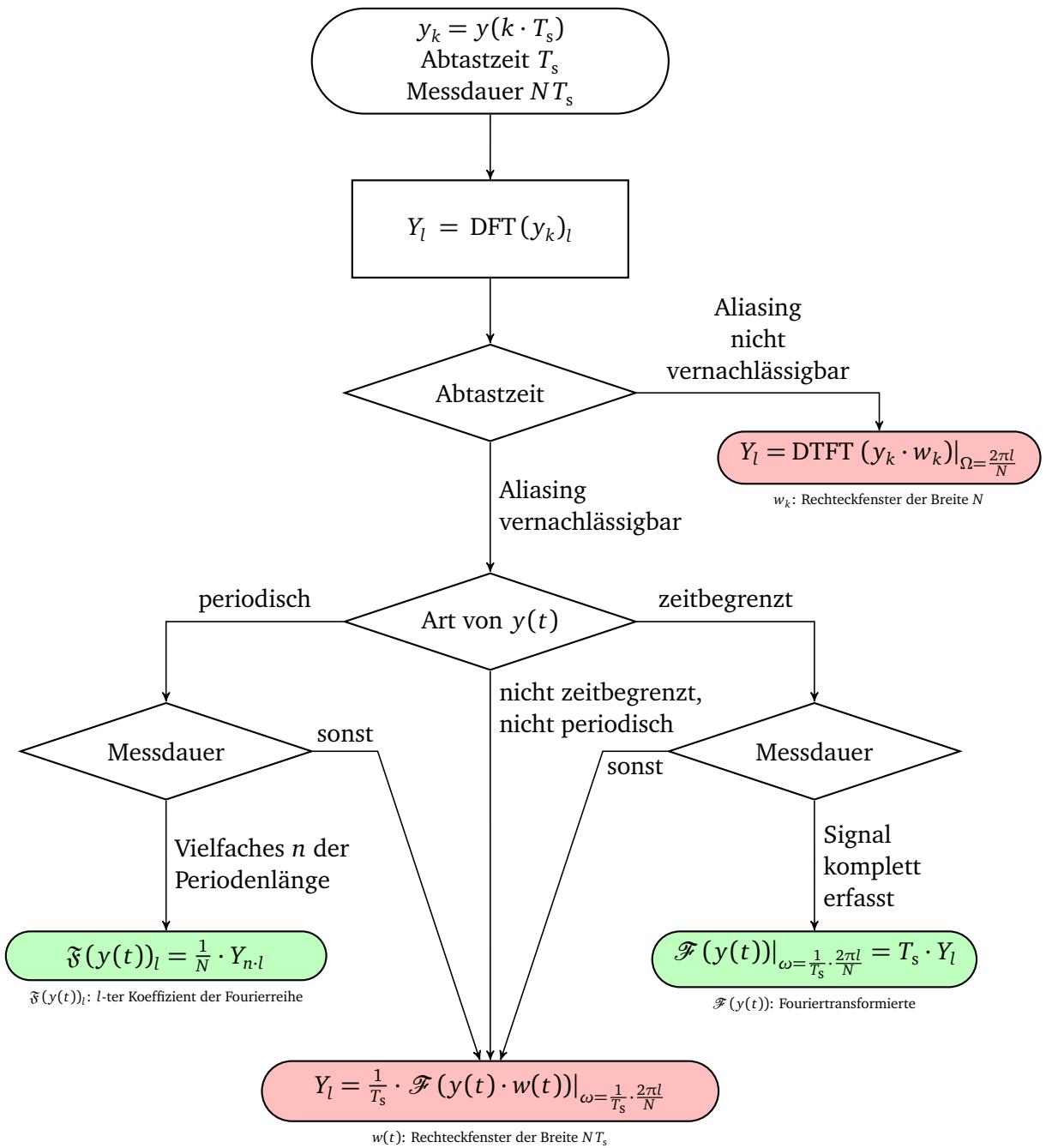


Abbildung 5.22: Interpretation der DFT

Bei der numerischen Integration ist dabei auf die Schrittweite zu achten. So entspricht die Integration über das einfache Eulerverfahren der Abtastung des Signales $y(t)$ mit einem Halteglied nullter Ordnung, die Integration über das Trapezverfahren ist ähnlich der Abtastung mit einem Halteglied erster Ordnung. Entsprechend kann auch Aliasing auftreten, welches aber durch eine geeignete Wahl der Schrittweite vernachlässigbar klein gehalten werden kann. Zumindest in einem vernünftigen Frequenzbereich.

Bei der analytischen Integration sollten die Eigenschaften der Transformationen bezüglich gerader und ungerader Funktionen sowie deren Linearität beachtet werden, um sich Arbeit zu sparen. Drückt man die Exponentialfunktionen in den Integralen durch die Eulerformel als $e^{-ja} = \cos(a) - j \cdot \sin(a)$ aus, so verschwinden bei geraden Funktionen die Sinusterme und bei ungeraden Funktionen die Kosinusterme. Zudem muss nur noch die Hälfte des ursprünglichen Integrationsbereichs ausgewertet werden. So ist es von Vorteil, die eigentliche Funktion auf Summen von (verschobenen) geraden und ungeraden Funktionen zurückzuführen.

Die Verschiebung von Funktionen im Zeitbereich entspricht einer Phasenverschiebung im Frequenzbereich,

$$\mathcal{F}(y(t - \tau))|_{\omega} = e^{-j\tau\omega} \cdot Y(\omega).$$

Beispiel Rechteckimpuls

Als Beispiel für die Berechnung wird der in Abbildung 5.23a dargestellte Rechteckimpuls, der hier schon häufig verwendet wurde, betrachtet.

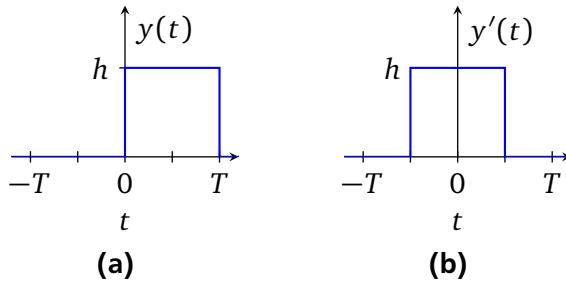


Abbildung 5.23

Für die einfachere Berechnung wird dazu zunächst der um $T/2$ nach links verschobene Rechteckimpuls $y'(t)$ betrachtet. Der Zusammenhang zwischen $y(t)$ und $y'(t)$ ist über

$$y(t) = y'(t - T/2)$$

gegeben, und so kann aus $Y'(\omega)$ dann über den Verschiebungssatz die eigentlich gesuchte Fouriertransformierte $Y(\omega)$ bestimmt werden.

Für $Y'(\omega)$ ergibt sich

$$\begin{aligned} Y'(\omega) &= \int_{-\infty}^{\infty} y'(t) \cdot e^{-j\omega t} dt \\ &= \int_{-T/2}^{T/2} h \cdot (\cos(\omega t) - j \cdot \sin(\omega t)) dt \\ &= 2h \cdot \int_0^{T/2} \cos(\omega t) dt \\ &= 2h \cdot \left[\frac{1}{\omega} \sin(\omega t) \right]_0^{T/2} = h \cdot T \cdot \frac{\sin\left(\frac{\omega T}{2}\right)}{\frac{\omega T}{2}}. \end{aligned}$$

Damit folgt

$$Y(\omega) = Y'(\omega) \cdot e^{-j\frac{\omega T}{2}} = h \cdot T \cdot \frac{\sin\left(\frac{\omega T}{2}\right)}{\frac{\omega T}{2}} \cdot e^{-j\frac{\omega T}{2}}.$$

Beispiel zusammengesetzte Funktion

Als zweites Beispiel wird die in Abbildung 5.24a gezeigte Funktion betrachtet.

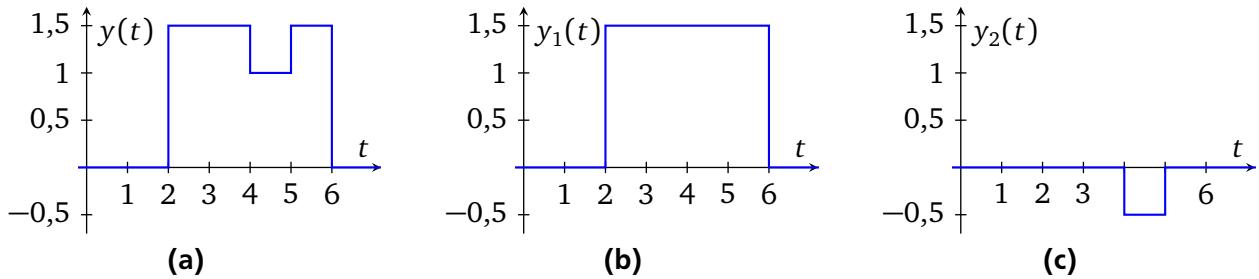


Abbildung 5.24

Diese kann als Summe

$$y(t) = y_1(t) + y_2(t)$$

von den in Abbildung 5.24b und (c) dargestellten verschobenen Rechteckimpulsen geschrieben werden. Aufgrund der Linearität der Fouriertransformation kann man direkt

$$\begin{aligned} Y(\omega) &= Y_1(\omega) + Y_2(\omega) \\ &= 1,5 \cdot 4 \cdot \frac{\sin\left(\frac{\omega \cdot 4}{2}\right)}{\frac{\omega \cdot 4}{2}} \cdot e^{-j\frac{\omega \cdot 4}{2}} - 0,5 \cdot 1 \cdot \frac{\sin\left(\frac{\omega \cdot 1}{2}\right)}{\frac{\omega \cdot 1}{2}} \cdot e^{-j\frac{\omega \cdot 1}{2}} \end{aligned}$$

hinschreiben.

5.6.2 DFT und DTFT

Die Werte der DFT nach Gl. (5.12),

$$Y_l = \text{DFT}(y_k)_l = \sum_{k=0}^{N-1} y_k e^{-j \frac{2\pi}{N} l k},$$

kann über

- die direkte Auswertung,
- die Fast-Fourier-Transformation (FFT) oder
- den Goertzel-Algorithmus

bestimmt werden.

Direkte Auswertung

Für jeden der $N/2$ notwendigen Frequenzpunkte sind N Multiplikationen einer reellen Zahl mit einer komplexen Zahl und $N-1$ komplexe Additionen durchzuführen. Der Aufwand wächst damit quadratisch mit der Anzahl der Stützstellen N an, $\mathcal{O}(N^2)$.

Dazu kommt, dass die N notwendigen Faktoren $e^{-j \frac{2\pi}{N} k'}$, $k' = 0, \dots, N-1$, bekannt sein müssen, was einen entsprechend hohen Speicherbedarf bedingt, oder einen entsprechend hohen Rechenaufwand, der den Aufwand der Multiplikationen übersteigen dürfte.

Fast-Fourier-Transformation

Unter dem Begriff Fast-Fourier-Transformationen (FFT) werden eine Vielzahl von Algorithmen zusammengefasst, die gewisse Symmetrieeigenschaften bei der Berechnung der DFT ausnutzen. Dadurch wächst der Aufwand mit $N \cdot \log N$ mit der Stützstellenanzahl, $\mathcal{O}(N \cdot \log N)$. Diese Algorithmen sind heute weit verbreitet und in vielen Softwarebibliotheken und -programmen integriert.

Für die reinen FFT-Algorithmen muss die Anzahl der Stützstellen N als Zweierpotenz 2^n darstellbar sein. Damit ist man diesbezüglich eingeschränkt, was ungünstig sein kann, wenn periodische Signale betrachtet werden, da bei diesen die Messdauer wie begründet möglichst genau einem ganzen Vielfachen der Periodendauer entsprechen sollte.

Es existieren jedoch auch Algorithmen, die die DFT beliebiger Länge effizient berechnen. So kann beispielsweise Matlab mit dem Befehl `fft` tatsächlich auch die DFT von Folgen bestimmen, deren Länge nicht als Zweierpotenz darstellbar ist.

Goertzel-Algorithmus

Der in Anhang B.3 vorgestellte Goertzel-Algorithmus berechnet die DFT an einem Frequenzpunkt l über einen Filter zweiter Ordnung, wobei bis zum letzten Schritt auch nur reelle Werte auftreten. Der Gesamtaufwand, d. h. der Aufwand zur Bestimmung der DFT an allen Frequenzpunkten ist dabei wie bei der direkten Auswertung $\mathcal{O}(N^2)$. Gegenüber der FFT kann er aber immer noch günstiger sein, wenn die DFT nur an wenigen Frequenzpunkten bestimmt werden soll. Gegenüber der direkten Auswertung liegt der Vorteil darin, dass nur einer der Koeffizienten $e^{-j \frac{2\pi}{N} \cdot k'}$ je Frequenzpunkt zu bestimmten oder zu speichern ist. Dafür liegen die Filterpole auf dem Einheitskreis, was zu Stabilitätsproblemen führen kann, wenn die Anzahl N sehr groß und die Zahlendarstellung zu ungenau ist.

Bestimmung der DTFT

Für zeitbegrenzte Signale kann man über die DTFT Werte des Spektrums für beliebige Frequenzpunkte bestimmen. Die DTFT kann dabei formal auch in die Form der DFT mit nicht-ganzzahligem l gebracht werden. Die direkte Auswertung und – mit einer kleinen Anpassung – der Goertzel-Algorithmus können auch mit nicht-ganzzahligen l durchgeführt werden. Man sollte sprachlich dabei aber darauf achten, dass man in diesem Fall die DTFT berechnet, und nicht die DFT. Letztere ist nur für ganzzahlige l definiert.

Muss aufgrund großer Datenmengen oder Bequemlichkeit ein FFT-Algorithmus verwendet werden, kann durch Zero-Padding mit einer geeigneten Anzahl an Nullen erreicht werden, den gewünschten Frequenzpunkt zu treffen.

Fazit

Zur Bestimmung der kompletten DFT sollte ein FFT-Algorithmus verwendet werden. Soll die DFT nur an wenigen Punkten ausgewertet werden, kann auch eine direkte Auswertung oder die Verwendung des Goertzel-Algorithmus sinnvoll sein. Letzteres insbesondere bei der Implementierung auf Mikrocontrollern.

6 Signale

In diesem Kapitel werden verschiedene stochastische und deterministische Signale behandelt. In den Diskussionen zu den Signalen wird häufiger von der Signalleistung oder der Signalenergie gesprochen werden. Diese Begriffe werden zunächst erörtert. Im Anschluss werden Signale behandelt, die häufig als Störungen auf ein System wirken, danach Signale, die zur Anregung des Systems bei der Identifikation geeignet sind. Dabei werden jeweils (pseudo-)stochastische und deterministische Signale berücksichtigt.

6.1 Signalleistung- und energie

Betrachtet man physikalische Größen wie beispielsweise Strom $i(t)$, Spannung $u(t)$ oder Geschwindigkeit $v(t)$, dann sind diese Größen in quadrierter Form häufig proportional zu einer Leistung oder einer Energie. So gilt für den Strom bzw. die Spannung an einem Widerstand R , dass die im Widerstand dissipierte Leistung $P(t) = R \cdot i^2(t) = \frac{1}{R} \cdot u^2(t)$ beträgt. Bei einem Kondensator ist die gespeicherte Energie mit $E(t) = \frac{1}{2} C u^2(t)$ proportional zur quadrierten Kondensatorspannung. Eine Masse mit der Geschwindigkeit $v(t)$ besitzt die kinetische Energie $E(t) = \frac{1}{2} m v^2(t)$ und ein linearer Dämpfer d dissipiert die Leistung $P(t) = d \cdot v^2(t)$.

Dies motiviert, auch bei einem allgemeinen Signal $y(t)$ von der Signalleistung $y^2(t)$ zu sprechen. In den wenigsten Fällen entspricht dies tatsächlich (alleine schon von der Einheit her) einer Leistung, jedoch könnte über eine geeignete Konstante zumindest die Einheit einer Leistung erreicht werden. Und in vielen Fällen kann $y^2(t)$ tatsächlich mit einer geeigneten Konstante auch als eine physikalische Leistung interpretiert werden.

Wird für allgemeine Signale die (zweiparametrische) Korrelationsfunktion für $t_1 = t_2 = t$ ausgewertet,

$$R_{yy}(t, t) = E\{y(t)y(t)\}_r ,$$

so entspricht diese dem Erwartungswert für die Signalleistung zum Zeitpunkt t .

Bei stationären Signalen entspricht dies der Auswertung der einparametrischen Korrelationsfunktion bei $\tau = 0$,

$$R_{yy}(0) = E\{y(t)y(t)\}_r .$$

Der Erwartungswert hängt der Voraussetzung nach nicht von t ab, und $R_{yy}(0)$ entspricht dem konstanten Erwartungswert der Signalleistung, d. h. der mittleren Signalleistung.

Für periodische Signale gilt

$$R_{yy}^P(0) = \frac{1}{NT} \cdot \int_{t_0}^{t_0+NT} y(t)y(t)dt ,$$

und $R_{yy}^P(0)$ gibt die durchschnittliche Signalleistung an.

Für zeitbegrenzte Signale wird die Impulskorrelationsfunktion angewendet, und es gilt

$$R_{yy}^I(0) = \int_{-\infty}^{\infty} y(t)y(t)dt .$$

$R_{yy}^I(0)$ gibt damit die Gesamtsignalenergie an.

Zu jeder einparametrischen Autokorrelationsfunktion kann das dazugehörige Autoleistungsdichtespektrum als dessen Fouriertransformierte definiert werden. Damit ist die Autokorrelationsfunktion auch über die inverse Fouriertransformation des Autoleistungsdichtespektrum bestimmbar,

$$R_{yy}(\tau) = \int_{-\infty}^{\infty} S_{yy}(\omega) e^{j\omega\tau} d\omega .$$

Für $\tau = 0$ gilt damit

$$R_{yy}(0) = \int_{-\infty}^{\infty} S_{yy}(\omega) e^0 d\omega = \int_{-\infty}^{\infty} S_{yy}(\omega) d\omega .$$

Das Integral des Autoleistungsdichtespektrum entspricht also der (durchschnittlichen) Signalleistung.

Filtert man das ursprüngliche Signal $y(t)$ mit einem idealen Bandpassfilter $H(\omega)$ für das

$$|H(\omega)| = \begin{cases} 1 & \text{für } \omega_a \leq |\omega| \leq \omega_b \\ 0 & \text{sonst} \end{cases}$$

gilt, so gilt für das gefilterte Signal $y'(t)$

$$S_{y'y'}(\omega) = |H(\omega)|^2 \cdot S_{yy}(\omega) = \begin{cases} S_{yy}(\omega) & \text{für } \omega_a \leq |\omega| \leq \omega_b \\ 0 & \text{sonst.} \end{cases}$$

Die Autokorrelationsfunktion bei $\tau = 0$ des gefilterten Signals

$$R_{y'y'}(0) = \int_{\omega_a}^{\omega_b} (S_{yy}(\omega) + S_{yy}(-\omega)) \cdot d\omega = 2 \int_{\omega_a}^{\omega_b} S_{yy}(\omega) \cdot d\omega$$

entspricht damit der Signalleistung des ursprünglichen Signals, welches nur durch die Signalanteile im Intervall $[\omega_a, \omega_b]$ entsteht.

So kann zum einen begründet werden, dass die Autoleistungsdichte nur Werte größer oder gleich Null annehmen kann. Gäbe es ein Intervall, indem die Autoleistungsdichte kleiner null wäre, so würde durch die beschriebene Filterung, die nur dieses Intervall durchlässt, ein Signal entstehen, welches eine negative durchschnittliche Leistung besäße. [Hänsler 2001, S. 152]

Zum anderen sind damit folgende Ausdrucksweisen gerechtfertigt: So wie bei der Fouriertransformierten eines Signals $y(t)$ von dessen Amplitudendichtespektrum gesprochen werden kann, so ist hier der Begriff (Auto)Leistungsdichtespektrum sinnvoll. Am Leistungsdichtespektrum kann abgelesen werden, in welchem Frequenzbereich ein Signal eine hohe oder niedrige Leistung¹ besitzt.

Zusammenhang Impuls-Leistungsdichtespektrum und Fouriertransformation

Bildet man das Produkt aus der konjugierten Fouriertransformierten und der Fouriertransformierten, so ergibt sich

$$\begin{aligned} Y^*(\omega) \cdot Y(\omega) &= \int_{-\infty}^{\infty} y^*(t) e^{j\omega t} dt \cdot \int_{-\infty}^{\infty} y(t) e^{-j\omega t} dt \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y(\tau) \cdot y(t) \cdot e^{-j\omega \cdot (t-\tau)} d\tau dt , \end{aligned}$$

¹ Bzw. Energie, im Fall der Impulsleistungsdichte. Der Zusammenhang $S_{y'y'}(\omega) = |H(\omega)|^2 \cdot S_{yy}(\omega)$ wurde zwar nur für den Fall stationärer Signale gezeigt, jedoch lässt sich ein gleicher Zusammenhang auch für die Impulsleistungsdichte zeigen.

wobei im zweiten Schritt $(y(t)e^{-j\omega t})^* = y^*(t)e^{j\omega t}$ verwendet und darüber hinaus die Reellwertigkeit von $y(t)$, $y^*(t) = y(t)$, ausgenutzt wurde.

Mit der Substitution $u = t - \tau$ im inneren Integral ergibt sich

$$Y^*(\omega) \cdot Y(\omega) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y(t-u) \cdot y(t) \cdot e^{-j\omega u} du dt ,$$

und weiter durch Vertauschen der Integrationsreihenfolge

$$Y^*(\omega) \cdot Y(\omega) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y(t-u) \cdot y(t) \cdot dt \cdot e^{-j\omega u} du .$$

Für zeitbegrenzte Signale entspricht das Integral $\int_{-\infty}^{\infty} y(t-u) \cdot y(t) \cdot dt$ der Impuls-Autokorrelationsfunktion $R_{yy}^I(\tau)$, so dass man

$$Y^*(\omega) \cdot Y(\omega) = |Y(\omega)|^2 = \int_{-\infty}^{\infty} R_{yy}^I(u) \cdot e^{-j\omega u} du = S_{yy}^I(\omega)$$

erhält.

Für zeitbegrenzte Signale entspricht das Quadrat des Betrages der Fouriertransformierten demnach der Impuls-Autoleistungsdichte.

6.2 Stochastische und deterministische Störungen

In diesem Abschnitt werden zunächst Rauschsignale diskutiert und dann verschiedene deterministische Störungen besprochen. Dabei steht die Frage im Mittelpunkt, in welchem Frequenzbereich die Störungen Einfluss nehmen, und – in Hinblick auf das folgende Kapitel – es wird untersucht, wie der Einfluss auf eine bestimmte Frequenz ω der Fouriertransformierten von der Messdauer T abhängt. Da später der Wert der Fouriertransformierten noch durch die Messdauer T geteilt wird, sind die Ergebnisse hier wie folgt zu deuten. Wenn der Einfluss linear mit T zunimmt, so hängt der Einfluss auf den späteren Schätzwert nicht von T ab. Nimmt der Einfluss auf die Fouriertransformierte schwächer als linear mit T zu, so sinkt der Einfluss auf den Schätzwert mit wachsender Messdauer.

6.2.1 Rauschen

Weißes Rauschen

Im Zeitdiskreten beschreibt „weißes Rauschen“ einen stochastischen Prozess $\{n_k\}_r$, der

- in sich unkorreliert ist, d. h. $E\{n_k n_l\}_r = E\{n_k\}_r \cdot E\{n_l\}_r$ für $k \neq l$,
- mittelwertfrei ist, $E\{n_k\}_r = 0$, und
- deren Variablen eine gleiche Verteilung besitzen, $\{n_k\}_r \sim \{n_l\}_r$.

Daraus folgt für die Korrelationsfunktion

$$R_{nn}(\kappa) = E\{n_k n_{k+\kappa}\}_r = \sigma_n^2 \cdot \delta_\kappa = \begin{cases} \sigma_n^2 & \kappa = 0 \\ 0 & \kappa \neq 0 \end{cases},$$

wobei δ_κ das Kroneckerdelta $\delta_\kappa = \begin{cases} 1 & \kappa = 0 \\ 0 & \kappa \neq 0 \end{cases}$ bezeichnet, und daraus das Autoleistungsdichespektrum

$$S_{nn}(\Omega) = \sum_{\kappa=-\infty}^{\infty} R_{nn}(\kappa) \cdot e^{-j\Omega\kappa} = \sigma_n^2.$$

Weißes Rauschen kann im Diskreten als eine Folge identisch verteilter, unabhängiger Zufallsvariablen aufgefasst werden. Übernimmt man diese Definition ins Zeitkontinuierliche, dann ergibt sich als Autokorrelationsfunktion

$$R_{xx}(\tau) = E\{x(t)x(t+\tau)\}_r = \begin{cases} \sigma_x^2 & \tau = 0 \\ 0 & \tau \neq 0. \end{cases}$$

Daraus würde dann aber auch

$$S_{xx}(\omega) = \int_{-\infty}^{\infty} R_{xx}(\tau) e^{-j\omega\tau} d\tau = 0$$

folgen.

Je höher die Abtastzeit bei einem diskreten weißen Rauschen ist, desto größer wird der Frequenzbereich, über den sich die Leistung verteilen muss. (Es ist immer nur ein Intervall von 2π Breite für Ω zu „füllen“, welches alle unterscheidbaren Frequenzen enthält. Durch die periodische Fortsetzung ist dann auch der Rest des Spektrums zwischen $(-\infty, \infty)$ gegeben.)

Bei einem kontinuierlichen Rauschen geht damit die Autoleistungsdichte gegen null. Ein solches Signal würde aber in keinem Frequenzbereich ein reales Rauschen sinnvoll annähern, und wäre für die mathematische Behandlung der kommenden Problemstellungen unbrauchbar – wenn es auch zu sehr einfachen Ergebnissen führen würde.

Es wird daher gefordert, dass sich ein weißes Rauschen dadurch auszeichnet, dass die Autoleistungsdichte über den gesamten Frequenzbereich $(-\infty, \infty)$ einen konstanten Wert annehmen soll,

$$S_{nn}(\omega) = S_0.$$

Eine solche Autoleistungsdichte wird durch die Autokorrelationsfunktion

$$R_{nn}(\tau) = S_0 \cdot \delta(\tau)$$

erzeugt. Diese ist überall null bis auf einen δ -Impuls bei $\tau = 0$. Da der Wert der Autokorrelationsfunktion bei $\tau = 0$ für mittelwertfreie Signale gerade der Varianz entspricht, bedeutet dies, dass kontinuierliches weißes Rauschen eine unendlich hohe Varianz besitzt.

Kontinuierliches weißes Rauschen ist damit (technisch) aus zwei Gründen nicht realisierbar. Zum einen ist das Signal zu zwei verschiedenen, aber beliebig nah aneinanderliegenden Zeitpunkten völlig unkoreliert, was eine unendlich hohe Änderungsgeschwindigkeit bedeuten würde. Zum anderen hätte dieses Rauschen unendlich hohe „Amplituden“ und damit eine unendlich hohe Leistung.²

Trotz dessen, dass kontinuierliches weißes Rauschen im Zeitbereich nicht darzustellen ist, ist es im Frequenzbereich ein gutes Modell für breitbandige Störungen, die sich über den gesamten Dynamikbereich des betrachteten Systems bzw. den betrachteten Frequenzbereich erstrecken.

Diskretes weißes Rauschen ist realisierbar.

² Physikalisch, aber nicht signaltheoretisch gesehen würde auch für die unendlich hohen Änderungen des Signalwertes in der Regel schon eine unendlich hohe Leistung notwendig sein, auch wenn die Varianz endlich wäre.

Farbiges Rauschen

Unter farbigen Rauschen versteht man Rauschprozesse, deren Autoleistungsdichtespektren nicht einen konstanten Wert über alle Frequenzen annehmen. In der Regel hat das farbige Rauschen eine hohe Leistung bei niedrigen Frequenzen und für hohe Frequenzen geht die Signalleistung gegen Null.

Allgemein gilt für die Autokorrelation eines stationären Ein- und Ausgangsprozesses eines LTI-Systems

$$S_{yy}(\omega) = |H(\omega)|^2 \cdot S_{uu}(\omega)$$

und damit kann aus weißem Rauschen ein farbiges Rauschen erzeugt werden,

$$S_{nn}(\omega) = |H_n(\omega)|^2 \cdot S_0 .$$

Geht man als einfachstes von einem PT₁-System mit $H_n(\omega) = \frac{1}{1+j\omega T_n}$ aus, so gilt

$$S_{nn}(\omega) = \frac{S_0}{1 + \omega^2 T_n^2} .$$

Aus dieser Autoleistungsdichte ergibt sich die Autokorrelationsfunktion

$$R_{nn}(\tau) = \frac{S_0}{2T_n} \cdot e^{-\frac{|\tau|}{T_n}} .$$

In Abbildung 6.1 ist das Leistungsdichtespektrum eines über ein PT₁-Filter aus weißem Rauschen erzeugtes farbiges Rauschen gezeigt. In den drei Teileabbildungen ist dabei jeweils eine andere Achseninteilung verwendet, um die verschiedenen Möglichkeiten der Darstellung aufzuzeigen. In Abbildung 6.2 ist die dazugehörige Autokorrelationsfunktion dargestellt.

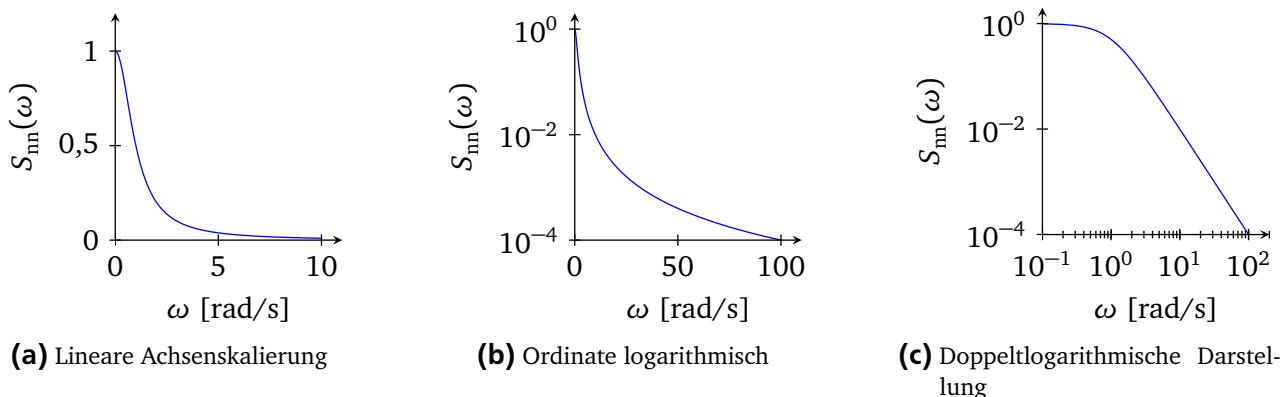


Abbildung 6.1: Leistungsdichtespektrum $S_{nn}(\omega)$ von farbigem Rauschen ($T_n = 1 \text{ s}$)

In Abbildung 6.3 ist weißes und farbiges Rauschen verglichen. Es sind zwei farbige Rauschsignale verwendet, wovon eines mit einer Filterung mit $T_n = 1 \text{ s}$ und eines mit $T_n = 10 \text{ s}$ erzeugt wurde.

In Abbildung 6.3a sind die Zeitverläufe einer simulierten Realisierung der Prozesse dargestellt. Dies ist nur für das farbige Rauschen möglich, da zeitkontinuierliches weißes Rauschen im Zeitbereich nicht sinnvoll dargestellt werden kann.

Die Leistungsdichtespektren sind in Abbildung 6.3b und die Autokorrelationsfunktionen in (b) gezeigt. Die Leistungsdichtespektren zeigen deutlich die Filtercharakteristik. Die Spitze der Autokorrelationen ist umso breiter, desto stärker die Filterung, d. h. desto größer die Filterzeitkonstante ist. Dies bedeutet, dass mit wachsender Zeitkonstante auch die Korrelation weiter entfernter Zeitpunkten zunimmt.

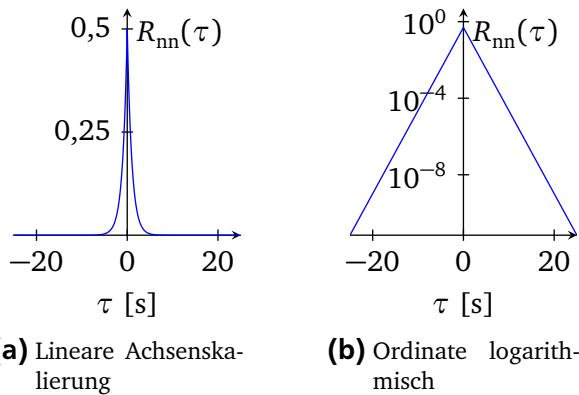


Abbildung 6.2: Autokorrelationsfunktion $R_{nn}(\tau)$ von farbigem Rauschen ($T_n = 1 \text{ s}$)

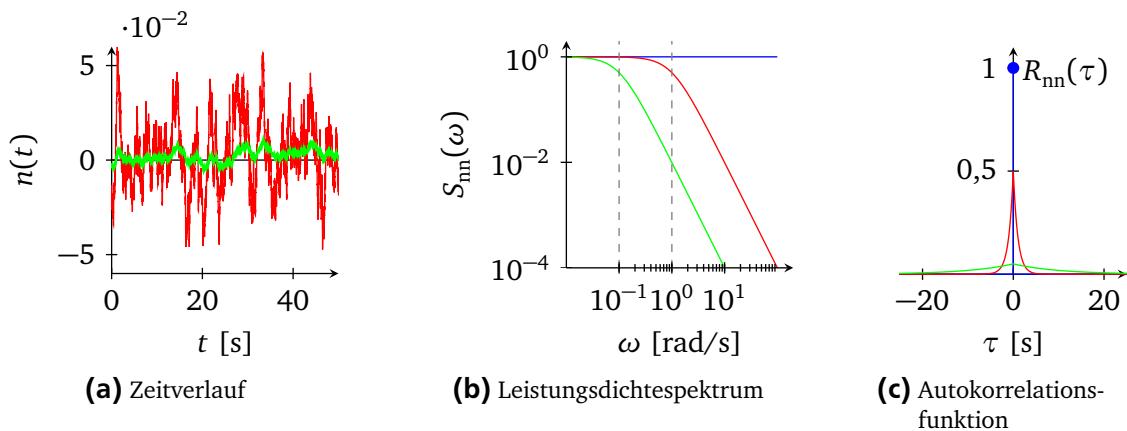


Abbildung 6.3: Vergleich von weißem und farbigem Rauschen (blau: weißes Rauschen, rot: $T_n = 1 \text{ s}$, grün: $T_n = 10 \text{ s}$)

Einfluss auf die Fouriertransformation

Es soll zu gestörten Daten

$$\tilde{y}(t) = y(t) + n(t)$$

die Fouriertransformierte bestimmt werden, wobei aufgrund der endlichen Messzeit T ein Rechteckfenster entsprechender Breite zu berücksichtigen ist. Aus der Linearität folgt

$$\tilde{Y}(\omega) = Y(\omega) + N(\omega),$$

d. h. der Einfluss des Rauschens kann unabhängig vom Nutzsignal $y(t)$ betrachtet werden.

Das Spektrum des Rauschens $N(\omega)$ ist eine komplexwertige Funktion, weshalb dessen Real- und Imaginärteil berücksichtigt werden muss. Bildet man mit einer Realisierung des Rauschprozesses $\{n(t)\}_r$ die Fouriertransformierte, erhält man eine Realisierung $N(\omega)$ des Prozesses $\{N(\omega)\}_r$,

$$N(\omega) = \int_0^T n(t) \cdot e^{-j\omega t} \cdot dt.$$

(Mit $N(\omega)$ ist hier die Fouriertransformierte unter Berücksichtigung eines Rechteckfensters mit der Breite T bezeichnet.)

Es gilt

$$\operatorname{Re}(N(\omega)) = \int_0^T n(t) \cdot \cos(\omega t) \cdot dt$$

und

$$\operatorname{Im}(N(\omega)) = \int_0^T n(t) \cdot \sin(\omega t) \cdot dt$$

für Real- bzw. Imaginärteil.

Das Rauschen wird als mittelwertfrei angenommen, $E\{n(t)\}_r = 0$, und damit ergibt sich

$$E\{\operatorname{Re}(N(\omega))\}_r = \int_0^T E\{n(t)\}_r \cdot \cos(\omega t) \cdot dt = 0$$

und

$$E\{\operatorname{Im}(N(\omega))\}_r = \int_0^T E\{n(t)\}_r \cdot \sin(\omega t) \cdot dt = 0.$$

Für die Varianz des Realteils ergibt sich damit

$$\begin{aligned} \operatorname{Var}\{\operatorname{Re}(N(\omega))\}_r &= E\{(\operatorname{Re}(N(\omega)))^2\}_r \\ &= E\left\{ \int_0^T n(t) \cdot \cos(\omega t) \cdot dt \cdot \int_0^T n(t') \cdot \cos(\omega t') \cdot dt' \right\}_r \\ &= E\left\{ \int_0^T \int_0^T n(t)n(t') \cdot \cos(\omega t) \cdot \cos(\omega t') \cdot dt' dt \right\}_r \\ &= \int_0^T \int_0^T E\{n(t)n(t')\}_r \cdot \cos(\omega t) \cdot \cos(\omega t') \cdot dt' dt \\ &= \int_0^T \int_0^T R_{nn}(t - t') \cdot \cos(\omega t) \cdot \cos(\omega t') \cdot dt' dt. \end{aligned}$$

Die letzte Zeile kann weiter zu

$$\operatorname{Var}\{\operatorname{Re}(N(\omega))\}_r = T \cdot \int_0^T R_{nn}(\nu) \cdot \left(\left(1 - \frac{\nu}{T}\right) \cdot \cos(\omega\nu) + \frac{1}{\omega T} \cdot (\sin(\omega \cdot (2T - \nu)) - \sin(\omega\nu)) \right) d\nu$$

umgeformt werden, was in [Eykhoff 1974, 532f] gezeigt und hier im Anhang C.1 dargestellt wird. Für den Imaginärteil gilt analog

$$\operatorname{Var}\{\operatorname{Im}(N(\omega))\}_r = T \cdot \int_0^T R_{nn}(\nu) \cdot \left(\left(1 - \frac{\nu}{T}\right) \cdot \cos(\omega\nu) - \frac{1}{\omega T} \cdot (\sin(\omega \cdot (2T - \nu)) - \sin(\omega\nu)) \right) d\nu.$$

Für das Quadrat des Betrages von $N(\omega)$ (was hier auch der Varianz der komplexen Zufallsvariable $\{N(\omega)\}_r$ entspricht) gilt

$$E\{|N(\omega)|^2\}_r = E\{N(\omega) \cdot N^*(\omega)\}_r,$$

was auf

$$E\{N(\omega) \cdot N^*(\omega)\}_r = E\{(\operatorname{Re}(N(\omega)))^2 + (\operatorname{Im}(N(\omega)))^2\}_r = E\{(\operatorname{Re}(N(\omega)))^2\}_r + E\{(\operatorname{Im}(N(\omega)))^2\}_r$$

führt. Da der Erwartungswert von Real- und Imaginärteil jeweils null ist, entspricht dies

$$E\{|N(\omega)|^2\}_r = \operatorname{Var}\{\operatorname{Re}(N(\omega))\}_r + \operatorname{Var}\{\operatorname{Im}(N(\omega))\}_r.$$

Setzt man die oben erhaltenen Ergebnisse für $\operatorname{Var}\{\operatorname{Re}(N(\omega))\}_r$ und $\operatorname{Var}\{\operatorname{Im}(N(\omega))\}_r$ hier ein, so erhält man

$$\begin{aligned} E\{|N(\omega)|^2\}_r &= 2 \cdot T \cdot \int_0^T R_{nn}(\nu) \cdot \left(1 - \frac{\nu}{T}\right) \cdot \cos(\omega\nu) \cdot d\nu \\ &= T \cdot \int_{-T}^T R_{nn}(\nu) \cdot \left(1 - \frac{|\nu|}{T}\right) \cdot \cos(\omega\nu) \cdot d\nu. \end{aligned}$$

Der letzte Schritt dient hier dazu, dass der Wert $\nu = 0$ nicht auf dem Rand des Integrationsgebiets liegt. Um die Auswirkung des Rauschens auf den Schätzfehler weiter präzisieren zu können, müssen Annahmen über das Rauschen bzw. dessen Autokorrelationsfunktion $R_{nn}(\nu)$ getroffen werden.

Wenn $n_y(t)$ weißes Rauschen ist, dann gilt $R_{nn}(\nu) = S_0 \cdot \delta(\nu)$ und man erhält

$$E\{|N(\omega)|^2\}_r = S_0 \cdot T.$$

Dies gilt qualitativ auch einzeln für die Varianzen von Real- und Imaginärteil. Damit verhalten sich die Varianzen proportional zur Messdauer T , und die Standardabweichungen als die Wurzeln der Varianz proportional zu \sqrt{T} . Wenn im Schätzverfahren Werte der Fouriertransformierten durch die Messdauer T geteilt werden, verhält sich die Standardabweichung des Messfehlers dann proportional zu $1/\sqrt{T}$.

Für den Sonderfall, dass die betrachtete Frequenz ω ein ganzes Vielfaches der durch die Messdauer T gegebenen Grundfrequenz $\frac{2\pi}{T}$ ist, sind Real- und Imaginärteil nicht korreliert. Ansonsten besteht eine Korrelation, die mit steigender Messdauer jedoch schnell abnimmt. (Siehe Anhang C.1.)

Bei farbigen Rauschen, dass über

$$S_{nn}(\omega) = \frac{S_0}{1 + \omega^2 T_n^2}$$

bzw.

$$R_{nn}(\tau) = \frac{S_0}{2T_n} \cdot e^{-\frac{|\tau|}{T_n}}$$

beschrieben wird, ergibt sich für $T \rightarrow \infty$ der Ausdruck³

$$E\{|N(\omega)|^2\}_r \approx S_0 \cdot T \cdot 2 \cdot \frac{1}{T_n^2 \omega^2 + 1}.$$

Dieses kann weiter als

$$E\{|N(\omega)|^2\}_r \approx T \cdot 2 \cdot S_{nn}(\omega)$$

³ Hier nicht berechnet. Siehe [Eykhoff 1974, S. 381], wobei die Berechnungen auch dort nicht ausgeführt sind.

geschrieben werden. Auch hier verhält sich die Varianz proportional zur Messdauer.

In Abbildung 6.4 sind die Beträge der Fouriertransformierten von farbigen Rauschen, welches durch Filterung durch ein PT_1 -Glied mit einer Zeitkonstante von $T_n = 1\text{ s}$ erzeugt ist, dargestellt. Dabei ist die Länge des Messfensters variiert. Der Prozess $\{N(\omega)\}_r$ ist zwar weder ergodisch noch stationär, dennoch erhält man anhand der Streuung der Werte über der Frequenz einen guten Eindruck der Streuung der Werte an einem Frequenzpunkt. (Eigentlich müsste man die Fouriertransformierte für viele Realisierungen bilden, und dann für einzelne Frequenzpunkte ein Histogramm plotten. Zudem ist dieses Histogramm mit den Quadraten der Beträge zu bilden, da im Allgemeinen $E\{|N(\omega)|\}_r \neq \sqrt{E\{|N(\omega)|^2\}_r}$ ist.)

Man erkennt, dass eine Vervierfachung der Messdauer von 100 auf 400 s zu einer ungefähren Verdopplung der ermittelten Beträge der Fouriertransformierten führt. Dies deckt sich mit dem Faktor T in $E\{|N(\omega)|^2\}_r$. Zu höheren Frequenzen hin nehmen die Werte ab, was sich auch mit dem angegebenen Zusammenhang deckt.

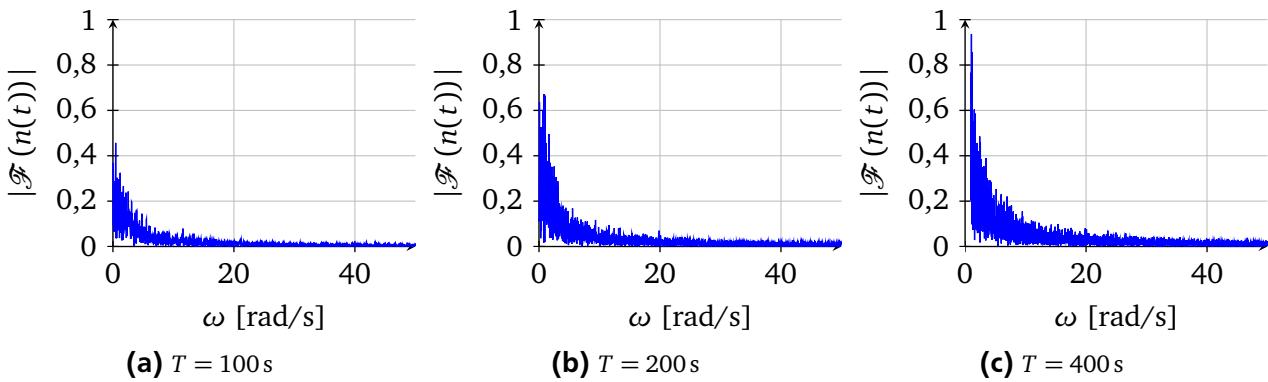


Abbildung 6.4: Fouriertransfomierte einer Realisierung eines farbigen Rauschens (PT_1 , $T_n = 1\text{ s}$)

6.2.2 Konstante Störung

Eine konstante Störung wird durch

$$d(t) = a$$

beschrieben. Könnte man die exakte Fouriertransformation bestimmen, dann würde eine konstante Störung $d(t) = a$ nur den Wert bei $\omega = 0$ beeinflussen,

$$\mathcal{F}(a)|_\omega = 2\pi \cdot a \cdot \delta(\omega).$$

Da in der Praxis jedoch immer nur ein endlicher Zeitraum zur Verfügung steht, wird die Störung immer mit einem Rechteckfenster $w_T(t)$ multipliziert werden. Damit wirkt die konstante Störung wie ein Rechteckimpuls, und es gilt

$$D_T(\omega) = \mathcal{F}(a \cdot w_T(t))|_\omega = a \cdot T \cdot \frac{\sin\left(\frac{\omega T}{2}\right)}{\frac{\omega T}{2}} \cdot e^{-j\frac{\omega T}{2}}.$$

Wird die Fouriertransformation (näherungsweise) über die DFT berechnet und dabei kein Zero-Padding angewendet, dann fallen die Werte, an denen $\mathcal{F}(a \cdot w_T(t)) = 0$ ist mit den Frequenzstützstellen zusammen, die mit der DFT berechnet werden. Die Störung wirkt sich dann (sichtbar) wieder nur auf $l = 0$ aus,

$$\text{DFT}(a \cdot w_T(t_k))_l = N \cdot a \cdot \delta_l = \begin{cases} N \cdot a & l = 0 \\ 0 & \text{sonst.} \end{cases}$$

(Hier ist N die Anzahl der Punkte der DFT.)

Allgemein gilt für den Betrag der Fouriertransformierten der Störung

$$\begin{aligned}|D_T(\omega)| &= a \cdot T \cdot \left| \frac{\sin\left(\frac{\omega T}{2}\right)}{\frac{\omega T}{2}} \cdot e^{-j\frac{\omega T}{2}} \right| \\&= a \cdot T \cdot \frac{\left|\sin\left(\frac{\omega T}{2}\right)\right|}{\left|\frac{\omega T}{2}\right|} \cdot \left|e^{-j\frac{\omega T}{2}}\right| \\&= a \cdot T \cdot \frac{\left|\sin\left(\frac{\omega T}{2}\right)\right|}{\frac{\omega T}{2}},\end{aligned}$$

positive Werte für a und ω vorausgesetzt. Da $|\sin(\cdot)| \leq 1$ kann

$$|D_T(\omega)| \leq a \cdot \frac{2}{\omega}$$

als obere Grenze angegeben werden.

In Abbildung 6.5a sind die Zeitsignale der Störung unter Berücksichtigung von drei Rechteckfenstern der Breite 1 s, 2 s und 4,5 s gezeigt. In Abbildung 6.5b sind dazu die Fouriertransformierten dargestellt. Zusätzlich ist gestrichelt die oben bestimmte obere Grenze der Störungen eingezzeichnet.

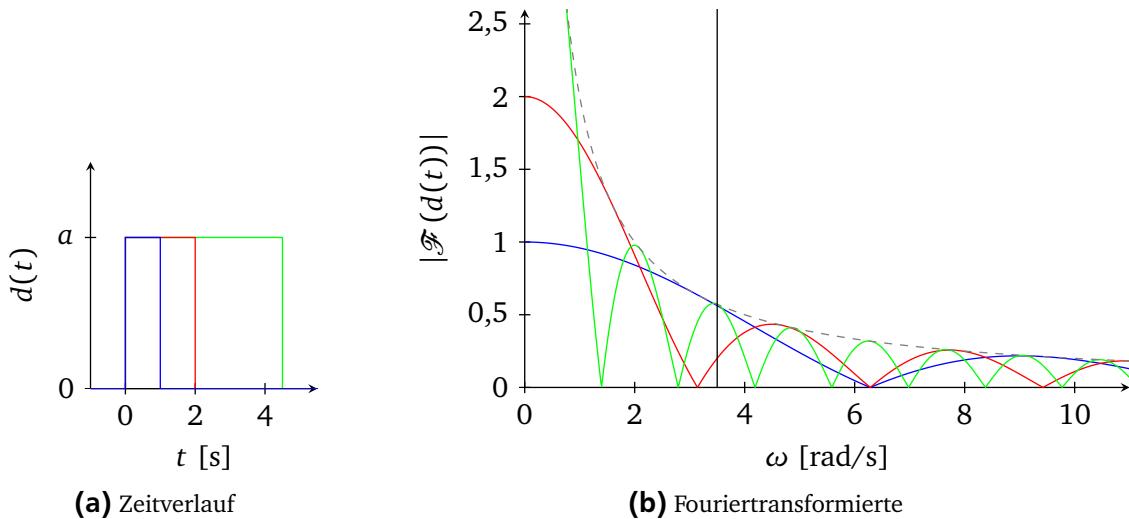


Abbildung 6.5: Konstante Störung bei endlicher Messdauer

Die obere Grenze hängt nicht von der Messdauer bzw. Fensterbreite T ab. Dies bedeutet, dass der Einfluss einer konstanten Störung auf den Schätzwert, für den in der Regel noch durch T geteilt wird, mit steigender Messdauer abnimmt.

Diese Aussage bezieht sich auf die obere Grenze. Tatsächlich kann es sein, dass bei steigernder Messdauer der Einfluss auf einen Frequenzpunkt sogar zunimmt. Dies ist in Abbildung 6.5b gezeigt. Dort ist der Frequenzpunkt $3,5 \frac{\text{rad}}{\text{s}}$ durch einen senkrechten Strich gekennzeichnet. Betrachtet man den Wert der Fouriertransformierten für die verschiedenen Fensterbreiten an diesem Punkt, so sieht man, dass der betreffende Wert der zu der Messdauer von zwei Sekunden gehört (rot) einen weniger als halb so großen Wert annimmt, wie die Fouriertransformierten zur den Fensterbreiten 1 s und vor allem 4,5 s.

Damit lässt sich festhalten, dass bei geeigneter Wahl der Messdauer der Einfluss einer konstanten Störung für bestimmte Messpunkte gezielt vermindert werden kann. Aber auch wenn die Messdauer nicht besonders beachtet wird, so fällt der maximal mögliche Fehler in der Regel mit $1/T$.

6.2.3 Rampenförmige Störung

Eine rampenförmige Störung wird durch

$$d(t) = at \cdot \sigma(t)$$

beschrieben. Dessen Fouriertransformierte bei endlicher Messdauer T ist

$$D(\omega) = \mathcal{F}(d(t) \cdot w_T(t))|_{\omega} = \int_0^T a \cdot t \cdot e^{-j\omega t} dt = a \cdot \left[\frac{1 + j\omega t}{\omega^2} \cdot e^{-j\omega t} \right]_0^T = a \cdot \frac{(1 + j\omega T) \cdot e^{-j\omega T} - 1}{\omega^2}.$$

Beispiele für Zeitverläufe sind in Abbildung 6.6a gezeigt. In Abbildung 6.6b sind die dazugehörigen Fouriertransformierten zu sehen.

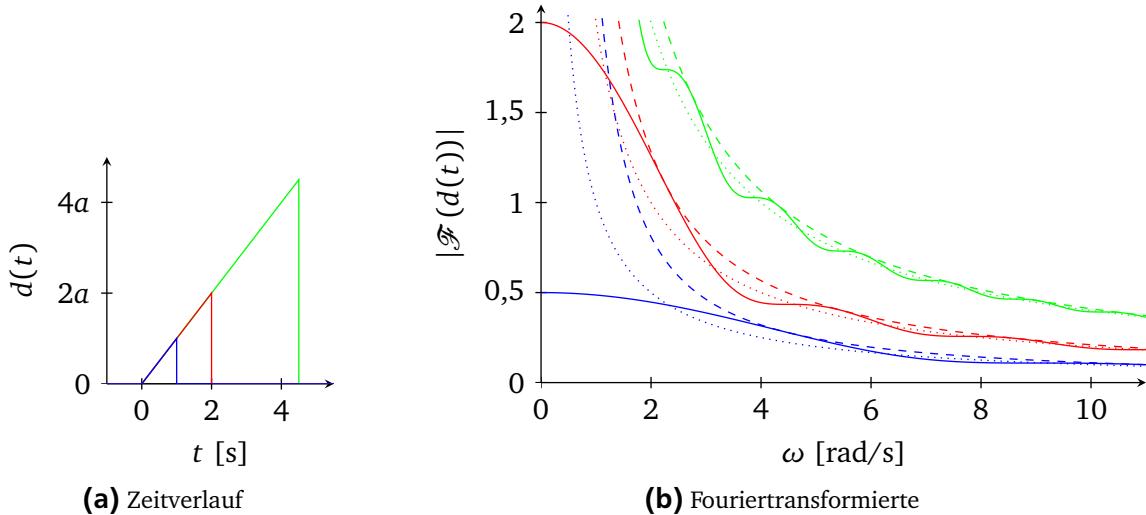


Abbildung 6.6: Rampenförmige Störung bei endlicher Messdauer

Für den Betrag der Fouriertransformierten gilt

$$\begin{aligned} |D(\omega)| &= \left| a \cdot \frac{(1 + j\omega T) \cdot e^{-j\omega T} - 1}{\omega^2} \right| \\ &= |a| \cdot \frac{|(1 + j\omega T) \cdot e^{-j\omega T} - 1|}{|\omega^2|} \\ &\leq a \cdot \frac{|(1 + j\omega T) \cdot e^{-j\omega T}| + 1}{\omega^2} \\ &= a \cdot \frac{|1 + j\omega T| + 1}{\omega^2} \\ &= a \cdot \frac{\sqrt{1 + \omega^2 T^2} + 1}{\omega^2}. \end{aligned}$$

Dabei ist im ersten Schritt von $|ab| = |a| \cdot |b|$ und $|1/a| = 1/|a|$ (a und b sind dabei beliebige komplexe Zahlen) Gebrauch gemacht. Im zweiten Schritt ist im Zähler die Dreiecksungleichung $|a + b| \leq |a| + |b|$ benutzt. Im dritten Schritt ist wieder $|ab| = |a| \cdot |b|$ sowie $|e^{ix}| = 1$ für alle reellen x angewendet.

In Abbildung 6.6b sind diese oberen Grenzen für jede Schrittweite gestrichelt eingezeichnet. Gepunktet ist die im folgenden berechnete Grenze eingetragen, die nur für $\omega \gg \frac{1}{T}$ gilt, jedoch auch einen guten Eindruck davon vermittelt, wie sich die Grenzen mit wachsendem T verhalten.

Für $\omega \gg \frac{1}{T}$ dominiert $\omega^2 T^2$ den Ausdruck unter der Wurzel, so dass $\sqrt{1 + \omega^2 T^2} \approx \omega T$ angenähert werden kann und im daraus entstehenden Zähler gilt wieder $\omega T + 1 \approx \omega T$, und damit

$$|D(\omega)| \leq a \cdot \frac{T}{\omega} \quad \text{für } \omega \gg \frac{1}{T}.$$

Wird eine Frequenz ω' betrachtet, die einem ganzen Vielfachen der durch die Fensterbreite T gegebenen Grundfrequenz entspricht,

$$\omega' = \frac{2\pi n}{T},$$

so gilt

$$\begin{aligned} D(\omega') &= a \cdot \frac{(1 + j \cdot \frac{2\pi n}{T} \cdot T) \cdot e^{-j \cdot \frac{2\pi n}{T} \cdot T} - 1}{\left(\frac{2\pi n}{T}\right)^2} \\ &= j \cdot a \cdot \frac{2\pi n}{\left(\frac{2\pi n}{T}\right)^2} = j \cdot a \cdot \frac{1}{\frac{2\pi n}{T} \cdot \frac{1}{T}} \\ &= j \cdot a \cdot \frac{T}{\omega'}. \end{aligned}$$

D. h. in diesem Sonderfall gilt

$$|D(\omega)| = a \cdot \frac{T}{\omega'}$$

sogar exakt.

Der Einfluss der Störungen nimmt linear mit T zu. D. h. durch ein späteres Teilen der Fouriertransformierten durch die Messdauer T ergibt sich ein konstanter Einfluss. Der Einfluss auf die geschätzten Werte fällt mit wachsender Messdauer nicht ab.

6.2.4 Sinusförmige Störung

Als letzte Störung wird eine sinusförmige Schwingung

$$d(t) = a \cdot \sin(\omega_d \cdot t + \varphi)$$

betrachtet. Aus der Amplitude a und der Phase φ lässt sich die halbe komplexe Amplitude \check{d} angeben. Zum Beispiel beträgt für $d(t) = a \cdot \sin(\omega_d \cdot t)$ die halbe komplexe Amplitude $\check{d} = -j \cdot \frac{a}{2}$.

Damit kann die Fouriertransformierte über die δ -Notation angegeben werden, was bei Berücksichtigung der Fensterung auf

$$\begin{aligned} D(\omega) &= \mathcal{F}(d(t) \cdot w_T(t))|_{\omega} \\ &= \check{d}^* \cdot T \cdot \underbrace{\frac{\sin\left(\frac{(\omega+\omega_d) \cdot T}{2}\right)}{\frac{(\omega+\omega_d) \cdot T}{2}} \cdot e^{-j \frac{(\omega+\omega_d) \cdot T}{2}}}_{D_l(\omega)} + \check{d} \cdot T \cdot \underbrace{\frac{\sin\left(\frac{(\omega-\omega_d) \cdot T}{2}\right)}{\frac{(\omega-\omega_d) \cdot T}{2}} \cdot e^{-j \frac{(\omega-\omega_d) \cdot T}{2}}}_{D_r(\omega)} \end{aligned}$$

führt. Nach der Dreiecksungleichung gilt

$$|D(\omega)| \leq |D_l(\omega)| + |D_r(\omega)|.$$

Für den Einfluss des linken Summanden gilt

$$|D_l(\omega)| = \left| \check{d}^* \cdot T \cdot \frac{\sin\left(\frac{(\omega+\omega_d) \cdot T}{2}\right)}{\frac{(\omega+\omega_d) \cdot T}{2}} \cdot e^{-j\frac{(\omega+\omega_d) \cdot T}{2}} \right| \\ = |\check{d}^*| \cdot T \cdot \frac{\left| \sin\left(\frac{(\omega+\omega_d) \cdot T}{2}\right) \right|}{\frac{|\omega+\omega_d| \cdot T}{2}} = 2 \cdot |\check{d}^*| \cdot \frac{\left| \sin\left(\frac{(\omega+\omega_d) \cdot T}{2}\right) \right|}{|\omega+\omega_d|},$$

woraus die obere Grenze

$$|D_l(\omega)| \leq 2 \cdot |\check{d}^*| \cdot \frac{1}{|\omega+\omega_d|}$$

folgt. Analog kann

$$|D_r(\omega)| \leq 2 \cdot |\check{d}| \cdot \frac{1}{|\omega-\omega_d|}$$

angegeben werden. Zusammen ergibt sich letztlich

$$|D(\omega)| \leq |D_l(\omega)| + |D_r(\omega)| \leq 2 \cdot |\check{d}| \cdot \left(\frac{1}{|\omega+\omega_d|} + \frac{1}{|\omega-\omega_d|} \right).$$

In Abbildung 6.7a sind wie für den Fall der konstanten und der rampenförmigen Störung die Zeitverläufe der sinusförmigen Störung unter Berücksichtigung verschiedener Messdauern gezeigt. In Abbildung 6.7b sind dazu die Fouriertransformierten dargestellt. Gestrichelt ist wieder die gerade bestimmte obere Grenze gezeigt.

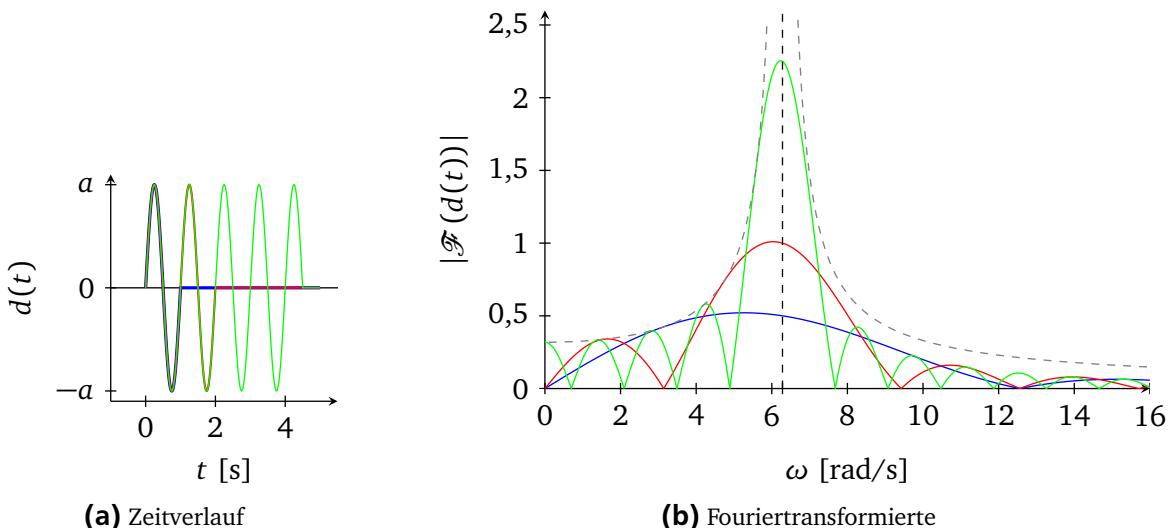


Abbildung 6.7: Sinusförmige Störung bei endlicher Messdauer

Der Einfluss der Störung nimmt umso mehr ab, desto weiter die betrachtete Frequenz von der Störfrequenz entfernt ist. Dieser Zusammenhang ist näherungsweise antiproportional. Wird eine ganzzahlige Anzahl an Perioden der Störung durch die Messung erfasst, dann fallen die Nullstellen der Fouriertransformierten der Störung mit Vielfachen der Grundfrequenz zusammen.

Wird die Fouriertransformierte zu einer bestimmten Frequenz ω' betrachtet und ist die Messdauer als ganzes Vielfaches von $\frac{2\pi}{\omega'}$ gewählt, dann haben alle die periodischen Störungen keinen Einfluss, deren Grundfrequenz ein ganzes Vielfaches von ω' ist. (Bzw. kann auch die betrachtete Frequenz ω' ein ganzes Vielfaches der Störfrequenz sein. In diesem Fall muss dann noch die Messdauer ein ganzes Vielfaches von $\frac{2\pi}{\omega_d}$ sein.) Auf die anderen Frequenzpunkte wirkt sich die Störung aus. Dabei zeigt die obere Grenze von $|D(\omega)|$ keine Abhängigkeit von der Messdauer T , so dass der Einfluss auf die Schätzwerte mit wachsender Messdauer abnehmen wird. Dies gilt mit Ausnahme der Frequenz der Störung selber. (An bzw. Nahe diesem Punkt ist die Abschätzung durch die oben angegebene obere Grenze zu konservativ, was auch in Abbildung 6.7b deutlich wird.)

6.3 Anregungssignale

Anregungssignale können bezüglich verschiedenster Kriterien bewertet oder ausgewählt werden. Im Wesentlichen geht es dabei aber darum, bei gegebenen Beschränkungen bezüglich der Eingangsgröße eine möglichst große Signalleistung in das zu identifizierende System einzubringen. Dabei soll diese Leistung möglichst auf einen vorgebbaren Frequenzbereich konzentriert sein.

Für die folgenden Betrachtungen wird die Eingangsgröße auf das Intervall $[-u_{\max}, u_{\max}]$ begrenzt. Die Signalparameter sind so gewählt, dass dieses Intervall möglichst gut ausgenutzt wird. Weitergehende Kriterien, wie z. B. eine maximale Änderung der Stellgröße, maximale Beschleunigung oder maximaler Ruck (Ableitung der Beschleunigung) werden hier nicht berücksichtigt. Ebenso wird nicht auf Grenzen, die möglicherweise durch das System gegeben sind (Beschränkung der Systemzustände) eingegangen, da hierzu ohnehin das System bekannt sein muss.

Für weitergehende Betrachtungen wird auf das vierte Kapitel in [Pintelon und Schoukens 2001] verwiesen, welchem auch vieles, was in der Diskussion zu den breitbandigen Anregungssignalen besprochen wird, entlehnt ist.

6.3.1 Signale zur Anregung einer Frequenz

In Abbildung 6.8a sind mit einer Sinus-, einer Rechteck- und einer Dreieckschwingung einfache periodische Signale gezeigt. Deren Fourierreihen sind in Abbildung 6.8 dargestellt.

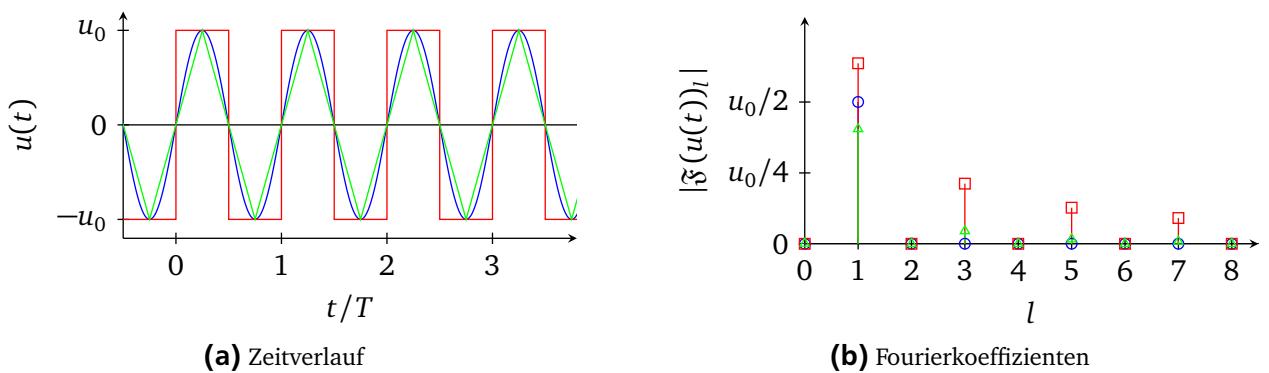


Abbildung 6.8: Periodische Anregungssignale

Nur die Sinusschwingung regt ausschließlich die Grundfrequenz an. Die anderen beiden Signale regen alle ungeraden Harmonischen an, wobei die entsprechenden Amplituden schnell (Rechteckimpuls) oder sehr schnell (Dreieckimpuls) abnehmen. Zur gezielten Anregung mehrerer Frequenzen eignet sich aber auch die Rechteckschwingung nicht, da die Harmonischen relativ schwach sind.

Die Motivation zur Verwendung einer Rechteck- oder Dreieckschwingung (oder auch einer Trapezschwingung) kann darin liegen, dass das betrachtete System oder die Messsoftware über keine Schnittstelle oder Möglichkeit verfügt, einen sauberen Sinus an den Systemeingang anzulegen. Zudem wird mit der Rechteckschwingung oder mit Trapezschwingungen mit steiler Flanke eine Überhöhung der Grundschwingung im Vergleich zur Sinusschwingung erreicht. Dies ist auch in Abbildung 6.8b zu sehen. Dabei ist aber auf der anderen Seite zu bedenken, dass mit den steilen Flanken und damit einhergehenden schnellen Änderungen der Systemzustände eher die Möglichkeit besteht, dass der lineare Bereich verlassen wird.

6.3.2 Impulsförmige Anregungssignale

Impulsförmige Anregungssignale regen nicht einzelne Frequenzen, sondern gewisse Frequenzbereiche an. Dabei ist der anzuregende Bereich jedoch nicht sehr frei vorzugeben, wie die folgenden Beispiele zeigen werden. Dafür sind impulsförmige Anregungssignale sehr einfach auf ein System aufzubringen. So können diese auch durchaus per Hand gesteuert werden.

Hat man die Möglichkeit über ein geeignetes Messsystem beliebige Signale aufzuschalten, dann sind die im folgenden Abschnitt vorgestellten breitbandigen Anregungssignale, zumindest für die Identifikation im Frequenzbereich, klar zu bevorzugen.

Rechteckimpuls

In Abbildung 6.9a sind Rechteckimpulse verschiedener Länge sowie eine Sprunganregung gezeigt. Die dazugehörigen Amplitudenspektren sind in Abbildung 6.9b zu sehen.

Es zeigt sich, dass die Anregung bei niedrigen Frequenzen am besten ist. Die erste Nullstelle der Amplitudendichte liegt bei $\frac{2\pi}{T'}$, wobei T' die Impulsbreite ist. D. h. für $T' = T$ (blau) liegt die erste Nullstelle bei $\frac{2\pi}{T}$, für $T' = 2T$ (rot) bei $0,5 \cdot \frac{2\pi}{T}$ und für $T' = 4T$ (grün) bei $0,25 \cdot \frac{2\pi}{T}$. Je breiter der Impuls, desto schmäler wird also der Bereich, indem eine nennenswerte Anregung stattfindet. Dafür steigt bei einem breiteren Impuls der Betrag der Amplitudendichte in diesem Frequenzbereich an, so dass in einen schmalen Frequenzbereich eine größere Signalleistung eingebracht wird.

Als Vergleich ist auch die Fouriertransformierte der Sprunganregung gezeigt. Diese besitzt keine Nullstelle, fällt aber von $\omega = 0$ aus dem unendlichen kommend sehr schnell ab, so dass nur sehr tiefe Frequenzen (theoretisch) angeregt werden. Bei diesem Signal wäre aber auch schon die Bestimmung der Fouriertransformierten aus den Messdaten aufwendiger, da im Allgemeinen weder ein zeitbegrenztes, noch ein periodisches Signal vorliegt.

Trapez- und Dreieckimpuls

In Abbildung 6.9c ist der Zeitverlauf des Rechteckimpulses mit dem eines Trapezimpuls und dem eines Dreieckimpuls gezeigt. In Abbildung 6.9d sind die dazugehörigen Fouriertransformierten verglichen. Es zeigt sich, dass je flacher die Flanken sind, desto weiter die erste Nullstelle der Amplitudendichte zu hohen Frequenzen hin wandert. Die Höhe der Anregung nimmt dabei ab. Dies ist auch von daher einsichtig, dass die Signalenergie, d. h. $\int u^2(t)dt$ abnimmt.

Mittelwertfreier Rechteckimpuls

In Abbildung 6.9e ist, neben dem Rechteckimpuls, ein mittelwertfreier Impuls gezeigt. Dessen Fouriertransformierte hat entsprechend für $\omega = 0$ den Wert null. (Abbildung 6.9f) Der Bereich der maximalen Anregung ist bei gleicher (Gesamt)impulslänge etwas zu höheren Frequenzen hin verschoben.

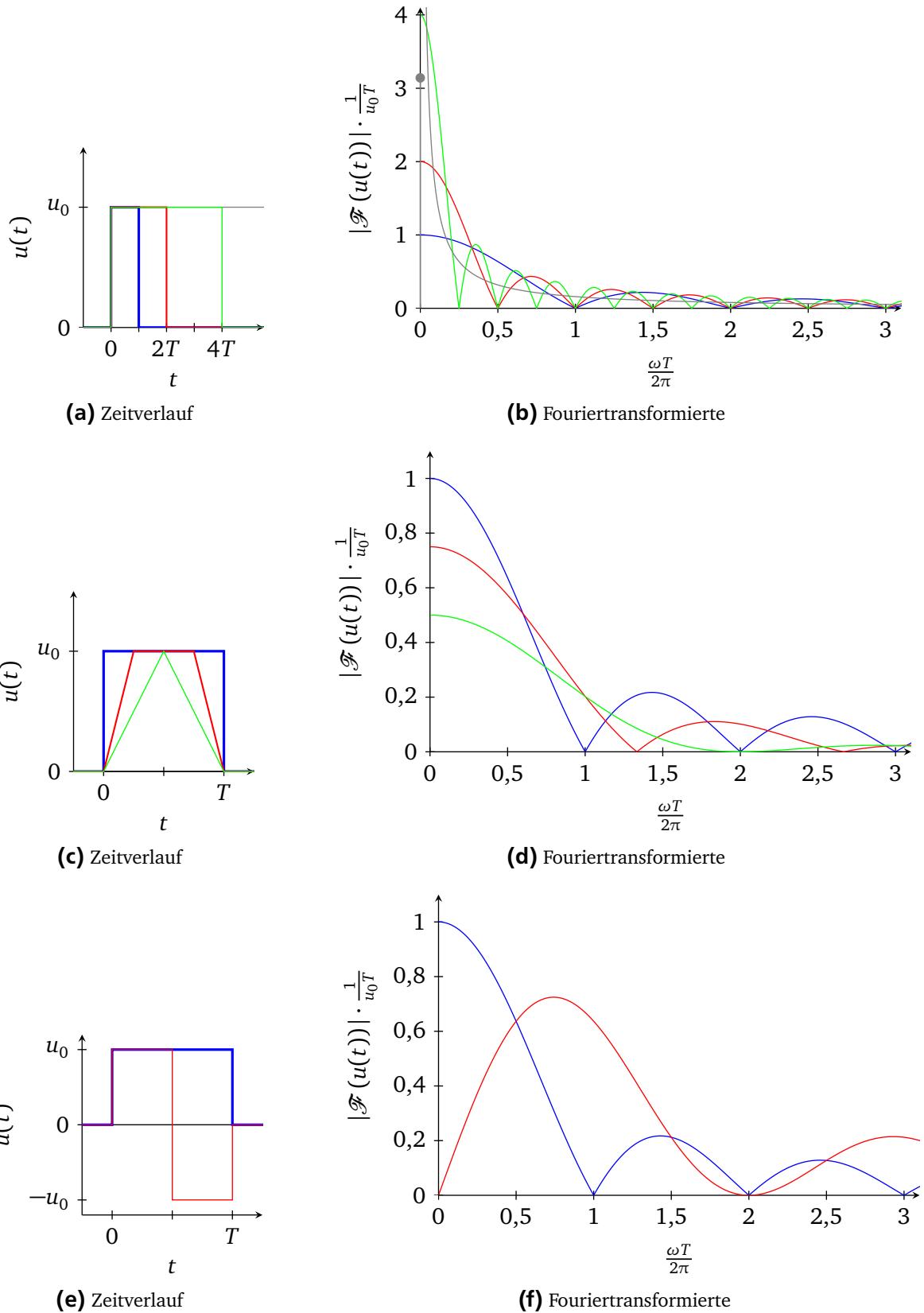


Abbildung 6.9: Impulsförmige Anregungen

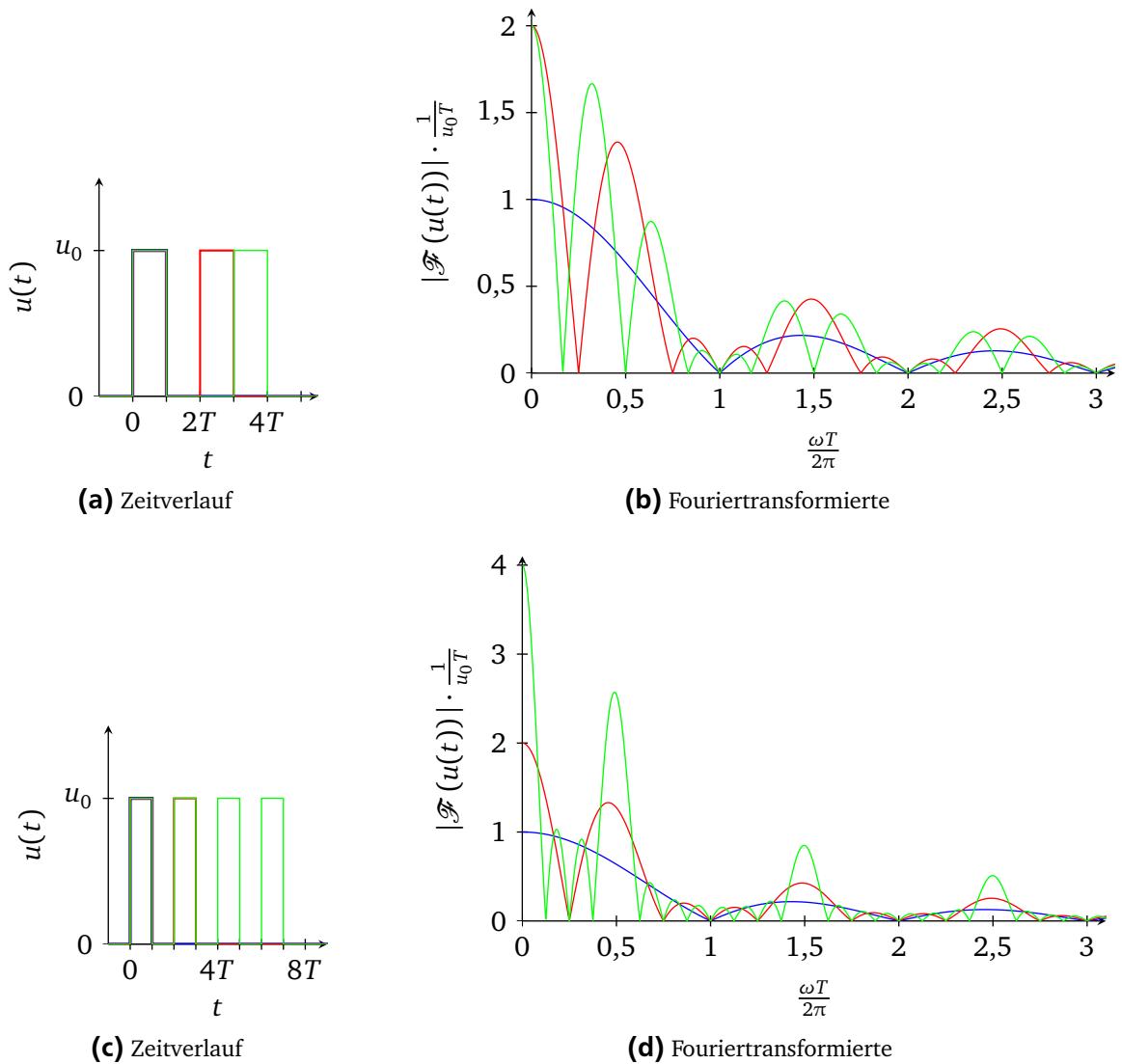


Abbildung 6.10: Impulsförmige Anregungen

Mehrere Impulse

Zuletzt sind in Abbildung 6.10 noch Beispiele gezeigt, in denen eine Folge mehrerer Impulse verwendet wird. In Abbildung 6.10a sind zwei doppelte Rechteckimpulse gezeigt, die sich in dem Abstand zueinander unterscheiden. Im einen Fall (rot) beträgt der Abstand genau eine Impulslänge (T), im zweiten Fall (grün) $2T$. In blau ist wieder der normale Rechteckimpuls zum Vergleich dargestellt.

Mit dem Abstand verändert sich auch das Spektrum, wie in Abbildung 6.10b zu sehen ist. Prinzipiell werden in beiden Fällen zweimal das Spektrum eines einzelnen Impulses addiert, wobei eines der Spektren über den Verschiebungssatz der Fouriertransformierten,

$$\mathcal{F}(y(t - \tau))|_{\omega} = e^{-j\tau\omega} \cdot Y(\omega),$$

mit $e^{-j2T\omega}$ bzw. $e^{-jT\omega}$ multipliziert ist. Dies führt an verschiedenen Stellen zu Verstärkungen und Auslöschen mit dem Ursprungsspektrum, die in den Amplitudengängen erkennbar sind. In jedem Fall entstehen neue Nullstellen, die näher an $\omega = 0$ liegen als bei einem einzelnen Impuls.

In Abbildung 6.10c sind Folgen mehrerer Rechteckimpulse gezeigt, nämlich mit zwei (rot) und vier (grün) einzelnen Impulsen. Auch hier entstehen im Spektrum durch die Überlagerung der Einzelspektren wieder Bereiche der Verstärkung, in denen die Amplitudendichte größer als die des einzelnen Rechteckimpulses ist, und Bereiche der Auslösung, wo diese kleiner ist und neue Nullstellen entstanden sind. (Abbildung 6.10d) Die Amplitudendichte kann damit in engen Bereichen gezielt vergrößert werden.

6.3.3 Breitbandige Anregungssignale

Die hier vorgestellten breitbandigen Anregungssignale, insbesondere Multisinus und PRBS, können entweder periodisch oder als zeitbegrenzte Signale verwendet werden. Werden diese als periodische Signale verwendet, dann werden die hier gezeigten Verläufe entsprechend wiederholt abgespielt und auf das System gegeben. Es wird dann gewartet, bis die transienten Anteile der Systemantwort abgeklungen sind, und auch die Systemantwort damit einen periodischen Verlauf zeigt. Dann werden über eine oder eine ganze Anzahl an Perioden die Daten aufgezeichnet und verarbeitet. In diesem Fall sind die Fourierkoeffizienten wesentlich um das Eingangssignal zu beschreiben.

Im Fall der zeitbegrenzten Anregung wird das Signal nur einmal auf das System gegeben. Hier ist die Fouriertransformation zur Beschreibung relevant. (Wählt man dieselben Frequenzstützstellen wie die, die sich im periodischen Fall zwangsweise ergeben, so sind die entsprechenden Werte bis auf die Skalierung gleich.)

Gleitsinus

Betrachtet man eine sinusförmige Schwingung mit der Frequenz ω und der Anfangsphase φ_0 ,

$$\sin(\omega t + \varphi_0),$$

dann kann man sich diese auch als den Sinus eines veränderlichen Winkels oder einer veränderlichen Phase $\varphi(t)$ auffassen,

$$\sin(\varphi(t)),$$

der sich aus der Integration der Frequenz über die Zeit ergibt,

$$\varphi(t) = \int_0^t \omega \cdot d\tau + \varphi_0.$$

Die Frequenz ω entspricht in dieser Interpretation der Winkelgeschwindigkeit des Phasenwinkels.

In dieser Betrachtungsweise kann die „Frequenz“ auch zeitvariant sein,

$$\varphi(t) = \int_0^t \omega(\tau) \cdot d\tau + \varphi_0 .$$

Man spricht teilweise von der Momentanfrequenz, wobei im Allgemeinen natürlich keine periodische Funktion vorliegt.

Setzt man

$$\omega(t) = \frac{\omega_b - \omega_a}{T_{ab}} \cdot t + \omega_a ,$$

so ergibt sich

$$\varphi(t) = \frac{\omega_b - \omega_a}{2T_{ab}} \cdot t^2 + \omega_a \cdot t + \varphi_0 .$$

Setzt man diesen Winkelverlauf in die Sinusfunktion ein, so erhält man einen (linearen) Gleitsinus

$$u(t) = a \cdot \sin \left(\left(\frac{\omega_b - \omega_a}{2T_{ab}} \cdot t + \omega_a \right) \cdot t + \varphi_a \right) .$$

Es gibt auch Varianten mit logarithmischer oder quadratischer Änderung der Momentanfrequenz, aber für die Systemidentifikation besitzt der lineare Gleitsinus gute Eigenschaften.

In Abbildung 6.11a ist der Zeitverlauf eines Gleitsinus, dessen Momentanfrequenz von 1 Hz bis 5 Hz innerhalb von $T_{ab} = 10$ s zunimmt dargestellt. (Der Unterschied beider Frequenzen ist hier relativ gering, eignet sich aber gut für die grafische Darstellung des Verlaufs.)

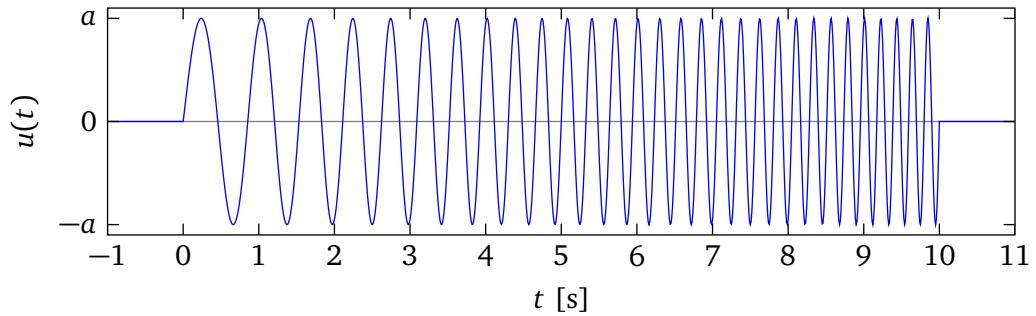
Um den Gleitsinus „stoßfrei“ zu beginnen, d. h. um einen Eingangsgrößensprung zu vermeiden, muss $\varphi_a = n \cdot \pi$ gelten. Das einfachste ist daher $\varphi_a = 0$ zu setzen.⁴ Zum stoßfreien Beenden muss $(\frac{1}{2} \cdot \omega_b + \frac{1}{2} \cdot \omega_a) \cdot T_{ab} + \varphi_a = n \cdot \pi$ gelten. Daraus folgt

$$T_{ab} = n \cdot \frac{2\pi}{\omega_a + \omega_b} \quad \text{für } \varphi_a = 0 .$$

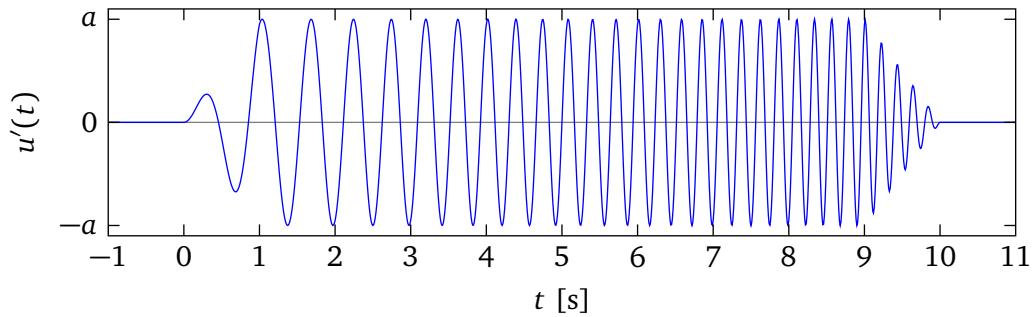
Ist auch ohne diese Überprüfung sicherzustellen, dass der Gleitsinus stetig beginnt und endet, bzw. soll auch Stetigkeit in der ersten Ableitung erreicht werden, um das zu identifizierende System zu schonen, kann der Gleitsinus auch ein- und ausgeblendet werden. Für den in Abbildung 6.11b gezeigten Gleitsinus ist dies erreicht, indem zu Beginn und Ende mit einer über eine Sekunde ansteigenden bzw. abfallenden Rampe multipliziert ist. Ist Stetigkeit in noch höherer Ableitungen gewünscht, so kann das Überblenden mit Hilfe eines Polynoms entsprechend höherer Ordnung realisiert werden.

In Abbildung 6.12 sind die Beträge der Fouriertransformationen verschiedener Gleitsinussignale dargestellt. In Abbildung 6.12a sind die Spektren zu den in Abbildung 6.11 gezeigten Zeitverläufen dargestellt. In Blau ist das Spektrum des Standardgleitsinus gezeichnet. Man erkennt, dass dieser ungefähr im Intervall $[\omega_a, \omega_b]$ sehr hohe Werte annimmt, und außerhalb sehr niedrige Werte. Mit diesen beiden Parametern lässt sich also gut der Frequenzbereich festlegen, indem eine Anregung erfolgen soll. (Auch wenn das Intervall $[\omega_a, \omega_b]$ etwas breiter gewählt sein sollte als der Frequenzbereich, indem eine gute

⁴ Bei Matlab kann ein Gleitsinus über die Funktion chirp bzw. den Block Chirp signal erzeugt werden. Dabei ist darauf zu achten, dass die Funktion chirp (nicht aber der Simulink-Block) von einem Kosinus ausgeht. Soll der Gleitsinus zum Zeitpunkt $t = 0$ bei null beginnen, muss daher die Phase noch um 90° verschoben werden.



(a) Standardform



(b) Ein- und ausgeblendeter Gleitsinus

Abbildung 6.11: Gleitsinus ($\omega_a = 2\pi \cdot 1 \frac{1}{s}$, $\omega_b = 2\pi \cdot 5 \frac{1}{s}$, $T_{ab} = 10 s$)

Anregung erforderlich ist.) In Rot ist das Spektrum des ein- und ausgeblendeten Gleitsinus dargestellt. Es zeigt sich, dass der Frequenzverlauf deutlich ruhiger ist, da viel von der Welligkeit von den in der ersten Ableitung unstetigen Übergängen zu Beginn und Ende des Gleitsinus erzeugt wird. Der Frequenzbereich indem eine gute Anregung erfolgt ist etwas schmäler geworden, insbesondere bei den niedrigen Frequenzen. Dies könnte durch die Verminderung von ω_a wieder ausgeglichen werden. Das Ein- und Ausblenden hat demnach keine negative Auswirkung auf das Spektrum des Gleitsinus.

Für das Spektrum in Abbildung 6.12b ist die Dauer T_{ab} auf 100 s erhöht, die Frequenzgrenzen aber unverändert gelassen. Die Frequenz der überlagerten „Schwingungen“ im Frequenzbereich nimmt dabei deutlich ab. Die Verminderung des gut angeregten Frequenzintervalls durch das Ein- und Ausblenden ist vernachlässigbar. Der Wert des Betrages der Fouriertransformierten innerhalb des gut angeregten Intervalls ist um etwa den Faktor $\sqrt{10} = 3,2$ erhöht. Dies ist von daher auch sinnvoll, da das Quadrat der Fouriertransformierten für die „Signalenergie“ steht, und sich diese mit der Dauer des Gleitsinus verzehnfacht hat.

Zuletzt ist die obere Frequenz ω_b von $2\pi \cdot 5 \frac{\text{rad}}{\text{s}}$ auf $2\pi \cdot 50 \frac{\text{rad}}{\text{s}}$ erhöht, wobei T_{ab} weiterhin 100 s beträgt. Abbildung 6.12c zeigt, dass das gut angeregte Frequenzintervall entsprechend verbreitert ist, wobei der Betrag wieder ungefähr auf den Wert aus Abbildung 6.12a entspricht. Zwar wird im Vergleich zu dem in dieser Abbildung gezeigten Fall zehnmal mehr Energie in das Signal gesteckt, allerdings muss sich diese Energie auch über einen ungefähr zehnmal so breiten Frequenzbereich verteilen.

Die Spektren zeigen auch, dass Aliasing kein Problem darstellt. Es muss zwar schneller als $2\omega_b$ abgetastet werden, jedoch nicht viel.

Schroeder-Multisinus

Theoretisch kann man beliebige Kombinationen aus L Sinusschwingungen mit den Frequenzen f_l und den Amplituden A_l auf ein System geben, und sich so ein beliebig fein aufgelöstes Amplitudenspek-

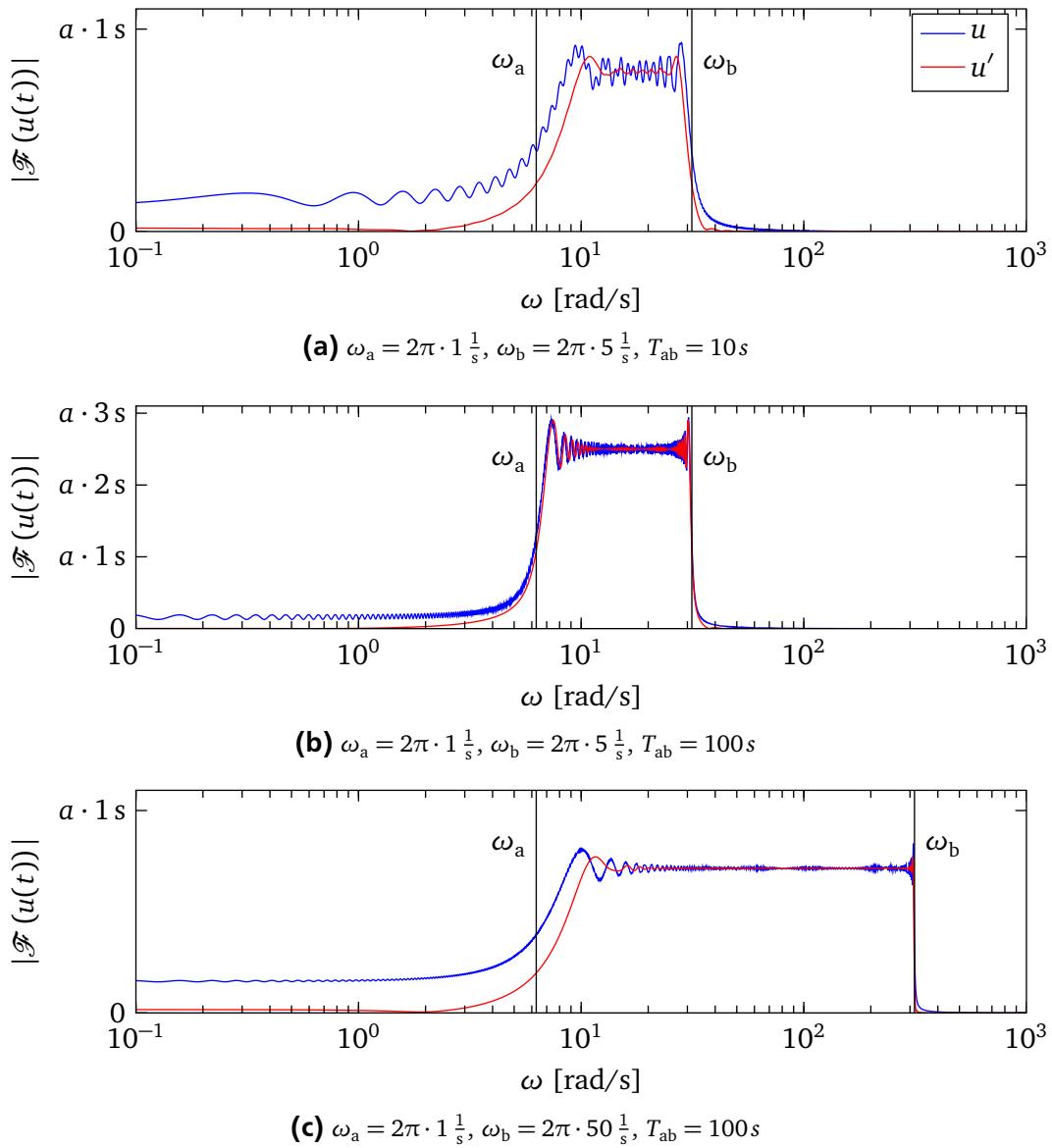


Abbildung 6.12: Fouriertransformierte verschiedener Gleitsinussignale

trum erzeugen. Praktisch steigt die Amplitude des so erzeugten Signals aber schnell an. Würde man die Schwingungen nach

$$u(t) = \sum_{l=1}^L A_l \cdot \cos(2\pi f_l t)$$

einfach addieren, so wäre der maximale Eingangswert $u_{\max} = \sum A_l$. Sollen $L = 100$ Frequenzen mit der gleichen Amplitude $A_l = A$ angeregt werden, so stünde für jede dieser Frequenzen nur ein Hundertstel des Stellbereiches zur Verfügung.

Eine deutliche Verbesserung kann erreicht werden, wenn die Phasen der L einzelnen Schwingungen geeignet zueinander verschoben werden. Dazu wird das Signal in der Form

$$u(t) = \sum_{l=1}^L A_l \cdot \cos(2\pi f_l t + \phi_l)$$

geschrieben, wobei man sich bei den Frequenzen f_l noch auf Vielfache einer Grundfrequenz f_0 , also $f_l = m_l f_0$, $m_l \in \mathbb{N}$, beschränkt. Die Phasen ϕ_l können dann beispielsweise die Schroeder-Phasen

$$\phi_l = \frac{-l \cdot (l-1) \cdot \pi}{L}$$

verwendet werden. Diese berücksichtigen zwar noch nicht die im einzelnen gewählten Amplituden und Frequenzen und sind daher im Allgemeinen nicht optimal in dem Sinne, dass der Maximalwert von $|u(t)|$ für gegebene f_l und A_l minimal ist, aber führen häufig zu guten Ergebnissen. In [Pintelon und Schoukens 2001] wird auch noch auf Verfahren zur weiteren Optimierung eingegangen.

Da f_l ein ganzes Vielfaches der Grundfrequenz f_0 ist, ist der Multisinus periodisch mit f_0 .

In Abbildung 6.13a ist der Zeitverlauf eines Multisinus mit den Parametern $f_0 = 1$, $f_l = l \cdot f_0$ und konstanten Amplituden $A_l = A = 1$ der Einzelschwingungen für $L = 100$ angeregte Frequenzen gezeigt. Der Maximalwert für $|u(t)|$ liegt bei einem Wert etwas unter 12. Dies ist deutlich weniger als 100, was sich als Maximalwert ergeben würde, wenn die Phasen alle Null gewählt würden.

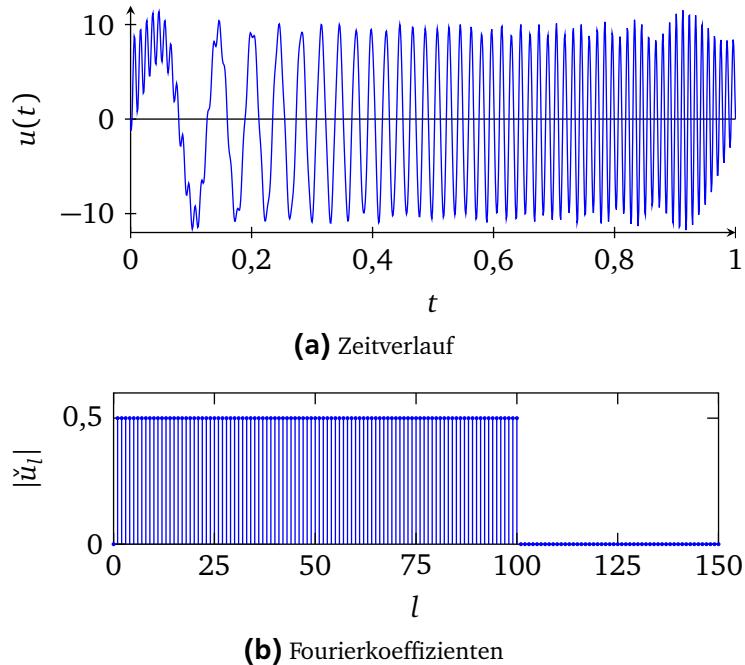


Abbildung 6.13: Schroeder-Multisinus

Die Fourierkoeffizienten des periodisch fortgesetzten Signals sind in Abbildung 6.13b dargestellt. Es zweigt sich der erwartete glatte Verlauf. Da die Amplituden alle zu eins gesetzt sind, ist der Betrag der halben komplexen Amplitude $|\check{u}_l|$ für alle Frequenzen die im Signal enthalten sind genau 0,5.

Anders als bei Gleitsinus ist hier nicht (einfach) sichergestellt, dass das Signal bei dem Wert null startet und endet. Da es aber periodisch ist, kann man – wenn es für die Anwendung wichtig ist, keine Stöße zu erzeugen – einfach bei einem Nulldurchgang beginnen, das Signal auf das System zu geben.

Soll der Multisinus nicht als periodisches Signal, sondern als zeitbegrenztes Signal verwendet werden, so muss anstelle der Fourierkoeffizienten die Fouriertransformation betrachtet werden. Für einen Frequenzausschnitt ist die Fouriertransformierte in Abbildung 6.14 gezeigt. Zusätzlich sind in dieser Abbildung noch die Werte der Fourierkoeffizienten dargestellt, die in den gezeigten Frequenzbereich fallen. An den Frequenzen, zu denen die Fourierkoeffizienten angegeben sind, fallen diese genau mit den entsprechenden Werten der Fouriertransformierten zusammen. (Dass dies hier ohne weitere Skalierung der Fall ist, liegt daran, dass die Signallänge genau 1 beträgt. Unabhängig von einer Skalierung ist wesentlich,

dass alle betreffenden Werte der Fouriertransformierten im gleichen Verhältnis wie die Fourierkoeffizienten stehen. In diesem Fall sind diese also alle gleich.)

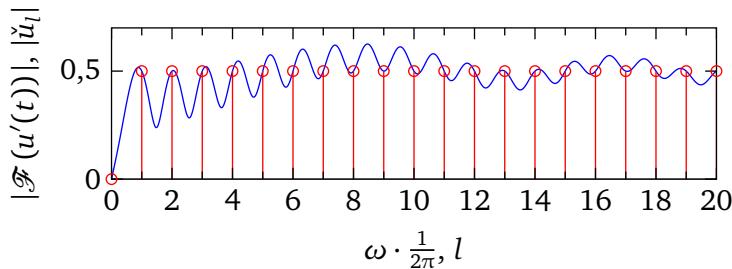


Abbildung 6.14: Schroeder-Multisinus

Zwischen diesen Stützstellen ist der Verlauf der Fouriertransformierten jedoch sehr wellig. Daher sollte, auch wenn der Multisinus als zeitbegrenztes Signal verwendet wird, sich auf die Auswertung der Frequenzstützstellen beschränkt werden, die ein ganzes Vielfaches der Grundfrequenz des (fortgesetzten) Multisinus sind. (Bzw. auf die Frequenzen, die auch Element der Frequenzen f_l sind, für die explizit eine Anregung im Multisinus vorhanden ist.)

Wird der Multisinus periodisch genutzt, so tritt kein Aliasing auf. Wird er als zeitbegrenztes Anregungs-signal verwendet, dann werden durch die (möglichen) Sprünge am Anfang und Ende auch höhere Frequenzen angeregt, aber bei einer vernünftigen Wahl der Abtastzeit bezogen auf die maximale Frequenz des Multisinus ist dies vernachlässigbar.

Pseudo-Rauschsignale

Die in Abschnitt 6.2.1 vorgestellten Rauschsignale können auch als Eingangssignale verwendet werden. Mit Digitalrechner werden dazu sogenannte Pseudo-Zufallszahlen über deterministische Zufallszahlen-generatoren ermittelt. Bei der Verwendung von modernen Programm paketen sollte auch davon ausgegangen werden können, dass die Qualität der so erzeugten „Zufallszahlen“ gut ist, was die Nachbildung der Eigenschaften wie Verteilung, Mittelwert und vor allem die Korrelation bzw. bei weißem Rauschen eher die Nicht-Korreliertheit angeht. (Davon kann man sich aber auch im Einzelfall durch die Analyse der erzeugten Zahlenfolgen versichern.)

Die Motivation dafür, Rauschen als Eingangsgröße zu nehmen, ist einmal dadurch gegeben, dass manche Identifikationsverfahren stationäre oder ergodische Ein- und Ausgangsgrößen benötigen. Dies sind Ver-fahren, die auf die Zusammenhänge wie $R_{uy}(\tau) = \int_{-\infty}^{\infty} g(\nu) \cdot R_{uu}(\tau - \nu) \cdot d\nu$ bzw. $S_{uy}(\omega) = H(\omega) \cdot S_{uu}(\omega)$ beruhen, die nur für stationäre Prozesse gelten. Zum anderen wurde in Abschnitt 6.2.1 gezeigt, dass weißes Rauschen und auch farbiges Rauschen in der Regel breite Frequenzbereiche anregen.

Jedoch hat sich auch gezeigt, dass die Leistungsdichte desto mehr annimmt, desto breiter der anzuregen-de Frequenzbereich ist. Der angeregte Frequenzbereich ergibt sich zum einen durch die Abtastzeit des Rauschens und zum anderen kann über einen Filter Einfluss auf das Leistungsdichtespektrum genom-men werden. (Dabei wird durch den Filter erreicht, dass die Amplitude des Rauschens vor dem Filter angehoben werden kann.)

Die Taktzeit des T_{ns} Rauschens muss nicht gleich der Abtastzeit T_s der Messung gewählt werden. Um die Signalleistung auf niedrigere Frequenzen zu konzentrieren sollte T_{ns} höher gewählt werden.

Um die Leistungsdichte zu erhöhen, wird häufig eine binäre Amplitude verwendet. Dies bedeutet, dass die Amplitude nur die beiden Werte $-u_0$ und u_0 annimmt. So ein binäres Rauschen kann über

$$n_b(t) = \begin{cases} u_0 & n(t) \geq 0 \\ -u_0 & n(t) < 0 \end{cases}$$

aus einem normalen Rauschprozess $n(t)$ gewonnen werden.

In Abbildung 6.15a ist in blau der Zeitverlauf gefilterten zeitdiskreten weißen Rauschens gezeigt. Die Abtastzeit beträgt $T_s = 1 \text{ ms}$ und als Filter ist ein Butterworth-Filter siebter Ordnung mit der Grenzfrequenz 100 Hz verwendet. Bei der gewählten Abtastzeit kann das Signal dann wieder im Zeitkontinuierlichen betrachtet werden. Für das hier gezeigte Signal sind zuletzt noch Werte außerhalb des Intervalls $[-u_{\max}, u_{\max}]$ auf die Intervallgrenzen gelegt. Ohne diese Maßnahme müsste das Signal sehr konservativ skaliert werden, um sicherzustellen, dass die Grenzen immer eingehalten werden. In rot ist das daraus erzeugte Binär-Rauschsignal dargestellt.

In Abbildung 6.15b sind die Leistungsdichtespektren der beiden Rauschsignale in den entsprechenden Farben gezeigt. Diese Spektren sind über die Methode von Welch (Matlab-Funktion `pwelch`) aus 100 s simulierten Signaldaten geschätzt. (Vernachlässigt man das Abschneiden der Werte außerhalb des zulässigen Intervalls für die Eingangsgröße, so kann das Leistungsdichtespektrum für das Nicht-Binär-Rauschsignal auch einfach berechnet werden.)

Es zeigt sich, dass im Fall des „normalen“ gefilterten Rauschsignals der Frequenzbereich, in dem das Signal Leistung zur Verfügung stellen soll, gut vorzugeben ist. Beim Binär-Rauschsignal ist die Signalleistung höher, was auf die bessere Ausnutzung der zur Verfügung stehenden Stellgröße zurückzuführen ist. Jedoch nimmt die Signalleistung zu höheren Frequenzen hin nicht so stark ab, wie durch den Filter eigentlich erwartet.

Zur Verwendung bei der Schätzung von Frequenzgängen über die die Fouriertransformation eignen sich die Rauschsignale nicht. Für diesen Zweck würden die Signale als zeitbegrenzte Signale auf das System geschaltet. Die sich ergebende Realisierung stellt dann für den einen Versuch die deterministische (tatsächliche) Eingangsgröße dar, dessen Amplitudenspektrum für die weitere Berechnung relevant ist. Die Amplitudenspektren eines jeweils 1 s Ausschnitt der beiden Rauschsignale sind in Abbildung 6.15c und (d) gezeigt. Es ist deutlich, dass es kaum möglich ist, in gegebenen Frequenzbereichen gezielt hohe Werte für die Amplitudendichte zu erhalten.

PRBS

Ist die Folge von binären Pseudo-Zufallszahlen periodisch und besitzt sie die in Abbildung 6.16 gezeigt Autokorrelationsfunktion, so spricht man von einem Pseudo-Binär-Zufallsprozess. (PRBS – pseudo random binary signal)

In [Pintelon und Schoukens 2001, S. 122] ist folgende Möglichkeit zitiert, um PRBS-Signale zu erzeugen. Es ist der Matlab-Code

```
Nn = 127;
u = -ones(Nn, 1);
u(mod((1:Nn).^2, Nn) + 1) = 1;
u(1) = 1;
```

angegeben. Mit diesem wird, unter der Voraussetzung, dass N_n in der Form $N_n = 4k - 1$ darstellbar und eine Primzahl ist, ein PRBS der Länge N_n erzeugt.

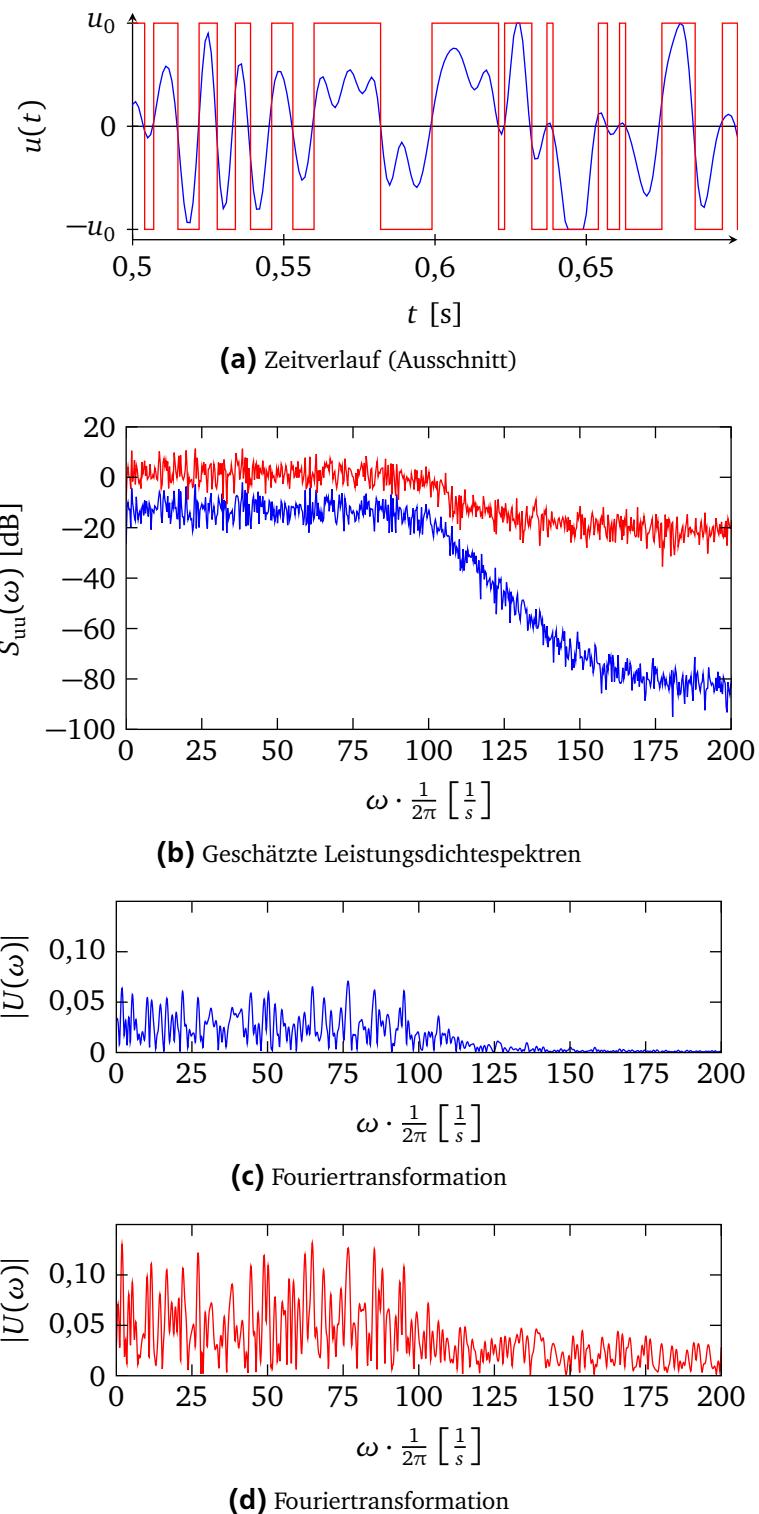


Abbildung 6.15: Gefiltertes Rauschen als Eingangsgröße

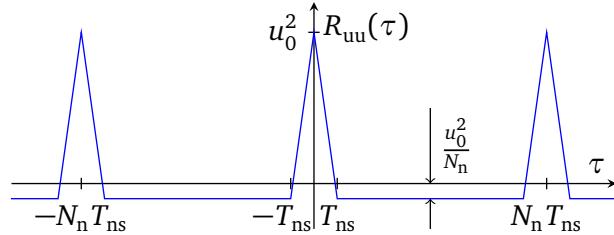


Abbildung 6.16: Autokorrelationsfunktion eines PRBS der Länge N_n und Taktzeit T_{ns}

Eine Variante von PRBS ist die MLBS (maximum length binary sequence). Hierbei wird eine Folge mit den genannten Eigenschaften über ein Schieberegister der Breite b Bits erzeugt. Es ergibt sich dann eine Länge $N_n = 2^b - 1$, was der Anzahl der möglichen Zustände von b Bits minus eins entspricht. (Der Zustand 0 darf nie vorkommen.) Darauf bezieht sich auch der Begriff „maximum length“.

In Abbildung 6.17 ist dies für ein Schieberegister der Breite sieben, was zu einer Länge $N_n = 2^7 - 1$ führt, gezeigt. Es werden die Bits 0 und 6 des Schieberegisters über eine XOR-Verknüpfung zurückgeführt.

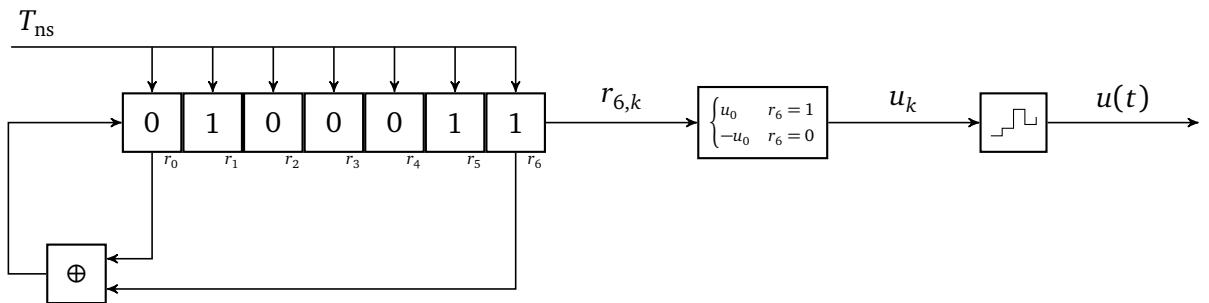


Abbildung 6.17: Erzeugung eines MLBS der Länge $N_n = 127$ über ein Schieberegister

Zur Erzeugung eines MLBS müssen mindestens zwei Bits zurückgeführt werden. Ein Bit ist dabei immer das letzte. (Ansonsten würde sich die effektive Länge des Schieberegisters entsprechend reduzieren.) In Tabelle 6.1 sind die Verknüpfungsregeln, die zu einem MLBS für verschiedene Registerlängen führen, aufgelistet. In vielen Fällen sind auch mehrere Varianten möglich, von denen hier aber nur jeweils eine gezeigt ist.

Der Zeitverlauf des sich mit dem Anfangszustand 1111111 ergebende MLBS der Länge $N_n = 127$ ist in Abbildung 6.18a gezeigt. Der Betrag der ersten 200 Fourierkoeffizienten ist in Abbildung 6.18b dargestellt.

Tabelle 6.1: MLBS

Bits	N_n	
6	63	$0 \oplus 5$
7	127	$0 \oplus 6$
8	255	$0 \oplus 1 \oplus 6 \oplus 7$
9	511	$3 \oplus 8$
10	1023	$2 \oplus 9$
11	2047	$1 \oplus 10$

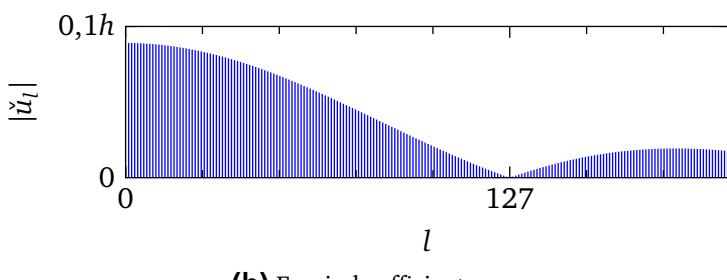
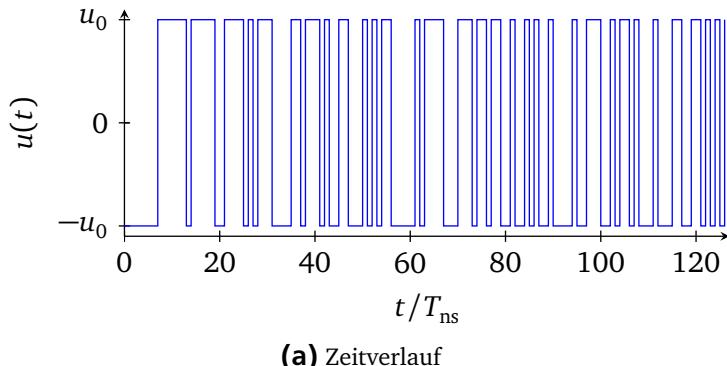


Abbildung 6.18: MLBS der Länge $N_n = 127$

Auf theoretischem Weg kann

$$|\check{u}_l| = \begin{cases} \frac{u_0}{N_n} & l = 0 \\ u_0 \cdot \frac{\sqrt{N_n + 1}}{N_n} \cdot \left| \frac{\sin \frac{l\pi}{N_n}}{\frac{l\pi}{N_n}} \right| & l > 0 \end{cases}$$

bestimmt werden.

Bei $l = N_n$ (hier 127) nimmt das Spektrum das erste Mal den Wert null an. Für eine konkrete Länge N_n und Taktzeit T_{ns} ergibt sich damit die Frequenz

$$\omega_{l=N_n} = N_n \cdot \omega_T = N_n \cdot \frac{2\pi}{N_n T_{ns}} = \frac{2\pi}{T_{ns}}.$$

Entsprechend ist die gesamte Abszisse in Abbildung 6.18b zu skalieren. Der Frequenzbereich der guten Anregung wird demnach mit wachsendem T_{ns} kleiner, dafür wird die Anregung bei gleicher Messdauer innerhalb des Bereiches dann stärker. In der Literatur findet sich als Anhaltswert $T_{ns} = 0,4 \cdot \frac{1}{f_{max}}$ für die Wahl der Taktzeit, wobei f_{max} die höchste Frequenz ist, die noch (gut) angeregt werden soll.

Die Messdauer und die Länge des PRBS sollten so aufeinander abgestimmt sein, dass immer vollständige Perioden erfasst werden. Ansonsten ist die Signalleistung nicht mehr einfach vorzugeben, wie Abbildung 6.19 zeigt. Dort sind die Fourierkoeffizienten des nach 126 Schritten abgeschnittenen MLBS (und dann periodisch fortgesetzte) MLBS gezeigt.

Für den Fall, dass das PRBS als zeitbegrenztes Signal verwendet werden soll, ist in Abbildung 6.20 ein Aschnitt der (kontinuierlichen) Fouriertransformierten des MLBS mit 127 Schritten gezeigt. Zusätzlich sind noch die Fourierkoeffizienten des entsprechenden Frequenzintervalls dargestellt. (Da eine Periode des MLBS hier auf eine Sekunde festgelegt ist, liegen die Zahlenwerte der Fourierkoeffizienten und die Fouriertransformierten bei den gemeinsamen Frequenzpunkten ohne weitere Skalierung genau übereinander.)

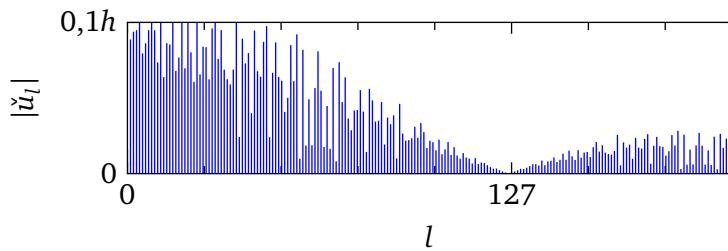


Abbildung 6.19: MLBS der Länge $N_n = 127$, nach $k = 126$ abgeschnitten

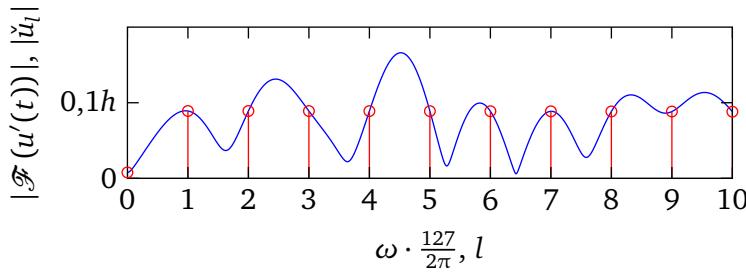


Abbildung 6.20: Fouriertransformierte eines MLBS der Länge $N_n = 127$

Es zeigt sich eine starke Welligkeit im kontinuierlichen Spektrum. Auch bei der Verwendung als zeibegrenztes Signal sollte man sich daher an die Frequenzstützstellen, die ein Vielfaches der Grundfrequenz der periodischen PRBS sind, halten.

Für die Identifikation zeitkontinuierlicher Modelle muss das Aliasing berücksichtigt werden. Denn durch die Sprünge enthält das Eingangssignal entsprechend hohe Frequenzen.

Wenn das Eingangssignal gemessen und aus den Messdaten die Fouriertransformierte (oder die Fourierkoeffizienten, wenn das PRBS periodisch verwendet wird) bestimmt werden soll, so muss mit einer entsprechend hohen Frequenz abgetastet werden. Vertraut man dem Stellglied und ist es nicht nötig die Eingangsgröße nochmals zu messen, so ist dies weniger kritisch, da offline die DFT des PRBS mit sehr hoher Abtastfrequenz bestimmt werden kann. Hierbei können keine reinen FFT-Algorithmen verwendet werden, da die Folgenlänge nicht als Zweierpotenz darstellbar ist.

Aliasing am Ausgang ist unkritisch, wenn die hohen Frequenzanteile durch das System ausreichend gedämpft werden.

Der Einfluss der Abtastzeit auf das Aliasing bei der Berechnung ist in Abbildung 6.21 illustriert. Hierzu ist die DFT des Verlaufs des MLBS mit verschiedenen Abtastzeiten bestimmt und mit dem tatsächlichen Spektrum, welches in Abbildung 6.18b gezeigt wurde verglichen. Zur Berechnung des tatsächlichen Spektrums ist das MLBS als Summe von 127 einzelnen, entsprechenden verschobenen positiven und negativen Impulsen geschrieben. Die Fouriertransformationen dieser Einzelimpulse sind dann addiert und noch durch 127 (Messdauer bei einer für den Vergleich festgelegten Abtastzeit von 1) geteilt, um die Fourierkoeffizienten zu erhalten.

Der Fehler ist nach

$$e_l = \frac{\hat{U}_l}{N} - \check{u}_l$$

berechnet, wobei \hat{U}_l die Werte der DFT mit der jeweils angegebenen Abtastzeit, N die sich damit ergebende Länge, und \check{u}_l der tatsächliche Wert der Fourierkoeffizienten ist.⁵

⁵ Bezuglich der Beträge, d. h. $\frac{|\hat{U}_l|}{N} - |\check{u}_l|$ ist der Fehler deutlich geringer. Der Fehler wirkt sich also primär auf die Phase der Schätzung aus.

In Abbildung 6.21a ist die Abtastzeit T_s zur Bestimmung der DFT gleich der Taktzeit T_{ns} des MLBS gesetzt. Es zeigt sich zweierlei. Erstens ist das PRBS mit der selben Takt- wie Abtastzeit im Zeitdiskreten eine sehr gute Näherung für ein weißes Rauschen. Zweitens können mit dieser Abtastzeit keine sinnvollen Werte des Spektrums im Zeitkontinuierlichen ermittelt werden. Der Fehler liegt in der Größenordnung der Referenzwerte. (Über $l = 63$ ist der Vergleich nicht sinnvoll, da bei 127 Punkten ohnehin nur die ersten 63 Punkte relevant sind.)

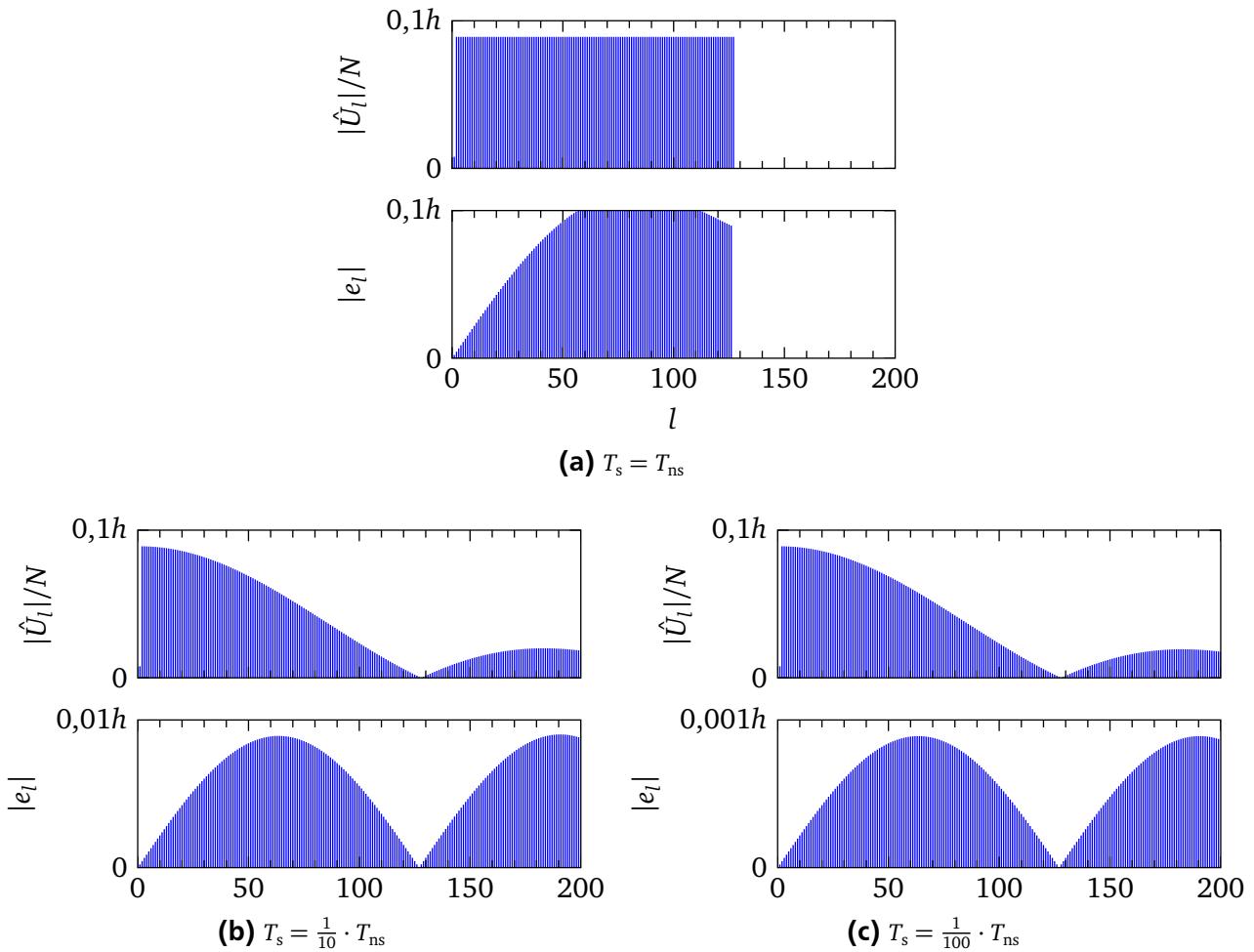


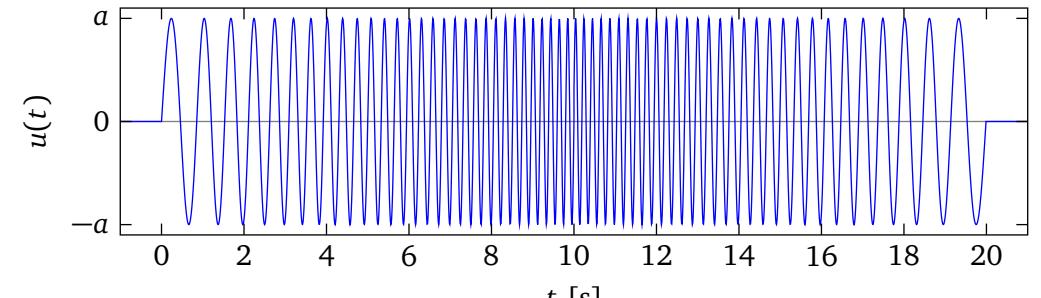
Abbildung 6.21: Zur Berechnung der Fourierkoeffizienten eines MLBS

Für $T_s = 0,1 \cdot T_{ns}$ ergibt sich ein Fehler von bis zu 10 % (Abbildung 6.21b), und bei $T_s = 0,01T_{ns}$ ein maximaler Fehler von 1 % (Abbildung 6.21c).

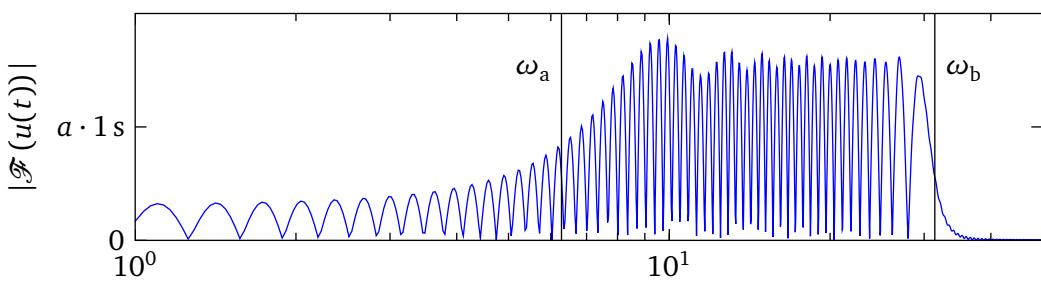
6.3.4 Kombination von Signalen

Abschließend ist noch zu erwähnen, dass die Kombination von zwei im Signalen, die jeweils ein vorteilhaftes Amplitudenspektrum besitzen, in der Regel nicht zu einem Signal führt, das wiederum ein gutes Amplitudendichtespektrum besitzt.

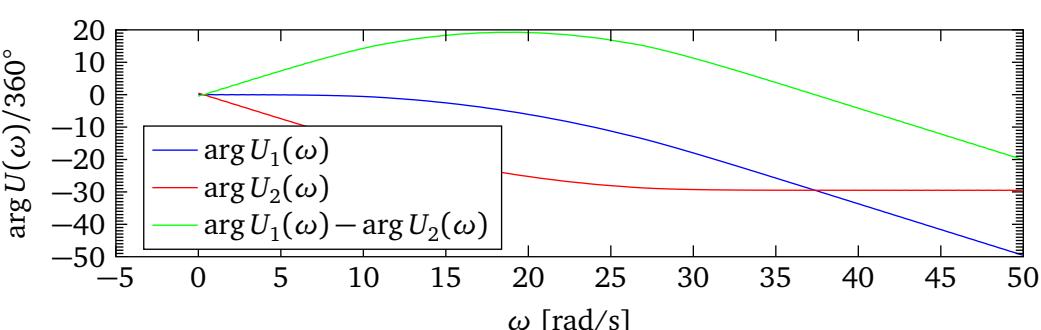
Dies wird an dem bidirektionalen Gleitsinus aus Abbildung 6.22a gezeigt. Dieser besteht aus dem in Abbildung 6.11 gezeigten Gleitsinus, dessen Momentanfrequenz nach den 10 s wieder linear zurückgefahren wird. Das Spektrum dieses Signals ist sehr wellig und eignet sich kaum für die Identifikation im Frequenzbereich, wie in Abbildung 6.22b zu sehen ist.



(a) Zeitverlauf



(b) Fouriertransformierte



(c) Zur Erklärung der Auslöschungen

Abbildung 6.22: Bidirektionaler Gleitsinus

Bezeichnet man den ursprünglichen Gleitsinus mit $u_1(t)$, so wird der bidirektionale Gleitsinus über

$$u(t) = u_1(t) + u_2(t)$$

beschrieben, wobei

$$u_2(t) = -u_1(-t - 20 \text{ s})$$

ist. D. h. der zweite Teil des Signals besteht aus einer rückwärtslaufenden, gedrehten und um 20 s nach rechts verschobenen Kopie des ersten Teils. Im Frequenzbereich gilt aufgrund der Linearität der Fouriertransformation

$$U(\omega) = U_1(\omega) + U_2(\omega).$$

Das Spiegeln und das Verschieben von Signalen im Zeitbereich beeinflusst nur das Argument der Fouriertransformierten. Der Betrag bleibt unverändert. Dies folgt aus Regeln zur Zeitverschiebung,

$$\mathcal{F}(y(t - \tau))|_{\omega} = e^{-j\tau\omega} \cdot Y(\omega),$$

und der Zeitskalierung,

$$\mathcal{F}(y(at))|_{\omega} = \frac{1}{|a|} \cdot Y\left(\frac{\omega}{a}\right),$$

der Fouriertransformation.

Hier folgt

$$U_2(\omega) = -U_1(-\omega) \cdot e^{-j \cdot 20s \cdot \omega} = -U_1^*(\omega) \cdot e^{-j \cdot 20s \cdot \omega}.$$

Da die Amplitudendichten von $U_1(\omega)$ und $U_2(\omega)$ identisch sind, tritt für die Frequenzen eine maximale Verstärkung auf, bei denen sich die Argumente von $U_1(\omega)$ und $U_2(\omega)$ um ein gerades Vielfaches von 180° unterscheiden. Bei einer Differenz von einem ungeraden Vielfachen von 180° tritt Auslöschung auf.

In Abbildung 6.22c sind die Argumente von $U_1(\omega)$ und $U_2(\omega)$ sowie deren Differenz dargestellt. Es wird deutlich, dass sich deren Differenz sehr schnell ändert. Dies führt zu den Verstärkungen und Auslöschungen, die in Abbildung 6.22b erkennbar sind.

Schon eine reine Zeitverschiebung (also keine Drehung der Funktion) wird praktisch immer dazu führen, dass die Differenz der Einzeltransformierten nicht konstant ist und häufig Auslöschungen auftreten. Diesen Effekt hat man schon bei den Rechteckimpulsen in Abbildung 6.10a bzw. (b) beobachten können.

7 Nichtparametrische Identifikation

Bei der nichtparametrischen Identifikation geht es im Wesentlichen darum, Frequenzgänge und Impulsantworten von Systemen zu bestimmen. Man kann beispielsweise auch die Messung einer Sprungantwort als nichtparametrische Identifikation betrachten, da auch die Sprungantwort ein nichtparametrisches Modell eines dynamischen Systems darstellt, jedoch hat diese für die Systemanalyse und den Reglerentwurf einen eher geringeren Wert.

Frequenzgang und Impulsantwort sind über die Fouriertransformation (bzw. deren Inversen) ineinander überführbar. Bei nichtparametrischen Modellen liegen die Daten jedoch punktweise (wenn man von analogen Datenschreibern absieht) und nur für einen begrenzten Zeit- bzw. Frequenzbereich vor, so dass diese Umrechnung nur näherungsweise möglich ist. Mit einer vernünftigen Wahl der Identifikationsparameter kann aber erreicht werden, dass diese Umrechnung ausreichend genau ist.

Da der Frequenzgang und die Impulsantwort damit in diesem Sinne äquivalent sind, wird hier schwerpunktmäßig nur die Bestimmung des Frequenzganges betrachtet. Auf die (direkte) Bestimmung der Impulsantwort wird am Ende kurz eingegangen und auf entsprechende Literatur verwiesen.

Hier wird davon ausgegangen, dass zeitkontinuierliche Systeme identifiziert werden sollen. Daher ist immer die „normale“ (zeitkontinuierliche) Fouriertransformation geschrieben. Voraussetzung hierfür ist, dass die – immer durchgeführte – Abtastung schnell genug erfolgt, um Aliasing auszuschließen. In diesem Fall wurde in Kapitel 5 gezeigt, dass die Werte, die über die DFT bestimmt werden, unter den gleichen Voraussetzungen der DTFT und der FT entsprechen. Da die Skalierung mit der Abtastzeit T_s , die notwendig ist, um die DTFT-Werte in FT-Werte „umzurechnen“, hier immer sowohl auf den Zähler als auch auf den Nenner angewandt wird, ändert diese nichts an den Werten der identifizierten Übertragungsfunktion. So ist die praktische Berechnung genau die gleiche, nur die Skalierung der Frequenzachse ändert sich.

Wird zu langsam abgetastet und ist das Shannon-Theorem nicht erfüllt, so kann nur der Frequenzgang eines zeitdiskreten Systems bestimmt werden. Das Vorgehen ist dabei völlig analog, lediglich müssen die in den Gleichungen vorkommenden fouriertransformierten Signalen durch deren DTFT-transformierte Varianten ersetzt werden. (Integrale über die Unabhängige t müssen entsprechend durch Summen über k ersetzt werden, aber dies ist für die numerische Berechnung ohnehin auch bei zeitkontinuierlichen Signalen nötig.)

7.1 Frequenzgang

Regt man ein (asymptotisch) stabiles System mit einer sinusförmigen Schwingung

$$u(t) = u_1 \cdot \sin(\omega t + \varphi_u)$$

an, so verbleibt nach dem Abklingen möglicher „Übergangsbewegungen“ am Ausgang ebenfalls eine sinusförmige Schwingung gleicher Frequenz, jedoch mit unterschiedlicher Amplitude und Phasenlage,

$$y(t) = y_1 \cdot \sin(\omega t + \varphi_y).$$

Das Verhältnis der beiden Amplituden sowie die Phasenverschiebung hängt von der Frequenz ω ab und wird durch den Frequenzgang $H(\omega)$ beschrieben. Der Frequenzgang ordnet jeder Frequenz ω eine kom-

plexen Zahl zu, deren Betrag das Amplitudenverhältnis und deren Argument (Winkel in der komplexen Ebene) der Phasenverschiebung entspricht,

$$|H(\omega)| = \frac{y_1}{u_1} \quad \text{und} \\ \arg(H(\omega)) = \varphi_y - \varphi_u = \Delta\varphi .$$

Damit lässt sich für den Ausgang nach Abklingen aller anderen Anteile auch

$$y(t) = |H(\omega)| \cdot u_1 \cdot \sin(\omega t + \varphi_u + \arg(H(\omega)))$$

schreiben.

Drückt man die periodischen Ein- und Ausgangssignale durch die komplexen Fourierkoeffizienten aus, so kann auch

$$H(l \cdot \omega_0) = \frac{\mathfrak{F}(y(t))_l}{\mathfrak{F}(u(t))_l}$$

geschrieben werden.

Verfolgt man den oben skizzierten Weg (sinusförmige Schwingung am Eingang, vernachlässigen aller abklingenden Anteile am Ausgang) im Laplace-Bereich, so zeigt sich, dass der Frequenzgang auch gerade der an $s = j\omega$ ausgewerteten Übertragungsfunktion entspricht¹ (siehe Skript SDRT I),

$$H(\omega) = G(s)|_{s=j\omega} = G(j\omega) .$$

(Dieser Zusammenhang gilt immer, auch wenn die Impulsantwort des zugrundeliegenden Systems aufgrund von Polen auf oder rechts der imaginären Achse nicht Fourier-transformierbar ist.)

Für Systeme mit abklingender Impulsantwort kann der Frequenzgang damit auch als Fouriertransformierte der Impulsantwort $g(t)$ geschrieben werden,

$$H(\omega) = \mathcal{F}(g(t))|_{\omega} .$$

Existiert die Fouriertransformierte für ein Paar aus Ein- und Ausgangssignal, dann ist der Frequenzgang auch über

$$H(\omega) = \frac{\mathcal{F}(y(t))|_{\omega}}{\mathcal{F}(u(t))|_{\omega}}$$

gegeben. Für alle technischen Signale kann die Übertragungsfunktion ihrer Definition nach als

$$H(\omega) = \left. \frac{\mathcal{L}(y(t))}{\mathcal{L}(u(t))} \right|_{s=j\omega}$$

angegeben werden.

7.2 Messaufbau

In Abbildung 7.1 ist ein Messaufbau dargestellt, wie er zur Identifikation verwendet werden kann.

¹ Tatsächlich wird der Frequenzgang als „Laplace-Transformierte bei $s = j\omega$ “ definiert [Föllinger 1977, S. 120]. Die Motivation für diese Definition liegt aber in der oben angegebenen Bedeutung.

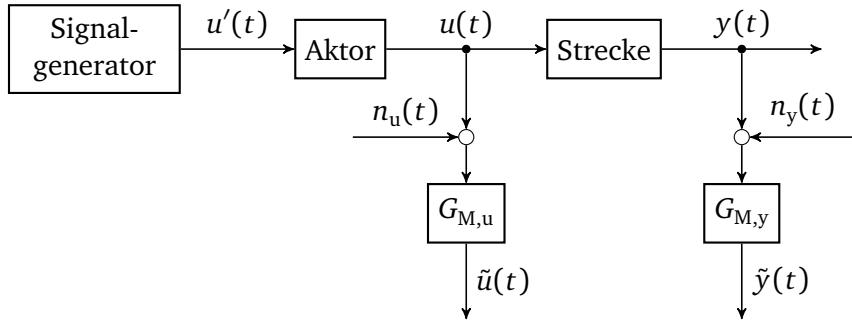


Abbildung 7.1: Messaufbau

Das Anregungssignal $u'(t)$ wird durch einen Signalgenerator erzeugt und auf den Aktor gegeben. Dieser erzeugt daraus das tatsächlich auf das System wirkende Anregungssignal $u(t)$, welches die Ausgangsgröße $y(t)$ hervorruft.

Wenn das Signal zwischen Aktor und Strecke nicht gemessen werden kann oder soll, so wird der Aktor der Strecke zugeschlagen. Häufig ist dies auch der gewünschte Fall, da beispielsweise für einen Reglerentwurf der Aktor ohnehin im Modell berücksichtigt werden muss. Wenn die Identifikation jedoch mit einem anderen als den später zu verwendeten Aktor erfolgt, oder wenn man auf eine Analyse der Strecke ohne Aktor aus ist, sollte das zu identifizierte Modell den Aktor möglichst nicht enthalten. (Wobei es auch möglich ist, zunächst Strecke und Aktor zusammen, und dann nur den Aktor zu identifizieren. Dann kann der Aktor aus dem Gesamtmodell wieder herausgerechnet werden.)

Mit $G_{M,u}(s)$ und $G_{M,y}(s)$ werden in Abbildung 7.1 die Filterungen des gemessenen Ein- bzw. Ausgangssignals bezeichnet. Im Folgenden soll kurz auf den Einfluss dieser Filter eingegangen werden.

Im Frequenzbereich lautet der eigentlich zu identifizierende Zusammenhang

$$Y(s) = G(s)U(s).$$

Unter Vernachlässigung des Rauschens gilt

$$\begin{aligned}\tilde{U}(s) &= G_{M,u}(s)U(s) \quad \text{und} \\ \tilde{Y}(s) &= G_{M,y}(s)Y(s).\end{aligned}$$

Damit folgt

$$\tilde{Y}(s) = G_{M,y}(s)G(s)U(s) = G_{M,y}(s)G(s)G_{M,u}^{-1}(s)\tilde{U}(s).$$

Wird für beide Messungen der selbe Filter $G_{M,u}(s) = G_{M,y}(s) = G_M(s)$ verwendet, so folgt daraus

$$\tilde{Y}(s) = G(s)\tilde{U}(s).$$

Mit den ungestörten Messgrößen gilt damit

$$H(\omega) = \frac{Y(j\omega)}{U(j\omega)} = \frac{\tilde{Y}(j\omega)}{\tilde{U}(j\omega)},$$

d. h. wenn $\tilde{u}(t)$ und $\tilde{y}(t)$ verwendet werden ergibt sich der selbe Frequenzgang wie bei $u(t)$ und $y(t)$.²

² Dies gilt auch noch im MIMO-Fall wenn für die Übertragungsmatrix $\mathbf{G}_M(s) = \mathbf{I} \cdot G_M(s)$ gilt, d. h. sich die Messungen nicht gegenseitig beeinflussen und das Messglied für alle Kanäle die gleiche Dynamik aufweist.

Häufig wird ein Anti-Aliasing-Filter vor der Abtastung eingesetzt. Dieses kann aber auch vernachlässigt werden, wenn die Abtastung schnell gegenüber der Systemdynamik erfolgt. Daneben können Filter notwendig sein, um hochfrequente Störungen und Messrauschen zu unterdrücken.

Wenn die Filterung der Messgröße am Ausgang nicht zu vernachlässigen ist, und der Eingang nicht (über die gleichen Filter) gemessen werden soll, dann kann auch eine (zusätzliche) Filterung im Rechner implementiert werden, um so künstlich eine (näherungsweise) gleiche Filterübertragungsfunktion in beiden Pfaden zu erhalten. Ist auch dies nicht möglich, so wird die Strecke mit dem Filter zusammen identifiziert.

Im weiteren wird die Filterung daher nicht mehr betrachtet. Unter „Strecke“ wird entweder die ursprüngliche Strecke oder die um den Aktor erweiterte Strecke verstanden.

7.3 Frequenzgangsschätzung über Amplitudenspektren (H0-Schätzung)

7.3.1 Ablesen aus dem Zeitschrieb

Die anschauliche Interpretation des Frequenzgangs legt es nahe, diesen frequenzweise wie folgt zu identifizieren. Es wird eine Sinusschwingung mit einer Frequenz ω_0 auf das System gegeben. Nachdem alle transienten Vorgänge abgeklungen sind, d. h. das System eingeschwungen ist, kann die Amplitude der Ausgangsgröße sowie der Phasenverzug zwischen Ein- und Ausgangsgröße gemessen werden. Dies ist in Abbildung 7.2a dargestellt. Das verwendete Beispielsystem ist der Drei-Massen-Schwinger, wobei hier und im Folgenden immer Simulationen gezeigt sind.

In schwarz ist die Eingangsgröße (Motormoment) dargestellt. Diese stellt eine Sinusschwingung mit der Amplitude $u_1 = 1 \text{ Nm}$ und der Frequenz $\omega_0 = \frac{2\pi}{T_0} = 2\pi \cdot 0,5 \frac{1}{\text{s}}$ dar. Dabei ist die Eingangs- und Ausgangsgröße immer als Abweichungen von einem Arbeitspunkt anzusehen, so dass die Annahme der Linearität hinreichend genau erfüllt ist. Für $t < 0$ ist die Eingangsgröße sowie die Ausgangsgröße null. Daher zeigt sich im blau dargestellten Verlauf der Ausgangsgröße (Winkelgeschwindigkeit der dritten Masse) ein Einschwingvorgang, welcher aber schnell abklingt. An der fünften „Periode“ des Ausgangssignals wird der Betrag der Amplitude $y_1 = 8,34 \text{ rad/s}$ abgelesen.

Für die Verstärkung des Systems bei der Frequenz $2\pi \cdot 0,5 \frac{1}{\text{s}} = \pi \text{ rad/s}$ ergibt sich damit

$$|H(\pi)| = \frac{y_1}{u_1} = 8,34 \frac{\text{rad/s}}{\text{Nm}},$$

bzw.

$$(|H(\pi)|)_{\text{dB}} = 20 \cdot \log_{10}(8,34) = 18,47$$

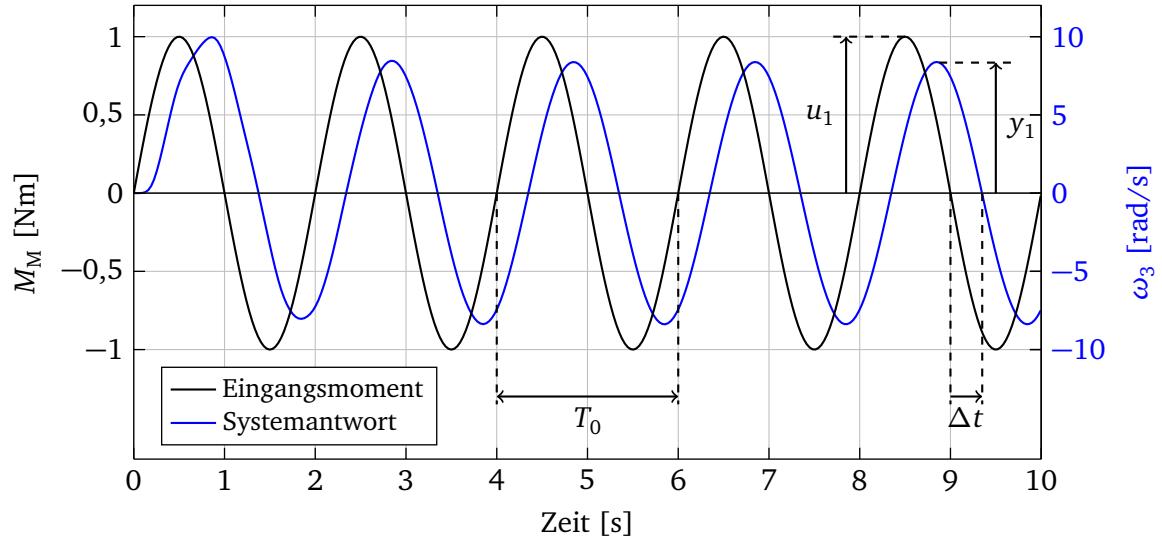
in dB angegeben.

Die Zeitdifferenz Δt zwischen den Nulldurchgängen der Ein- und Ausgangsgröße beträgt nach dem Einschwingen 0,35 s. Damit erhält man den Phasenwinkel

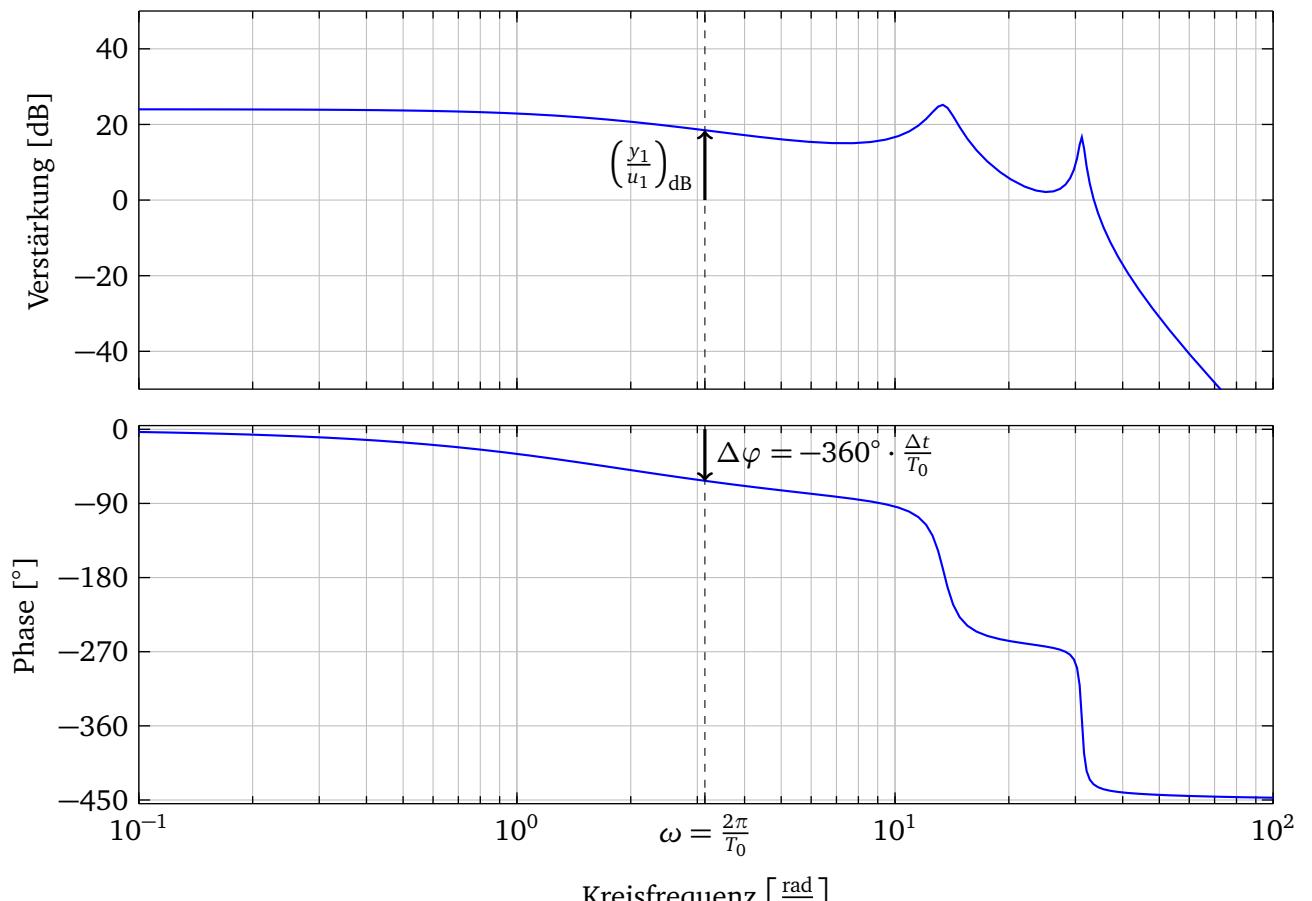
$$\arg H(\pi) = \Delta\varphi = -360^\circ \cdot \frac{\Delta t}{T_0} = -63^\circ.$$

Diese ermittelte Verstärkung und Phasenverschiebung für die Frequenz $\omega = \pi$ sind in Abbildung 7.2b eingetragen.

Sind die Messdaten verrauscht wird es schnell problematisch, die nötigen Werte aus den Zeitschrieben abzulesen. Dies wird in Abbildung 7.3 deutlich. Darüberhinaus lässt sich dieses „Ablesen“ nicht ohne weiteres automatisieren.



(a) Zeitverlauf von Eingangs- und Ausgangsgröße (um Arbeitspunkt)



(b) Frequenzgang

Abbildung 7.2: Punktweise Messung des Frequenzganges

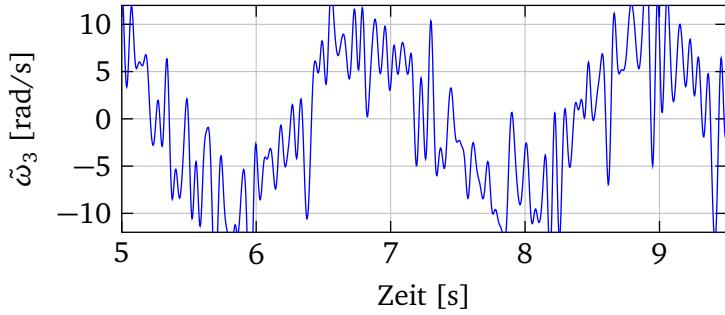


Abbildung 7.3: Verrauschter Ausgangsgrößenverlauf

7.3.2 Schätzung über Fourierkoeffizienten

Da die Frequenz des Ausgangsgrößenverlaufs bekannt ist, wäre es sinnvoll einen Sinusverlauf mit dieser Frequenz optimal auf die in Abbildung 7.3 dargestellten Messdaten anzupassen. Genau dies macht die Fourierreihenentwicklung, denn diese zerlegt eine Periode eines Signals in eine Summe von Sinusschwingungen. Dabei sind diese verschiedenen Anteile orthogonal zueinander, d. h. auch wenn das Signal dann nur mit der ersten Harmonischen alleine beschrieben wird, liegt der durch die Vernachlässigung der höheren Harmonischen entstehende Fehler orthogonal zu der ersten Harmonischen. D. h. jede Änderung des dazugehörigen Fourierkoeffizienten \hat{y}_1 erhöht den Fehler.

Da nur die verrauschten Messdaten $\tilde{y}(t)$ bekannt sind, kann hier lediglich ein Schätzwert

$$\hat{\tilde{y}}_1 = \widehat{\mathfrak{F}(y(t))}_1 = \frac{1}{T_0} \int_0^{T_0} \tilde{y}(t) \cdot e^{-j\omega_0 t} dt$$

für den Fourierkoeffizienten der Grundharmonischen bestimmt werden. Um den Einfluss des Rauschens auf diesen Schätzwert möglichst zu verringern, können mehrere Fourierkoeffizienten aus den Messdaten bestimmt und diese gemittelt werden. Liegen diese Intervalle dazu direkt hintereinander, so kann dies auch als

$$\hat{\tilde{y}}_1 = \widehat{\mathfrak{F}(y(t))}_1 = \frac{1}{N \cdot T_0} \int_0^{N \cdot T_0} \tilde{y}(t) \cdot e^{-j\omega_0 t} dt$$

geschrieben werden. (Man könnte es auch als die N -te Harmonische bezüglich der Grundperiode $N \cdot T_0$ bzw. der Grundfrequenz ω_0/N auffassen.)

Analog ergibt sich für die Schätzung des Fourierkoeffizienten des Eingangs

$$\hat{\tilde{u}}_1 = \widehat{\mathfrak{F}(u(t))}_1 = \frac{1}{N \cdot T_0} \int_0^{N \cdot T_0} \tilde{u}(t) \cdot e^{-j\omega_0 t} dt$$

und damit erhält man

$$\hat{H}(\omega_0) = \frac{\hat{\tilde{y}}_1}{\hat{\tilde{u}}_1} = \frac{\frac{1}{N \cdot T_0} \int_0^{N \cdot T_0} \tilde{y}(t) \cdot e^{-j\omega_0 t} dt}{\frac{1}{N \cdot T_0} \int_0^{N \cdot T_0} \tilde{u}(t) \cdot e^{-j\omega_0 t} dt} \quad (7.1)$$

als Schätzwert für den Wert des Frequenzganges bei ω_0 .

Die Effektivität der Schätzung der Grundharmonischen über die Fourierreihenentwicklung wird in Abbildung 7.4 deutlich. Dort ist in grau die verrauschte Messung gezeigt. Der tatsächliche Ausgangsgrößenverlauf ist in blau dargestellt. In rot (durchgezogen), orange (gestrichelt) und gelb (strich-punktiert) ist

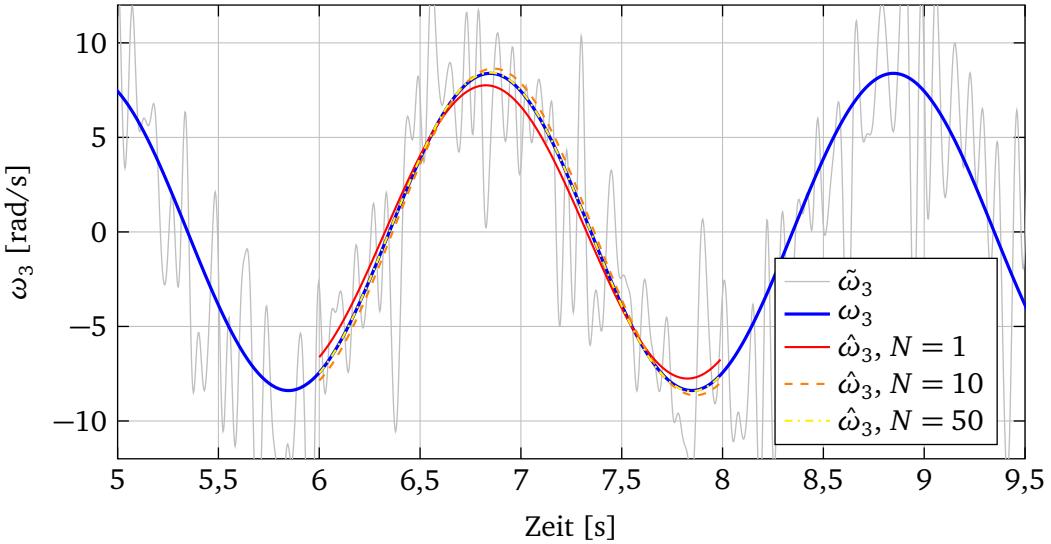


Abbildung 7.4: Schätzung der Grundharmozischen des Ausgangsgrößenverlaufs

eine Periode des geschätzten Verlaufs gezeigt, die aufgrund der Berücksichtigung einer, zehn bzw. fünfzig Perioden bestimmt sind.

Aufgrund des hohen Rauschens zeigt sich bei der Schätzung über $N = 1$ noch eine deutlich sichtbare Abweichung. Aber auch dieser Wert ist im Allgemeinen noch besser als jede per Hand bestimmte Näherung. Bei $N = 10$ gemittelten Perioden hat sich die Abweichungen deutlich reduziert, und bei $N = 50$ ist in dieser Darstellung kein Unterschied mehr zwischen dem tatsächlichen und dem geschätzten Verlauf auszumachen.

Sonderfall Eingang exakt

Kann die Eingangsgröße direkt vorgegeben werden (keine Totzeiten, keine Aktordynamik, keine Störungen zwischen Generator und Strecke), dann kann der entsprechende Koeffizient \check{u}_1 schon vor der eigentlichen Messung bestimmt werden. Wird für den Eingang

$$u(t) = u_1 \cdot \sin(\omega_0 t)$$

gewählt, dann ist

$$\check{u}_1 = \mathfrak{F}(u(t))_1 = -j \cdot \frac{1}{2} \cdot u_1 .$$

Damit erhält man

$$\hat{H}(\omega_0) = j \cdot \frac{2}{u_1} \cdot \frac{1}{NT_0} \cdot \int_0^{NT_0} \tilde{y}(t) \cdot e^{-j\omega_0 t} \cdot dt , \quad (7.2)$$

was noch in Real- und Imaginärteil aufgespalten werden kann,

$$\begin{aligned} \operatorname{Re}(\hat{H}(\omega_0)) &= \frac{\frac{1}{NT_0} \cdot \int_0^{NT_0} \tilde{y}(t) \cdot \sin(\omega_0 t) \cdot dt}{\frac{1}{2} u_1} \\ \operatorname{Im}(\hat{H}(\omega_0)) &= \frac{\frac{1}{NT_0} \cdot \int_0^{NT_0} \tilde{y}(t) \cdot \cos(\omega_0 t) \cdot dt}{\frac{1}{2} u_1} . \end{aligned}$$

Alternativ kann man dies auch als

$$\operatorname{Re}(\hat{H}(\omega_0)) = \frac{2}{u_1^2} \cdot \frac{1}{NT_0} \cdot \int_0^{NT_0} \tilde{y}(t) \cdot u_1 \cdot \sin(\omega_0 t) \cdot dt$$

$$\operatorname{Im}(\hat{H}(\omega_0)) = \frac{2}{u_1^2} \cdot \frac{1}{NT_0} \cdot \int_0^{NT_0} \tilde{y}(t) \cdot u_1 \cdot \cos(\omega_0 t) \cdot dt$$

schreiben. Diese Form ist in Abbildung 7.5 als Blockschaltbild dargestellt. In dieser Darstellung wird der Algorithmus auch als „orthogonale Korrelation“ bezeichnet. (Dieser Name röhrt von einer anderen Herangehensweise an den Algorithmus her.)

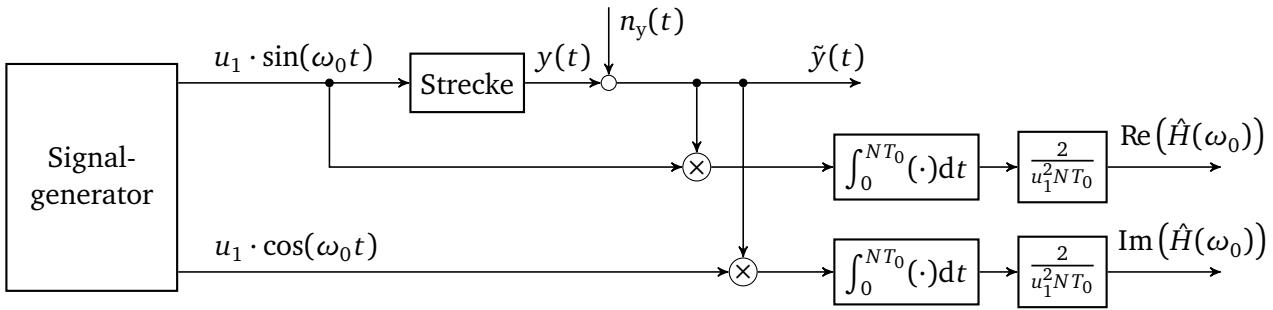


Abbildung 7.5: Orthogonale Korrelation

Sollte eine Aktordynamik vorhanden sein, so wird diese bei der Identifikation nach Abbildung 7.5 immer der Strecke zugeschlagen und mit dieser zusammen identifiziert. Dasselbe gilt für die Messkette am Ausgang.

Beispiel zur Identifikation mit monofrequentem Anregungssignal

Im Folgenden wird die Identifikation auf die beschriebene Art an dem Beispielsystem betrachtet. Das in Abbildung 7.4 dargestellte Rauschen diente zur Demonstration der Güte der Schätzung der Fourierkoefzienten. Für das gegebene System würde ein solches starkes Rauschen aber keine gute Identifikation zulassen. Für die in Abbildung 7.4 verwendete Frequenz von 0,5 Hz ist die Ausgangsgröße zwar noch gut zu schätzen, jedoch besitzt das System dort auch noch eine hohe Verstärkung. Bei 4 Hz (25 rad/s) ist die Systemverstärkung um fast 20 dB, also den Faktor 10 geringer. Da hier die Amplitude der Anregung sowie das Leistungsspektrum des Rauschens über den gesamten relevanten Frequenzbereich (näherungsweise) konstant sind, gewinnt das Rauschen damit deutlich an Einfluss. Entsprechend wären sehr viele Mittelungen bzw. Perioden nötig, um die stochastische Störung ausreichend gut zu filtern.

Für alle folgenden Simulationen wird daher ein Rauschen verwendet, welches um ungefähr den Faktor drei geringer ist. Dies ist beispielhaft in Abbildung 7.6 dargestellt und in vielen Fällen auch realistischer.

Die Ergebnisse, d. h. der Frequenzgang, der Identifikation sind in Abbildung 7.7 dargestellt. In blau ist der tatsächliche Frequenzgang als Referenz gegeben.

Es wurde der Frequenzgang für einhundert gleichmäßig verteilte Frequenzpunkte von 0,1 Hz bis 7,5 Hz geschätzt. Dabei wurde eine Periode ($N = 1$, schwarze Punkte), zehn Perioden (grüne Punkte) und fünfzig Perioden (orangene Dreiecke) verwendet.

Für jede Frequenz ist eine Einschwingzeit von 15 s abgewartet worden. Dies ist etwas mehr als die 6 s bei dem Beispiel oben, da bei Frequenzen geringerer Verstärkung der Einschwingvorgang länger relevant ist. Dies liegt daran, dass zum einen die Systemantwort bei hohen Frequenzen schwächer wird, und damit der Fehler durch die Einschwingvorgänge stärker ins Gewicht fällt. Zum anderen ändert sich auch die Stärke der Anfangsstörung, die abklingen muss.

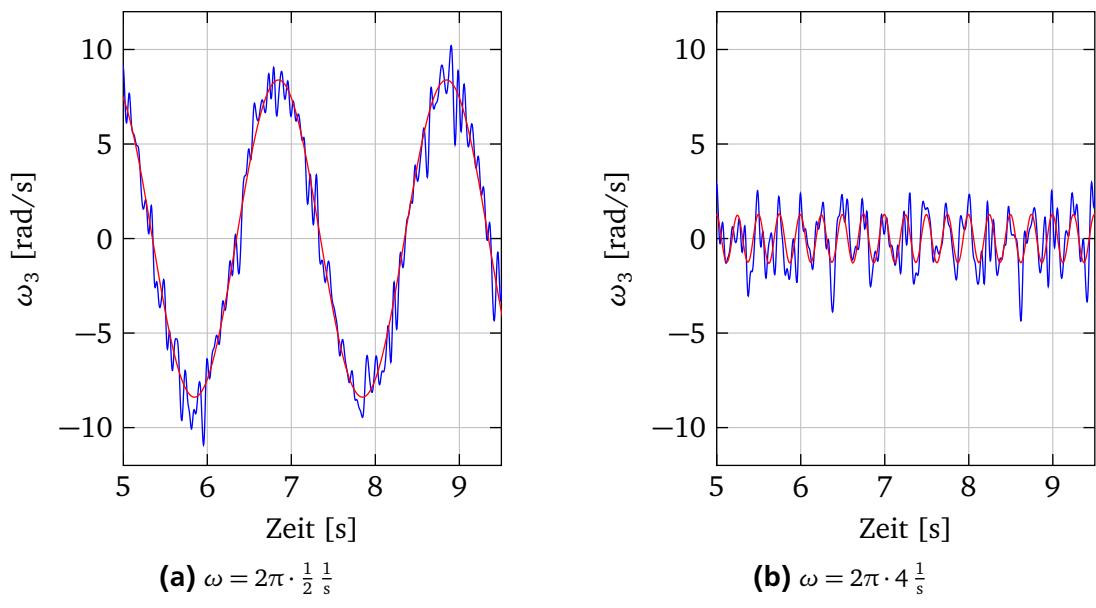


Abbildung 7.6: Ausgangssignal für zwei Frequenzpunkte

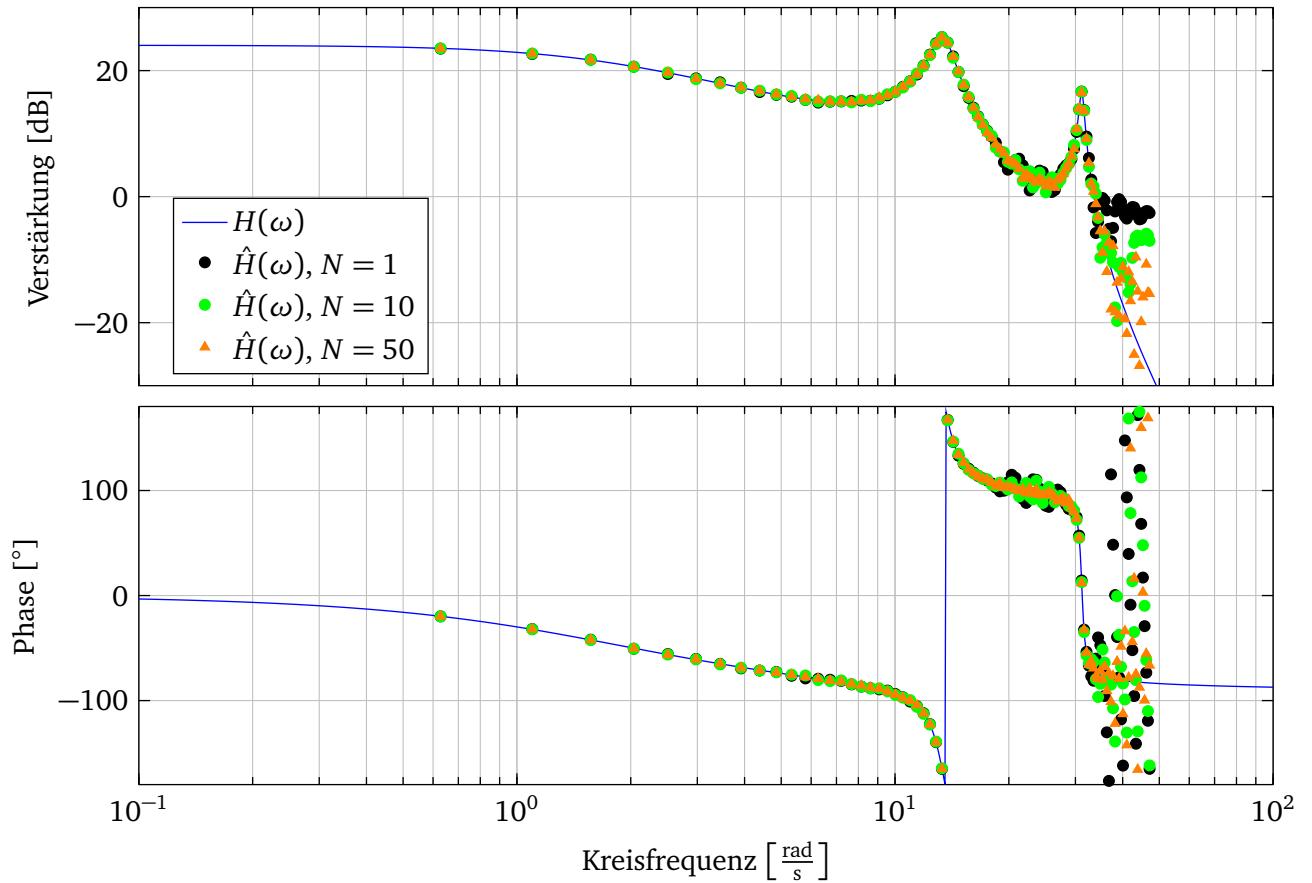


Abbildung 7.7: Geschätzter Frequenzgang bei monofrequenter Sinusanregung

In den Frequenzbereichen, in denen die Systemverstärkung und damit die Ausgangsgröße groß ist, werden für alle N gute Ergebnisse erreicht. In dem Bereich schwacher Verstärkung zwischen den beiden Resonanzfrequenzen zeigen die Schätzungen mit $N = 1$ und $N = 10$ schon deutlichere Streuungen, wobei diese bei $N = 1$ etwas stärker ausgeprägt sind. Mit $N = 50$ ist auch hier eine sehr gute Schätzung möglich. Für Frequenzen oberhalb der zweiten Resonanz, für die die Systemverstärkung unter 0 dB abfällt, wird der Schätzwert im Wesentlichen von dem Messrauschen dominiert. Der Frequenzbereich einer sinnvollen Schätzung wird bei $N = 1$ als erstes verlassen. Mit $N = 10$ und $N = 50$ ergeben sich hier bis zu einer Systemverstärkung von ca. -10 dB sinnvolle Werte.

Für den Fall, dass jeweils nur eine Periode gemessen wird, wird die Gesamtmesszeit im Prinzip nur von der Einschwingdauer bestimmt. Es ergeben sich ca. 26 Minuten. (Genauer 1563 Sekunden, davon sind $100 \cdot 15$ s = 1500 s die kumulierte Einschwingdauer.) Bei der Messung von $N = 10$ Perioden verlängert sich die Messdauer auf 36 Minuten, bei $N = 50$ auf 78 Minuten. Je größer die Anzahl der zu messenden Perioden N , desto mehr fallen die niedrigen Frequenzen ins Gewicht. Bei $N = 50$ fallen alleine 500 s (ohne Einschwingen) für die Messung am ersten Frequenzpunkt von 0,1 Hz an. Da bei niedrigen Frequenzen die Verstärkung in der Regel aber auch hoch ist, wäre es hier eine Möglichkeit, die Anzahl N für die kleinen Frequenzen geringer zu halten.

Ein Nachteil der frequenzpunktweisen Messung ist die sehr lange Messdauer. Dafür erhält man in der Regel auch sehr genaue Werte. Weiter besteht die Möglichkeit, die Parameter zur Identifikation in einfacher Weise während den Messungen zu verändern und auf die aktuelle Situation anzupassen. So kann nicht unbedingt für alle Frequenzen die maximale Stellamplitude genutzt werden, da der Systemausgang oder Systemzustände den zulässigen oder den linearen Bereich verlassen. Für alle Frequenzen dieselbe Eingangsamplitude zu verwenden führt jedoch zu ungenauen Schätzwerten in Bereichen geringer Verstärkung. Daher sollte die Eingangsamplitude in jedem Schritt aufgrund der im letzten Schritt erreichten Ausgangsamplitude angepasst werden. Auch kann die Frequenzschrittweite adaptiv in Abhängigkeit der Veränderung der Ausgangsamplituden angepasst werden. In Resonanzbereichen, wo starke Änderungen vorliegen sollten viele Punkte gesetzt werden. In anderen Bereichen weniger.

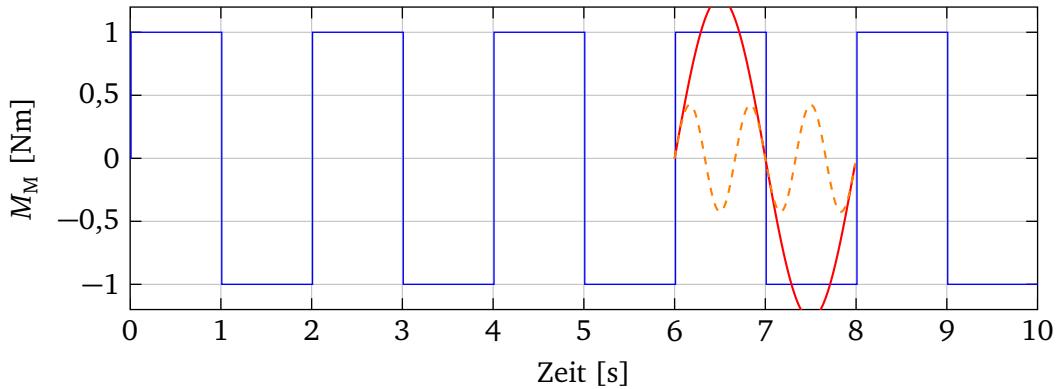
Anregung über andere periodische Signale

Ist es nicht möglich (genaue) Sinusanregungen zu realisieren, oder soll der Aktor nicht mitidentifiziert werden, so dass die auf das System wirkende Eingangsgröße gemessen wird, so kann die Schätzung nach Gl. (7.1) dennoch genauso verwendet werden. Dies folgt aus der Linearität der Fourierreihenentwicklung. In Abbildung 7.8 ist als Eingangssignal eine Rechteckschwingung verwendet. Das Eingangssignal ist in Abbildung 7.8a gezeigt, die (ungestörte) Systemantwort in Abbildung 7.8b.

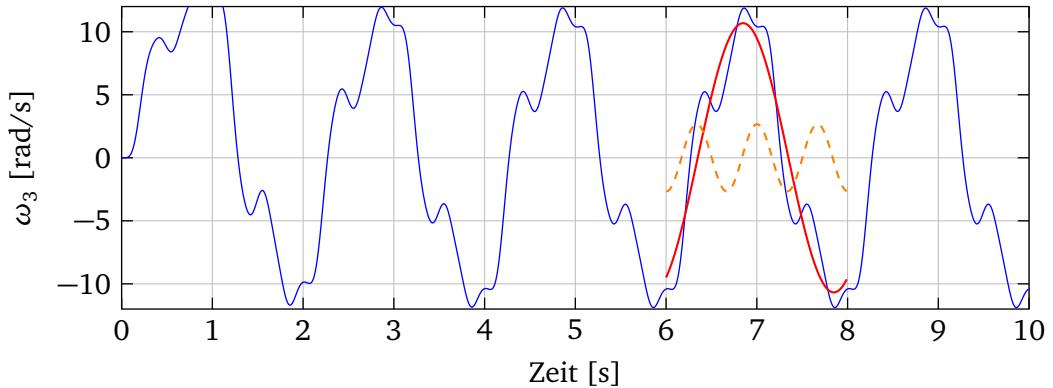
Es ist hier genauso wie im Fall der Sinusanregung das Abklingen der transienten Vorgänge abzuwarten. Vom Ein- und Ausgangssignal wird dann von $t = 6$ s bis 8 s, also über eine Periode, die zur Grundfrequenz gehörenden Fourierkoeffizienten \check{u}_1 bzw. \check{y}_1 bestimmt. Die dazu gehörenden Zeitverläufe sind in den Graphen in rot eingezeichnet. Aus diesen kann dann einfach der dazugehörige Werte $H(\omega_0)$ berechnet werden. Ist eines oder beide Signale verrauscht, so werden wieder Schätzwerte für \check{u}_1 sowie \check{y}_1 und daraus dann ein Schätzwert für $H(\omega_0)$ bestimmt.

Bei einer Rechteckschwingung wird neben der Grundharmonischen z. B. auch die dritte Harmonische zur Frequenz $3\omega_0$ angeregt. Aus den Daten kann ebenfalls für diese die Fourierkoeffizienten, \check{u}_3 und \check{y}_3 , berechnet werden. Die dazugehörigen Schwingungen sind gestrichelt und in orange eingezeichnet. Somit kann auch $H(3\omega_0)$ bestimmt werden.

Jedoch nehmen die Amplituden der höheren angeregten Frequenzen schnell ab, so dass eine Rechteckschwingung im Wesentlichen nur zu Identifikation der Grundfrequenz geeignet ist.



(a) Eingangsgröße



(b) Ausgangsgröße

Abbildung 7.8: Anregung mit Rechteckschwingung

Breitbandige periodische Signale

Geht man den eingeschlagenen Weg weiter, so bietet sich

$$\hat{H}(\omega_l) = \frac{\hat{y}_l}{\hat{u}_l} = \frac{\frac{1}{N \cdot T_0} \int_0^{N \cdot T_0} \tilde{y}(t) \cdot e^{-j \cdot l \cdot \omega_0 t} dt}{\frac{1}{N \cdot T_0} \int_0^{N \cdot T_0} \tilde{u}(t) \cdot e^{-j \cdot l \cdot \omega_0 t} dt}$$

als allgemeiner Schätzer bei der Anregung mit periodischen Signalen an. Dieser Schätzer basiert wieder auf dem Quotienten der Fourierkoeffizienten, jedoch wird hier nicht nur der Frequenzgang bei der Grundfrequenz $\omega_1 = 1 \cdot \omega_0$ bestimmt, sondern bei allen Harmonischen $\omega_l = l \cdot \omega_0$, für die der Zähler, d. h. die Eingangsanregung, hinreichend groß ist. (Was unter „hinreichend groß“ zu verstehen ist, hängt neben den Anforderungen an die Schätzgüte und dem Betrag des Fourierkoeffizienten \hat{u}_l auch von der Systemverstärkung und der Leistungsdichte des Rauschens an der jeweiligen Frequenz ab.)

Das Eingangssignal $u(t)$ sollte eine möglichst gute Anregung ermöglichen, d. h. es sollte hohe Beträge für die Fourierkoeffizienten in dem interessierenden Frequenzbereich besitzen. Beispiele für solche Signale sind im Kapitel 6 vorgestellt worden.

Im Weiteren wird hier als Beispiel der periodisch verwendete Gleitsinus verwendet. Die dabei gemachten Beobachtungen lassen sich auf alle periodischen Anregungssignale sinngemäß übertragen.

Der Gleitsinus wird wie in Abbildung 7.9a dargestellt direkt hintereinander auf das System gegeben. Die ungestörte Systemantwort ist in Abbildung 7.9b gezeigt. Die ersten beiden Perioden dienen dabei zum Einschwingen des Systems und werden für die Identifikation nicht verwendet.³

³ Aufgrund des sanften Beginns ist das Einschwingen hier auch deutlich früher abgeschlossen. Jedoch sollte hier zum Vergleich mit der Einzelfrequenzmessung die gleiche Dauer von 15 s verwendet werden. Das Intervall hätte damit auch

Ab der dritten Periode werden die Daten über eine ganze Anzahl an Perioden aufgezeichnet. In Abbildung 7.9 ist nur eine Periode verwendet, das entsprechende Intervall ist gelb hinterlegt.⁴

Die Eigenschaften des Gleitsinus sind wie folgt gewählt. Die Grenzfrequenzen sind $\omega_a = 2\pi \cdot 0$ und $\omega_b = 2\pi \cdot 8$ Hz, um von möglichst niedrigen Frequenzen bis ungefähr 7,5 Hz (47 rad/s) eine gute Anregung zu erhalten. Der Gleitsinus wird über 0,1 s linear ausgeblendet. Bei der Wahl von $\omega_a = 0$ ist der Beginn schon sanft, und eine Verminderung der Anregung bei den niedrigen Frequenzen durch ein Einblenden nicht wünschenswert oder nötig. Die Frequenzauflösung $\Delta f = \frac{1}{T_{ab}}$ sollte hier 0,1 Hz nicht überschreiten, um die Resonanzen noch genügend aufzulösen. (Genau genommen ist dies hier schon relativ knapp gewählt.) D. h. $T_{ab} \geq 10$ s. Der Wert von $T_{ab} = 12$ s wurde gewählt, da dies der nächste Wert ist, bei dem der Gleitsinus „phasenrichtig“ periodisch fortgesetzt wird. (Zu einem besseren ästhetischen Empfinden bringt dies noch eine etwas geringere Belastung der Mechanik sowie ein etwas glatteres Spektrum, da die Unstetigkeiten in den höheren Ableitungen geringer sind.)

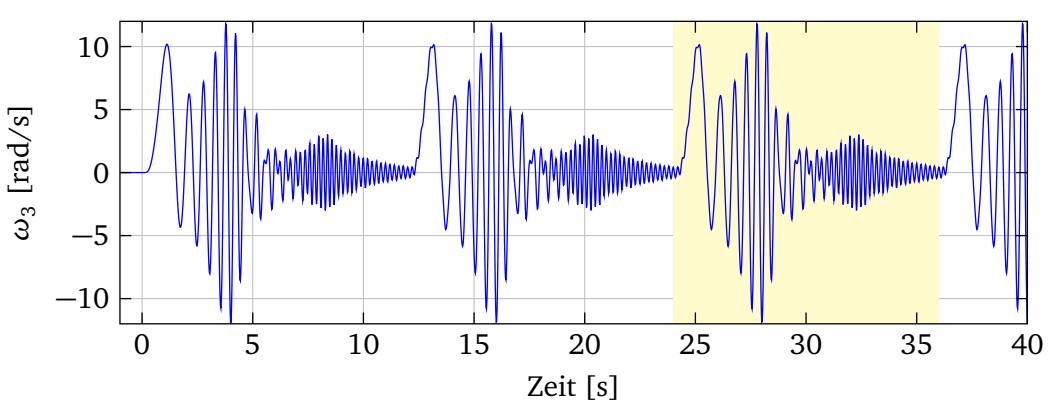
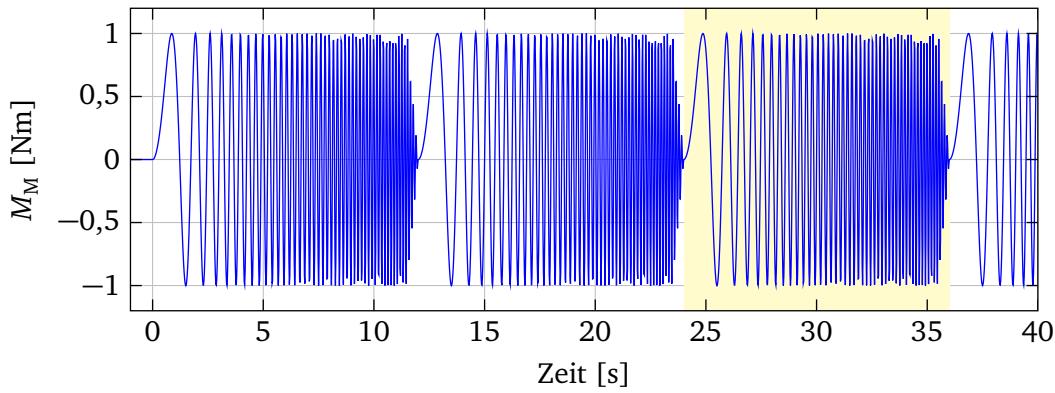


Abbildung 7.9: Anregung mit periodischem Gleitsinus

Die Ergebnisse der darauf basierenden Schätzung über eine, zehn und fünfzig Perioden unter Berücksichtigung des gleichen Rauschens wie oben sind in Abbildung 7.10 dargestellt. Dabei sind die Werte für $l = 1, \dots, 90$, also die ersten 90 Harmonischen gezeigt.

Dafür ist die Messzeit deutlich geringer. Diese beträgt hier nur 36 Sekunden, 144 Sekunden bzw. 624 Sekunden. Bei Berücksichtigung nur einer Periode ($N = 1$) ist das Ergebnis auch bei niedrigeren Frequenzen schon etwas ungenau. Mit $N = 10$ und $N = 50$ werden aber zumindest die Resonanzen und der Bereich niedriger Frequenzen schon gut geschätzt.

direkt von 15 s bis 27 s gehen können. Mathematisch macht dies keinen Unterschied. Lediglich aus rein optischen Gründen wurde der Beginn des Intervalls zur Auswertung auf 24 s gelegt.

⁴ Es könnte natürlich auch direkt nach dem Ende der dritten Periode das Eingangssignal abgeschaltet werden.

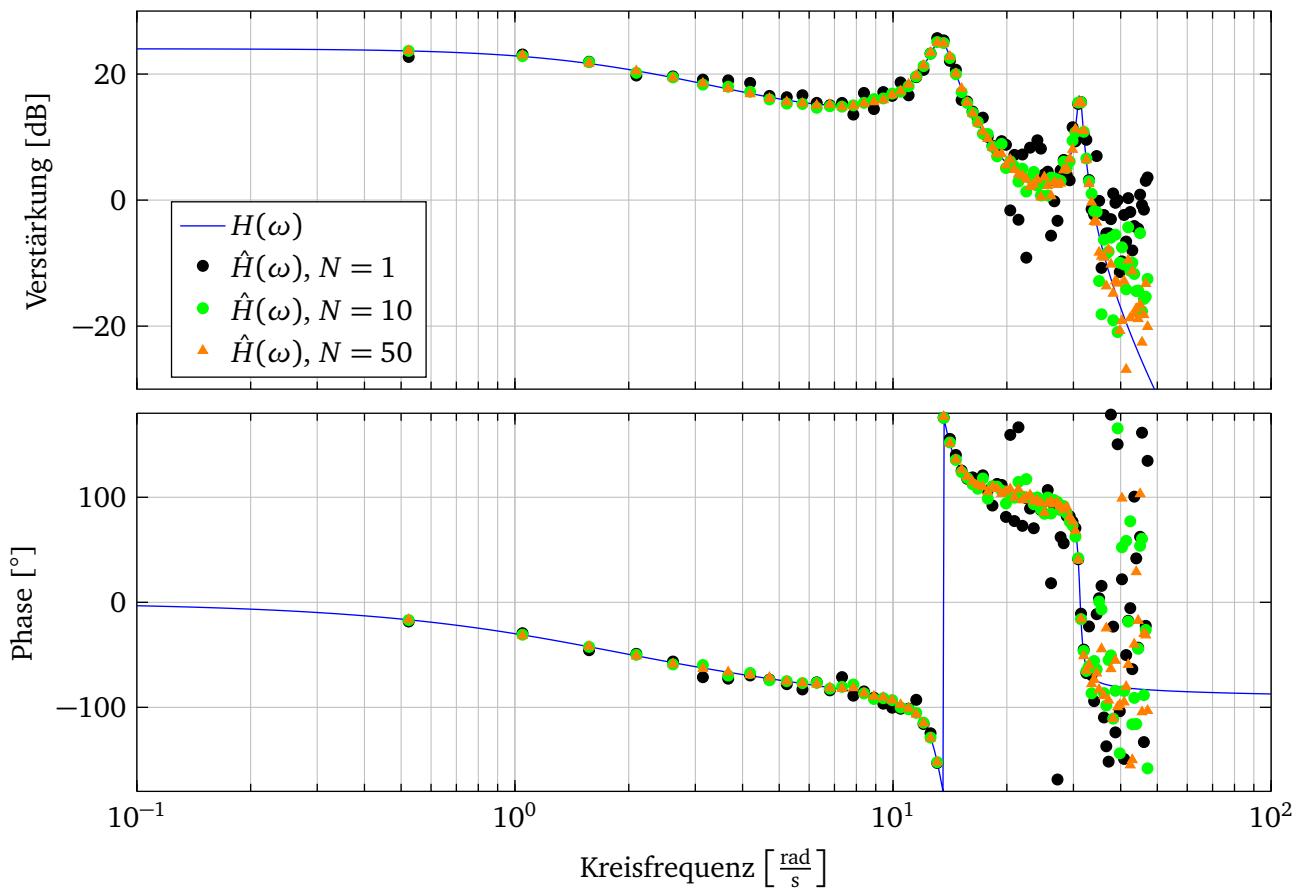


Abbildung 7.10: Frequenzgangsschätzung mit periodischem Gleitsinus

Die Nettomesszeit bei $N = 50$ beträgt hier 600 Sekunden. Dies entspricht etwa der Nettomesszeit bei der Einzelpunktmessung mit $N = 10$. (36 Minuten minus die Einschwingzeit von 1500 Sekunden ergibt 660 Sekunden.) Damit ist die eingebrachte mittlere Signalleistung in beiden Fällen ähnlich. Für die mittleren Frequenzen entspricht sich die eingebrachte Signalenergie daher ungefähr in diesen beiden Fällen. Bei niedrigen Frequenzen ist weniger Signalenergie in das System eingebracht, bei hohen Frequenzen mehr. (Im Fall des Gleitsinus ist die Gesamtsignalenergie im Frequenzband dabei etwas geringer als bei der Einzelanregung, da auch Frequenzen oberhalb der hier dargestellten maximalen Frequenz angeregt werden.)

7.3.3 Schätzung über Fouriertransformierte

Der letzte Schritt führt von den periodischen, multifrequenten Signalen und die Identifikation über die Fourierkoeffizienten zu den zeitbegrenzten Signalen. In diesem Fall beruht die Schätzung auf den Fouriertransformierten,

$$\hat{H}(\omega) = \frac{\hat{Y}(\omega)}{\hat{U}(\omega)}.$$

Da bei „sauber“ gefensterten periodischen Signalen eine Verbindung zu der Fouriertransformierten aufgezeigt werden kann, können alle bisher genannten Fälle als Identifikation mit der Fouriertransformierten aufgefasst werden. Damit braucht der Einfluss von Fehlern nur für die Schätzung über die Fouriertransformierten behandelt werden.

Berücksichtigung von mehreren Einzelmessungen

Werden R Messungen durchgeführt, so können diese auf verschiedene Arten zusammengeführt werden. Zum einen kann zu jeder Messung r der Schätzwert

$$\hat{H}_r(\omega) = \frac{\tilde{Y}_r(\omega)}{\tilde{U}_r(\omega)}$$

mit den aus den gemessenen Daten bestimmten Amplitudendichten bestimmt werden. Diese Schätzwerte können dann gemittelt werden,

$$\hat{H}(\omega) = \frac{1}{R} \cdot \sum_r \hat{H}_r(\omega).$$

Dieses Vorgehen sollte allerdings nur dann genutzt werden, wenn die Eingangsgröße gar nicht oder nur sehr gering verrauscht ist. Dann hat dies jedoch den Vorteil, dass die Eingangsgrößenverläufe der einzelnen Messungen verschieden sein können. Zudem kann über

$$\begin{aligned}\hat{\sigma}_{\text{Re}(\hat{H}_r)}^2(\omega) &= \frac{1}{R-1} \cdot \sum_r (\text{Re}(\hat{H}_r(\omega)) - \text{Re}(\hat{H}(\omega)))^2 \quad \text{und} \\ \hat{\sigma}_{\text{Im}(\hat{H}_r)}^2(\omega) &= \frac{1}{R-1} \cdot \sum_r (\text{Im}(\hat{H}_r(\omega)) - \text{Im}(\hat{H}(\omega)))^2\end{aligned}$$

die Varianz der Schätzwerte $\{\text{Re}(\hat{H}_r(\omega))\}_r$ und $\{\text{Im}(\hat{H}_r(\omega))\}_r$ geschätzt werden (wovon hierfür davon ausgegangen wird, dass diese für alle r gleich sind). Damit können dann Schätzwerte für die entsprechenden Varianzen

$$\begin{aligned}\hat{\sigma}_{\text{Re}(\hat{H})}^2(\omega) &= \frac{1}{R} \cdot \hat{\sigma}_{\text{Re}(\hat{H}_r)}^2(\omega) \quad \text{und} \\ \hat{\sigma}_{\text{Im}(\hat{H})}^2(\omega) &= \frac{1}{R} \cdot \hat{\sigma}_{\text{Im}(\hat{H}_r)}^2(\omega)\end{aligned}$$

des gemittelten Schätzwertes $\{\hat{H}(\omega)\}_r$ angegeben werden.

Ist auch die gemessene Eingangsgröße (zu stark) verrauscht, dann sollte die Mittelung von Zähler und Nenner einzeln erfolgen, d. h.

$$\hat{H}(\omega) = \frac{\hat{Y}(\omega)}{\hat{U}(\omega)}$$

mit

$$\begin{aligned}\hat{Y}(\omega) &= \frac{1}{R} \cdot \sum_r \tilde{Y}_r(\omega) \quad \text{und} \\ \hat{U}(\omega) &= \frac{1}{R} \cdot \sum_r \tilde{U}_r(\omega).\end{aligned}$$

Alternativ kann die Mittelung auch mit den Zeitdaten erfolgen,

$$\begin{aligned}\hat{Y}(\omega) &= \mathcal{F} \left(\frac{1}{R} \cdot \sum_r \tilde{y}_r(t) \right) \Big|_{\omega} \quad \text{und} \\ \hat{U}(\omega) &= \mathcal{F} \left(\frac{1}{R} \cdot \sum_r \tilde{u}_r(t) \right) \Big|_{\omega}.\end{aligned}$$

Aufgrund der Linearität der Fouriertransformation sind beide Varianten bezüglich des Ergebnisses äquivalent. Bei der Mittelung im Frequenzbereich könnten wieder die Varianzen (und Kovarianzen) des Real- und Imaginärteil der Einzelwerte $\{\tilde{Y}_r(\omega)\}_r$ und $\{\tilde{U}_r(\omega)\}_r$ und damit letztlich auch die Varianzen der gemittelten Werte $\{\hat{Y}(\omega)\}_r$ und $\{\hat{U}(\omega)\}_r$ geschätzt werden. Aus diesen kann man dann wieder einen Näherungswert für die Varianz von $\{\hat{H}(\omega)\}_r$ erhalten.

Bei der separaten Mittelung von Zähler und Nenner muss immer der gleiche Eingangsgrößenverlauf verwendet werden. Wenn dies nicht sichergestellt werden kann, und Störungen bei der Eingangsmessung die Mittelung der Quotienten nicht ratsam erscheinen lässt, können der H1- oder H2-Schätzer, die später vorgestellt werden, verwendet werden.

Beispiel zur Identifikation mit zeitbegrenzten Signalen – Gleitsinus

In Abbildung 7.11a ist ein zeitbegrenzter Gleitsinus dargestellt. Der mit diesem erzeugte (ungestörte) Ausgangsgrößenverlauf ist in Abbildung 7.11b gezeigt. Der gelb hinterlegte Bereich gibt wieder das Intervall an, welches für die Identifikation verwendet wird. Nach Beenden des Gleitsinus wird hier noch fünf Sekunden gewartet, bis die Systemantwort weitestgehend abgeklungen ist.

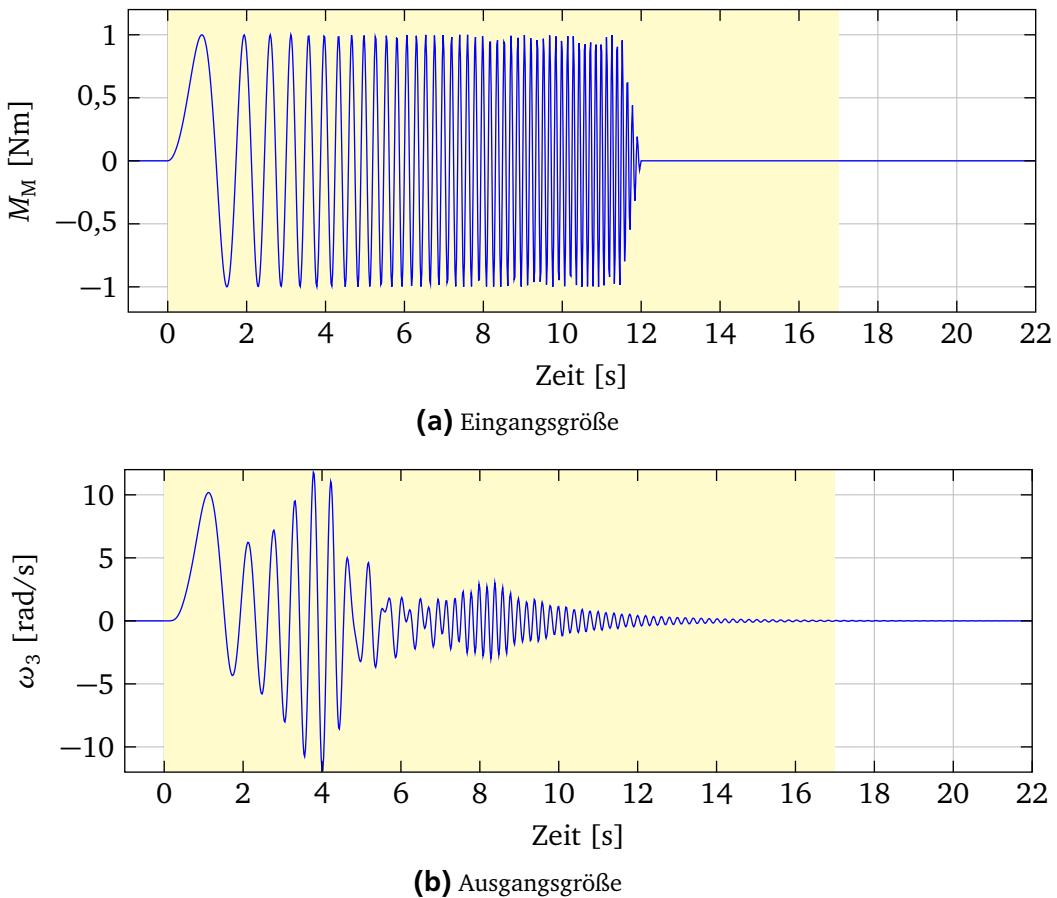


Abbildung 7.11: Anregung mit zeitbegrenztem Gleitsinus

Die Identifikation ist wieder mit dem selben Rauschen wie oben simuliert. Die Ergebnisse sind in Abbildung 7.12 dargestellt. Die Frequenzauflösung ist etwas feiner als zuvor, da man hier nicht an das durch die Grundfrequenz eines periodischen Signals gebunden ist, sondern mit Zeropadding auch feinere Auflösungen erreichen kann.

Die Ergebnisse zeigen eine ein wenig höhere Streuung im Vergleich zu der Anregung über den periodischen Gleitsinus. Die Messzeit ist dagegen etwas länger. Zwar muss im Falle des periodischen Gleitsinus

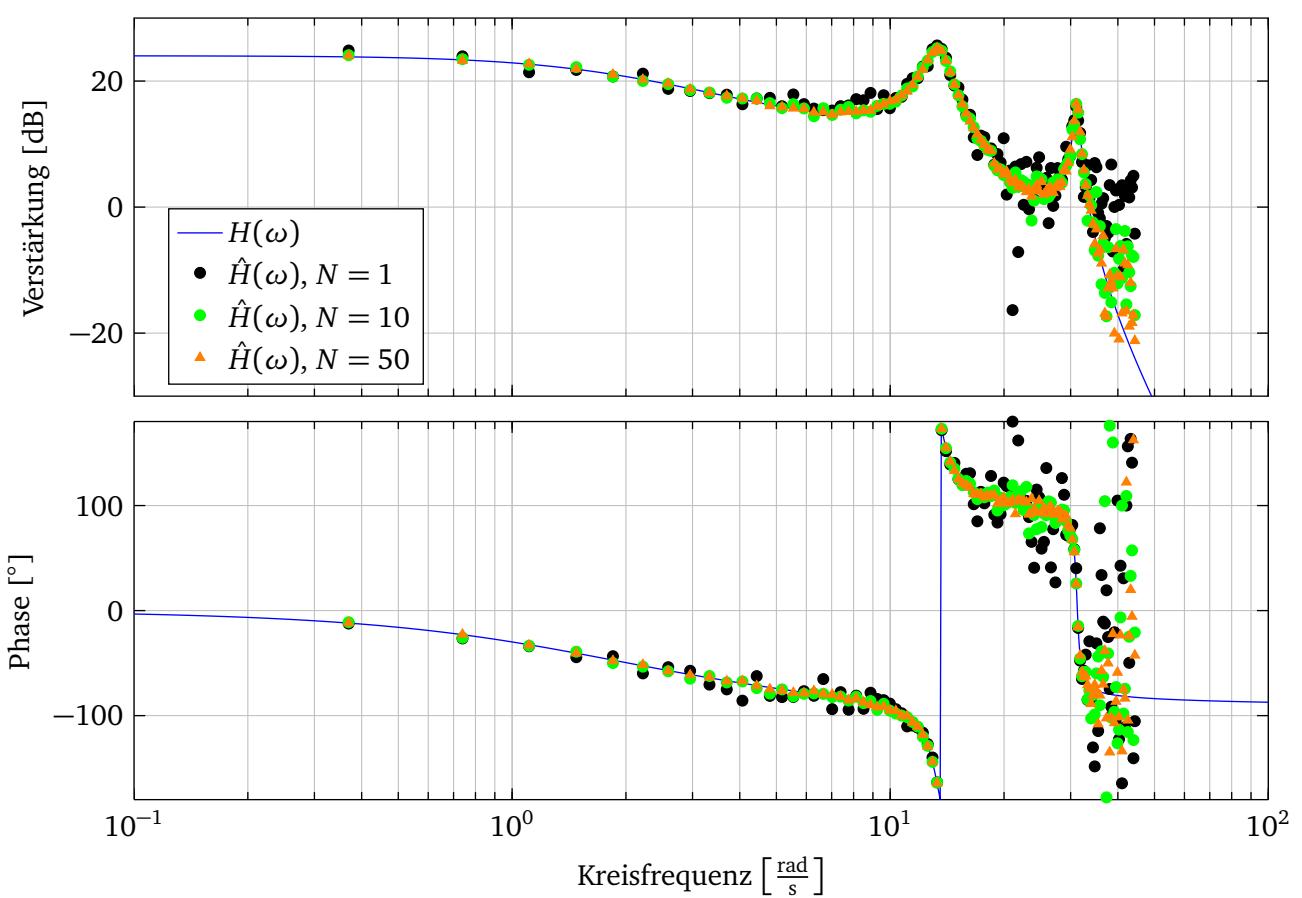


Abbildung 7.12: Frequenzgangsschätzung mit zeitbegrenztem Gleitsinus

das Einschwingen einmal abgewartet werden. Dies entfällt bei der Anregung über einen zeitbegrenzten Gleitsinus. Dafür muss hier N mal gewartet werden, bis das System wieder ausgeschwungen ist.

Beispiel zur Identifikation mit zeitbegrenzten Signalen – Rechteckimpuls

Als zweites zeitbegrenztes Signal wird ein Rechteckimpuls betrachtet. Das System soll bis ca. 7,5 Hz ausreichend angeregt werden. Daher wird die Impulsbreite so gewählt, dass die erste Nullstelle dessen Amplitudenspektrums bei 15 Hz liegt. Es ergibt sich eine Impulsbreite von 67 ms. Der Impuls ist in Abbildung 7.13a und die Ausgangsgröße (ungestört und gemessen) in Abbildung 7.13b gezeigt.

Nach Ende des Impulses wird noch 10 s gemessen. Wie diese Zeit hier gewählt wurde, wird später noch begründet. Zunächst ist in Abbildung 7.14 das Ergebnis der Identifikation gezeigt.

Es wird deutlich, dass mit dieser Kombination aus Eingangssignal und Rauschen kaum eine sinnvolle Schätzung des Frequenzganges möglich ist. Selbst bei $N = 50$ Messungen geht die zweite Resonanz völlig im Rauschen unter.

Daher sind in Abbildung 7.15 die Ergebnisse bei einem geringeren Rauschen gezeigt. Der Eingangsimpuls bleibt dabei unverändert. In Abbildung 7.15a ist die gemessene Ausgangsgröße mit dem geringeren Rauschen dargestellt. Abbildung 7.15b zeigt, dass dann auch mit dem Rechteckimpuls eine sinnvolle Identifikation möglich ist.

Hier könnte noch versucht werden, die Impulshöhe zu vergrößern. Dies wäre möglich, wenn die Begrenzung auf 1 Nm nicht aufgrund einer begrenzten Eingangsgröße, sondern aufgrund dessen gewählt wurde, dass die Ausgangsgröße einen bestimmten Bereich nicht überschreiten soll. Da aufgrund der

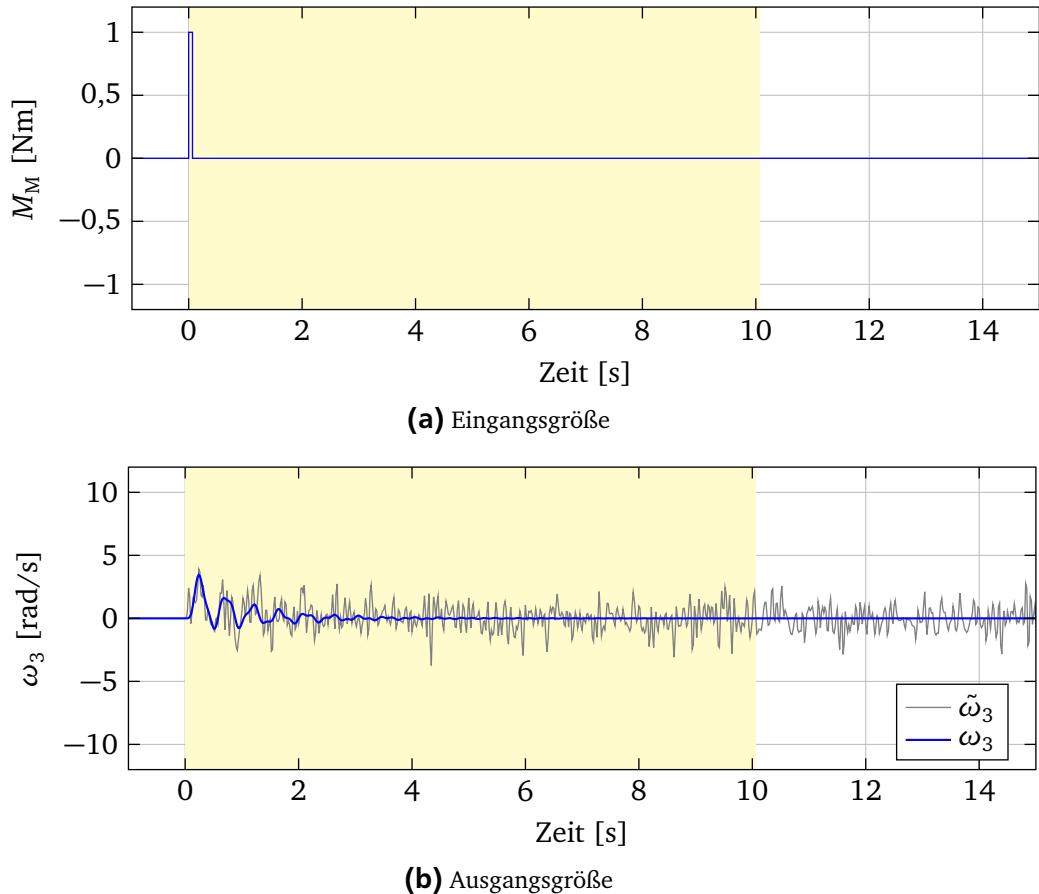


Abbildung 7.13: Anregung mit Rechteckimpuls

geringeren Anregung bei einem Impuls sich auch die Ausgangsgröße viel weniger ändert, könnte der Impuls entsprechend angehoben werden.

Zur Wahl der Wartezeit T_{post} von 10 s, die dem System nach Ende des Gleitsinus hier gegeben wurde um auszuschwingen, sind in Abbildung 7.16 Schätzwerte für $H(\omega)$ gezeigt, die sich ohne Rauschen für verschiedene Wartezeiten ergeben.

D. h. bei einer Wartezeit von 10 s hat der Schätzer theoretisch die Möglichkeit, die Werte des Frequenzgangs bis ca. 7,5 Hz genau zu schätzen. Betrachtet man jedoch die geschätzten Frequenzgänge mit Rauscheinfluss, so wird deutlich, dass hier andere Probleme vorliegen als das zu frühe Abschneiden der Messung. Schneidet man die Messungen – bei vorhandenem Rauschen, das ist hier nicht dargestellt – früher ab, so werden die Ergebnisse im unteren Frequenzbereich, der von der Abschneidproblematik weniger betroffen ist, besser. Dies wird später nochmal im Zusammenhang mit der Diskussion des Rauscheinflusses genauer diskutiert. Kurz gesagt liegt es daran, dass die längere Messdauer nicht mehr zur Informationsmenge bezüglich der unteren Frequenzen beiträgt, dafür aber das Rauschen weiter zu Fehlern bei allen Frequenzen beiträgt.

Dies hat aber weniger mit dem Gegensatz „niedrige Frequenzen – hohe Frequenzen“ zu tun. Bei dem vorliegenden System ist lediglich das Polpaar, welches zu der höheren Eigenfrequenz gehört, deutlich schwächer gedämpft. D. h. dies ist keine allgemeine Aussage, dass das Abschneiden bei einer Impulsanregung verstärkt zu Fehlern bei hohen Frequenzen führt. Eher zu Fehlern im Frequenzbereich der schwächer gedämpften Eigenschwingungen.

Praktisch, d. h. wenn man nicht die Möglichkeit hat das ungestörte System zu simulieren, steht man damit vor dem Problem, die Wartezeit gar nicht genau festlegen zu können. Das Rauschen mag die

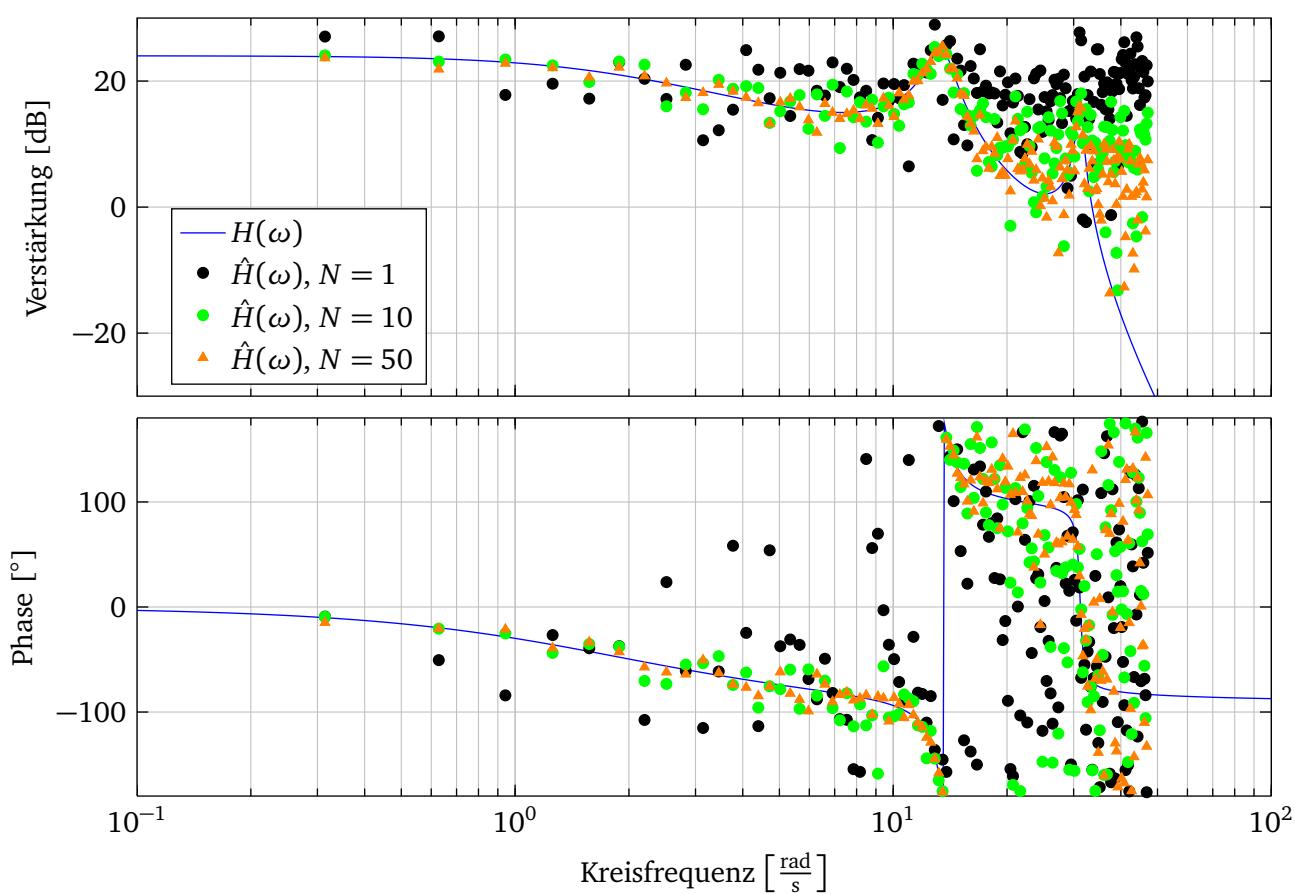
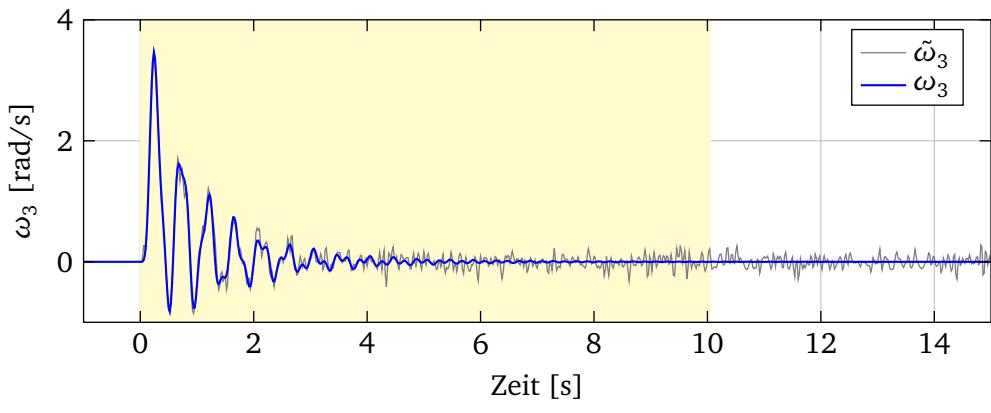


Abbildung 7.14: Frequenzgangsschätzung mit Rechteckimpuls

eigentliche Systemantwort schon lange verdecken, obwohl die Systemantwort noch einen wichtigen Beitrag leistet. Misst man zu kurz, so erhält man einen systematischen Fehler in der Schätzung. Misst man zu lange, so wächst der Einfluss des Rauschens und der Störungen auf die Schätzwerte.



(a) Ausgangsgröße

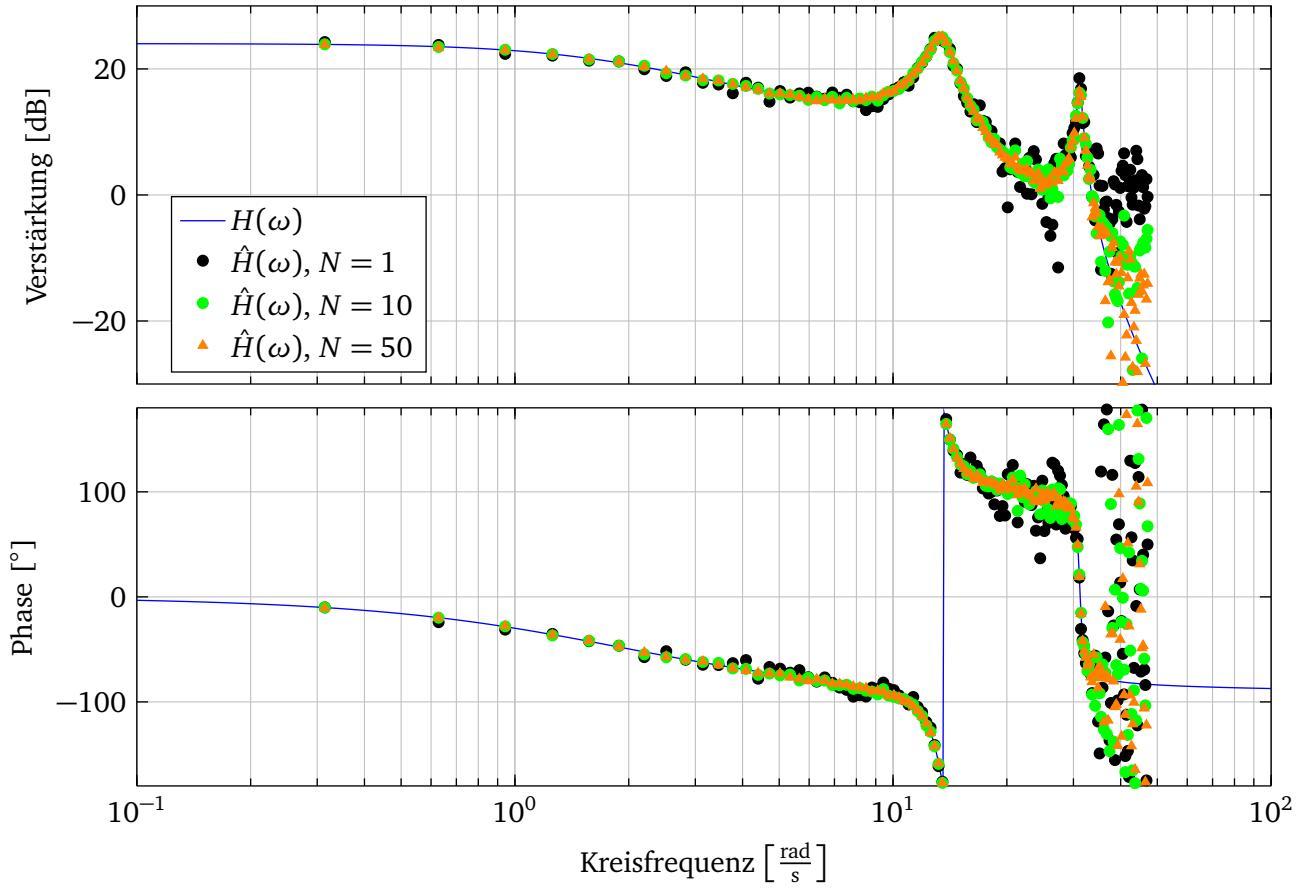


Abbildung 7.15: Anregung mit Rechteckimpuls, weniger Rauschen

Frequenzgangsschätzung mit Rechteckimpuls

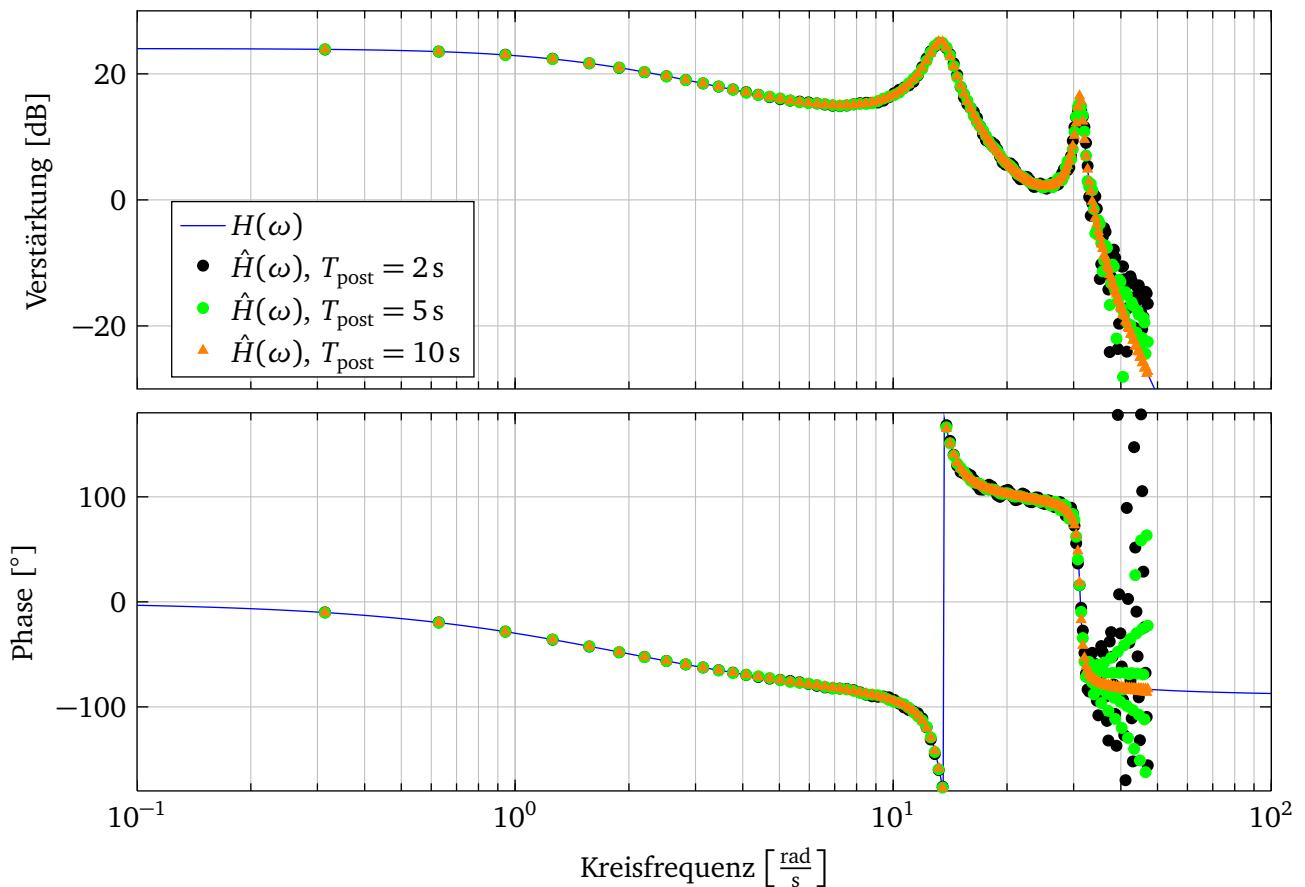


Abbildung 7.16: Frequenzgangsschätzung mit Rechteckimpuls, Vergleich verschiedener Wartezeiten, ohne Rauschen

Wirkung der Fensterung von Ein- und Ausgangsdaten

Die Rechteckfensterung der gemessenen Ein- und Ausgangsdaten kann nicht verhindert werden, da jede praktische Messung endlich ist. Bei periodischen Signalen sollte möglichst ein ganzes Vielfaches der Grundperiode erfasst werden, und bei zeitbegrenzten Signalen sollte gewartet werden, dass das Ausgangssignal bis auf vernachlässigbare Werte abgeklungen ist.

Diese Fensterung kann in der Darstellung des Messaufbaus explizit berücksichtigt werden, siehe Abbildung 7.17.

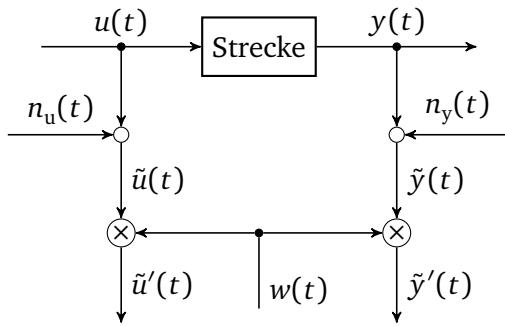


Abbildung 7.17: Fensterung der Messdaten

Die Multiplikation im Zeitbereich entspricht im Frequenzbereich der Faltung, und im Allgemeinen gilt

$$\frac{\tilde{Y}(\omega) * W(\omega)}{\tilde{U}(\omega) * W(\omega)} \neq \frac{\tilde{Y}(\omega)}{\tilde{U}(\omega)},$$

wobei „*“ die Faltungsoperation bezeichnet. D. h. auch wenn für Ein- und Ausgang das gleiche Fenster verwendet wird, fällt dieses nicht aus der Schätzung heraus. (Dies haben auch schon die Beispiele aus Abbildung 7.16 gezeigt, bei denen verschiedene Messdauern, d. h. Rechteckfenster verschiedener Breite, verglichen sind.)

In diesem Abschnitt soll die Wirkung anderer Fenster als das Rechteckfenster betrachtet werden. Zum einen wird das schon kurz im Kapitel 5 vorgestellte Hann-Fenster betrachtet und zum anderen das Tukey-Fenster

$$w(t) = \begin{cases} \frac{1}{2} \cdot \left(1 + \cos\left(\frac{2\pi}{r} \cdot \left(\frac{t}{T} - \frac{r}{2}\right)\right)\right) & 0 \leq \frac{t}{T} < \frac{r}{2} \\ 1 & \frac{r}{2} \leq \frac{t}{T} \leq 1 - \frac{r}{2} \\ \frac{1}{2} \cdot \left(1 + \cos\left(\frac{2\pi}{r} \cdot \left(\frac{t}{T} - 1 + \frac{r}{2}\right)\right)\right) & 1 - \frac{r}{2} \leq \frac{t}{T} \leq 1 \\ 0 & \text{sonst.} \end{cases}$$

Das Tukey-Fenster nimmt ähnlich wie das Trapezfenster in der Mitte konstant den Wert eins an und fällt nur in den Randbereichen ab. Während dieser Abfall beim Trapezfenster linear ist, wird beim Tukey-Fenster ein kosinusförmiger Verlauf verwendet, so dass die Übergänge noch etwas sanfter sind. Der Parameter r gibt beim Tukey-Fenster an, welcher Anteil der Gesamtbreite für beide Übergänge zur Verfügung steht. D. h. bei dem hier verwendeten $r = 0,2$ dauert der Anstieg 10 % der Gesamtdauer und der Abstieg ebenfalls.⁵

In Abbildung 7.18a sind die Fensterfunktionen dargestellt. Für die Beispiele zur Fensterung sind die Parameter zur Identifikation etwas verändert, um die Effekte deutlicher zu machen. So läuft der Gleitsinus mit 24 s doppelt so lang und auch nur bis 6 Hz. Die Messung wird mit dem Ende des Gleitsinus direkt beendet.

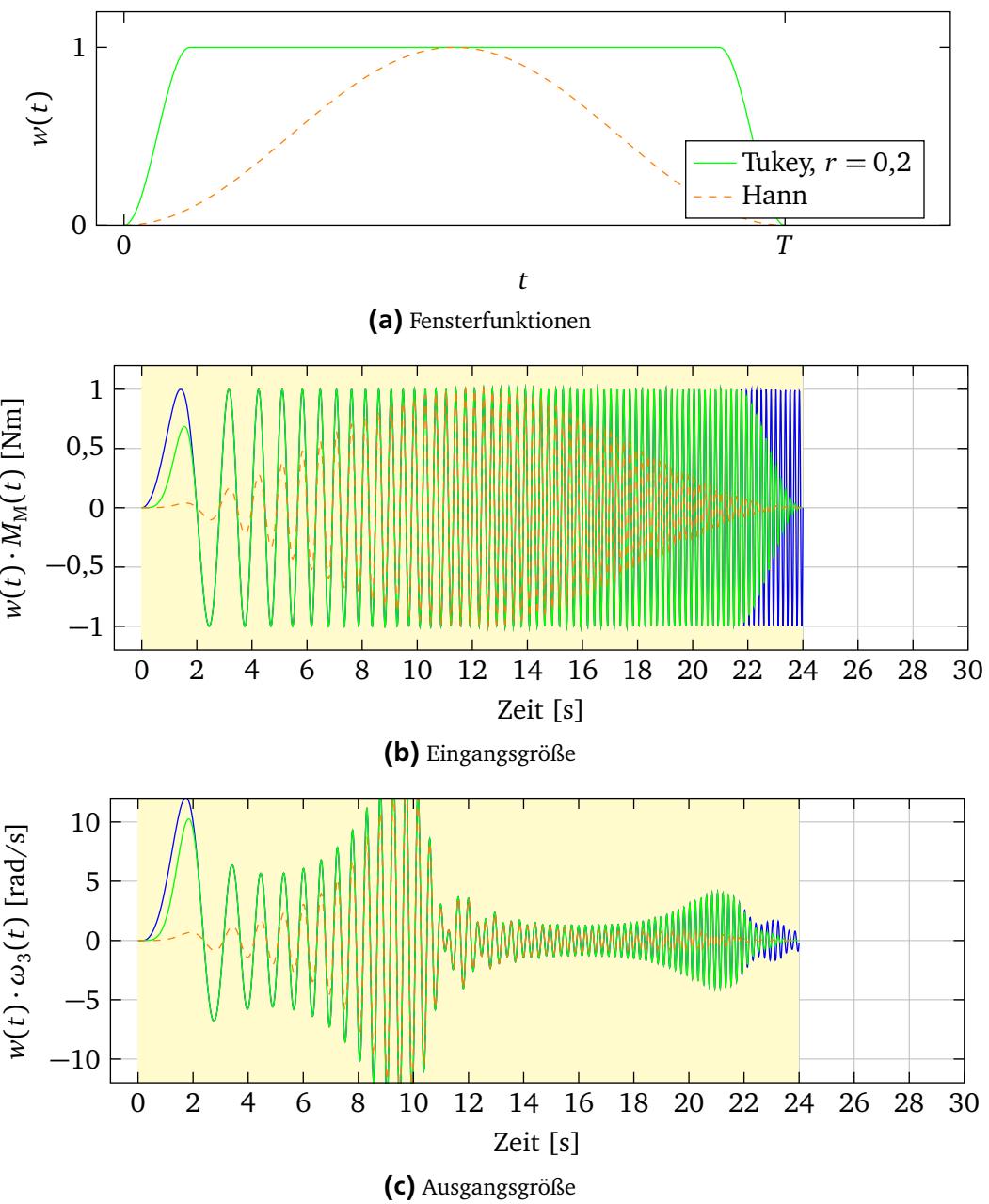


Abbildung 7.18: Anregung mit zeitbegrenztem Gleitsinus, verschiedene Fensterungen

Das (gefensterte) Eingangssignal und Ausgangssignal ist in Abbildung 7.18b bzw. (c) für beide Fensterfunktionen gezeigt.

Die Ergebnisse der Identifikation über eine Messung, jedoch ohne Rauschen, sind in Abbildung 7.19 gezeigt. Dabei ist neben dem Betrag der geschätzten Frequenzgänge hier nicht die Phase, sondern der relative Fehler der geschätzten Beträge dargestellt. Im Fall des Rechteckfensters (schwarze Punkte) zeigen sich bei niedrigeren Frequenzen kleinere Abweichungen. Bei Frequenzen ab 20 rad/s nehmen die Fehler dann deutlich zu. Dies ist darauf zurückzuführen, dass die Messung zu früh abgebrochen wurde.

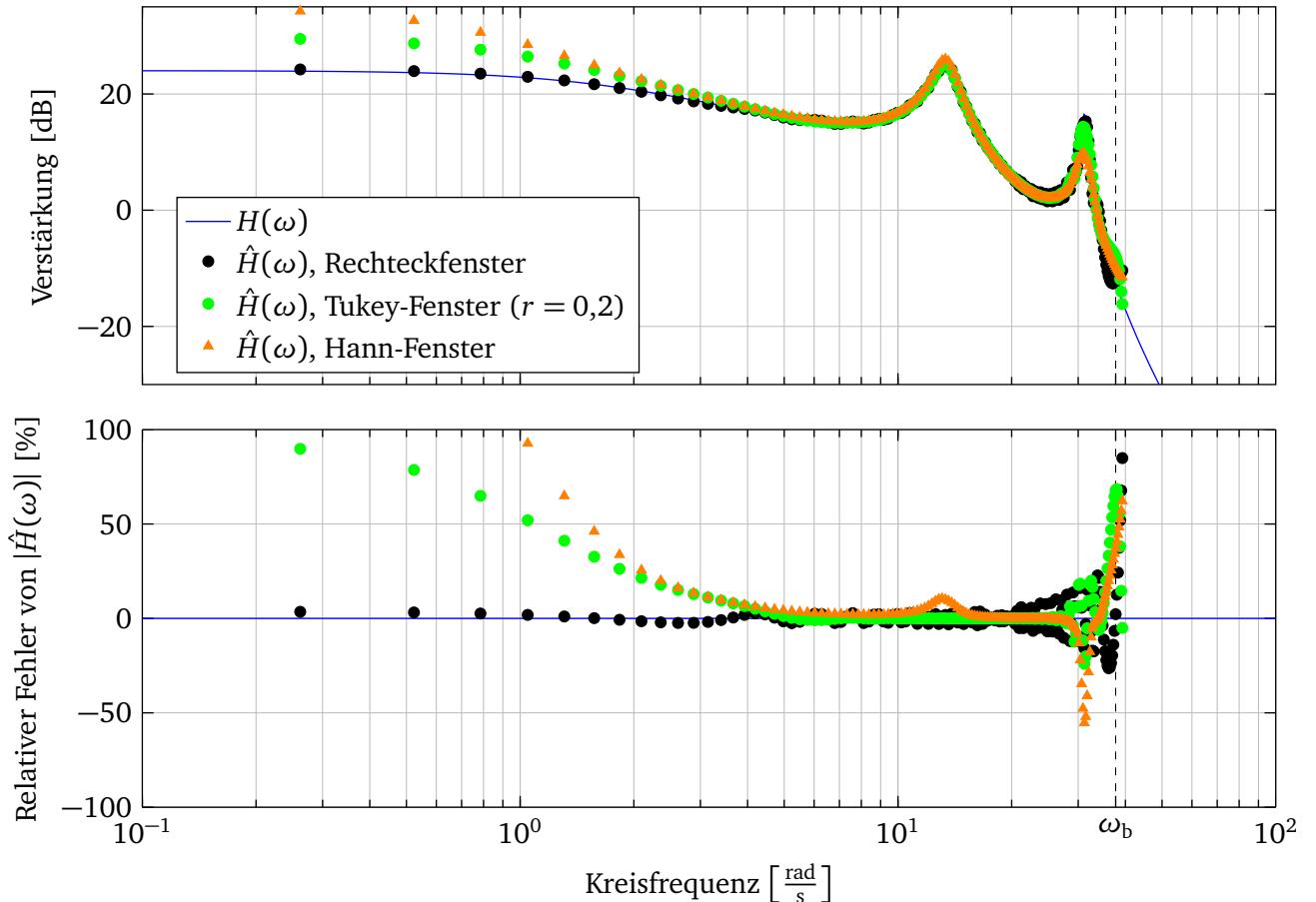


Abbildung 7.19: Frequenzgangsschätzung mit zeitbegrenztem Gleitsinus und verschiedenen Fensterfunktionen, ohne Rauschen

Beim Tukey-Fenster (grüne Punkte) wird zu Beginn der Frequenzgang betragsmäßig deutlich überschätzt. Dafür ist der Fehler im mittleren Frequenzbereich etwas geringer als bei Rechteckfenster. Bei hohen Frequenzen nimmt auch beim Tukey-Fenster der Fehler dann wieder zu, allerdings etwas später als beim Rechteckfenster.

Die mit dem Hann-Fenster durchgeführte Schätzung (orangene Dreiecke) überschätzt den Betrag des Frequenzgangs bei niedrigen Frequenzen deutlich. Im Bereich der zweiten Resonanz wird der Frequenzgang dann unterschätzt. Im Bereich um die Endfrequenz ω_b des Gleitsinus wird dann wieder überschätzt.

Diese Wirkung kann man insbesondere beim Gleitsinus gut anschaulich begründen, da hier eine relative gute Korrelation von Messzeit und angeregter Frequenz vorliegt. Zunächst werden die niedrigeren Frequenzen angeregt. Die Systemantwort darauf folgt etwas verzögert. Daher wird die Eingangsgröße mit einem kleineren Fensterwert multipliziert als die dazugehörige Systemantwort, und der Frequenzgang somit für diese Frequenzen überschätzt. Ab $t = 3$ s nimmt der Wert der Fensterfunktion beim

⁵ Das Hann-Fenster kann auch als Tukey-Fenster mit $r = 1$ aufgefasst werden.

Hann-Fenster wieder ab, d. h. für Frequenzen ab $\frac{6}{2} \text{ Hz} = 18,8 \text{ rad/s}$ gilt das Umgekehrte. Die Systemantwort, die der Erregung nacheilt, wird mit einem geringeren Fensterwert als die dazugehörige Anregung multipliziert und damit der Frequenzgang betragsmäßig unterschätzt. (Bei den ganz hohen Frequenzen gilt der näherungsmäßige Zusammenhang zwischen Messzeit und angeregtem Frequenzbereich nicht mehr.⁶⁾

Bei einem Gleitsinus kann die Wirkung noch einigermaßen abgeschätzt werden. Bei Signalen, bei denen die Zuordnung von Messzeit und Frequenz nicht so offensichtlich ist, ist dies nicht ohne weiteres möglich.

An dem Beispiel erkennt man, dass eine Fensterung zu einer Verfälschung des geschätzten Frequenzganges führt. Man kann aber auch den Effekten, die durch eine zu kurze Messdauer (bzw. einem zu kurzen Rechteckfenster) entstehen etwas „entgegen“wirken. Dabei versucht man aber letztlich, einen Fehler durch einen anderen zu kompensieren, so dass dies mit Vorsicht zu verwenden ist. Speziell zur Verminderung des negativen Effektes eines zu frühen Abschneidens der Messung werden auch einseitige Fenster verwendet. Wenn eine Fensterung vorgenommen wird, sollte für das Ein- und Ausgangssignal das gleiche Fenster verwendet werden. Wenn die Änderung des Fensterwertes langsam bezüglich der Systemzeitkonstanten ist, hält dies den Einfluss in Grenzen.

Ein weiterer Einsatzzweck für die Fenster wird später im Zusammenhang mit periodischen Störungen diskutiert.

Bei kurzen impulsförmigen Signalen sollten keinesfalls (zweiseitige) Fenster verwendet werden, was das in Abbildung 7.20 gezeigt Beispiel zeigt. Hier werden alle in der Eingangsgröße vorhandenen Frequenzen, die der Schätzer „sieht“, durch das Fenster deutlich verringert, während der verzögerte Ausgang kaum oder gar nicht gedämpft wird. Folglich wird der gesamte Frequenzgang völlig überschätzt.

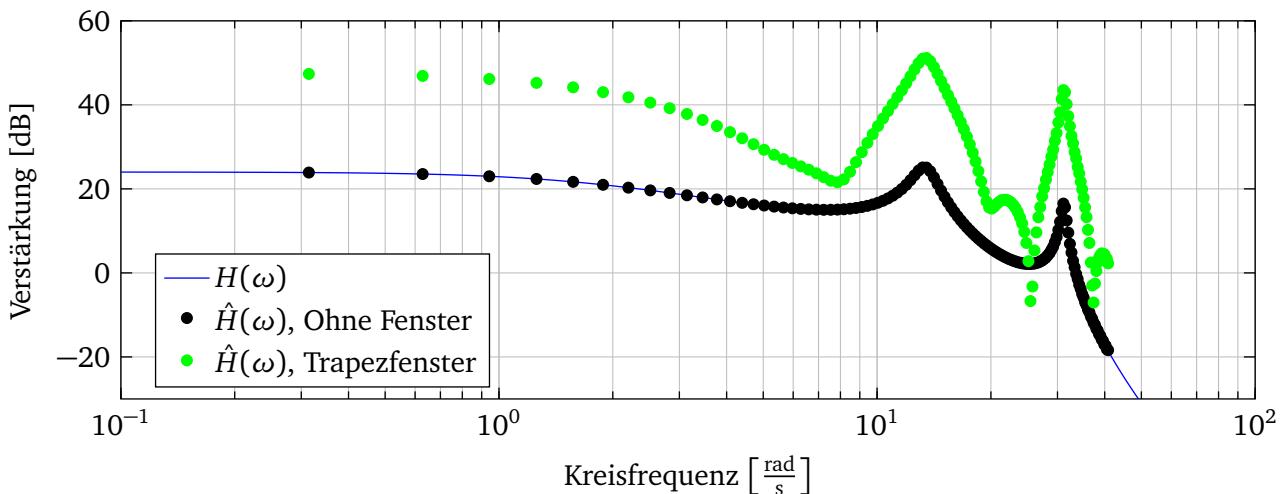


Abbildung 7.20: Anregung mit Impuls, Fensterung

7.3.4 Einfluss von Störungen

Es wird hier von der in Abbildung 7.21 gezeigten Konfiguration ausgegangen, bei der lediglich auf die Messung der Ausgangsgröße eine Störung wirkt. Damit kann auch auf den Index „y“ verzichtet werden, und die Ausgangsstörung wird kurz als $n(t)$ bezeichnet.

⁶ Vergleicht man die Spektren des normalen Gleitsinus mit dem ein- und ausgebenden Gleitsinus, so sieht man, dass auch die zeitlich „außen“ liegenden Bereiche einen Einfluss auf das gesamte Spektrum haben. Diese Zuordnung „Messzeit – Frequenz“ ist damit nur eine Näherung.

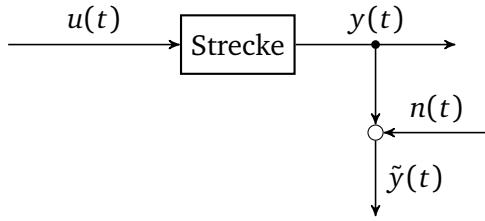


Abbildung 7.21: Betrachteter Fehler

Liegen Rauschen oder Störungen am Ausgang vor, d. h. $\tilde{y}(t) = y(t) + n(t)$, so ergibt sich für den geschätzten Frequenzgang $\hat{H}(\omega)$ bei einer Realisierung

$$\hat{H}(\omega) = \frac{\hat{Y}(\omega)}{U(\omega)} = \frac{Y(\omega)}{U(\omega)} + \frac{N(\omega)}{U(\omega)}.$$

Dabei ist $N(\omega)$ die unter Berücksichtigung des Messfensters der Breite T bestimmte Amplitudendichte. Der Schätzfehler $\Delta H(\omega)$ wird als

$$\Delta H(\omega) = \hat{H}(\omega) - H(\omega)$$

definiert. Damit gilt

$$\Delta H(\omega) = \frac{N(\omega)}{U(\omega)}.$$

Zunächst werden aus der letzten Gleichung für die Diskussion des Einflusses der Störungen geeignete Zusammenhänge abgeleitet. Dies wird getrennt für den Fall deterministischer und stochastischer Störungen gemacht. Im Anschluss daran werden die Einflüsse diskutiert.

Deterministische Störungen

Es wird der Betrag

$$|\Delta H(\omega)| = \frac{|D(\omega)|}{|U(\omega)|}$$

des Schätzfehlers betrachtet, wobei anstelle von $N(\omega)$ noch, analog zu Kapitel 6, $D(\omega)$ geschrieben ist. Alternativ – um eine dem stochastischen Fall ähnlichere Form zu erhalten – kann auch das Quadrat der Abweichung betrachtet werden,

$$|\Delta H(\omega)|^2 = \frac{|D(\omega)|^2}{|U(\omega)|^2}.$$

Stochastische Störungen

Zunächst wird der Ausdruck des Schätzfehlers mit $U^*(\omega)$ erweitert,

$$\Delta H(\omega) = \frac{N(\omega)U^*(\omega)}{|U(\omega)|^2},$$

um einfach dessen Real- und Imaginärteil angeben zu können,

$$\begin{aligned}\operatorname{Re}(\Delta H(\omega)) &= \frac{\operatorname{Re}(N(\omega)) \cdot \operatorname{Re}(U(\omega)) + \operatorname{Im}(N(\omega)) \cdot \operatorname{Im}(U(\omega))}{|U(\omega)|^2} \\ \operatorname{Im}(\Delta H(\omega)) &= \frac{\operatorname{Im}(N(\omega)) \cdot \operatorname{Re}(U(\omega)) - \operatorname{Re}(N(\omega)) \cdot \operatorname{Im}(U(\omega))}{|U(\omega)|^2}.\end{aligned}$$

In Kapitel 6 wurde gezeigt, dass $E\{\operatorname{Re}(N(\omega))\}_r = 0$ und $E\{\operatorname{Im}(N(\omega))\}_r = 0$ gilt, wenn das Rauschen mittelwertfrei ist. Damit folgt sofort

$$E\{\operatorname{Re}(\Delta \hat{H}(\omega))\}_r = 0$$

und

$$E\{\operatorname{Im}(\Delta \hat{H}(\omega))\}_r = 0.$$

D.h. $E\{\operatorname{Re}(\hat{H}(\omega))\}_r$ und $E\{\operatorname{Im}(\hat{H}(\omega))\}_r$, die Schätzer für den Real- sowie den Imaginärteil von $H(\omega)$, sind damit erwartungstreu. (Man spricht auch davon, dass die komplexe Zufallsvariable $\{\Delta \hat{H}(\omega)\}_r$ den Erwartungswert null hat, bzw. das $\{\hat{H}(\omega)\}_r$ erwartungstreu ist.)

Unter der Annahme, dass $E\{\operatorname{Re}(N(\omega))\}_r$ und $E\{\operatorname{Im}(N(\omega))\}_r$ unkorreliert sind, folgt

$$\operatorname{Var}\{\operatorname{Re}(\Delta \hat{H}(\omega))\}_r = \frac{(\operatorname{Re}(U(\omega)))^2}{|U(\omega)|^4} \cdot \operatorname{Var}\{\operatorname{Re}(N(\omega))\}_r + \frac{(\operatorname{Im}(U(\omega)))^2}{|U(\omega)|^4} \cdot \operatorname{Var}\{\operatorname{Im}(N(\omega))\}_r.$$

Diese Annahme ist für alle Frequenzen ω , die ein Vielfaches der Grundfrequenz $\frac{2\pi}{T}$ sind, exakt erfüllt (siehe Anhang C.1). Dies sind alle Frequenzen, die über die DFT berechnet werden, wenn kein Zero-padding verwendet wird. Für alle anderen Frequenzen ist diese Annahme zumindest asymptotisch mit steigender Messdauer T erfüllt.

Der erste Summand nimmt seinen größten Wert an, wenn $U(\omega)$ rein reell ist, und damit $(\operatorname{Re}(U(\omega)))^2 = |U(\omega)|^2$ gilt. Der zweite Summand wird maximal, wenn $U(\omega)$ rein imaginär ist. Damit kann die Abschätzung

$$\operatorname{Var}\{\operatorname{Re}(\Delta \hat{H}(\omega))\}_r \leq \frac{|U(\omega)|^2}{|U(\omega)|^4} \cdot \operatorname{Var}\{\operatorname{Re}(N(\omega))\}_r + \frac{|U(\omega)|^2}{|U(\omega)|^4} \cdot \operatorname{Var}\{\operatorname{Im}(N(\omega))\}_r$$

angegebenen werden. Daraus folgt

$$\operatorname{Var}\{\operatorname{Re}(\Delta \hat{H}(\omega))\}_r \leq \frac{1}{|U(\omega)|^2} \cdot (\operatorname{Var}\{\operatorname{Re}(N(\omega))\}_r + \operatorname{Var}\{\operatorname{Im}(N(\omega))\}_r)$$

bzw.

$$\operatorname{Var}\{\operatorname{Re}(\Delta \hat{H}(\omega))\}_r \leq \frac{E\{|N(\omega)|^2\}_r}{|U(\omega)|^2}.$$

Analog erhält man über

$$\operatorname{Var}\{\operatorname{Im}(\Delta \hat{H}(\omega))\}_r = \frac{(\operatorname{Re}(U(\omega)))^2}{|U(\omega)|^4} \cdot \operatorname{Var}\{\operatorname{Im}(N(\omega))\}_r + \frac{(\operatorname{Im}(U(\omega)))^2}{|U(\omega)|^4} \cdot \operatorname{Var}\{\operatorname{Re}(N(\omega))\}_r$$

die Abschätzung

$$\text{Var}\{\text{Im}(\Delta\hat{H}(\omega))\}_r \leq \frac{1}{|U(\omega)|^2} \cdot (\text{Var}\{\text{Re}(N(\omega))\}_r + \text{Var}\{\text{Im}(N(\omega))\}_r)$$

bzw.

$$\text{Var}\{\text{Im}(\Delta\hat{H}(\omega))\}_r \leq \frac{\text{E}\{|N(\omega)|^2\}_r}{|U(\omega)|^2}.$$

Wie auch in Kapitel 6 für $\{N(\omega)\}_r$ kann der Erwartungswert für $\{|\Delta\hat{H}|^2\}_r$ angegeben werden. Hier ergibt sich

$$\begin{aligned} \text{E}\{|\Delta\hat{H}(\omega)|^2\}_r &= \text{Var}\{\text{Re}(\Delta\hat{H}(\omega))\}_r + \text{Var}\{\text{Im}(\Delta\hat{H}(\omega))\}_r \\ &= \frac{(\text{Re}(U(\omega)))^2}{|U(\omega)|^4} \cdot \text{Var}\{\text{Re}(N(\omega))\}_r + \frac{(\text{Im}(U(\omega)))^2}{|U(\omega)|^4} \cdot \text{Var}\{\text{Im}(N(\omega))\}_r \\ &\quad + \frac{(\text{Re}(U(\omega)))^2}{|U(\omega)|^4} \cdot \text{Var}\{\text{Im}(N(\omega))\}_r + \frac{(\text{Im}(U(\omega)))^2}{|U(\omega)|^4} \cdot \text{Var}\{\text{Re}(N(\omega))\}_r \\ &= \frac{(\text{Re}(U(\omega)))^2 + (\text{Im}(U(\omega)))^2}{|U(\omega)|^4} \cdot \text{Var}\{\text{Re}(N(\omega))\}_r \\ &\quad + \frac{(\text{Re}(U(\omega)))^2 + (\text{Im}(U(\omega)))^2}{|U(\omega)|^4} \cdot \text{Var}\{\text{Im}(N(\omega))\}_r \end{aligned}$$

und weiter, da $(\text{Re}(U(\omega)))^2 + (\text{Im}(U(\omega)))^2 = |U(\omega)|^2$ ist,

$$\begin{aligned} \text{E}\{|\Delta\hat{H}(\omega)|^2\}_r &= \frac{\text{Var}\{\text{Re}(N(\omega))\}_r + \text{Var}\{\text{Im}(N(\omega))\}_r}{|U(\omega)|^2} \\ &= \frac{\text{E}\{|N(\omega)|^2\}_r}{|U(\omega)|^2}. \end{aligned}$$

Anders als für den Real- und Imaginärteil tritt hier eine Gleichheit und keine Abschätzung auf. Der Erwartungswert $\text{E}\{|\Delta\hat{H}(\omega)|^2\}_r$ stellt auch die Varianz der komplexen Zufallsvariable $\{\hat{H}(\omega)\}_r$ dar.

Diskussion der Störungen

Im Fall deterministischer Störungen gilt

$$|\Delta H(\omega)|^2 = \frac{|D(\omega)|^2}{|U(\omega)|^2}$$

und im Fall stochastischer Störungen

$$\text{E}\{|\Delta H(\omega)|^2\}_r = \frac{\text{E}\{|N(\omega)|^2\}_r}{|U(\omega)|^2}.$$

Der Einfluss einer Störung ist also dann gering, wenn der Betrag der Amplitudendichte $U(\omega)$ des Eingangssignals groß, und der Betrag der Amplitudendichte $D(\omega)$ der Störung bzw. dessen Erwartungswert klein ist.

Ist das Eingangssignal periodisch mit der Periodendauer T_0 und beträgt die Messdauer T ein ganzes Vielfaches dieser Periodendauer, $T = NT_0$, so gilt

$$U(l\omega_0) = NT_0 \cdot \check{u}_l = T \cdot \check{u}_l$$

für den Wert der Fouriertransformierten an den zu den Harmonischen gehörenden Frequenzstellen $l\omega_0$. D. h. für periodische Signale wächst $|U(\omega)|$ proportional zur Messdauer, bzw. $|U(\omega)|^2$ proportional zu T^2 .

Gleiches gilt für zeitbegrenzte Signale, die aus einer (endlichen) Aneinanderreihung eines Grundsignals bestehen, an den Frequenzen, die Vielfachen der Grundfrequenz entsprechen. Für die Frequenzen zwischen den Vielfachen der Grundfrequenz wächst die Amplitudendichte mit wachsendem T jedoch weniger bzw. nimmt letztlich sogar ab.

Wird bei einem Gleitsinus jedoch die Messzeit vergrößert, indem T_{ab} vergrößert wird, so wächst lediglich $|U(\omega)|^2$ proportional zu T , und $|U(\omega)|$ damit nur wie \sqrt{T} .

Bei zeitbegrenzten Signalen wird nach Ende des Signales noch weitergemessen, um die Systemantwort vollständig zu erfassen. In dieser Zeit erhöht sich der Wert von $|U(\omega)|$ nicht weiter, wohl aber der entsprechende Wert der Störung. Daher sollte diese Phase so kurz wie möglich gehalten werden.

Konstante Störung

Eine konstante Störung verfälscht bei periodischen Signalen immer nur den Wert bei $\omega = 0$, da alle anderen Frequenzen, die für ein periodisches Signal sinnvoll ausgewertet werden können (Zeropadding macht hier keinen Sinn) gerade mit Nullstellen der Amplitudendichte der Störung zusammenfallen.

Bei zeitbegrenzten Signalen können auch Frequenzen dazwischen ausgewertet werden. In Kapitel 6 wurde für $|D(\omega)|$ dabei eine obere Grenze angegeben, die nicht von der Messzeit T abhängt. Damit verbessert sich das Signal-Rausch-Verhältnis mit steigender Messdauer bei allen Eingangssignalen.

Periodische Störung

Es ist anschaulich, dass periodische Störungen das Ergebnis bei der Frequenz der Störung immer verfälschen, egal wie lange gemessen wird. Für alle anderen Frequenzen kann eine obere Grenze für den Wert von $|D(\omega)|$ angegeben werden, die unabhängig von T ist. Damit sinkt der (worst-case) Einfluss bei periodischen Anregungssignalen mit $1/T$, beim gedehnten Gleitsinus mit $1/\sqrt{T}$. (Bei Frequenzen nahe der Störfrequenz ist diese Abschätzung für $|D(\omega)|$ jedoch sehr konservativ. Dort sinkt der Einfluss tatsächlich langsamer.) Bei Frequenzen, die ein ganzes Vielfaches der Störfrequenz sind bzw. die als ganzes Vielfaches der Störfrequenz ausgedrückt werden können, wirkt sich die Störung nicht aus, solange die Messdauer ein ganzes Vielfaches der betrachteten Frequenz sowie der Störfrequenz beträgt.

Rampenförmige Störung

Bei einer rampenförmigen Störung gilt $|D(\omega)| \sim T$. Bei periodischen Anregungssignalen ist der Einfluss damit unabhängig von der Messdauer. Beim gedehnten Gleitsinus wächst der Einfluss mit \sqrt{T} , d. h. die Schätzung wird mit wachsender Messdauer schlechter!

Stochastische Störungen

Bei weißem Rauschen gilt

$$\text{E}\{|N(\omega)|^2\}_r = S_0 T ,$$

und für farbiges Rauschen, welches mit einem PT₁-Filter aus weißem Rauschen erzeugt ist, gilt

$$\text{E}\{|N(\omega)|^2\}_r \approx T \cdot 2 \cdot S_{nn}(\omega).$$

Bei periodischen Eingangssignalen oder zeitbegrenzten Eingangssignalen, bei denen ein Grundsignal öfters wiederholt wird, sinkt der Einfluss des Rauschens damit mit $1/T$ bezüglich der Varianz bzw. $1/\sqrt{T}$ bezüglich der Standardabweichung. Beim gedehnten Gleitsinus bleibt das Signal-Rausch-Verhältnis dagegen konstant!

Vergleich von periodischen und zeitbegrenzten Signalen

Die verschiedenen Varianten, ein gegebenen Signalverlauf als Anregung zur Identifikation zu verwenden ist am Beispiel eines Gleitsinus sind in Abbildung 7.22a gezeigt. Grundsignal ist dabei ein Gleisinus von $\omega_a = 0$ nach $\omega_b = 2\pi \cdot 1 \frac{1}{s}$ mit der Übergangszeit von $T_{ab} = 10\text{s}$. Dieses T_{ab} wird hier auch als „Grundperiode“ bezeichnet, auch wenn dies vom Begriff her eigentlich nur bei echten periodischen Signalen korrekt wäre.

Als Ein- oder Ausschwingzeit wird hier immer von $T_{post} = 10\text{s}$ ausgegangen. Eine zusätzliche Vor- und Nachbereitungszeit T_{prep} je Messung (z. B. für den Upload der Messdaten vom Messgerät auf einen Rechner und die (Re-)Initialisierung der Messung), die in der Praxis auch leicht die eigentliche Messdauer überschreiten kann, ist in dieser Darstellung nicht berücksichtigt.

Im obersten Plot von Abbildung 7.22a ist in dunkelblau ein einzelner Gleitsinus der Länge $T_{ab} = 10\text{s}$ dargestellt. In hellblau ist die zehnfache Wiederholung (das erste Vorkommen mitgezählt) dieses Signals gezeigt. Der Betrag des zu dem einfachen Gleitsinus gehörenden Amplitudendichtespektrums ist in Abbildung 7.22b ebenfalls in dunkelblau eingezeichnet.

Wird die Länge des Gleitsinus von $T_{ab} = 10\text{s}$ auf $T_{ab} = 100\text{s}$ verlängert („gedehnt“), so verzehnfacht sich dabei auch ungefähr die Signalenergie. Daraus folgt, dass sich die Amplitudendichte, die der Wurzel der Impuls-Autoleistungsdichte entspricht, $S_{uu}^I(\omega) = |U(\omega)|^2$, sich ungefähr um den Faktor $\sqrt{10} = 3,16$ vergrößert, bzw. gilt ungefähr $|U(\omega)| \sim \sqrt{T}$. Da das Signal nicht einfach eine Skalierung des ursprünglichen Signals ist, ändert sich auch die Form der Amplitudendichte etwas. Aber im Mittel stimmt diese Aussage. Dies kann auch in den Graphen nachvollzogen werden. Der gedehnte Gleitsinus ist im zweiten Plot von Abbildung 7.22a gezeigt. In Abbildung 7.22b ist dessen Amplitudendichtespektrum in rot dargestellt. Zum Vergleich ist in grau auch das $\sqrt{10}$ -fache des Spektrums des einfachen, kurzen Gleitsinus dargestellt.

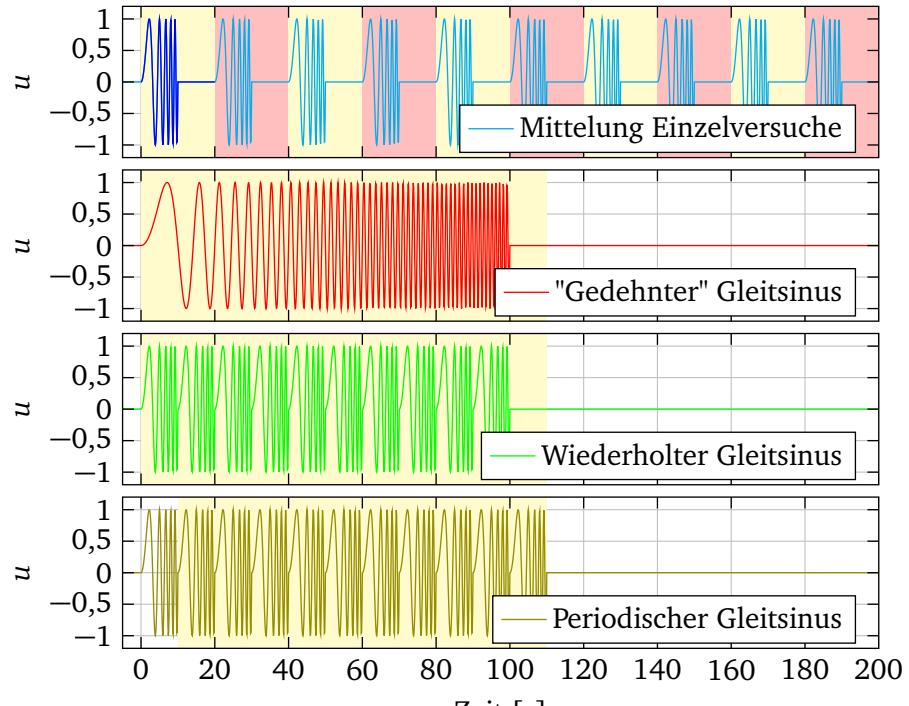
Daraus folgt, dass der Einfluss von stochastischen Störungen, für die auch $E\{|N(\omega)|^2\}_r \sim T$ gilt, nicht vermindert werden kann, indem der Gleitsinus „gedehnt“ wird, indem T_{ab} erhöht wird.

Verwendet man einen Gleitsinus der Länge $T_{ab} = 10\text{s}$ jedoch als periodisches Anregungssignal (unterster Plot in Abbildung 7.22a⁷), so bleiben die Fourierkoeffizienten dabei natürlich unverändert. Damit steigt die Signalleistung wegen $U(l\omega_0) = NT_0 \cdot \ddot{u}_l = T \cdot \ddot{u}_l$ mit T^2 an.

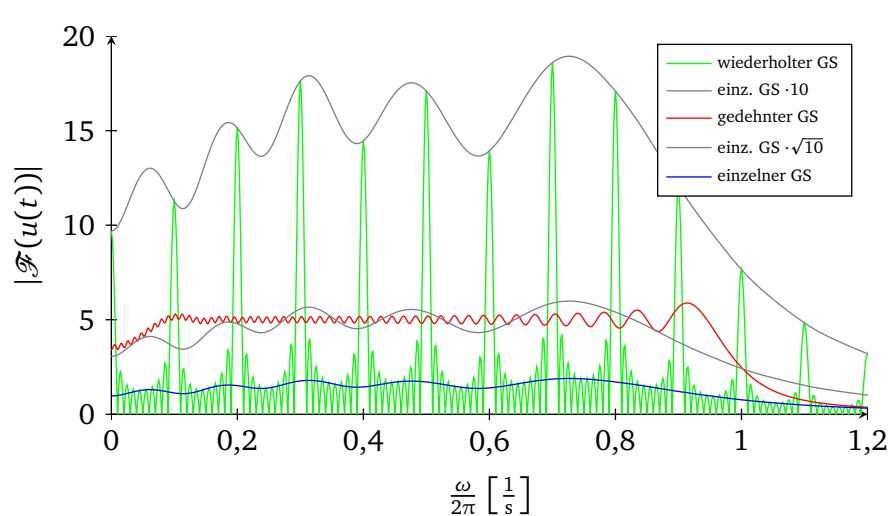
Dies ist jedoch kein Widerspruch. Vergleicht man die Fouriertransformierte eines einfachen Gleitsinus mit der eines zehnfachen (zeitbegrenzten) Gleitsinus, der im dritten Plot von Abbildung 7.22a dargestellt ist, so sind die Amplitudendichten an den Frequenzen, die den Vielfachen der Grundfrequenz eines Gleitsinus entsprechen auch genau zehnmal so groß, siehe Abbildung 7.22b. Die Leistungsdichte (oder hier eigentlich die Energiedichte) hat sich an diesen Frequenzen damit verhundertfacht. Dafür treten zwischen diesen Frequenzen auch Auslöschungen aus, an denen die Leistungsdichte geringer wird. Das Integral über die Leistungsdichte ergibt letztlich genau den zehnfachen Wert des ursprünglichen einfachen Gleitsinus.

Der Gleitsinus müsste daher auch nicht unbedingt als periodisches Signal verwendet werden. Genauso gut können zehn hintereinandergehängte Gleitsinusse auch als zeitbegrenztes Eingangssignal verwendet

⁷ Es ist zu beachten, dass hier keine Ausschwing- sondern eine Einschwingzeit abgewartet werden muss. Das Eingangssignal kann direkt nach der Erfassung des letzten Abastpunktes abgeschaltet werden. Dies ist hier auch gemacht, um die gesamte Messdauer besser zu veranschaulichen.



(a) Zeitverläufe



(b) Amplitudenspektren

Abbildung 7.22: Zu den verschiedenen Anregungsvarianten

werden. In diesem Fall können aber nicht mehr alle Frequenzen genutzt werden, sondern nur noch die Vielfachen der Grundfrequenz (bzw. Frequenzen, die nahe an diesen liegen). Es ist zwar nicht verboten, die Frequenzen dazwischen zu nutzen, nur ist die Anregung dort dann sehr schlecht.

Von der Messzeit macht dies keinen großen Unterschied. Werden die hintereinandergehängten Gleitsinusse als periodisches Signal verwendet, dann muss zu Beginn der Messung gewartet werden, bis die transienten Vorgänge des Systems abgeklungen sind. Bei der Verwendung als zeitbegrenztes Signal wird mit der Messung sofort begonnen, dafür muss nach Ende des letzten Gleitsinus so lange gewartet werden, bis das System zur Ruhe gekommen ist. In beiden Fällen entspricht dies in etwa der Zeit, für die die Impulsantwort des Systems wesentliche Anteile besitzt.

Bei der Verwendung von zehn aufeinanderfolgenden Gleitsinusen als zeitbegrenztes Signal muss allerdings genau die auszuwertenden Frequenzpunkte beachtet werden. Das Spektrum ist im Wesentlichen nur bei den Vielfachen der Grundfrequenz eines einzelnen Gleitsinus brauchbar. Dadurch, dass dem System aber Zeit gegeben werden sollte, nach dem letzten Gleitsinus noch auszuschwingen, wird im Allgemeinen die Messzeit kein Vielfaches der Dauer eines einzelnen Gleitsinus sein. Damit werden aber auch nicht die „guten“ Frequenzen mit der DFT getroffen. Das Signal muss also noch mit Nullen aufgefüllt werden, um auf ein Vielfaches zu kommen. (Einfach die Messung länger laufen lassen sollte man wegen der Abnahme des Signal-Rausch-Verhältnisses nicht.) Von den berechneten DFT sind dann nur die zu der Grundfrequenz eines Gleitsinus gehörenden Frequenzpunkte zu verwenden. (Wird beispielhaft das Messsignal auf die doppelte Länge der zehn Gleitsinusen aufgefüllt, so wird dann nur jeder zwanzigste Wert der DFT verwendet.) Es wird aber weniger fehleranfällig sein, den Gleitsinus einfach als periodisches Signal zu verwenden.

Die Länge des Grund-Gleitsinus sollte dabei möglichst kurz gewählt werden. Die untere Grenze der (sinnvoll) möglichen Länge ist dabei über die geforderte Frequenzauflösung gegeben.

Für die zeitbegrenzte Anregung kann man auch, anstatt die zehn Gleitsinusse in einer Messung auf das System zu geben, zehn Messungen mit nur einem Gleitsinus durchführen. Dabei vergrößert sich aber die Messdauer, da jedesmal das Ausschwingen des Systems abgewartet werden muss. Dafür können dann auch Werte für andere Frequenzen berechnet werden. Zudem können die Varianzen der Schätzwerte bestimmt werden.

Durch die Mittelung von R Messungen mit der (gleichen) Varianz $\text{Var}\{\hat{H}_r(\omega)\}_r$ gilt für die Varianz des gemittelten Wertes $\{\hat{H}(\omega)\}_r$

$$\text{Var}\{\Delta H(\omega)\}_r = \frac{1}{R} \cdot \text{Var}\{\hat{H}_r(\omega)\}_r,$$

sofern die Einzelschätzwerte $\{\hat{H}_r(\omega)\}_r$ unkorreliert sind. Dies ist eine vernünftige Annahme, wenn die Leistung des Rauschens im Wesentlichen oberhalb den Frequenzen wirkt, dessen Periodendauern im Bereich der Dauer einer Einzelmessung liegen. D. h. auch dann fällt die Varianz proportional zu $1/T$.

Der Vorteil der letzten Variante gegenüber dem gedeckten Gleitsinus wird also nicht dadurch erreicht, dass die Anregung auf bestimmte Frequenzen konzentriert wird, sondern dass es stochastisch günstiger ist, Einzelmessungen zu mitteln.

Wahl des Anregungssignals

- Die „Grundperiode“ des Anregungssignals (z. B. die Laufzeit T_{ab} eines Gleitsinus) sollte so klein wie möglich gewählt werden, wobei dies durch die gewünschte Frequenzauflösung gegeben ist.
- Es sollten so viele Grundperioden wie möglich gemessen werden.
- Bei der Anregung mit zeitbegrenzten Signalen können die Grundperioden entweder in einer Messung zusammengefasst, oder aber in Einzelmessungen auf das System gegeben werden. Von der

ersten Möglichkeit sollte zugunsten einer periodischen Verwendung des Anregungssignals abgesehen werden. Bei der zweiten Möglichkeit verlängert sich die Messdauer und verringert sich das Signal-Rausch-Verhältnis dann dadurch, dass jedesmal das Abklingen der Systemantwort abgewartet werden muss. Dafür kann über die Einzelmessungen die Varianz geschätzt werden.

- Bei zeitbegrenzten Signalen sollte die Messung so früh wie möglich abgebrochen werden, um das Signal-Rausch-Verhältnis hoch zu halten. Allerdings muss so lange gemessen werden, bis alle wesentlichen Anteile der Systemantwort abgeklungen sind. Bei Rauschen ist dies praktisch auch schwer zu sehen. Bricht man zu früh ab, so macht man einen systematischen Fehler. Bricht man zu spät ab, so wächst der stochastische Fehler weiter an. Letzterer wird bei der Analyse der Varianz berücksichtigt, ersterer nicht.

7.3.5 Umgang mit periodischen Störungen

In Abbildung 7.23 ist die gemessene Ausgangsgröße bei einer starken periodischen Störung gezeigt. Die eigentliche Systemantwort auf den Gleitsinus ist kaum noch zu erkennen.

Nutzt man dieses Signal zur Identifikation, so ergeben sich (mit einer Messung, ohne Rauschen) die in Abbildung 7.23b mit schwarzen Punkten dargestellten Schätzwerte, wenn die gleichen Parameter wie bisher verwendet werden.

Der eigentliche Verlauf des Frequenzganges ist zwar noch zu erkennen, jedoch von dem Spektrum des am Ausgang zusätzlich gemessenen und gefensterten Sinus überlagert. Der damit erhaltene Frequenzgang eignet sich nur bedingt für weitere Schritte.

Wenn es möglich ist, die Messdauer auf die periodische Störung abzustimmen, so lässt sich deren Einfluss deutlich reduzieren. Mit den grünen Punkten und den orangenen Dreiecken sind in Abbildung 7.23b Schätzwerte eingezeichnet, die erhalten werden, wenn die Messdauer ungefähr ein ganzes Vielfaches der Störfrequenz beträgt. Dabei ist in orange die Messdauer verwendet, die am nächsten an ein ganzes Vielfaches herankommt. In grün ist im Vergleich zu orange ein Abtastpunkt mehr hinzugekommen.

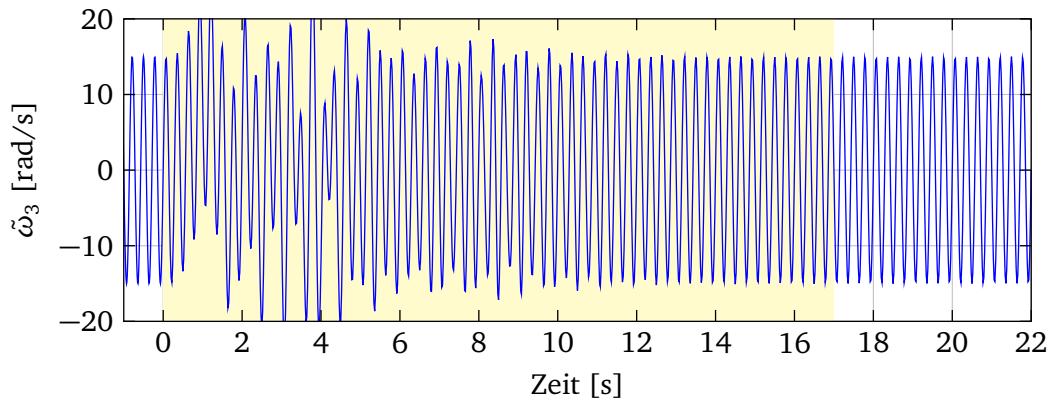
Bei den orangenen Schätzwerten ist nur ein einziger Wert sichtbar falsch geschätzt, nämlich der genau bei der Störfrequenz. Bei der grünen Schätzung sind auch um die Störfrequenz und bei ganz hohen Frequenzen noch ein paar schlechtere Werte dabei, aber im Vergleich zu den schwarzen Werten ist dies noch immer eine extreme Verbesserung.⁸

Es ist also sinnvoll, bei periodischen Störungen die Messdauer möglichst gut auf diese anzupassen. Zwei Gründe können dem aber entgegenstehen. Zum einen kann es bei mehreren periodischen Störungen sein, dass die Messdauer, die sich aus dem kleinsten gemeinsamen Vielfachen der Periodendauern dieser Störungen ergibt, zu groß wird, und zum anderen können Störungen, deren Frequenz über die Messdauer leicht variiert, im Spektrum etwas verschleifen.

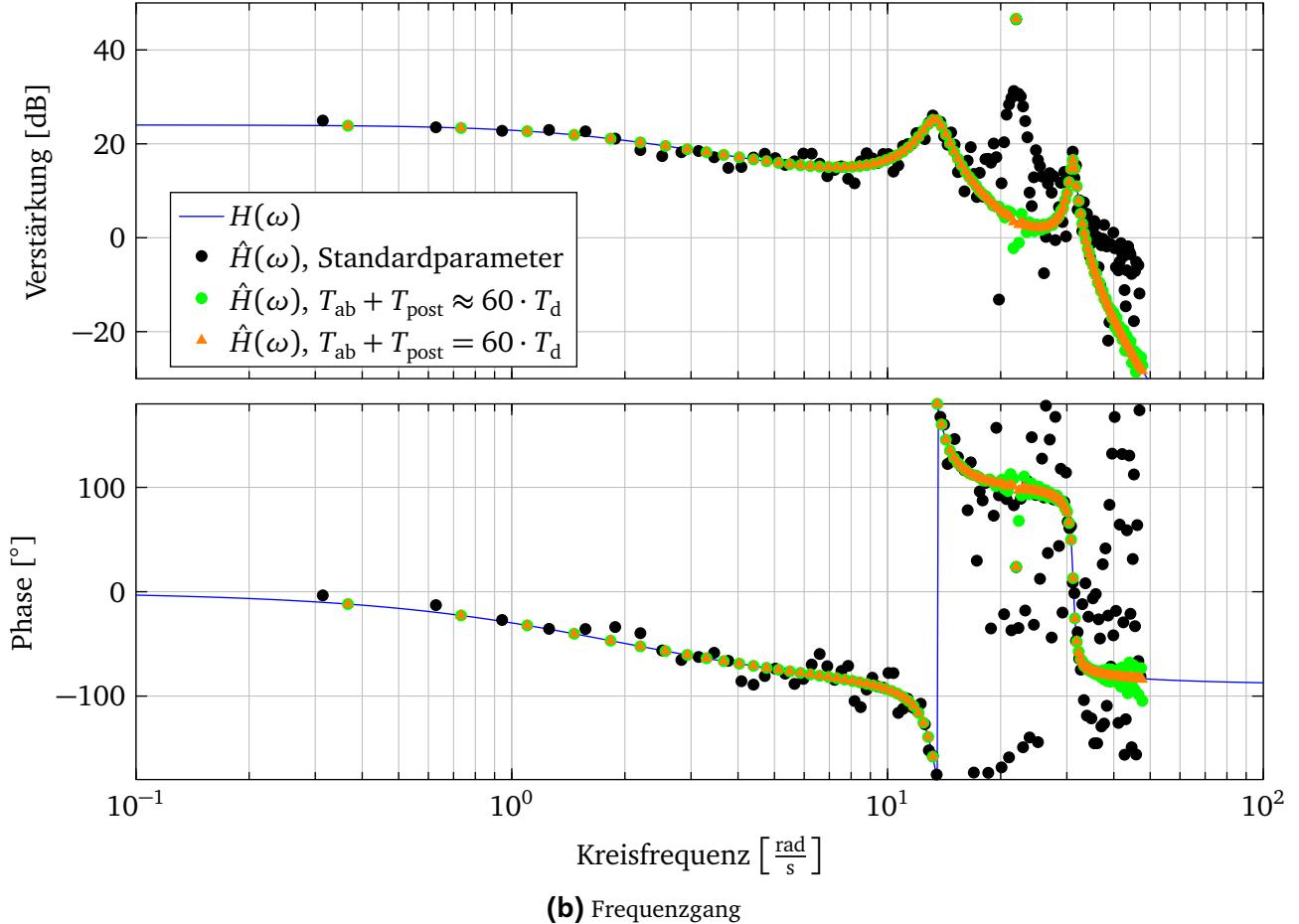
In diesem Fall kann eine Fensterung der gemessenen Ein- und Ausgangsdaten helfen.

In Abbildung 7.24 ist die Wirkung des Tukey-Fensters mit $r = 0,2$ und des Hann-Fensters im Vergleich zu dem Rechteckfenster demonstriert. Dabei ist bei dem Tukey-Fenster noch das Anregungssignal erst zwei Sekunden nach dem Starten der Messung gestartet, so dass die Fensterung möglichst außerhalb der Nutzdaten von Ein- und Ausgangsgrößenverlauf erfolgt. Daher wird der Frequenzgang mit dem Tukey-Fenster in allen Frequenzbereichen gut geschätzt (abgesehen von dem Einfluss der periodischen Störung) während beim Hann-Fenster der Frequenzgang bei niedrigen Frequenzen betragsmäßig über- und bei hohen Frequenzen unterschätzt wird.

⁸ Würde man hier mit Zeropadding die Auflösung erhöhen, würde man allerdings auch wieder die Störungen sehen können.



(a) Ausgangsgröße



(b) Frequenzgang

Abbildung 7.23: Anregung mit zeitbegrenztem Gleitsinus, starke periodische Störung, ohne Rauschen

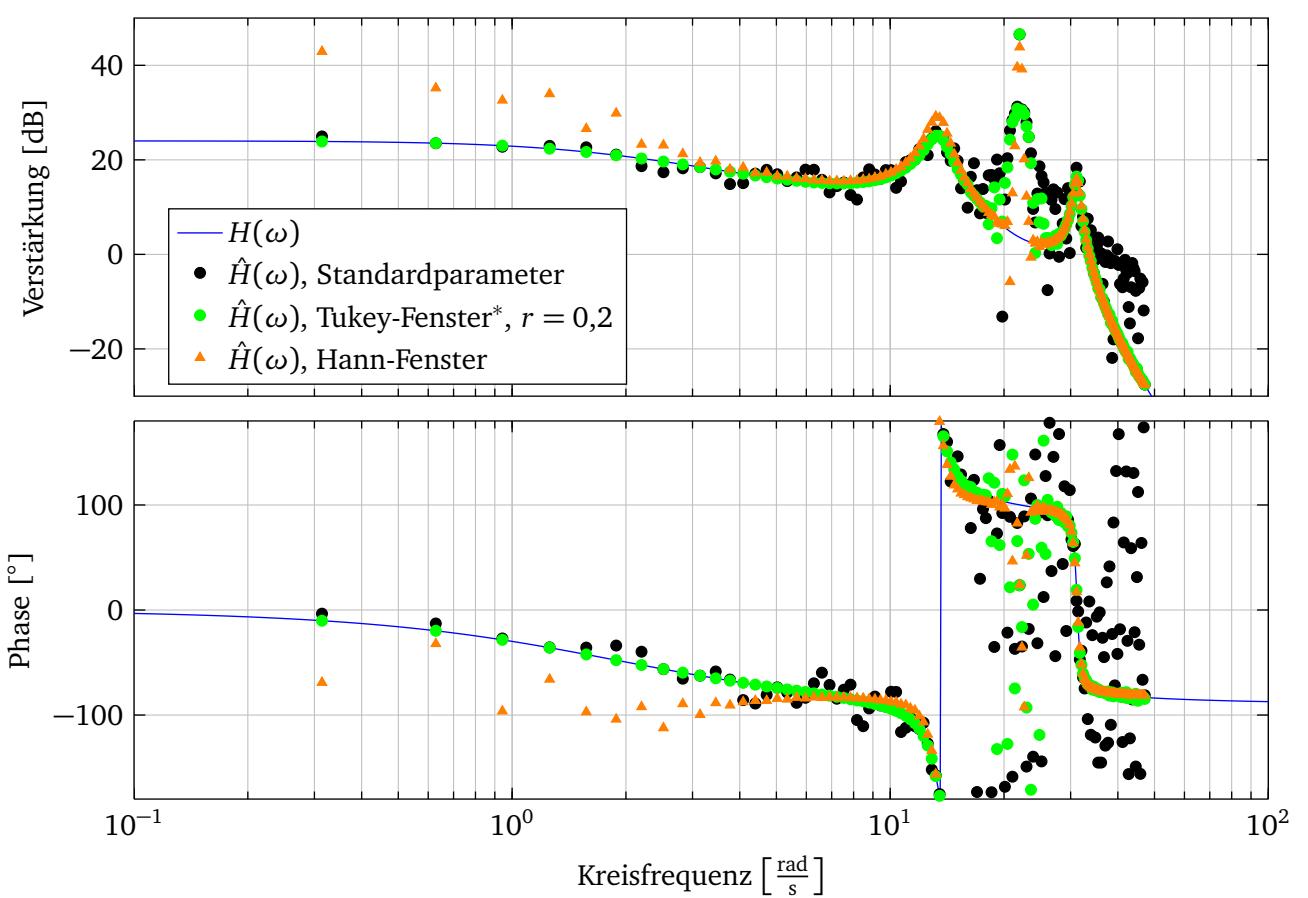


Abbildung 7.24: Anregung mit zeitbegrenztem Gleitsinus, starke periodische Störung, ohne Rauschen

Der Bereich, in dem die Störung erkennbar ist, ist bei der Fensterung zwar etwas breiter als in den Fällen, in denen die Messdauer auf die Periodenlänge der Störung (näherungsweise) abgestimmt ist, allerdings ist außerhalb dieses Bereiches keine sichtbare Abweichung durch die Störung mehr vorhanden. (Die Abweichungen beim Hann-Fenster entstehen unabhängig von der Störung durch die Fensterung an sich.)

Besonders die Fensterung mit dem Tukey-Fenster und einem verzögerten Beginn der Anregung ist damit gut geeignet, den Störeinfluss auf einen engen Frequenzbereich zu begrenzen. Der Vorteil gegenüber der angepassten Messzeit ist dabei, dass die Fensterung unabhängig von der Störfrequenz wirkt und auch eine andere Frequenzauflösung durch Zeropadding berechnet werden kann. Die längere Messzeit, die notwendig ist, das Nutzsignal nicht von den Fensterflanken erfassen zu lassen, verschlechtert dafür wieder das Signal-Rausch-Verhältnis.

7.4 Frequenzgangsschätzung über Impuls-Leistungsdichten (H1- und H2-Schätzung)

7.4.1 H0-Schätzung

Die im vorherigen Abschnitt behandelte Schätzmethode

$$\hat{H}_0(\omega) = \frac{\hat{Y}(\omega)}{\hat{U}(\omega)}$$

wird auch als H0-Schätzung bezeichnet.

Wird über mehrere Messungen gemittelt, so kann dies anhand der Quotienten oder für Zähler und Nenner getrennt geschehen. Die Mittelung der Quotienten ist dabei nur sinnvoll, wenn die gemessenen Eingangsgrößen nicht oder nur sehr gering gestört sind. Die getrennte Mittelung von Zähler und Nenner ist allerdings problematisch, wenn verschiedene Eingangsgrößenverläufe in den Messungen verwendet werden, etwa weil diese per Hand aufgeschaltet sind, oder weil mit im Betrieb auftretenden Verläufen gearbeitet werden muss. Ebenfalls problematisch ist es, wenn die Messungen nicht sauber synchronisiert sind, d. h. wenn der Messbeginn und der Start des – ansonsten gleichen – Eingangsgrößenverlaufs unterschiedlich zueinander verschoben sind.

In diesem Fall treten Auslöschungen im Frequenzbereich auf, die die erreichbare Güte der Schätzung reduzieren und für manche Frequenzen eine Schätzung unmöglich machen. Dies ist in Abbildung 7.25 an einem Beispiel verdeutlicht.

Es werden zwei Messungen mit einer zeitbegrenzten Gleitsinusanregung durchgeführt, wobei die Anregung bei der zweiten Messung bezogen auf den Messbeginn 0,5 s später startet. Die addierten Eingangsgrößenverläufe aus beiden Messungen sind in Abbildung 7.25a gezeigt. (Für die Schätzung macht es keinen Unterschied, ob im Zeit- oder Frequenzbereich gemittelt wird.)

Damit ergibt sich effektiv das in Abbildung 7.25b in blau gezeigte Amplitudendichte der Anregung. Bei synchronisierter Anregung hätte sich das rote Spektrum ergeben, welches einfach dem doppelten einfachen Spektrum entspricht. Durch die Zeitverschiebung treten demnach Auslöschungen auf. Entsprechend ergeben sich die in Abbildung 7.25c in schwarz dargestellten Schätzwerte für den Frequenzgang. In den Bereichen um die Auslöschungen im Anregungsspektrum treten sehr hohe Schätzfehler auf.

Eigenschaften

Die Eigenschaften dieses Schätzer wurde für den Fall, dass keine Störung am Eingang vorliegt schon ausführlich diskutiert. Da der H1- bzw. H2-Schätzer besonders dann wichtig sein kann, wenn auch die Messung der Eingangsgröße gestört ist, wird auf diesen Fall hier kurz eingegangen.

Um die Anzahl der Subindizes zu verringern, wird der Ausgangsfehler $n_y(t)$ im Weiteren mit $n(t)$ und der Eingangsfehler $n_u(t)$ mit $m(t)$ abgekürzt.

Damit gilt

$$\tilde{u}(t) = u(t) + m(t) \quad \text{und} \quad \tilde{y}(t) = y(t) + n(t)$$

bzw.

$$\tilde{U}(\omega) = U(\omega) + M(\omega) \quad \text{und} \quad \tilde{Y}(\omega) = Y(\omega) + N(\omega)$$

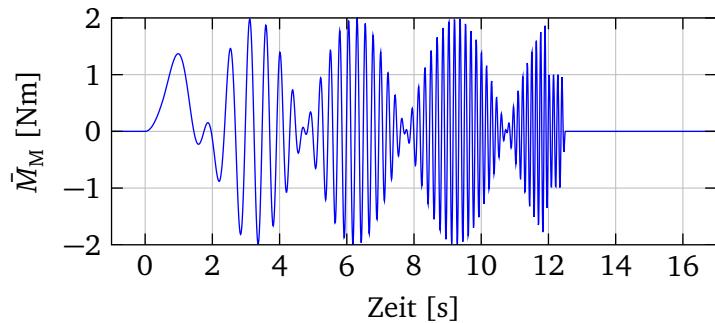
nach Anwendung der Fouriertransformation.

Für den relativen Schätzfehler gilt, wenn $\hat{U}(\omega) = \tilde{U}(\omega)$ und $\hat{Y}(\omega) = \tilde{Y}(\omega)$ gesetzt wird,

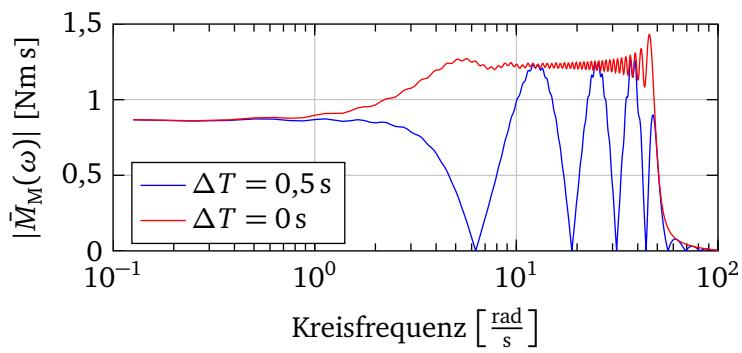
$$\Delta \hat{H}_{0,\text{rel}}(\omega) = \frac{\hat{H}_0(\omega) - H(\omega)}{H(\omega)} = \frac{\frac{Y(\omega) + N(\omega)}{U(\omega) + M(\omega)} - \frac{Y(\omega)}{U(\omega)}}{\frac{Y(\omega)}{U(\omega)}} = \frac{\frac{Y(\omega) + N(\omega)}{Y(\omega)} - \frac{U(\omega) + M(\omega)}{U(\omega)}}{\frac{U(\omega) + M(\omega)}{U(\omega)}},$$

wobei im letzten Schritt mit $\frac{U(\omega) + M(\omega)}{Y(\omega)}$ erweitert wurde. Aus dem letzten Ausdruck folgt dann direkt

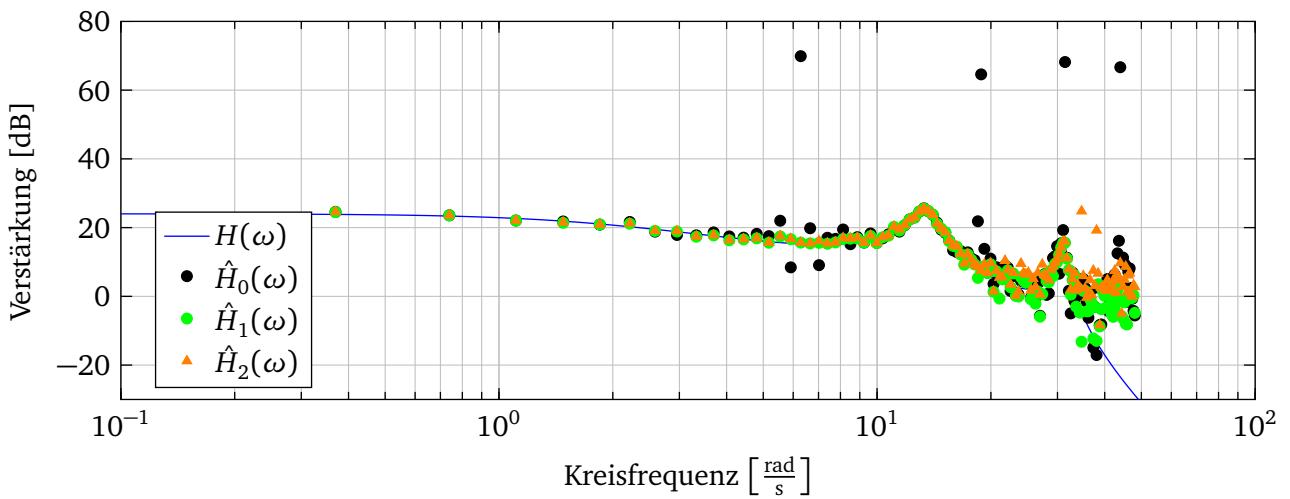
$$\Delta \hat{H}_{0,\text{rel}}(\omega) = \frac{\hat{H}_0(\omega) - H(\omega)}{H(\omega)} = \frac{\frac{N(\omega)}{Y(\omega)} - \frac{M(\omega)}{U(\omega)}}{1 + \frac{M(\omega)}{U(\omega)}}.$$



(a) Addierte Eingangsgröße für $\Delta T = 0,5\text{ s}$



(b) Amplitudendichte der addierten Eingangsgröße



(c) Geschätzter Frequenzgang, mit Rauschen, $N = 2$

Abbildung 7.25: Mittelung von zwei Messungen mit verschobenen Eingangsgrößenverläufen

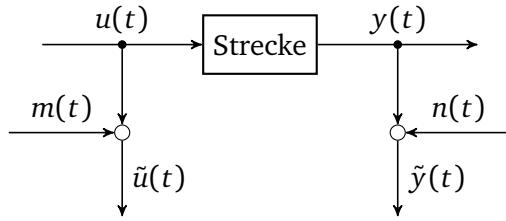


Abbildung 7.26: Benennung der Größen

Wenn der Eingangsfehler gering gegenüber den Eingang ist, d. h. $|M(\omega)/U(\omega)| \ll 1$, dann gilt

$$\Delta \hat{H}_{0,\text{rel}}(\omega) = \frac{\hat{H}_0(\omega) - H(\omega)}{H(\omega)} \approx \frac{N(\omega)}{Y(\omega)} - \frac{M(\omega)}{U(\omega)}.$$

Dies kann man wieder in Real- und Imaginärteil aufteilen und den Erwartungswert bilden. Es lässt sich aber an der obenstehenden Gleichung schon erkennen, dass der H_0 -Schätzer dann biasfrei ist, wenn sowohl das Ausgangsrauschen als auch das Eingangsrauschen mittelwertfrei sind.

7.4.2 H1- und H2-Schätzung

Die H1- und H2-Schätzung wird hier für deterministische, zeitbegrenzte Anregungssignale formuliert. Damit ergeben sich etwas andere Formulierungen, als wenn man nur von stationären stochastischen Signalen ausgeht.

Nimmt man von dem Frequenzbereichszusammenhang $Y(\omega) = H(\omega) \cdot U(\omega)$ das konjugiert-komplexe und multipliziert man dann beide Seiten mit $U(\omega)$, so erhält man

$$Y^*(\omega)U(\omega) = H^*(\omega) \cdot U^*(\omega)U(\omega)$$

und damit

$$H^*(\omega) = \frac{Y^*(\omega)U(\omega)}{U^*(\omega)U(\omega)}.$$

Im Nenner steht immer die positive reelle Zahl $U^*(\omega)U(\omega) = |U(\omega)|^2$. Für das Produkt im Zähler gilt

$$Y^*(\omega)U(\omega) = |Y^*(\omega)| \cdot |U(\omega)| \cdot e^{j \cdot (\arg U(\omega) - \arg Y(\omega))} = |Y^*(\omega)| \cdot |U(\omega)| \cdot e^{-j \cdot \arg H(\omega)}.$$

Die Phase des Zählers hängt also nicht von dem Eingangssignal, sondern nur von dem System selber ab. Damit können in dieser Form der Zähler und der Nenner problemlos für sich gemittelt werden, ohne das Auslöschungen des Nutzanteils der Signale auftreten können.

Damit gilt mit $S_{yu}^I(\omega) = Y^*(\omega)U(\omega)$ und $S_{uu}^I(\omega) = U^*(\omega)U(\omega)$ auch

$$H^*(\omega) = \frac{S_{yu}^I(\omega)}{S_{uu}^I(\omega)}.$$

Da der Zusammenhang $S_{yu}^I(\omega) = H^*(\omega)S_{uu}^I(\omega)$ linear in $S_{uu}^I(\omega)$ ist, gilt bei R (ungestörten) Messungen auch

$$H^*(\omega) = \frac{\bar{S}_{yu}^I(\omega)}{\bar{S}_{uu}^I(\omega)},$$

wenn $\bar{S}_{yu}^I(\omega)$ und $\bar{S}_{uu}^I(\omega)$ die mittleren Impuls-Leistungsdichten

$$\bar{S}_{yu}^I(\omega) = \frac{1}{R} \cdot \sum_r Y_r^*(\omega) U_r(\omega) \quad \text{bzw.}$$

$$\bar{S}_{uu}^I(\omega) = \frac{1}{R} \cdot \sum_r U_r^*(\omega) U_r(\omega)$$

sind.

Praktisch muss man die idealen mittleren Impuls-Leistungsdichten $\bar{S}_{yu}^I(\omega)$ und $\bar{S}_{uu}^I(\omega)$ durch die Schätzungen

$$\hat{\bar{S}}_{yu}^I(\omega) = \frac{1}{R} \cdot \sum_r \tilde{Y}_r^*(\omega) \tilde{U}_r(\omega) \quad \text{bzw.}$$

$$\hat{\bar{S}}_{uu}^I(\omega) = \frac{1}{R} \cdot \sum_r \tilde{U}_r^*(\omega) \tilde{U}_r(\omega)$$

ersetzen, was dann auf den H1-Schätzer

$$\hat{H}_1^*(\omega) = \frac{\hat{\bar{S}}_{yu}^I(\omega)}{\hat{\bar{S}}_{uu}^I(\omega)}$$

führt.

Multipliziert man anstelle von $U(\omega)$ mit $Y(\omega)$, so ergibt sich über

$$Y^*(\omega)Y(\omega) = H^*(\omega) \cdot U^*(\omega)Y(\omega)$$

und

$$H^*(\omega) = \frac{S_{yy}^I(\omega)}{S_{uy}^I(\omega)}$$

die Schätzgleichung

$$\hat{H}_2^*(\omega) = \frac{\hat{\bar{S}}_{yy}^I(\omega)}{\hat{\bar{S}}_{uy}^I(\omega)}$$

des H2-Schätzers.

Eigenschaften

Für die mittlere geschätzte Kreuzleistungsdichte $\hat{\bar{S}}_{yu}^I(\omega)$ gilt

$$\begin{aligned} E\{\hat{\bar{S}}_{yu}^I(\omega)\}_r &= \frac{1}{R} \cdot \sum_r (E\{Y_r^*(\omega)U_r(\omega)\}_r + E\{Y_r^*(\omega)M_r(\omega)\}_r + E\{N_r^*(\omega)U_r(\omega)\}_r + E\{N_r^*(\omega)M_r(\omega)\}_r) \\ &= \frac{1}{R} \cdot \sum_r Y_r^*(\omega)U_r(\omega) = \bar{S}_{yu}^I(\omega), \end{aligned}$$

wobei $\{n(t)\}_r$ und $\{m(t)\}_r$ wieder als mittelwertfrei angenommen sind. Analog ergibt sich

$$E\{\hat{\bar{S}}_{uy}^I(\omega)\}_r = \bar{S}_{uy}^I(\omega).$$

Für die geschätzte mittlere Autoleistungsdichte $\hat{\bar{S}}_{uu}^I(\omega)$ gilt

$$\begin{aligned} E\{\hat{\bar{S}}_{uu}^I(\omega)\}_r &= \frac{1}{R} \cdot \sum_r (E\{U_r^*(\omega)U_r(\omega)\}_r + E\{U_r^*(\omega)M_r(\omega)\}_r + E\{M_r^*(\omega)U_r(\omega)\}_r + E\{M_r^*(\omega)M_r(\omega)\}_r) \\ &= \frac{1}{R} \cdot \sum_r U_r^*(\omega)U_r(\omega) + \frac{1}{R} \cdot \sum_r E\{M_r^*(\omega)M_r(\omega)\}_r \\ &= \bar{S}_{uu}^I(\omega) + \frac{1}{R} \cdot \sum_r E\{M_r^*(\omega)M_r(\omega)\}_r. \end{aligned}$$

Der Ausdruck im Erwartungswert entspricht bis auf einen Faktor dem Periodogramm

$$\hat{S}_{mm,PG}(\omega) = \frac{1}{T} \cdot M_r^*(\omega)M_r(\omega),$$

welches ein (schlechter) Schätzer für die Leistungsdichte des stationären Prozesses $\{m(t)\}_r$ ist.⁹ Dessen Erwartungswert

$$S_{mm,PG}(\omega) = E\{\hat{S}_{mm,PG}(\omega)\}_r = \frac{1}{T} \cdot E\{M_r^*(\omega)M_r(\omega)\}_r$$

ist unabhängig von der Messung r , so dass sich letztlich

$$E\{\hat{\bar{S}}_{uu}^I(\omega)\}_r = \bar{S}_{uu}^I(\omega) + T \cdot S_{mm,PG}(\omega)$$

ergibt. Analog erhält man

$$E\{\hat{\bar{S}}_{yy}^I(\omega)\}_r = \bar{S}_{yy}^I(\omega) + T \cdot S_{nn,PG}(\omega).$$

Die Varianz eines Schätzwertes $\{\hat{S}_{mm,PG}(\omega)\}_r$ ist unabhängig von der Messzeit T . Über die Einzelmessungen R reduziert sich jedoch dessen Varianz, und für $R \rightarrow \infty$ geht diese gegen Null. Damit konvergieren $\hat{\bar{S}}_{uu}^I(\omega)$ und $\hat{\bar{S}}_{yy}^I(\omega)$ im quadratischen Sinne, wenn auch nicht gegen die eigentlich gesuchten mittleren Leistungsdichten.

Für den H_1 - und H_2 -Schätzer ergibt sich damit

$$\plim_{R \rightarrow \infty} \{\hat{H}_1^*(\omega)\}_r = \frac{\bar{S}_{yu}^I(\omega)}{\bar{S}_{uu}^I(\omega) + T \cdot S_{mm,PG}(\omega)} = H^*(\omega) \cdot \frac{1}{1 + \frac{T \cdot S_{mm,PG}(\omega)}{\bar{S}_{uu}^I(\omega)}} \quad (7.3)$$

und

$$\plim_{R \rightarrow \infty} \{\hat{H}_2^*(\omega)\}_r = \frac{\bar{S}_{yy}^I(\omega) + T \cdot S_{nn,PG}(\omega)}{\bar{S}_{uy}^I(\omega)} = H^*(\omega) + \frac{T \cdot S_{nn,PG}(\omega)}{\bar{S}_{uy}^I(\omega)} = H^*(\omega) \cdot \left(1 + \frac{T \cdot S_{nn,PG}(\omega)}{\bar{S}_{yy}^I(\omega)} \right), \quad (7.4)$$

⁹ $\hat{S}_{mm,PG}(\omega)$ ist für endliche Messzeiten T lediglich asymptotisch biasfrei, und die Varianz nimmt mit wachsendem T nicht ab, d. h. es ist ein nicht konsistenter Schätzer für die Leistungsdichte.

wobei die Konvergenz für Real- und Imaginärteil getrennt zu verstehen ist. Da Autoleistungsdichtespektren nie negativ sind, und auch mit einem Periodogramm nie negativ geschätzt werden können, folgt daraus, dass

$$\operatorname{plim}_{R \rightarrow \infty} \{|\hat{H}_1^*(\omega)|\}_r \leq |H^*(\omega)|$$

und

$$\operatorname{plim}_{R \rightarrow \infty} \{|\hat{H}_2^*(\omega)|\}_r \geq |H^*(\omega)|$$

gilt. Mit der H_1 -Schätzung wird also der wahre Frequenzgang betragmäßig unterschätzt und mit der H_2 -Schätzung überschätzt. Die Phasen werden in beiden Fällen korrekt geschätzt. Diese Aussagen gelten aber nur asymptotisch für $R \rightarrow \infty$! Praktisch kann dieses Verhalten jedoch häufig auch schon für wenige Messungen beobachtet werden.

Es ist hierbei in Kauf zu nehmen, dass die Autoleistungsdichten der stochastischen Anteile lediglich über ein Periodogramm geschätzt werden, da die stochastischen Anteile in den Zeitdaten nicht von den deterministischen Anteilen zu trennen sind. Keinesfalls können hier Methoden aus der Signalverarbeitung zur Schätzung der Leistungsdichten für *stationäre* Signale verwendet werden, wie die Bartlett- oder Welch-Methode. Diese auf die sicher nicht stationären Signale $\tilde{u}(t)$ oder $\tilde{y}(t)$ anzuwenden ist nicht möglich.

Die genannten Methoden zur Bestimmung des Leistungsdichtespektrums von stationären Signalen beruhen unter Anderem darauf, dass die Zeitdaten in kleinere Intervalle zerlegt werden, für die dann zunächst die Periodogramme bestimmt werden. Diese Periodogramme werden dann gemittelt, womit sich die Varianz insgesamt verringert. Aus diesen Erkenntnissen über die Eigenschaften des Periodogramms können hier entsprechende Maßnahmen abgeleitet werden. Kurz ausgedrückt ergibt sich, dass auch bei der H_1 - und H_2 -Schätzung lieber mehr kurze einzelne Messungen erfolgen sollten, die dann gemittelt werden, als wenige lange.

Aus den Ausdrücken

$$\frac{T \cdot S_{mm,PG}(\omega)}{\bar{S}_{uu}^I(\omega)}$$

und

$$\frac{T \cdot S_{nn,PG}(\omega)}{\bar{S}_{yy}^I(\omega)}$$

in Gl. (7.3) und (7.4) ergibt sich auch die selbe Konsequenz wie für den H_0 -Schätzer. $\bar{S}_{uu}^I(\omega)$ und damit $|U(\omega)|^2$ muss mehr als proportional zur Messzeit T wachsen, damit der Bias sinkt. (Streng genommen ist hier noch zu beachten, dass für kleine Messzeiten auch $S_{nn,PG}(\omega)$ einen Bias besitzen kann, dessen Wirkung in dieser Aussage nicht berücksichtigt ist.)

Kohärenz

Neben der eigentlichen Schätzung des Frequenzganges, welches mit der H_1 - oder H_2 -Methode nur dann sinnvoll ist, wenn die Messungen nicht synchronisiert werden können und der Eingang ebenfalls gestört ist, führt die Betrachtung dieser beiden Schätzer noch auf die sogenannte Kohärenz, welche auch dann eine interessante Größe ist, wenn die eigentliche Schätzung mit der H_0 -Methode erfolgt.

Die Kohärenz ist hier als

$$\gamma_{yu}^2(\omega) = \frac{H_1^*(\omega)}{H_2^*(\omega)}$$

definiert, wobei mit $H_1^*(\omega)$ und $H_2^*(\omega)$ die jeweiligen Grenzwerte für $R \rightarrow \infty$ gemeint sind,

$$\gamma_{yu}^2(\omega) = \text{plim}_{R \rightarrow \infty} \frac{\hat{H}_1^*(\omega)}{\hat{H}_2^*(\omega)} = \text{plim}_{R \rightarrow \infty} \frac{\hat{S}_{yu}^I(\omega) \hat{S}_{uy}^I(\omega)}{\hat{S}_{yy}^I(\omega) \hat{S}_{uu}^I(\omega)}.$$

Damit hängt hier $\gamma_{yu}^2(\omega)$ auch von der Messzeit T ab, da zum einen $S_{nn,PG}(\omega)$ selber von der Messzeit abhängt, und zum anderen dieser Wert nochmals in Gl. (7.3) und (7.4) mit der Messzeit multipliziert wird.¹⁰

Da mit dem H1-Schätzer der wahre Frequenzgang unterschätzt und der H2-Schätzer den wahren Frequenzgang überschätzt, gilt

$$0 \leq \gamma_{yu}^2(\omega) \leq 1.$$

Damit erlaubt die Kohärenz, den Störeinfluss zu bewerten. Ist die Kohärenz groß, so ist davon auszugehen, dass nur geringe Störungen auf die Messgrößen wirken und die Schätzung damit gut ist. Ist die Kohärenz klein, so weist dies auf größere Störungen hin.

Da man nicht unendlich Messungen durchführen kann, kann auch nur ein Schätzwert

$$\hat{\gamma}_{yu}^2(\omega) = \frac{\hat{H}_1^*(\omega)}{\hat{H}_2^*(\omega)}$$

für die Kohärenz bestimmt werden.

Ist $R = 1$, liegt also nur eine Messung vor, so ergeben der H0-, H1- und H2-Schätzer alle das gleiche Ergebnis. Der Schätzwert für die Kohärenz ist in diesem Fall konsequenter Weise

$$\hat{\gamma}_{yu}^2(\omega) = \frac{\tilde{Y}^*(\omega) \tilde{U}(\omega) \cdot \tilde{U}^*(\omega) \tilde{Y}(\omega)}{\tilde{Y}^*(\omega) \tilde{Y}(\omega) \cdot \tilde{U}^*(\omega) \tilde{U}(\omega)} = 1.$$

Bei nur einer Messung lässt sich also mit der Kohärenz keine Aussage über die Güte der Ergebnisse treffen.¹¹

Aber schon bei zwei Messungen kann zum Teil durch die Kohärenz eine Abschätzung der Schätzgüte erfolgen. Dies ist in Abbildung 7.27a gezeigt. Da in diesem Beispiel der Eingangsgrößenverlauf für alle Messungen gleich und ideal synchronisiert ist und nur der Ausgang verrauscht ist, ergeben der H0-Schätzer und der H1-Schätzer das gleiche Ergebnis. Die Werte des H2-Schätzers (grün) liegen in den Frequenzbereichen guter Anregung über denen des H1-Schätzers und in den Bereichen schlechter Anregung, in denen das Rauschen mehr Einfluss erhält, erwartungsgemäß immer über denen des H1-Schätzers. Die sich damit ergebenden Werte für die Kohärenz sind im unteren Graphen dieser Abbildung dargestellt. Man kann zwar nicht sagen, dass jeder Punkt, für den die Kohärenz einen großen Wert annimmt auch einen guten Schätzwert liefert, da auch aufgrund der stochastischen Streuung hohe Werte der Kohärenz entstehen können. Jedoch lassen sich prinzipiell die Bereiche mit guten und schlechten Schätzwerten erkennen.

¹⁰ Werden nur stationäre Signale zugelassen, so ist $\gamma_{yu}^2(\omega) = \frac{S_{yu}(\omega) S_{uy}(\omega)}{S_{uu}(\omega) S_{yy}(\omega)}$ und nur von den Eigenschaften der Signale, nicht aber der Messzeit abhängig.

¹¹ Wenn nur ergodische Prozesse zugelassen werden, und die Schätzung der Leistungsdichten nicht über das Periodogramm erfolgt, gilt dies nicht.

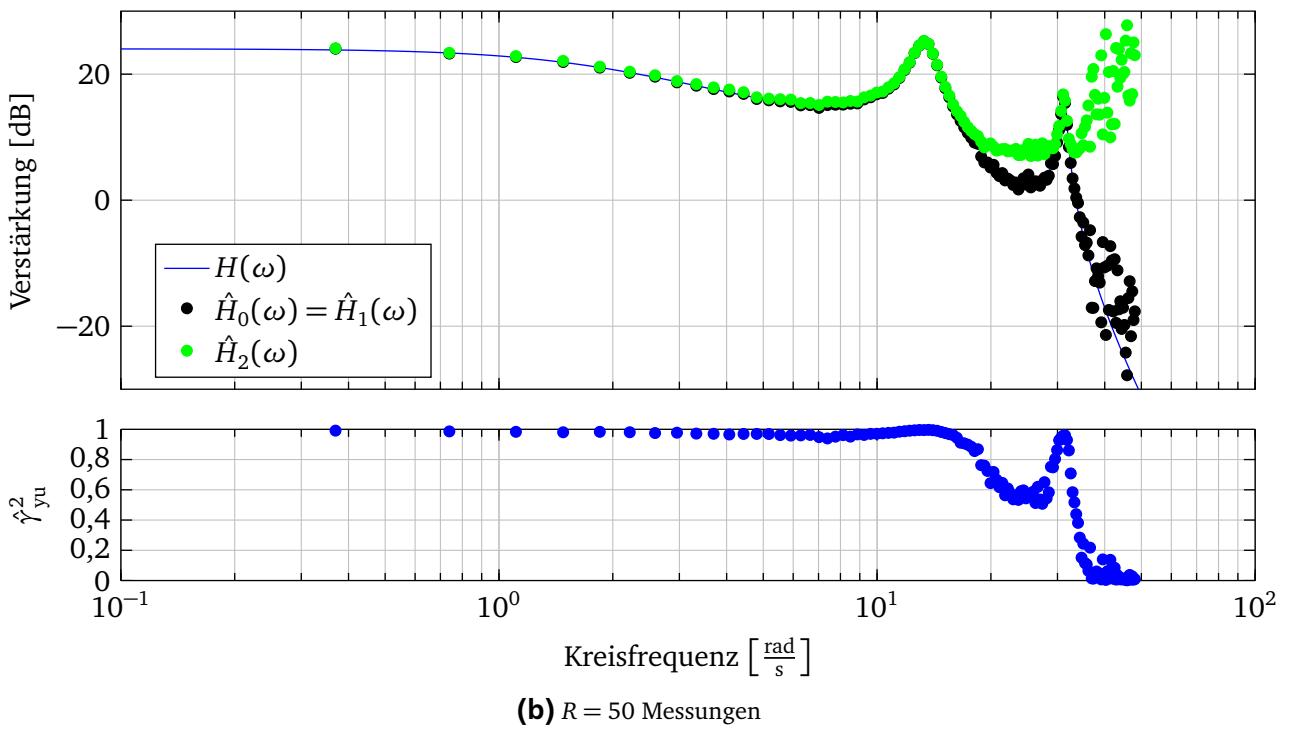
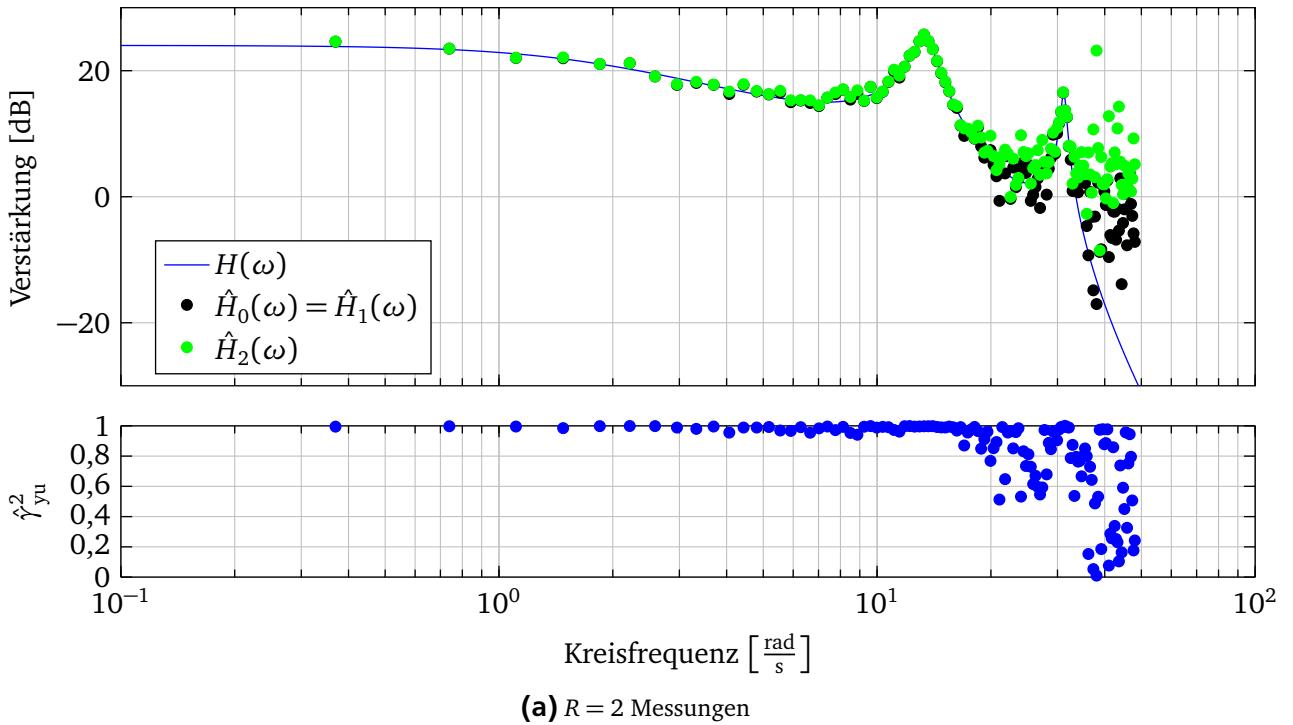


Abbildung 7.27: Schätzung mit Rauschen

In Abbildung 7.27b ist die Anzahl der Messungen auf fünfzig erhöht. Der H0- und H1-Schätzer nähert sich auch in den Bereichen, in denen das Rauschen dominiert dem tatsächlichen Frequenzgang an. Der H2-Schätzer überschätzt hier deutlich, und zwar umso mehr, desto geringer der Wert der Impulsleistungsdichte der Systemantwort auf die Eingangsgröße ist. Es zeigt sich also dessen Bias.

Die Kohärenz kann nicht nur stochastische Störungen, sondern auch deterministische Störungen „erkennen“. Dazu ist es jedoch notwendig, dass sich Nutzsignal und Störung bei verschiedenen Messungen unterschiedlich wirken.

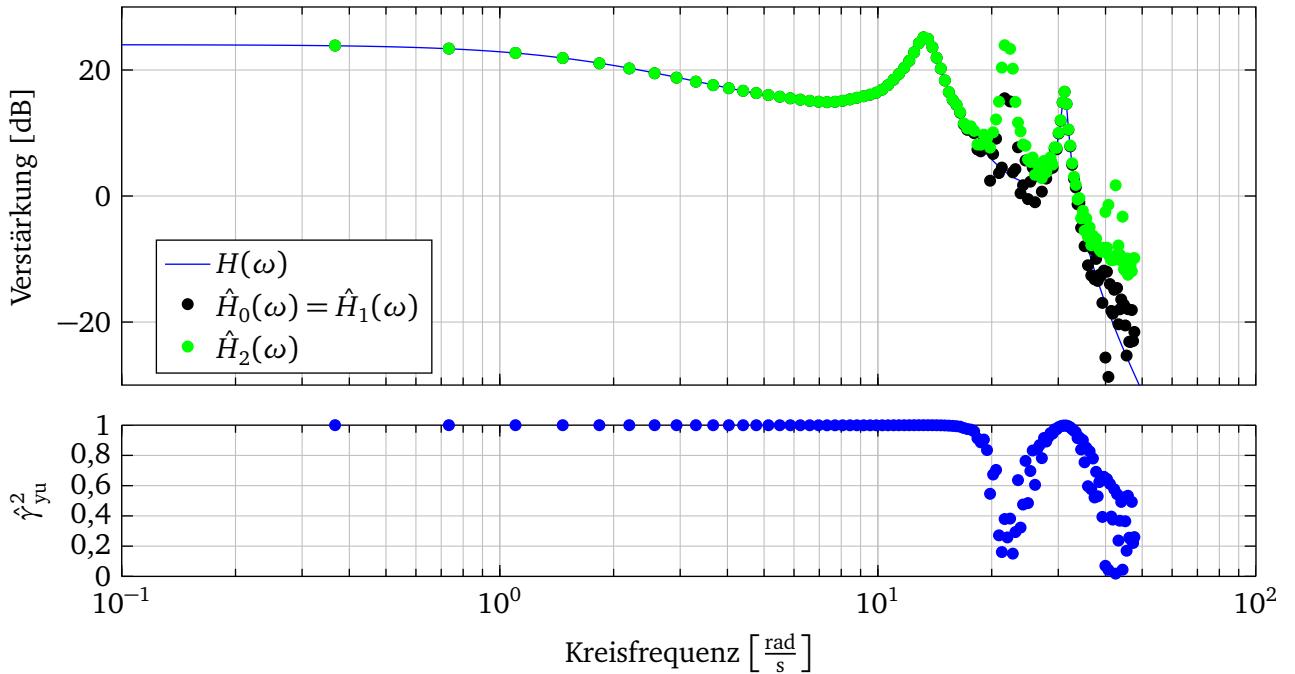


Abbildung 7.28: Schätzung mit starker periodischer Störung, ohne Rauschen, $N = 2$ Messungen

Wenn beispielsweise im Fall der starken periodischen Störung die Messung immer synchron mit dem Beginn des Gleitsinus sowie bei immer demselben Phasenwinkel der Störung beginnt, dann schlägt sich dies nicht in einem niedrigen Wert der Kohärenz nieder. Auch wenn es für den Betrachter offensichtlich ist, dass hier eine Störung vorliegt, so kann das mathematisch nicht automatisch entschieden werden. Aus Sicht der Schätzverfahren antwortet das System auf die gleiche Anregung immer gleich.

Eine solche Störung wird durch die Kohärenz erkannt, wenn der Phasenwinkel der Störung bezüglich dem Beginn des Gleitsinus variiert, oder wenn verschiedene Anregungssignale verwendet werden. Dies ist in Abbildung 7.28 gezeigt. Dort sind zwei Messungen bei Vorhandensein einer starken periodischen Störung (ohne Rauschen) durchgeführt, wobei die Phasenlage der Störung bezüglich der Anregung unterschiedlich ist. Der Einbruch der Kohärenz deutet an, dass die „zusätzliche Resonanz“, die in dem geschätzten Frequenzgang erscheint, nicht vertrauenswürdig ist.

7.5 Glättung des geschätzten Frequenzgangs

Bisher wurden für gegebene Frequenzpunkte ω_l der Frequenzgang geschätzt. Liegen L Frequenzpunkte vor, so erhält man auch L unabhängige Schätzwerte $\hat{H}(\omega_l)$. Diese haben miteinander zunächst nichts zu tun, da – bis auf die Linearität – keine Annahmen über das System getroffen sind. Damit sind aber auch viele Mittelungen notwendig, um eine geringe Varianz der Schätzwerte zu erreichen.

Die praktische Erfahrung zeigt aber, dass der Frequenzgang von realen Systemen in der Regel glatt ist. Dieses Wissen ist noch nicht eingeflossen. Als Folge dessen ist es sinnvoll möglich, den Frequenzgang zu glätten.

[Ljung 1999, Kap. 6.4]

7.6 Mehrgrößensysteme

Im Frequenzbereich gilt für ein MIMO-System mit p Ein- und q Ausgängen

$$\begin{bmatrix} Y_1(\omega) \\ \vdots \\ Y_q(\omega) \end{bmatrix} = \begin{bmatrix} H_{11}(\omega) & \cdots & H_{1p}(\omega) \\ \vdots & \ddots & \vdots \\ H_{q1}(\omega) & \cdots & H_{qp}(\omega) \end{bmatrix} \begin{bmatrix} U_1(\omega) \\ \vdots \\ U_p(\omega) \end{bmatrix} .$$

Fasst man die Fouriertransformierten der Eingangsgrößen und Ausgangsgrößen einer Messung r in je einen Vektor

$$\mathbf{U}_r^T(\omega) = [U_{1,r}(\omega) \ U_{2,r}(\omega) \ \cdots \ U_{p,r}(\omega)] \quad \text{bzw.}$$

$$\mathbf{Y}_r^T(\omega) = [Y_{1,r}(\omega) \ Y_{2,r}(\omega) \ \cdots \ Y_{q,r}(\omega)]$$

zusammen, so kann kompakter

$$\mathbf{Y}_r(\omega) = \begin{bmatrix} H_{11}(\omega) & \cdots & H_{1p}(\omega) \\ \vdots & \ddots & \vdots \\ H_{q1}(\omega) & \cdots & H_{qp}(\omega) \end{bmatrix} \cdot \mathbf{U}_r(\omega) = \mathbf{H}(\omega) \cdot \mathbf{U}_r(\omega) \quad (7.5)$$

geschrieben werden.

H0-Schätzung

Betrachtet man den Zusammenhang (7.5) für insgesamt R durchgeführte Messungen r , so kann dies in der Form

$$[\mathbf{Y}_1(\omega) \ \cdots \ \mathbf{Y}_R(\omega)] = \mathbf{H}(\omega) \cdot [\mathbf{U}_1(\omega) \ \cdots \ \mathbf{U}_R(\omega)]$$

geschrieben werden. Mit

$$\mathbf{U}(\omega) = [\mathbf{U}_1(\omega) \ \cdots \ \mathbf{U}_R(\omega)] \quad \text{und}$$

$$\mathbf{Y}(\omega) = [\mathbf{Y}_1(\omega) \ \cdots \ \mathbf{Y}_R(\omega)]$$

ergibt sich noch kürzer

$$\mathbf{Y}(\omega) = \mathbf{H}(\omega) \cdot \mathbf{U}(\omega) . \quad (7.6)$$

Liegen $R = p$ Messungen vor, und sind die Eingangsgrößen geeignet gewählt, so dass die in diesem Fall quadratische Matrix $\mathbf{U}(\omega)$ vollen Rang besitzt und damit invertierbar ist, dann gilt

$$\mathbf{H}(\omega) = \mathbf{Y}(\omega) \cdot \mathbf{U}^{-1}(\omega) .$$

Liegen weniger als p Messungen vor, so reichen die Informationen nicht aus, um $\mathbf{H}(\omega)$ zu bestimmen. Ist $R > p$ und sind die Eingänge so gewählt, dass $\mathbf{U}(\omega)$ vollen Zeilenrang hat, dann kann Gl. (7.6) zunächst

von rechts mit $\mathbf{U}^T(\omega)$ und dann weiter von rechts mit $(\mathbf{U}(\omega)\mathbf{U}^T(\omega))^{-1}$ multipliziert werden. Damit erhält man

$$\mathbf{H}(\omega) = \mathbf{Y}(\omega) \cdot \mathbf{U}^T(\omega) \cdot (\mathbf{U}(\omega)\mathbf{U}^T(\omega))^{-1}.$$

Der Fall, dass $\mathbf{U}(\omega)$ invertierbar ist, ist darin auch enthalten.

Für die praktische Berechnung sind die ungestörten Ein- und Ausgangsgrößen nicht bekannt, so dass auf die gemessenen Zeitverläufe und die daraus ermittelten Amplitudendichten zurückgegriffen werden muss. Als Schätzgleichung ergibt sich dann

$$\hat{\mathbf{H}}_0(\omega) = \tilde{\mathbf{Y}}(\omega) \cdot \tilde{\mathbf{U}}^{-1}(\omega)$$

bzw.

$$\hat{\mathbf{H}}_0(\omega) = \tilde{\mathbf{Y}}(\omega) \cdot \tilde{\mathbf{U}}^T(\omega) \cdot (\tilde{\mathbf{U}}(\omega)\tilde{\mathbf{U}}^T(\omega))^{-1}. \quad (7.7)$$

Werden in Gl. (7.6) gemessene Daten verwendet, so ist diese Gleichung für $R > p$ im Allgemeinen nicht genau zu erfüllen. Die Frequenzgangsmatrix $\hat{\mathbf{H}}_0(\omega)$ die sich über die Schätzgleichung (7.7) ergibt, minimiert den dabei auftretenden Fehler im quadratischen Sinne.

H1-Schätzung

Wird von Gl. (7.5) das konjugiert-komplexe genommen und dann von rechts mit $\mathbf{U}_r^T(\omega)$ multipliziert, so erhält man

$$\mathbf{Y}_r^*(\omega)\mathbf{U}_r^T(\omega) = \mathbf{H}^*(\omega) \cdot \mathbf{U}_r^*(\omega)\mathbf{U}_r^T(\omega). \quad (7.8)$$

Mit

$$\begin{aligned}\mathbf{S}_{uu,r}^I(\omega) &= \mathbf{U}_r^*(\omega)\mathbf{U}_r^T(\omega) \\ \mathbf{S}_{yu,r}^I(\omega) &= \mathbf{Y}_r^*(\omega)\mathbf{U}_r^T(\omega)\end{aligned}$$

kann dies auch als

$$\mathbf{S}_{yu,r}^I(\omega) = \mathbf{H}^*(\omega)\mathbf{S}_{uu,r}^I(\omega)$$

geschrieben werden. Liegen mehrere Messungen vor, so können diese Impuls-Leistungsdichten addiert werden,

$$\begin{aligned}\bar{\mathbf{S}}_{uu}^I(\omega) &= \sum_r \mathbf{S}_{uu,r}^I(\omega) \quad \text{bzw.} \\ \bar{\mathbf{S}}_{yu}^I(\omega) &= \sum_r \mathbf{S}_{yu,r}^I(\omega),\end{aligned}$$

womit auch

$$\bar{\mathbf{S}}_{yu}^I(\omega) = \mathbf{H}^*(\omega)\bar{\mathbf{S}}_{uu}^I(\omega)$$

geschrieben werden kann.

$\bar{\mathbf{S}}_{uu}^I(\omega)$ ist dabei immer eine quadratische Matrix. Liegen genügend Messungen vor (mindestens p) und sind die Eingänge genügend verschieden angeregt, so ist $\bar{\mathbf{S}}_{uu}^I(\omega)$ invertierbar und es gilt

$$\mathbf{H}^*(\omega) = \bar{\mathbf{S}}_{yu}^I(\omega)\bar{\mathbf{S}}_{uu}^{I,-1}(\omega).$$

Damit erhält man die Schätzgleichung

$$\hat{\mathbf{H}}_1^*(\omega) = \hat{\mathbf{S}}_{yu}^I(\omega)\hat{\mathbf{S}}_{uu}^{I,-1}(\omega).$$

H2-Schätzung

Multipliziert man in Gl. (7.8) anstelle von $\mathbf{U}_r^T(\omega)$ mit $\mathbf{Y}_r^T(\omega)$ so erhält man

$$\mathbf{Y}_r^*(\omega)\mathbf{Y}_r^T(\omega) = \mathbf{H}^*(\omega) \cdot \mathbf{U}_r^*(\omega)\mathbf{Y}_r^T(\omega)$$

bzw.

$$\mathbf{S}_{yy,r}^I(\omega) = \mathbf{H}^*(\omega)\mathbf{S}_{uy,r}^I(\omega)$$

mit

$$\mathbf{S}_{yy,r}^I(\omega) = \mathbf{Y}_r^*(\omega)\mathbf{Y}_r^T(\omega)$$

$$\mathbf{S}_{uy,r}^I(\omega) = \mathbf{U}_r^*(\omega)\mathbf{Y}_r^T(\omega).$$

Summiert man wieder die Impulsleistungsdichten über verschiedene Messungen auf,

$$\bar{\mathbf{S}}_{yy}^I(\omega) = \sum_r \mathbf{S}_{yy,r}^I(\omega)$$

$$\bar{\mathbf{S}}_{uy}^I(\omega) = \sum_r \mathbf{S}_{uy,r}^I(\omega),$$

so gilt auch

$$\bar{\mathbf{S}}_{yy}^I(\omega) = \mathbf{H}^*(\omega)\bar{\mathbf{S}}_{uy}^I(\omega).$$

Ist die Anzahl der Eingänge gleich der der Ausgänge, so ist bei geeigneter Wahl der Anregung die Matrix $\bar{\mathbf{S}}_{uy}^I(\omega)$ invertierbar, und diese Gleichung kann direkt nach $\mathbf{H}^*(\omega)$ aufgelöst werden. Sind mehr Ausgänge vorhanden, und ist die Anregung so gewählt, dass $\bar{\mathbf{S}}_{uy}^I(\omega)$ vollen Zeilenrang hat, dann kann mit $\bar{\mathbf{S}}_{uy}^{I,T}(\omega)$ von rechts multipliziert werden,

$$\bar{\mathbf{S}}_{yy}^I(\omega)\bar{\mathbf{S}}_{uy}^{I,T}(\omega) = \mathbf{H}^*(\omega)\bar{\mathbf{S}}_{uy}^I(\omega)\bar{\mathbf{S}}_{uy}^{I,T}(\omega),$$

und damit weiter

$$\mathbf{H}^*(\omega) = \bar{\mathbf{S}}_{yy}^I(\omega)\bar{\mathbf{S}}_{uy}^{I,T}(\omega) \cdot (\bar{\mathbf{S}}_{uy}^I(\omega)\bar{\mathbf{S}}_{uy}^{I,T}(\omega))^{-1}$$

angegeben werden.

Daraus ergibt sich bei Verwendung der gemessenen Größen die Schätzgleichung

$$\hat{\mathbf{H}}_2^*(\omega) = \hat{\mathbf{S}}_{yy}^I(\omega)\hat{\mathbf{S}}_{uy}^{I,T}(\omega) \cdot (\hat{\mathbf{S}}_{uy}^I(\omega)\hat{\mathbf{S}}_{uy}^{I,T}(\omega))^{-1}.$$

Wahl der Anregung

Die prinzipielle Wahl des Anregungssignal folgt den gleichen Überlegungen wie im Eingrößenfall. Hinzu kommt hier die Entscheidung, wie die einzelnen Eingänge zueinander angeregt werden sollen.

Im Grunde kann für jeden Eingang ein völlig anderes Anregungssignal gewählt werden. Der Einfachheit halber wird hier aber davon ausgegangen, dass immer der gleiche Zeitverlauf $u(t)$ für alle Messungen und alle Eingänge verwendet wird. Für jede Messung r wird dieses Signal $u(t)$ dann über den Vektor λ_r auf die einzelnen Systemeingänge verteilt,

$$\mathbf{u}_r(t) = \begin{bmatrix} u_{1,r}(t) \\ \vdots \\ u_{p,r}(t) \end{bmatrix} = \lambda_r u(t).$$

Die „Verteilungsvektoren“ λ_r aller Messungen können in einer Matrix

$$\Lambda = [\lambda_1 \ \lambda_2 \ \cdots \ \lambda_R]$$

gesammelt werden.

Da aufgrund der Linearität der Fouriertransformation $\mathbf{U}_r(\omega) = \Lambda_r \cdot U(\omega)$ und $\mathbf{U}(\omega) = \Lambda \cdot U(\omega)$ ist, kann aufgrund der Betrachtung von Λ darauf geschlossen werden, ob die oben angegebenen Matrizen die Rangbedingungen einhalten. Damit die Identifikation anhand der oben angegebenen Schätzgleichungen für den H0-, H1- und H2-Schätzer möglich ist, muss Λ den Rang p , d. h. vollen Zeilenrang besitzen. Anschaulich bedeutet das, dass die Eingänge auf mindestens p verschiedene Arten angeregt sein müssen.

Im einfachsten Fall wird bei der ersten Messung nur der erste Eingang, bei der zweiten Messung nur der zweite Eingang usw. angeregt. In diesem Fall wäre Λ die Einheitsmatrix,

$$\Lambda = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix}.$$

Durch eine Anregung von mehreren Eingängen gleichzeitig kann erreicht werden, dass die gesamte in das System eingebrachte Signalenergie größer und damit die Schätzung besser wird.

Wenn nur die Begrenzung der Stellgröße maßgeblich ist, so wäre für ein System mit drei Eingängen und bei drei Messungen

$$\Lambda = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \end{bmatrix}$$

eine geeignete Wahl. Diese Matrix hat vollen Rang und kann damit verwendet werden. Hierbei wird bei jeder Messung jeder Eingang voll angeregt, womit sich die Gesamt signalenergie deutlich erhöht.

Auch wenn es praktisch nicht immer möglich ist, alle Eingänge gleichzeitig voll auszulenken, so lassen sich in der Regel doch Aufteilungen finden, die einen höheren Energieeintrag als die einfache Einpunktanregung besitzen.

7.7 Weitere Verfahren zur nichtparametrischen Identifikation

Bisher wurden Verfahren betrachtet, die deterministische Eingangssignale verwenden. Abschließend werden hier noch Verfahren erwähnt, die auf der Anregung mit stochastischen (oder pseudostochastischen) Signalen beruhen. Dabei wird wieder vom SISO-Fall ausgegangen.

H1- und H2-Schätzer mit stochastischer Erregung

Im Kapitel 4 wurde für stationäre Signale der Zusammenhang

$$S_{yu} = H^*(\omega)S_{uu}(\omega)$$

angegeben. Hieraus lässt sich direkt

$$\hat{H}_1^*(\omega) = \frac{\hat{S}_{yu}(\omega)}{\hat{S}_{uu}(\omega)}$$

als Schätzer angeben. Für den H2-Schätzer ergibt sich

$$\hat{H}_2^*(\omega) = \frac{\hat{S}_{yy}(\omega)}{\hat{S}_{uy}(\omega)}.$$

Anders als im Fall der deterministischen Anregung treten hier „echte“ Leistungsdichtspektren auf, die mit allen Methoden, die die Signalverarbeitung dazu bereithält, geschätzt werden können.

Die Eigenschaften der Schätzer sind dabei wie oben schon besprochen. Beide Schätzer besitzen bei vorhandenem Rauschen einen Bias, der auch für beliebig lange Messdauern nicht verschwindet. Der H1-Schätzer unterschätzt bei Rauschen am Eingang den Frequenzgang, während der H2-Schätzer bei Rauschen am Ausgang diesen überschätzt.

Schätzung der Gewichtsfunktion über Korrelationsfunktionen

Ein anderer Ansatz geht von der Systemgleichung bezüglich der Korrelationsfunktionen aus. Im Zeitkontinuierlichen lautet diese

$$R_{uy}(\tau) = \int_0^\infty g(\nu) \cdot R_{uu}(\tau - \nu) \cdot d\nu,$$

wobei die für die praktische Anwendung noch zu diskretisieren wäre, und im Zeitdiskreten

$$R_{uy,k} = \sum_{k'=0}^{\infty} g_{k'} \cdot R_{uu,k-k'}.$$

Hierbei ist schon berücksichtigt, dass bei realen Systemen $g(t) = 0$ für $t < 0$ bzw. $g_k = 0$ für $k < 0$ gilt.

Hierbei müssen in einem ersten Schritt zunächst Schätzwerte $\hat{R}_{uu,k}$ und $\hat{R}_{yu,k}$ für die Korrelationsfunktionen aus den Messdaten ermittelt werden.

Dann muss aus diesen Werten die Gewichtsfunktion berechnet werden. Es ist dabei festzulegen, wie viele Elemente der Gewichtsfunktionsfolge g_k bestimmt werden sollen. Diese Anzahl wird hier mit K bezeichnet. (Nach dem Diskretisieren des Integrals im zeitkontinuierlichen Fall ergibt sich prinzipiell das gleiche Vorgehen.) Dann wird die Gleichung für N Werte von $R_{uy,k}$ ausgeschrieben, z.B. von $k = 0$ bis $k = N - 1$,

$$\begin{aligned} R_{uy,0} &= g_0 \cdot R_{uu,0} + g_1 \cdot R_{uu,-1} + g_2 \cdot R_{uu,-2} + \cdots + g_{K-1} \cdot R_{uu,-(K-1)} \\ R_{uy,1} &= g_0 \cdot R_{uu,1} + g_1 \cdot R_{uu,0} + g_2 \cdot R_{uu,-3} + \cdots + g_{K-1} \cdot R_{uu,1-(K-1)} \\ &\vdots \\ R_{uy,N-1} &= g_0 \cdot R_{uu,N-1} + g_1 \cdot R_{uu,N-2} + g_2 \cdot R_{uu,N-3} + \cdots + g_K \cdot R_{uu,N-1-(K-1)}. \end{aligned}$$

Dies stellt ein lineares Gleichungssystem in den gesuchten Parametern g_k dar. Wenn $N = K$ ist, dann kann dieses (bei ausreichender Anregung, d.h. wenn die N Gleichungen linear unabhängig sind) gelöst werden. Für die Schätzung muss dabei wieder mit den geschätzten Korrelationsfunktionswerten gearbeitet werden. Wird dann $N > K$ gewählt, so kann die entstehende Gleichung optimal im Sinne der kleinsten Quadrate gelöst werden, und es ergibt sich eine Filterwirkung. Dieses Vorgehen die Gewichtsfunktion zu bestimmen wird auch als Entfaltung bezeichnet.

Wird als Eingangssignal weißes Rauschen, oder ein Signal das die Eigenschaften von weißem Rauschen hinreichend genau imitiert, verwendet, so vereinfacht sich die Schätzung nochmals deutlich. Bei weißem Rauschen gilt nämlich

$$R_{uy}(\tau) = \int_0^\infty g(\nu) \cdot S_0 \cdot \delta(\tau - \nu) \cdot d\nu = S_0 \cdot g(\tau),$$

bzw.

$$R_{uy,k} = \sum_{k'=0}^{\infty} g_{k'} \cdot \sigma_u^2 \cdot \delta_{k-k'} = \sigma_u^2 \cdot g_{k'} .$$

In diesem Fall ergibt sich die gesuchte Gewichtsfunktion durch eine einfache Skalierung der geschätzten Kreuzkorrelationsfunktion,

$$\hat{g}(t) = \frac{1}{S_0} \cdot \hat{R}_{uy}(t)$$

bzw.

$$\hat{g}_k = \frac{1}{\sigma_u^2} \cdot \hat{R}_{yu,k} .$$

Diese Verfahren sind beispielsweise in [Isermann und Münchhof 2011, Kap. 6 und 7] beschrieben.

8 Methode der kleinsten Quadrate

Thema dieses und der folgenden Kapitel ist die parametrische Identifikation. Ein wesentliches Werkzeug hierzu ist die Methode der kleinsten Quadrate. Diese wird in diesem Kapitel eingeführt. Um zunächst die wesentlichen Eigenschaften dieser Methode vorzustellen, wird diese anhand von statischen Zusammenhängen motiviert und besprochen.

Die Anwendung auf statische Zusammenhänge ist jedoch auch praktisch für viele Fälle relevant und bei weitem mehr als eine akademische Fingerübung, was auch die Beispiele zeigen.

Im darauffolgenden Kapitel wird die Methode der kleinsten Quadrate dann zur Ermittlung von Übertragungsfunktionen dynamischer Systeme angewendet. Dabei können im Allgemeinen nicht mehr alle Annahmen, die der Methode der kleinsten Quadrate zugrunde liegen, erfüllt werden. Auf die Konsequenzen wird eingegangen und es werden Maßnahmen besprochen, diese zu umgehen.

8.1 Beispiele

Coulomb'sche Reibung

Beim Aufstellen der Gleichungen zu linearen Systemen kann nur die geschwindigkeitsproportionale, viskose Reibung d berücksichtigt werden. Häufig zeigen mechanische Systeme jedoch auch einen deutlichen Coulomb'schen Reibungsanteil, d. h. der nur von der Richtung der Bewegung, jedoch nicht dem Betrag der Geschwindigkeit abhängt. Berücksichtigt man in einem einfachen Reibmodell die Coulomb'sche und die viskose Reibung, so ergibt sich für ein rotatorisches System das gesamte Reibmoment zu

$$M_R(t) = M_C \cdot \text{sign}(\omega(t)) + d \cdot \omega(t).$$

Geht man zunächst davon aus, dass die Geschwindigkeit $\omega(t)$ immer positiv ist, so kann auch einfacher

$$M_R(t) = M_C + d \cdot \omega(t)$$

geschrieben werden. Auch hierbei handelt es sich aber um einen nichtlinearen, nämlich affinen Zusammenhang. Stellt man ein parametrisches Modell des dynamischen Systems auf, welches diesem Reibmodell gehorcht, so gibt es zwei Möglichkeiten. Entweder ist der Arbeitspunkt bei einer so hohen Geschwindigkeit gewählt, dass sich das Vorzeichen der Geschwindigkeit nie ändert. In diesem Fall hat das Coulomb'sche Reibmoment lediglich einen Einfluss auf das Moment des Arbeitspunktes. Für das dynamische Verhalten fällt es heraus. Wird die Geschwindigkeit während der Identifikation Null, oder treten gar Wechsel der Bewegungsrichtung auf, so ist, wenn der Wert der Coulomb'schen Reibung nicht vernachlässigbar ist, keine sinnvolle Identifikation basierend auf linearen Ansätzen möglich.

In Abbildung 8.1 ist in blau der beispielhafte Verlauf des Reibmoments dargestellt über der (Relativ-)Geschwindigkeit ω , wenn das oben angegebene Modell verwendet wird. Für die Geschwindigkeit $\omega = 0$ befindet sich das System im Haftbereich, bei dem das Moment M_R jeden Wert bis zu einem maximalen Haftmoment annehmen kann. Unterscheidet sich das maximale Haftmoment von dem Coulomb'schen Reibmoment, so ergibt sich ein Übergangsbereich. Dieser kann beispielsweise nach Stribeck modelliert werden, was in Abbildung 8.1 in grau eingezeichnet ist.

Zur Simulation oder zum Entwurf von Reglern (Vorsteuerungen), die die Coulomb'sche Reibung berücksichtigen, kann es jedoch wünschenswert sein, diese zu identifizieren. Dies kann unabhängig von der

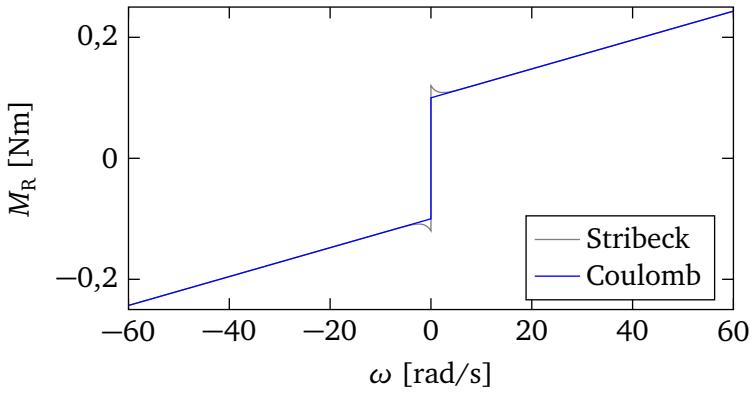


Abbildung 8.1: Reibmomente nach Stribeck-Modell und Coulomb'scher Reibung, jeweils mit viskoser Reibung

Identifikation des dynamischen Verhaltens erfolgen und führt, bei entsprechender Wahl der Kriterien, auf ein statisches LS-Problem.

Eine Möglichkeit, die Parameter M_C und d des Zusammenhangs $M_R(t) = M_C + d \cdot \omega(t)$ zu bestimmen ist, ein konstantes Moment $M(t)$ mit dem Motor zu erzeugen. Ist das Moment höher als das maximale Haftmoment, so wird sich, nach einem Beschleunigungs- und Einschwingvorgang, eine konstante Geschwindigkeit einstellen. Da dann nur noch die Reibungskräfte wirken, entspricht das Moment $M(t)$ im stationären Zustand gerade $M_R(t)$.

Misst man die stationäre Geschwindigkeit, die sich bei der Wahl eines konstanten Moments $M_{k'}$ einstellt, so erhält man einen Messpunkt

$$(M_{k'}, \tilde{\omega}_{k'}) .$$

Dabei wird davon ausgegangen, dass das Moment exakt vorgegeben werden kann.

Nach der Geschwindigkeit umgestellt lautet der tatsächliche Zusammenhang

$$\omega_{k'} = \frac{1}{d} \cdot M_{k'} - \frac{1}{d} \cdot M_C ,$$

wobei die Größen hier nicht mehr als zeitabhängige, sondern als abhängige der Messungsnummer k' angegeben sind.

In Tabelle 8.1 sind simulierte „Messdaten“ aufgelistet.¹ Diese sind in Abbildung 8.2 auch grafisch dargestellt, wobei zu beachten ist, dass die Abzissen- und Ordinatenachse im Vergleich zu Abbildung 8.1 ihre Bedeutung getauscht haben. Der Grund dafür ist der, dass das Moment als Eingangs- und damit unabhängige Größe der Identifikation auf der Abszissenachse aufgetragen werden sollte.

Tabelle 8.1: Simulierte „Messdaten“ zur Bestimmung der Coulomb'schen Reibung

k'	1	2	3	4	5	6	7	8	9	10
$M_{k'} [\text{Nm}]$	0,13	0,14	0,15	0,16	0,17	0,18	0,19	0,20	0,21	0,22
$\tilde{\omega}_{k'} [\text{rad/s}]$	13,1040	18,5890	18,6851	25,9949	29,6403	32,2026	37,2655	42,2305	49,6551	53,0349

Es sind zehn Messungen durchgeführt, wobei das minimale Moment so groß gewählt wurde, dass der Übergangsbereich, indem das Coulomb'sche Modell nicht ausreichend gut gilt, nicht erfasst wird.

Ziel ist es nun, eine Gerade durch diese Messdaten zu legen. In Abbildung 8.3 ist beispielhaft eine „Kandidat“ für eine solche Gerade eingezeichnet, wobei es sicherlich keine gute Lösung ist.

¹ Die wahren Parameter für dieses Beispiel sind $M_C = 0,1 \text{ Nm}$ und $d = 0,00239 \text{ Nm}/(\text{rad/s})$

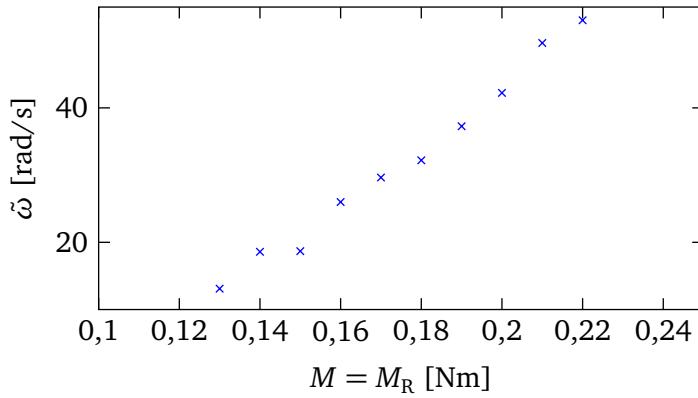


Abbildung 8.2: Gemessene Geschwindigkeiten im stationären Zustand für verschiedene Eingangsmodelle

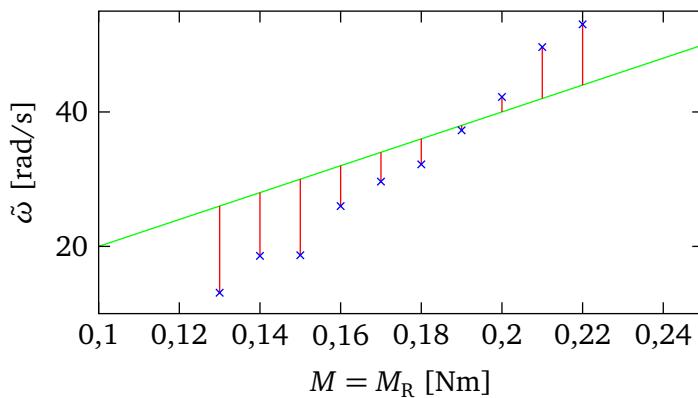


Abbildung 8.3: Kandidat für Gerade mit Abweichungen

Bezeichnet man die Geschwindigkeitswerte, die die grüne Gerade für die Momente $M_{k'}$ annimmt, mit $\hat{\omega}_{k'}$, dann kann man die Abweichungen

$$\hat{e}_{k'} = \tilde{\omega}_{k'} - \hat{\omega}_{k'}$$

definieren. Diese sind in Abbildung 8.3 in rot eingezeichnet.

Es ist sicherlich ein sinnvoller Ansatz, die Gerade so zu wählen, dass die Abweichungen $\hat{e}_{k'}$ in einer gewissen Weise minimiert werden. Ein solcher Ansatz kann durch eine Gütekoeffizienten der Art

$$V = \sum_{k'} v(\hat{e}_{k'})$$

ausgedrückt werden, die minimiert werden soll. Dabei ist $v(\cdot)$ eine nicht-negative Funktion, die zumindest von der Abweichung abhängt. (Prinzipiell kann diese aber auch noch die Messungsnummer k' oder den Wert der Eingangsgröße $M_{k'}$ oder weitere Größen berücksichtigen.) Einfache Funktionen wären z. B. der Betrag $v(\hat{e}_{k'}) = |\hat{e}_{k'}|$ oder das Quadrat $v(\hat{e}_{k'}) = \hat{e}_{k'}^2$ der Abweichung.

In der Regel wird an dieser Stelle die quadratische Abweichung gewählt, so dass das Kriterium

$$V = \sum_{k'} \hat{e}_{k'}^2$$

zu minimieren ist. Die große Beliebtheit dieses Ansatzes liegt an dem pragmatischen Grund, dass hierfür eine analytische Lösung einfach angegeben werden kann, wenn das zugrundeliegende Modell (d. h. hier die Gerade) bezüglich der zu schätzenden Parameter linear ist.

Um für das vorliegende Beispiel auf eine Form zu kommen, in welche die Parameter linear eingehen, werden die neuen Parameter

$$d' = \frac{1}{d}$$

$$M'_C = -\frac{1}{d} \cdot M_C$$

eingeführt, mit denen der Zusammenhang

$$\omega_{k'} = d' \cdot M_{k'} + M'_C$$

lautet. (Die Linearität bezüglich der Parameter wird in der später eingeführten Vektorschreibweise deutlicher.)

Geht man nicht von den tatsächlichen Parametern d' und M'_C , sondern von geschätzten Parameterwerten \hat{d}' und \hat{M}'_C aus, so erhält man in Abhängigkeit des Eingangsmomentes $M_{k'}$ die dazugehörige geschätzte Geschwindigkeit

$$\hat{\omega}_{k'} = \hat{d}' \cdot M_{k'} + \hat{M}'_C.$$

Die Parameterwerte \hat{d}' und \hat{M}'_C sind jetzt so zu wählen, dass die Summe der quadrierten Schätzfehler

$$V(\hat{M}'_C, \hat{d}') = \sum_{k'} \hat{e}_{k'}^2 = \sum_{k'} (\tilde{\omega}_{k'} - \hat{\omega}_{k'})^2$$

minimiert werden. Formal kann dies als

$$[\hat{M}'_C \quad \hat{d}']^T = \arg \min_{[\hat{M}'_C \quad \hat{d}']^T} V(\hat{M}'_C, \hat{d}')$$

geschrieben werden.

Der Algorithmus zur Berechnung der Schätzwerte \hat{M}'_C und \hat{d}' wird später besprochen. Aus diesen kann dann auch

$$\hat{d} = \frac{1}{\hat{d}'}$$

$$\hat{M}_C = -\hat{M}'_C \cdot \hat{d}$$

angegeben werden.

Durchlaufofen

Als zweites Beispiel wird der Aufheizvorgang in einem Durchlaufofen betrachtet. Dabei wird Wärmegut mit einer Durchlaufgeschwindigkeit v durch einen Ofen mit einer Ofentemperatur von T_O gefördert. Dabei heizt sich das Wärmegut auf. Die Endtemperatur hängt dabei zum einen von den (wählbaren) Eingangsgrößen Ofentemperatur und Durchlaufgeschwindigkeit, und zum anderen von nicht vorgebbaren Größen wie die Anfangstemperatur des Wärmegutes und dessen Beschaffenheit ab.

Hier wird nur der stationäre Zusammenhang betrachtet. D. h. ändert sich die Ofentemperatur oder ändert man die Durchlaufgeschwindigkeit, so wird vor der Messung der Ausgangstemperatur des Wärmeguts immer gewartet, bis sich ein stationärer Zusammenhang eingestellt hat.

Geht man davon aus, dass sich die nicht vorgebbaren Größen wie die Anfangstemperatur des Wärmegutes nicht ändern, so hängt die Ausgangstemperatur y des Wärmeguts in nichtlinearer Weise von der Ofentemperatur und der Durchlaufgeschwindigkeit ab,

$$y_{k'} = y(T_{O,k'}, v_{k'}) .$$

Ein einfacher (affiner) Ansatz zur Beschreibung der Abhängigkeit der Temperatur y des Wärmegutes nach dem Ofen in Abhängigkeit von Ofentemperatur und Durchlaufgeschwindigkeit wäre

$$y_{k'} = T_0 + k_T \cdot T_{O,k'} + k_v \cdot v_{k'} ,$$

wobei zu erwarten ist, dass $k_T > 0$ und $k_v < 0$ ist. D. h. bei steigender Ofentemperatur sollte sich auch die Temperatur des Wärmeguts erhöhen und bei einer Erhöhung der Durchlaufgeschwindigkeit und damit der Verminderung der Aufenthaltsdauer im Ofen sollte die Temperatur des Wärmeguts abnehmen.

Dies ist natürlich ein sehr einfacher Ansatz, und sicherlich nur in einem bestimmten, engen Wertebereich von $v_{k'}$ und $T_{O,k'}$ gültig.

Eine Möglichkeit komplexeren Ansätze zu finden ist es ggf., eine physikalische Modellierung durchzuführen und damit die prinzipiellen Abhängigkeiten aufzuzeigen. Alternativ können, ohne einen physikalischen Hintergrund, rein anhand von Messdaten verschiedene Ansätze untersucht werden. So ein Vorgehen könnte z. B. auf einen Ansatz der Form

$$y_{k'} = T_0 + k_T \cdot T_{O,k'} + k_v \cdot v_{k'} + k_{TT} \cdot T_{O,k'}^2 + k_{vT} \cdot v_{k'} T_{O,k'} .$$

Führen. Beide Ansätze zeichnen sich dadurch aus, dass sie linear *in den Parametern* sind.

Weder das eine noch das andere soll hier jedoch vertieft werden. Für die anschauliche Diskussion der Eigenschaften des LS-Verfahrens ist schon der einfache, bezüglich der Eingangsgrößen $T_{O,k'}$ und $v_{k'}$ affine Ansatz sehr gut geeignet.

8.2 Allgemeine Formulierung

8.2.1 Prozess

Bei dem Beispiel der Reibung liegt mit dem Moment $M_{k'}$ eine physikalische Eingangsgröße vor. Es sind aber zwei Parameter, d' und M'_C zu identifizieren. Beim Beispiel des Durchlaufofens liegen zwei physikalische Eingangsgrößen, $T_{O,k'}$ und $v_{k'}$ vor. Je nach Ansatz sind aber drei bzw. fünf Parameter zu identifizieren.

Zur Anwendung des (linearen) LS-Verfahrens ist der Zusammenhang in eine bezüglich der Parameter lineare Form zu schreiben. Zu jedem der p Parameter gehört damit eine „Eingangsgröße“, die hier als $\psi_{j,k'}$, $j = 1, \dots, p$, bezeichnet wird.

Es muss also zwischen den physikalischen Eingängen u_i und den Eingangsgrößen ψ_j des LS-Problems unterschieden werden. Auch ist es günstig, zwischen dem physikalischen Ausgang y und der Ausgangsgröße η des LS-Problems zu unterscheiden.²

Damit kann der betrachtete Zusammenhang immer in der Form

$$\eta_{k'} = \theta_1 \psi_{1,k'} + \theta_2 \psi_{2,k'} + \dots + \theta_p \psi_{p,k'}$$

² Letzteres insbesondere, wenn später dynamische Systeme behandelt werden.

angegeben werden. Dabei sind die ψ_j des LS-Problems im Allgemeinen nichtlineare Funktionen der physikalischen Eingangsgrößen³ u_i ,

$$\psi_{j,k'} = \psi_j(u_{1,k'}, \dots, u_{p',k'}) .$$

Für den Ausgang gilt entsprechend

$$y_{k'} = y(\eta_{k'}) .$$

Fasst man die Eingangsgrößen sowie die Parameter in einen Eingangsvektor bzw. einen Parametervektor zusammen,

$$\begin{aligned} \Psi_{k'}^T &= [\psi_{1,k'} \ \psi_{2,k'} \ \dots \ \psi_{p,k'}] \quad \text{und} \\ \Theta &= [\theta_1 \ \theta_2 \ \dots \ \theta_p]^T , \end{aligned}$$

so kann man den Zusammenhang auch kompakt als

$$\eta_{k'} = \Psi_{k'}^T \Theta$$

schreiben.

Dieser Zusammenhang ist auch in Abbildung 8.4 dargestellt.

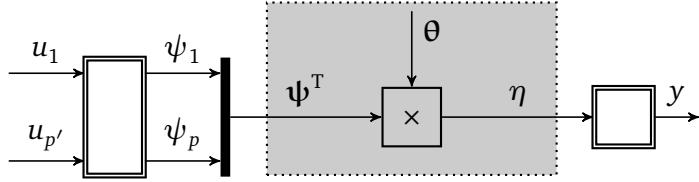


Abbildung 8.4: Größen des (statischen) LS-Problems

Im Weiteren wird immer von dem grau hinterlegten „LS-Prozess“ ausgegangen. Es ist daher zu beachten, dass sich alle Annahmen und daraus abgeleiteten stochastischen Eigenschaften zunächst nur auf die die Größen des „LS-Prozess“ beziehen. Was dies bezüglich der tatsächlichen Ein- und Ausgangsgrößen bedeutet, ist für den jeweils vorliegenden Fall zu prüfen.

Beispiel Coulomb'sche Reibung

Bei statischen Zusammenhängen ist häufig einer der zu schätzenden Parameter eine Konstante. Dessen Beitrag hängt von keiner physikalischen Eingangsgröße ab und entsprechend ist die dazugehörige „LS-Eingangsgröße“ $\psi_{1,k'} = 1$. (Die Konstante wird, wenn vorhanden, hier immer als erster Parameter, $j = 1$, geschrieben.)

Der Prozess zur Coulomb'schen Reibung lautet in der Standardformulierung

$$\begin{aligned} \Theta &= [M'_C \ d']^T \\ \Psi_{k'}^T &= [1 \ M_{k'}] \\ \eta_{k'} &= \omega_{k'} . \end{aligned}$$

³ Hier wird die Anzahl der physikalischen Parameter mit p' und nicht mit p bezeichnet. In den LS-Kapiteln ist mit p immer die Anzahl der zu identifizierenden Parameter gemeint!

Beispiel Durchlaufofen

Der einfache Ansatz lautet

$$\begin{aligned}\boldsymbol{\Theta} &= [T_0 \ k_T \ k_v]^T \\ \boldsymbol{\Psi}_{k'}^T &= [1 \ T_{0,k'} \ v_{k'}] \\ \eta_{k'} &= y_{k'}.\end{aligned}$$

Der etwas kompliziertere Ansatz würde über

$$\begin{aligned}\boldsymbol{\Theta} &= [T_0 \ k_T \ k_v \ k_{TT} \ k_{VT}]^T \\ \boldsymbol{\Psi}_{k'}^T &= [1 \ T_{0,k'} \ v_{k'} \ T_{0,k'}^2 \ v_{k'} T_{0,k'}] \\ \eta_{k'} &= y_{k'}\end{aligned}$$

beschrieben werden.

8.2.2 LS-Problem

Basierend auf dem Prozess

$$\begin{aligned}\eta_{k'} &= \theta_1 \psi_{1,k'} + \theta_2 \psi_{2,k'} + \cdots + \theta_p \psi_{p,k'} \\ &= \boldsymbol{\Psi}_{k'}^T \boldsymbol{\Theta}\end{aligned}\tag{8.1}$$

soll jetzt das LS-Problem mathematisch formuliert werden.

Ziel ist es, die Parameter $\boldsymbol{\Theta}$ durch Messungen zu bestimmen. Dabei sind die Messungen im Allgemeinen durch ein Rauschen $n_{\eta,k'}$ gestört, d. h. für die gemessene Ausgangsgröße gilt

$$\begin{aligned}\tilde{\eta}_{k'} &= \theta_1 \psi_{1,k'} + \theta_2 \psi_{2,k'} + \cdots + \theta_p \psi_{p,k'} + n_{\eta,k'} \\ &= \boldsymbol{\Psi}_{k'}^T \boldsymbol{\Theta} + n_{\eta,k'}.\end{aligned}$$

Dies entspricht der in Abbildung 8.5a gezeigten Struktur. Solange nicht explizit etwas anderes gesagt wird, wird davon ausgegangen, dass die Eingangsgrößen $\psi_{j,k'}$ exakt gemessen werden können. D. h. es wird *nicht* von dem in Abbildung 8.5b dargestellten Error-In-Variables-Ansatz ausgegangen.

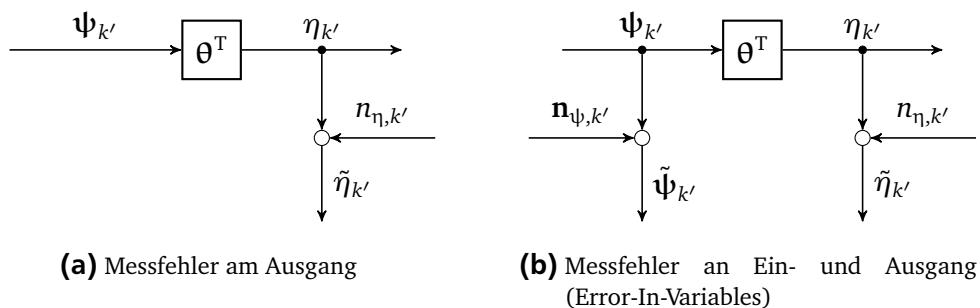


Abbildung 8.5

Die geschätzten Parameter werden mit einem „Dach“ gekennzeichnet, d. h. $\hat{\theta}_1, \dots, \hat{\theta}_p$. Berechnet man mit diesen Werten den Prozessausgang, so erhält man den Schätzwert

$$\begin{aligned}\hat{\eta}_{k'} &= \hat{\theta}_1 \psi_{1,k'} + \hat{\theta}_2 \psi_{2,k'} + \cdots + \hat{\theta}_p \psi_{p,k'} \\ &= \boldsymbol{\Psi}_{k'}^T \hat{\boldsymbol{\Theta}}.\end{aligned}$$

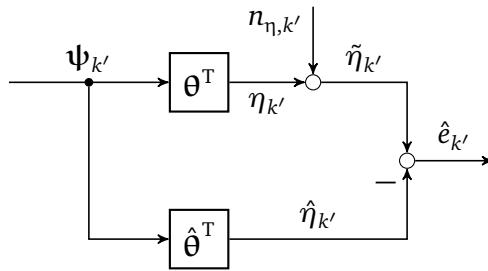


Abbildung 8.6: Schätzfehler

Der Schätzfehler

$$\hat{e}_{k'} = \tilde{\eta}_{k'} - \hat{\eta}_{k'}$$

gibt die Abweichung zwischen gemessenem und geschätztem Ausgangswert an. (Siehe Abbildung 8.6.) Ziel ist es nun, anhand von k Messungen⁴ $k' = 1, \dots, k$ die Schätzwerte $\hat{\theta}$ so zu bestimmen, dass die Summe der quadrierten Schätzfehler minimal wird, es soll also das Gütemaß

$$V = \sum_{k'=1}^k \hat{e}_{k'}^2$$

minimiert werden.

Fasst man die k Gleichungen

$$\begin{aligned}\eta_1 &= \theta_1 \psi_{1,1} + \theta_2 \psi_{2,1} + \dots + \theta_p \psi_{p,1} = \Psi_1^T \theta \\ \eta_2 &= \theta_1 \psi_{1,2} + \theta_2 \psi_{2,2} + \dots + \theta_p \psi_{p,2} = \Psi_2^T \theta \\ &\vdots \\ \eta_k &= \theta_1 \psi_{1,k} + \theta_2 \psi_{2,k} + \dots + \theta_p \psi_{p,k} = \Psi_k^T \theta\end{aligned}$$

zusammen, so kann dies auch als

$$\eta_k = \Psi_k \theta$$

mit

$$\eta_k = \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_k \end{bmatrix} \quad \text{und} \quad \Psi_k = \begin{bmatrix} \Psi_1^T \\ \vdots \\ \Psi_k^T \end{bmatrix} = \begin{bmatrix} \psi_{1,1} & \psi_{2,1} & \dots & \psi_{p,1} \\ \vdots & \vdots & & \vdots \\ \psi_{1,k} & \psi_{2,k} & \dots & \psi_{p,k} \end{bmatrix}$$

geschrieben werden. Entsprechend werden auch die Schätzwerte

$$\hat{\eta}_k = \Psi_k \hat{\theta}$$

und die Messwerte

$$\tilde{\eta}_k = \Psi_k \theta + n_{\eta,k}$$

⁴ Die Messungen werden bei $k' = 1$ beginnend gezählt!

in

$$\hat{\eta}_k = \begin{bmatrix} \hat{\eta}_1 \\ \vdots \\ \hat{\eta}_k \end{bmatrix}, \quad \tilde{\eta}_k = \begin{bmatrix} \tilde{\eta}_1 \\ \vdots \\ \tilde{\eta}_k \end{bmatrix} \quad \text{und} \quad \mathbf{n}_{\eta,k} = \begin{bmatrix} n_{\eta,1} \\ \vdots \\ n_{\eta,k} \end{bmatrix}$$

zusammengefasst.

Mit dem Vektor der Schätzfehler $\hat{\mathbf{e}}_k = \tilde{\eta}_k - \hat{\eta}_k$ kann das Gütemaß auch

$$V = \sum_{k'=1}^k \hat{e}_{k'}^2 = \hat{\mathbf{e}}_k^T \hat{\mathbf{e}}_k$$

geschrieben werden. Für den gesuchten Parametervektor $\hat{\theta}$ gilt, dass dieser das Gütemaß V minimieren soll,

$$\hat{\theta} = \arg \min_{\theta} V(\hat{\theta}).$$

8.3 Lösung des LS-Problems

8.3.1 Schätzgleichung

Setzt man $\hat{\mathbf{e}}_k = \tilde{\eta}_k - \hat{\eta}_k$ in $V = \hat{\mathbf{e}}_k^T \hat{\mathbf{e}}_k$ ein, so erhält man⁵

$$\begin{aligned} V &= (\tilde{\eta} - \Psi \hat{\theta})^T \cdot (\tilde{\eta} - \Psi \hat{\theta}) \\ &= \tilde{\eta}^T \tilde{\eta} - \tilde{\eta}^T \Psi \hat{\theta} - (\Psi \hat{\theta})^T \tilde{\eta} + (\Psi \hat{\theta})^T \Psi \hat{\theta} \\ &= \tilde{\eta}^T \tilde{\eta} - \tilde{\eta}^T \Psi \hat{\theta} - \hat{\theta}^T \Psi^T \tilde{\eta} + \hat{\theta}^T \Psi^T \Psi \hat{\theta}. \end{aligned}$$

Eine notwendige Bedingung für das Minimum ist, dass die Ableitungen von V bezüglich aller Parameter θ_j null ist,

$$\frac{dV(\hat{\theta}_1, \dots, \hat{\theta}_p)}{d\hat{\theta}_j} \stackrel{!}{=} 0 \quad \forall j = 1, \dots, p.$$

Dies kann auch als Ableitung nach der vektorwertigen Unabhängigen $\hat{\theta}$ geschrieben werden,

$$\frac{dV(\hat{\theta})}{d\hat{\theta}} = \begin{bmatrix} \frac{dV}{d\hat{\theta}_1} & \frac{dV}{d\hat{\theta}_2} & \dots & \frac{dV}{d\hat{\theta}_p} \end{bmatrix} \stackrel{!}{=} \mathbf{0}^T.$$

⁵ Der Übersichtlichkeit halber ist der Index k hier weggelassen. Dies wird auch im Weiteren meist gemacht, wenn nicht der „Prozesscharakter“ hervorgehoben werden soll.

Ableitungen nach Vektoren

Prinzipiell ist es einem freigestellt, wie man die Ableitung einer skalaren Funktion nach vektorwertigen Variablen definiert. So kann $\frac{dV(\mathbf{x})}{d\mathbf{x}}$ als stehender oder als liegender Vektor festgelegt werden. Beide Varianten werden in der Literatur auch verwendet. Es wird in diesem Skript dem liegenden Vektor den Vorzug gegeben,

$$\frac{dV(\mathbf{x})}{d\mathbf{x}} = \begin{bmatrix} \frac{dV}{dx_1} & \cdots & \frac{dV}{dx_n} \end{bmatrix},$$

da hier vektorwertige Variablen in der Regel von rechts multipliziert werden, und damit

$$\Delta V = \frac{dV(\mathbf{x})}{d\mathbf{x}} \cdot \Delta \mathbf{x}$$

mit dem stehenden Vektor $\Delta \mathbf{x}$ geschrieben werden kann.

Um nun die Ableitung der skalarwerten Funktion $\mathbf{x}^T \mathbf{b}$ zu bestimmen, kann das Vektorprodukt als Summe $\sum_j x_j b_j$ geschrieben werden, und diese dann „klassisch“ nach den einzelnen x_k abgeleitet werden,

$$\frac{d\mathbf{x}^T \mathbf{b}}{d\mathbf{x}} = \begin{bmatrix} \frac{d \sum_j x_j b_j}{dx_1} & \cdots & \frac{d \sum_j x_j b_j}{dx_n} \end{bmatrix} = \begin{bmatrix} b_1 & \cdots & b_n \end{bmatrix} = \mathbf{b}^T.$$

Mit dem gleichen Vorgehen sieht man auch, dass ebenfalls

$$\frac{d\mathbf{b}^T \mathbf{x}}{d\mathbf{x}} = \mathbf{b}^T$$

gilt.

Auch um die Ableitung $\mathbf{x}^T \mathbf{A} \mathbf{x}$ nach \mathbf{x} zu bestimmen, kann im Grunde gleich vorgegangen werden. D. h. es kann $\mathbf{x}^T \mathbf{A} \mathbf{x} = \sum_i \sum_j x_i a_{ij} x_j$ geschrieben und dann die einzelnen Ableitungen $\frac{d \sum_i \sum_j x_i a_{ij} x_j}{dx_k}$ bestimmt werden. Dies ist jedoch deutlich aufwendiger als oben.

Alternativ kann folgende Argumentation herangezogen werden. In $\mathbf{x}^T \mathbf{A} \mathbf{x}$ wird der Vektor \mathbf{x} durch $\mathbf{x}_0 + \delta \mathbf{x}$ ersetzt. Dies führt auf

$$\begin{aligned} (\mathbf{x}_0^T + \delta \mathbf{x}^T) \cdot \mathbf{A} \cdot (\mathbf{x}_0 + \delta \mathbf{x}) &= \mathbf{x}_0^T \mathbf{A} \mathbf{x}_0 + \mathbf{x}_0^T \mathbf{A} \delta \mathbf{x} + \delta \mathbf{x}^T \mathbf{A} \mathbf{x}_0 + \delta \mathbf{x}^T \mathbf{A} \delta \mathbf{x} \\ &= \mathbf{x}_0^T \mathbf{A} \mathbf{x}_0 + \mathbf{x}_0^T \mathbf{A} \delta \mathbf{x} + \mathbf{x}_0^T \mathbf{A}^T \delta \mathbf{x} + \delta \mathbf{x}^T \mathbf{A} \delta \mathbf{x} \\ &= \mathbf{x}_0^T \mathbf{A} \mathbf{x}_0 + (\mathbf{x}_0^T \mathbf{A} + \mathbf{x}_0^T \mathbf{A}^T) \cdot \delta \mathbf{x} + \delta \mathbf{x}^T \mathbf{A} \delta \mathbf{x}. \end{aligned}$$

Es ist damit die Taylorreihenentwicklung der Funktion $\mathbf{x}^T \mathbf{A} \mathbf{x}$ um den Punkt \mathbf{x}_0 erhalten. Der erste Summand stellt den Wert am Punkt \mathbf{x}_0 dar. Der zweite Summand beschreibt die Änderung des Funktionswert, die linear von der Abweichung zu \mathbf{x}_0 abhängt und der dritte Summand beschreibt eine quadratisch abhängige Änderung. Weitere Terme existieren hier nicht, da es sich um eine quadratische Funktion handelt. Die (erste) Ableitung beschreibt gerade den linearen Term, so dass

$$\frac{d\mathbf{x}^T \mathbf{A} \mathbf{x}}{d\mathbf{x}} = \mathbf{x}^T \cdot (\mathbf{A} + \mathbf{A}^T)$$

gilt. Wenn \mathbf{A} symmetrisch ist, also $\mathbf{A}^T = \mathbf{A}$ gilt, dann vereinfacht dies sich noch zu

$$\frac{d\mathbf{x}^T \mathbf{A} \mathbf{x}}{d\mathbf{x}} = 2\mathbf{x}^T \mathbf{A}.$$

Zusammenfassend gelten damit die Regeln

$$\frac{d\mathbf{x}^T \mathbf{b}}{d\mathbf{x}} = \mathbf{b}^T, \quad \frac{d\mathbf{b}^T \mathbf{x}}{d\mathbf{x}} = \mathbf{b}^T \quad \text{und} \quad \frac{d\mathbf{x}^T \mathbf{A} \mathbf{x}}{d\mathbf{x}} = 2\mathbf{x}^T \mathbf{A} \quad \text{wenn } \mathbf{A}^T = \mathbf{A}.$$

Schätzgleichung aus der notwendigen Bedingung

Mit den gerade dargestellten Regeln zur Ableitung skalarer Funktionen nach vektorwertigen Argumenten kann die erste Ableitung nach dem geschätzten Parametervektor $\hat{\theta}$ als

$$\begin{aligned}\frac{dV}{d\hat{\theta}} &= -\frac{d\tilde{\eta}^T \Psi \hat{\theta}}{d\hat{\theta}} - \frac{d\theta^T \Psi^T \tilde{\eta}}{d\hat{\theta}} + \frac{d\theta^T \Psi^T \Psi \hat{\theta}}{d\hat{\theta}} \\ &= -\tilde{\eta}^T \Psi - \tilde{\eta}^T \Psi + 2\hat{\theta}^T \Psi^T \Psi \\ &= -2\tilde{\eta}^T \Psi + 2\hat{\theta}^T \Psi^T \Psi\end{aligned}$$

bestimmt werden. Die notwendige Bedingung für das Minimum ist

$$\frac{dV}{d\hat{\theta}} \stackrel{!}{=} \mathbf{0}^T.$$

Besitzt Ψ vollen Spaltenrang, wovon ab einer bestimmten Anzahl k an Messungen bei vernünftiger Anregung ausgegangen werden kann, dann ist das Minimum eindeutig und $\frac{dV}{d\hat{\theta}} \stackrel{!}{=} \mathbf{0}^T$ damit eine hinreichende Bedingung.⁶ Man erhält dann

$$\hat{\theta}^T \cdot \Psi^T \Psi = \tilde{\eta}^T \Psi.$$

bzw. nach Transponieren

$$\Psi^T \Psi \cdot \hat{\theta} = \Psi^T \tilde{\eta}. \quad (8.2)$$

Diese Gleichungen werden als Normalgleichungen bezeichnet. Ist $\Psi^T \Psi$ invertierbar, was genau dann der Fall ist, wenn Ψ vollen Spaltenrang besitzt, so kann diese Gleichung formal nach

$$\hat{\theta} = (\Psi^T \Psi)^{-1} \Psi^T \tilde{\eta} \quad (8.3)$$

gelöst werden. Da aus numerischer Sicht die einfache Inversion nicht zu empfehlen ist (dies wird später noch ausführlicher diskutiert), dient diese Gleichung in erster Linie nicht zur tatsächlichen Berechnung der Schätzwerte, sondern zur Diskussion der Eigenschaften des LS-Schätzers.

Mit der Matrix

$$\mathbf{L}_k = \Psi_k^T \Psi_k$$

und der Abkürzung

$$\mathbf{r}_k = \Psi_k^T \tilde{\eta}_k$$

lässt sich die Schätzgleichung auch als

$$\hat{\theta}_k = \mathbf{L}_k^{-1} \mathbf{r}_k \quad (8.4)$$

schreiben.

⁶ Formal: Die Hessematrix $\frac{d^2 V}{d\hat{\theta}^2} = \Psi^T \Psi$ ist positiv definit, wenn Ψ vollen Spaltenrang hat.

Anschaulich: Wenn Ψ vollen Spaltenrang hat, dann ist $V = (\tilde{\eta} - \Psi \hat{\theta})^T \cdot (\tilde{\eta} - \Psi \hat{\theta})$ in alle Richtungen von $\hat{\theta}$ positiv unbeschränkt, es muss sich bei $\frac{dV}{d\hat{\theta}} \stackrel{!}{=} \mathbf{0}^T$ um das Minimum handeln.

Die Matrix \mathbf{L}_k ist symmetrisch und mindestens positiv semidefinit, d. h. es gilt $\mathbf{x}^T \mathbf{L}_k \mathbf{x} \geq 0$ für alle Vektoren \mathbf{x} .⁷ Hat Ψ vollen Spaltenrang, dann ist \mathbf{L}_k positiv definit, d. h. $\mathbf{x}^T \mathbf{L}_k \mathbf{x} > 0$ für alle $\mathbf{x} \neq \mathbf{0}$.⁸

Ist \mathbf{L}_k symmetrisch und positiv definit, so ist auch die Inverse \mathbf{L}_k^{-1} symmetrisch⁹ und positiv definit¹⁰. Aus der positiven Definitheit folgt auch, dass alle Diagonaleinträge der Matrizen \mathbf{L}_k und \mathbf{L}_k^{-1} positiv sind¹¹.

Die Matrix \mathbf{L}_k sowie der Vektor \mathbf{r}_k lassen sich nach

$$\mathbf{L}_k = \Psi_k^T \Psi_k = [\Psi_1 \ \cdots \ \Psi_k] \cdot \begin{bmatrix} \Psi_1^T \\ \vdots \\ \Psi_k^T \end{bmatrix} = \sum_{k'=1}^k \Psi_{k'} \Psi_{k'}^T$$

und

$$\mathbf{r}_k = \Psi_k^T \tilde{\eta}_k = [\Psi_1 \ \cdots \ \Psi_k] \cdot \begin{bmatrix} \tilde{\eta}_1 \\ \vdots \\ \tilde{\eta}_k \end{bmatrix} = \sum_{k'=1}^k \Psi_{k'} \tilde{\eta}_{k'}$$

auch als Summe über die entsprechenden Daten der einzelnen Messungen schreiben. Von dieser Darstellung wird später insbesondere bei Beispielen zu theoretischen Überlegungen Gebrauch gemacht.

Beispiel Reibung

Für das Beispiel der Identifikation des Coulomb'schen Rauschens ergeben sich mit

$$\Psi = \begin{bmatrix} 1 & 0,13 \\ 1 & 0,14 \\ 1 & 0,15 \\ 1 & 0,16 \\ 1 & 0,17 \\ 1 & 0,18 \\ 1 & 0,19 \\ 1 & 0,20 \\ 1 & 0,21 \\ 1 & 0,22 \end{bmatrix} \quad \text{und} \quad \tilde{\eta} = \begin{bmatrix} 13,1040 \\ 18,5890 \\ 18,6851 \\ 25,9949 \\ 29,6403 \\ 32,2026 \\ 37,2655 \\ 42,2305 \\ 49,6551 \\ 53,0349 \end{bmatrix}$$

die Schätzwerte

$$\hat{\theta} = [-45,4839 \ 442,9949]^T$$

und damit

$$\hat{M}_C = 0,1027 \text{ Nm} \quad \text{und} \quad \hat{d} = 0,00226 \text{ Nm/(rad/s)}.$$

In Abbildung 8.7 ist in grün die mit diesen Parametern bestimmte Gerade eingezeichnet, wobei die Abweichungen wieder in rot markiert sind. Zum Vergleich ist in hellgrau die „wahre“ Gerade eingezeichnet.

⁷ Jede Matrix $\mathbf{M}^T \mathbf{M}$ ist positiv semidefinit, da $\mathbf{x}^T \mathbf{M}^T \mathbf{M} \mathbf{x} = \mathbf{y}^T \mathbf{y}$ und $\mathbf{y}^T \mathbf{y} \geq 0 \ \forall \mathbf{y}$.

⁸ Besitzt \mathbf{M} vollen Spaltenrang, dann gilt $\mathbf{M} \mathbf{x} \neq 0$ für alle $\mathbf{x} \neq 0$. D. h., $\mathbf{x}^T \mathbf{M}^T \mathbf{M} \mathbf{x} > 0 \ \forall \mathbf{x} \neq 0$.

⁹ $\mathbf{I} = (\mathbf{L} \cdot \mathbf{L}^{-1})^T = (\mathbf{L}^{-1})^T \cdot \mathbf{L}^T = (\mathbf{L}^{-1})^T \cdot \mathbf{L} \Rightarrow (\mathbf{L}^{-1})^T = \mathbf{L}^{-1}$.

¹⁰ Es gelte $\mathbf{x}^T \mathbf{L} \mathbf{x} \geq 0$. Mit $\mathbf{x} = \mathbf{L}^{-1} \mathbf{y}$ gilt für alle \mathbf{y} , dass $\mathbf{x}^T \mathbf{L} \mathbf{x} = \mathbf{y}^T \mathbf{L}^{-1} \mathbf{L} \mathbf{L}^{-1} \mathbf{y} = \mathbf{y}^T \mathbf{L}^{-1} \mathbf{y} \geq 0$.

¹¹ Mit dem i -ten Einheitsvektor \mathbf{e}_i gilt $\mathbf{e}_i^T \mathbf{L} \mathbf{e}_i = (\mathbf{L})_{ii} \geq 0$.

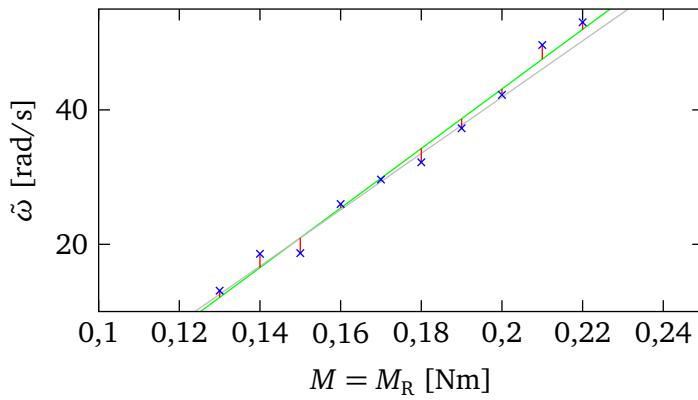


Abbildung 8.7: Optimale Gerade im Sinne der kleinsten Summe der quadrierten Fehler

8.3.2 Eigenschaften

Zur Herleitung der Schätzgleichung $\hat{\theta} = (\Psi^T \Psi)^{-1} \Psi^T \tilde{\eta}$ wurden keine Annahmen über den Messfehler $n_{\eta,k'}$ getroffen. Das LS-Verfahren stellt unabhängig von dem Rauschen sicher, dass die Summe der quadratischen Abweichungen minimiert wird. D. h. es kann prinzipiell auch ohne irgendwelche stochastische Überlegungen eingesetzt werden. Dennoch ist es, insbesondere wenn mit dem Modell Vorhersagen der Ausgangsgröße bei anderen als den zur Identifikation verwendeten Eingangsdaten gemacht werden sollen (was im Kontext der Systemidentifikation praktisch immer der Fall ist), interessant zu wissen, wie gut oder sicher die Parameter geschätzt werden.

Um weitere Eigenschaften, insbesondere bezüglich der geschätzten Parameter $\hat{\theta}$ angeben zu können, müssen Annahmen getroffen werden. Die folgende Aufzählung fasst die „klassischen“ Annahmen zusammen. Anschließend werden verschiedene Eigenschaften vorgestellt, für die zum Teil jedoch nicht alle unten genannten Annahmen zutreffen müssen. [Wooldridge 2009, 803ff]

Die Herleitungen bzw. Beweise zu den meisten der Eigenschaften werden nach dieser Übersicht gezeigt.

Annahmen

A1: $\eta = \Psi^T \theta$:

Das ungestörte System wird durch die Gleichung (8.1) vollständig beschrieben.

A2: $E\{\eta_r\}_r = \mathbf{0}$ und nicht mit $\{\Psi\}_r$ korreliert ($E\{\eta_r|\Psi\}_r = \mathbf{0}$):

Das Rauschen hat den Erwartungswert Null und ist nicht mit den Eingangsgrößen korreliert.

A3: $\text{rg}(\Psi) = p$ (Ψ besitzt vollen Spaltenrang):

Es existieren keine exakten linearen Abhängigkeiten unter den Eingangsgrößen.

A4: $\text{Cov}(\{\eta_r\}_r, \{\eta_r\}_r) = \sigma_\eta^2 \cdot \mathbf{I}$:

Die Varianz des Rauschens ist konstant σ_η^2 und das Rauschen ist unkorreliert

A5: $\{n_{\eta,k}\}_r \sim N(0, \sigma_\eta^2)$:

Das Rauschen ist normalverteilt

Wie aus der Annahme A2 ersichtlich, ist auch der Fall zugelassen, dass die Eingangswerte Ψ eine Realisierung einer Zufallsvariablen ist. In diesem Fall ist auch η eine Zufallsvariable. Dies bedeutet aber nicht, dass damit der Fall eines Messrauschen auf der Eingangsgröße abgedeckt ist! Es wird immer noch davon ausgegangen, dass die Eingangsgrößen exakt messbar bzw. bekannt sind. Der bedingte Erwartungswert $E\{\eta_r|\Psi\}_r = \mathbf{0}$ drückt aus, dass das Rauschen für jeden möglichen Wert von $\{\Psi\}_r$ mittelwertfrei ist.

Die wesentlichen Eigenschaften von Schätzern sind Bias, Konsistenz und Effizienz.

Biasfreiheit

Unter Bias wird die Abweichung des Erwartungswerts der Schätzwerte $\hat{\theta}_k$ vom tatsächlichen Wert θ verstanden, $b = E\{\hat{\theta}_k - \theta\}_r$.

Unter den Annahmen A1, A2 und A3 ist der Schätzwert $\hat{\theta}_k$ biasfrei, $E\{\hat{\theta}_k\}_r = \theta$. Dies gilt für jedes beliebige k und nicht nur für $k \rightarrow \infty$.

Konsistenz

Ein Schätzer wird als (schwach) konsistent bezeichnet, wenn der Schätzwert für $k \rightarrow \infty$ schwach gegen den wahren Wert konvergiert,

$$\text{plim}_{k \rightarrow \infty} \{\hat{\theta}_k\}_r = \theta .$$

Ein Schätzer wird als „konsistent im quadratischen Mittel“ bezeichnet, wenn

$$\lim_{k \rightarrow \infty} E\{(\hat{\theta}_k - \theta) \cdot (\hat{\theta}_k - \theta)^T\}_r = \mathbf{0} ,$$

also

$$\text{l.i.m.}_{k \rightarrow \infty} \{\hat{\theta}_k\}_r = \theta$$

gilt.

Ein konsistenter Schätzer ist (zumindest) asymptotisch biasfrei und die Varianz geht für $k \rightarrow \infty$ gegen Null. (Zumindest aus praktischer Sicht. Mathematisch muss der Erwartungswert bei einem (schwach) konsistenten Schätzer nicht existieren, und damit kann der Bias und die Varianz nicht angegeben werden.)

Unter den Annahmen A1 bis A4 und einer dauerhaften Anregung (dazu später) ist der LS-Schätzer konsistent.

Effizienz

Unter Effizienz wird die Eigenschaft verstanden, dass ein Schätzer den minimalen quadratischen Fehler innerhalb einer bestimmten Gruppe von Schätzern besitzt.

Unter den Annahmen A1 bis A4 besitzt der LS-Schätzer die geringste Varianz¹² aller linearen, biasfreien Schätzer (Gauß-Markov-Theorem, [Wooldridge 2009, S. 806]). Dabei werden unter linearen Schätzern alle Schätzer verstanden, deren Schätzgleichung in der Form $\hat{\theta} = \mathbf{M}(\Psi) \cdot \tilde{\eta}$ geschrieben werden kann, wobei die Einträge der Matrix \mathbf{M} Funktionen der Eingangsgrößen sein können, aber unabhängig von den gemessenen Ausgangsgrößen $\tilde{\eta}$ sein müssen.

Wird die Annahme A5 (normalverteiltes Rauschen) hinzugenommen, so lässt sich zeigen, dass die Varianz des LS-Schätzers gerade der Cramer-Rao-Grenze entspricht, welche die geringste Varianz aller (nicht nur linearen) biasfreien Schätzer darstellt. Somit kann kein biasfreier Schätzer besser als der LS-Schätzer sein und dieser ist damit effizient. [Wooldridge 2009, 808f] ¹³

Varianz der Schätzwerte

Unter den Annahmen A1 bis A4 gilt für die Kovarianzmatrix $\mathbf{P} = \text{Cov}(\{\hat{\theta}\}_r, \{\hat{\theta}\}_r)$ der Schätzwerte

$$\mathbf{P} = \sigma_\eta^2 \cdot \mathbf{L}^{-1} . \tag{8.5}$$

Die Varianz der Schätzwerte ist eine wesentliche Größe, so dass auf diese noch öfters eingegangen werden wird. Die prinzipiellen Einflussfaktoren auf die Varianz werden in Abschnitt 8.4 diskutiert.

¹² Dies bedeutet, dass die Varianz jedes Schätzwertes θ_i minimal ist.

¹³ Für die Herleitung der Cramer-Rao-Grenze siehe z. B. [Hänsler 2001, 434ff].

Verteilung der Schätzwerte

Unter den Annahmen A1 bis A5 sind die Schätzwerte $\hat{\theta}$ ebenfalls normalverteilt.

Herleitungen der stochastischen Eigenschaften

Erwartungswert

Für den Erwartungswert gilt allgemein

$$\begin{aligned} E\{\hat{\theta}\}_r &= E\{(\Psi^T \Psi)^{-1} \Psi^T \tilde{\eta}\}_r \\ &= E\{(\Psi^T \Psi)^{-1} \Psi^T (\eta + n_\eta)\}_r \\ &= E\{(\Psi^T \Psi)^{-1} \Psi^T \eta + (\Psi^T \Psi)^{-1} \Psi^T n_\eta\}_r. \end{aligned}$$

Der erste Summand $(\Psi^T \Psi)^{-1} \Psi^T \eta$ ist genau θ und keine Zufallsvariable¹⁴, womit man

$$E\{\hat{\theta}\}_r = \theta + E\{(\Psi^T \Psi)^{-1} \Psi^T n_\eta\}_r$$

erhält. Unter der Annahme, dass die Einträge in Ψ und das Messrauschen n_η nicht korreliert sind, kann der Erwartungswert für das Produkt $(\Psi^T \Psi)^{-1} \Psi^T n_\eta$ aufgeteilt werden, so dass sich

$$E\{\hat{\theta}\}_r = \theta + E\{(\Psi^T \Psi)^{-1} \Psi^T\}_r \cdot E\{n_\eta\}_r$$

ergibt. (Die Existenz des Erwartungswertes $E\{(\Psi^T \Psi)^{-1} \Psi^T\}_r$ vorausgesetzt.) Ist weiter das Messrauschen mittelwertfrei, d. h. $E\{n_\eta\}_r = \mathbf{0}$, so gilt

$$E\{\hat{\theta}\}_r = \theta.$$

Unter den Voraussetzungen, dass

- der reale Prozess durch den angenommenen Zusammenhang beschrieben wird ($\eta = \Psi \theta$),
- die Eingangsgrößen und das Rauschen nicht korreliert sind ($E\{(\Psi^T \Psi)^{-1} \Psi^T n_\eta\}_r = E\{(\Psi^T \Psi)^{-1} \Psi^T\}_r \cdot E\{n_\eta\}_r$) und
- das Rauschen mittelwertfrei ist ($E\{n_\eta\}_r = \mathbf{0}$),

ist der Least-Squares-Schätzer demnach biasfrei.

Varianz bei deterministischen Eingängen

Die Kovarianzmatrix der geschätzten Parameter ist definiert als

$$\mathbf{P} = E\left\{(\hat{\theta} - E\{\hat{\theta}\}_r) \cdot (\hat{\theta} - E\{\hat{\theta}\}_r)^T\right\}_r.$$

Unter den oben gemachten Voraussetzungen gilt $E\{\hat{\theta}\}_r = \theta$, und damit

$$\begin{aligned} \mathbf{P} &= E\left\{((\Psi^T \Psi)^{-1} \Psi^T \tilde{\eta} - \theta) \cdot ((\Psi^T \Psi)^{-1} \Psi^T \tilde{\eta} - \theta)^T\right\}_r \\ &= E\left\{((\Psi^T \Psi)^{-1} \Psi^T \tilde{\eta} + (\Psi^T \Psi)^{-1} \Psi^T n_\eta - \theta) \cdot ((\Psi^T \Psi)^{-1} \Psi^T \tilde{\eta} + (\Psi^T \Psi)^{-1} \Psi^T n_\eta - \theta)^T\right\}_r \\ &= E\left\{((\Psi^T \Psi)^{-1} \Psi^T n_\eta) \cdot ((\Psi^T \Psi)^{-1} \Psi^T n_\eta)^T\right\}_r \\ &= E\left\{(\Psi^T \Psi)^{-1} \Psi^T n_\eta n_\eta^T \Psi \cdot ((\Psi^T \Psi)^{-1})^T\right\}_r. \end{aligned} \tag{8.6}$$

¹⁴ Dies ist auch der Fall, wenn die Eingänge zufällig sein können, $\{\Psi\}_r$. Dann sind zwar auch $(\Psi^T \Psi)^{-1} \Psi^T$ und η Zufallsvariablen, aber deren Produkt ist für jede Realisierung (sofern die Inverse existiert) θ .

Auch hier wurde im zweiten Schritt wieder davon Gebrauch gemacht, dass auch wenn die Eingangsgrößen Zufallsvariablen sind für alle Realisierungen $(\Psi^T \Psi)^{-1} \Psi^T \eta - \theta = 0$ gilt.

Die Eingangsmatrix Ψ wird als deterministisch angenommen, so dass diese aus dem Erwartungswert herausgezogen werden kann. Außerdem gilt $((\Psi^T \Psi)^{-1})^T = (\Psi^T \Psi)^{-1}$, da $\Psi^T \Psi$ symmetrisch ist. Somit erhält man

$$\mathbf{P} = (\Psi^T \Psi)^{-1} \Psi^T \cdot E\{\mathbf{n}_\eta \mathbf{n}_\eta^T\}_r \cdot \Psi \cdot (\Psi^T \Psi)^{-1}.$$

Nimmt man weiter an, dass $\{n_{\eta,k'}\}_r$ unkorreliert ist und für alle k' dieselbe Varianz σ_η^2 besitzt, so ist $E\{\mathbf{n}_\eta \mathbf{n}_\eta^T\}_r = \sigma_\eta^2 \cdot \mathbf{I}$. Damit ergibt sich die Varianz der geschätzten Parameter zu

$$\mathbf{P} = \sigma_\eta^2 \cdot (\Psi^T \Psi)^{-1}. \quad (8.7)$$

Varianz, wenn Eingänge auch Zufallsvariablen sind

Mit der Abkürzung $\mathbf{M} = \Psi \cdot (\Psi^T \Psi)^{-1}$ kann Gl. (8.6) auch als

$$(\mathbf{P})_{qp} = E\left\{ \sum_{i=1}^k (\mathbf{M}^T)_{qi} \sum_{j=1}^k (\mathbf{n}_\eta \mathbf{n}_\eta^T)_{ij} (\mathbf{M})_{jp} \right\}_r = E\left\{ \sum_{i=1}^k \sum_{j=1}^k (\mathbf{M}^T)_{qi} (\mathbf{n}_\eta \mathbf{n}_\eta^T)_{ij} (\mathbf{M})_{jp} \right\}_r$$

geschrieben werden. Die Erwartungswertbildung kann immer in die Summe gezogen werden, und nimmt man weiter die Unkorreliertheit von $\{\mathbf{M}\}_r$ und $\{\mathbf{n}_\eta\}_r$ an, so folgt

$$(\mathbf{P})_{qp} = \sum_{i=1}^k \sum_{j=1}^k E\{(\mathbf{M}^T)_{qi} (\mathbf{M})_{jp}\}_r \cdot E\{(\mathbf{n}_\eta \mathbf{n}_\eta^T)_{ij}\}_r.$$

Unter der Annahme, dass $\{n_{\eta,k'}\}_r$ unkorreliert ist und für alle k' dieselbe Varianz σ_η^2 besitzt,

$$E\{(\mathbf{n}_\eta \mathbf{n}_\eta^T)_{ij}\}_r = \begin{cases} \sigma_\eta^2 & i = j \\ 0 & i \neq j, \end{cases}$$

vereinfacht sich dies weiter zu

$$(\mathbf{P})_{qp} = \sigma_\eta^2 \cdot \sum_{i=1}^k E\{(\mathbf{M}^T)_{qi} (\mathbf{M})_{ip}\}_r,$$

bzw.

$$\mathbf{P} = \sigma_\eta^2 \cdot E\{\mathbf{M}^T \mathbf{M}\}_r = \sigma_\eta^2 \cdot E\{(\Psi^T \Psi)^{-1}\}_r.$$

Erwartungswert und Varianz über bedingte Erwartungswerte

Eine weitere Möglichkeit mit zufälligen Eingangsgrößen umzugehen, die in der Literatur auch häufig angewendet wird, ist es mit bedingten Erwartungswerten zu arbeiten.

Für die Betrachtung des Bias wird damit zunächst der bedingte Erwartungswert

$$E\{\hat{\theta}|\Psi\}_r = E\{(\Psi^T \Psi)^{-1} \Psi^T \tilde{\eta} | \Psi\}_r$$

betrachtet. Bezuglich der bedingten Erwartungswerte $E\{\cdot|\Psi\}_r$ kann Ψ wie eine Konstante behandelt werden. Damit erhält man

$$\begin{aligned} E\{\hat{\theta}|\Psi\}_r &= \Theta + E\{(\Psi^T\Psi)^{-1}\Psi^T\mathbf{n}_\eta | \Psi\}_r \\ &= \Theta + (\Psi^T\Psi)^{-1}\Psi^T \cdot E\{\mathbf{n}_\eta|\Psi\}_r. \end{aligned}$$

Aufgrund der Voraussetzung A2, $E\{\mathbf{n}_\eta|\Psi\}_r = \mathbf{0}$, gilt damit

$$E\{\hat{\theta}|\Psi\}_r = \Theta.$$

Dieser bedingte Erwartungswert hängt nicht mehr von Ψ ab, d. h. es gilt für alle Realisierungen von $\{\Psi\}_r$, dass der Schätzer biasfrei ist. Man kann dies auch als

$$E\{E\{\hat{\theta}|\Psi\}_r\}_r = \Theta$$

schreiben. (Im Allgemeinen ist ein bedingter Erwartungswert eine Zufallsvariable. Hier ist jedoch schon der bedingte Erwartungswert eine Konstante, so dass die unbedingte Erwartungswertbildung äußerst einfach auszuführen ist.)

Es soll hier nochmals betont werden, dass dies nicht den Fall von Messrauschen am Eingang (Abbildung 8.5b) einschließt! In diesem Fall müsste der Erwartungswert¹⁵ von

$$E\{\hat{\theta}\}_r = E\{(\tilde{\Psi}^T\tilde{\Psi})^{-1}\tilde{\Psi}^T\tilde{\eta}\}_r$$

gebildet werden. Dies kann wieder zu

$$E\{\hat{\theta}\}_r = E\{(\tilde{\Psi}^T\tilde{\Psi})^{-1}\tilde{\Psi}^T\eta\}_r + E\{(\tilde{\Psi}^T\tilde{\Psi})^{-1}\tilde{\Psi}^T\mathbf{n}_\eta\}_r$$

umgeformt werden. Der zweite Summand wird wieder Null (vorausgesetzt, das auf den Eingang wirkende Rauschen ist nicht mit dem Ausgangsrauschen korreliert), allerdings kann der erste Summand nicht durch Θ ersetzt werden, so dass man

$$E\{\hat{\theta}\}_r = E\{(\tilde{\Psi}^T\tilde{\Psi})^{-1}\tilde{\Psi}^T\eta\}_r$$

erhält. Im Allgemeinen wird der Erwartungswert auf der rechten Seite nicht dem wahren Parameterwert Θ entsprechen.

Im Fall der Varianz wird entsprechend zunächst nach der bedingten Varianz

$$\mathbf{P}_{|\Psi} = E\left\{(\hat{\theta} - E\{\hat{\theta}\}_r) \cdot (\hat{\theta} - E\{\hat{\theta}\}_r)^T \mid \Psi\right\}_r$$

gefragt. Unter sonst gleichen Annahmen wie oben erhält man

$$\begin{aligned} \mathbf{P}_{|\Psi} &= E\left\{((\Psi^T\Psi)^{-1}\Psi^T\tilde{\eta} - \Theta) \cdot ((\Psi^T\Psi)^{-1}\Psi^T\tilde{\eta} - \Theta)^T \mid \Psi\right\}_r \\ &= E\left\{(\Psi^T\Psi)^{-1}\Psi^T\mathbf{n}_\eta\mathbf{n}_\eta^T\Psi \cdot ((\Psi^T\Psi)^{-1})^T \mid \Psi\right\}_r \\ &= (\Psi^T\Psi)^{-1}\Psi^T \cdot E\{\mathbf{n}_\eta\mathbf{n}_\eta^T \mid \Psi\}_r \cdot \Psi \cdot (\Psi^T\Psi)^{-1} \\ &= \sigma_\eta^2 \cdot (\Psi^T\Psi)^{-1}. \end{aligned}$$

Diese so berechnete Kovarianzmatrix $\mathbf{P}_{|\Psi}$ ist immer noch eine Zufallsvariable, da diese Ergebnis eines bedingten Erwartungswertes ist. Bildet man von $\{\mathbf{P}_{|\Psi}\}_r$ den unbedingten Erwartungswert, so erhält man den deterministischen Wert

$$\mathbf{P} = E\{\mathbf{P}_{|\Psi}\}_r = \sigma_\eta^2 \cdot E\{(\Psi^T\Psi)^{-1}\}_r,$$

der mit dem oben erhaltenen Ergebnis übereinstimmt.

Anschaulich bedeutet das Ergebnis, dass die Varianz der Schätzwerte von den Werten der Eingangsgrößen abhängen, und die Varianz der Schätzwerte im Allgemeinen auch nur dann angegeben werden kann, wenn die Werte der Eingangsgrößen bekannt sind.

¹⁵ Bedingt oder unbedingt macht hier für das zu Zeigende keinen großen Unterschied.

Konsistenz

Die Konsistenz des Schätzers wird später diskutiert.

Effizienz bezüglich der Klasse linearer biasfreier Schätzer

Unter einem linearen Schätzer wird ein Schätzer der Form

$$\hat{\theta} = M\tilde{\eta}$$

verstanden, bei dem der Schätzwert linear von den gemessenen Ausgangsdaten $\tilde{\eta}$ abhängt. Die Matrix M darf dabei natürlich nicht von $\tilde{\eta}$, aber (durchaus auch in nichtlinearer Form) von den Eingangsgrößen Ψ abhängen. Für das LS-Verfahren gilt z. B. $M = (\Psi^T \Psi)^{-1} \Psi^T$.

Es soll jetzt die Varianz des LS-Schätzers

$$\hat{\theta} = (\Psi^T \Psi)^{-1} \Psi^T \tilde{\eta}$$

mit der eines allgemeinen linearen Schätzers

$$\hat{\theta}' = M' \tilde{\eta}$$

verglichen werden, wobei nur biasfreie lineare Schätzer betrachtet werden.¹⁶

Die folgende Herleitung des Gauß-Markov-Theorems ist [Wooldridge 2009, S. 806] entnommen. Zunächst wird der Erwartungswert eines allgemeinen linearen Schätzers gebildet,

$$\begin{aligned} E\{\hat{\theta}'\}_r &= M' \cdot E\{\eta + n_\eta\}_r \\ &= M' \eta + M' \cdot E\{n_\eta\}_r \\ &= M' \Psi \theta . \end{aligned}$$

Daraus lässt sich ablesen, dass für einen biasfreien linearen Schätzer

$$M' \Psi = I$$

gelten muss.

Für die Varianz des allgemeinen linearen Schätzers gilt

$$\text{Var}\{\hat{\theta}'\}_r = M' \cdot \text{Var}\{n_\eta\}_r \cdot M'^T = \sigma_\eta^2 \cdot M' M'^T ,$$

wobei wieder von $\text{Var}\{n_\eta\}_r = \sigma_\eta^2 \cdot I$ ausgegangen wird. Bildet man die Differenz der Varianzen des allgemeinen linearen Schätzers und des LS-Schätzers, so erhält man

$$\begin{aligned} \text{Var}\{\hat{\theta}'\}_r - \text{Var}\{\hat{\theta}\}_r &= \sigma_\eta^2 \cdot (M' M'^T - (\Psi^T \Psi)^{-1}) \\ &= \sigma_\eta^2 \cdot (M' M'^T - \underbrace{M' \Psi}_{I} \cdot (\Psi^T \Psi)^{-1} \cdot \underbrace{\Psi^T M'}_{I}) \\ &= \sigma_\eta^2 \cdot M' \cdot (I - \Psi \cdot (\Psi^T \Psi)^{-1} \cdot \Psi^T) \cdot M'^T \\ &= \sigma_\eta^2 \cdot M' \cdot \Pi_{\Psi^\perp} \cdot M'^T , \end{aligned}$$

¹⁶ In irgendeiner Form muss eine Beschränkung erfolgen. Würden alle linearen Schätzer zugelassen werden, dann könnte z. B. auch $\hat{\theta}' = 0\tilde{\eta}$ gewählt werden. Dieser „Schätzer“ besitzt eine Varianz von null, ist aber offensichtlich unbrauchbar.

wobei im zweiten Schritt von der Gleichheit $\mathbf{M}'\Psi = \mathbf{I}$ für biasfreie Schätzer Gebrauch gemacht wurde. Im letzten Schritt wurde die Abkürzung $\Pi_{\Psi^\perp} = \mathbf{I} - \Psi \cdot (\Psi^T \Psi)^{-1} \cdot \Psi^T$ verwendet.

Es gilt $\mathbf{M}' \cdot \Pi_{\Psi^\perp} \cdot \mathbf{M}'^T = \mathbf{M}' \cdot \Pi_{\Psi^\perp} \cdot \Pi_{\Psi^\perp}^T \cdot \mathbf{M}'^T$, d. h. $\mathbf{M}' \cdot \Pi_{\Psi^\perp} \cdot \mathbf{M}'^T$ ist positiv semidefinit. Damit gilt für jede Linearkombination $\gamma^T \theta$ der Parameter

$$\text{Var}\{\gamma^T \hat{\theta}'\}_r - \text{Var}\{\gamma^T \hat{\theta}\}_r = \sigma_\eta^2 \cdot \gamma^T \mathbf{M}' \cdot \Pi_{\Psi^\perp} \cdot \mathbf{M}'^T \gamma \geq 0.$$

Insbesondere gilt damit

$$\text{Var}\{\hat{\theta}_i\}_r \leq \text{Var}\{\hat{\theta}'_i\}_r.$$

Damit ist bewiesen, dass der LS-Schätzer der beste lineare biasfreie Schätzer ist. Diese Eigenschaft wird häufig als BLUE (best linear unbiased estimator) bezeichnet. Dabei bezieht sich das „beste“ auf den mittleren quadratischen Fehler, MSE. Da nach der Voraussetzung hier nur biasfreie Schätzer betrachtet werden, entspricht die Varianz gerade dem MSE.

Es kann aber durchaus sein, dass ein anderer linearer Schätzer zwar biasbehaftet ist, dafür aber eine geringere Varianz besitzt, so dass auch der MSE geringer ist.

Effizienz bezüglich der Klasse biasfreier Schätzer und Normalverteilung der Schätzwerte

Der Beweis der allgemeinen Effizienz würde den Rahmen der Vorlesung sprengen. Ebenso wird nicht bewiesen, dass die Schätzwerte der Parameter normalverteilt sind, wenn das Messrauschen normalverteilt ist.

8.3.3 Schätzung der Varianz des Messrauschens

Die Varianz des Messrauschens ist nicht immer bekannt. Um in diesem Fall dennoch die Varianz der Schätzwerte abschätzen zu können, wofür ja der Wert der Varianz des Messrauschens bekannt sein muss, kann die Varianz des Messrauschens aus den Messwerten geschätzt werden. Und auch wenn die Varianz des Messrauschens bekannt ist, kann es sinnvoll sein, den Schätzwert für diese Varianz zu bestimmen. Dieser kann z. B. darüber Aufschluss geben, ob der verwendete Ansatz $\eta = \Psi^T \theta$ das vorliegende System sinnvoll beschreiben kann, oder er kann zu Diagnosezwecken verwendet werden.

Die Herleitung folgt hier [Wooldridge 2009, S. 807].

Der geschätzte Messfehler \hat{n}_η entspricht dem Schätzfehler \hat{e} , es gilt also

$$\hat{n}_\eta = \tilde{\eta} - \hat{\eta} = \tilde{\eta} - \Psi \hat{\theta} = \tilde{\eta} - \Psi \cdot (\Psi^T \Psi)^{-1} \cdot \Psi^T \tilde{\eta} = (\mathbf{I} - \Psi \cdot (\Psi^T \Psi)^{-1} \cdot \Psi^T) \cdot \tilde{\eta} = \Pi_{\Psi^\perp} \tilde{\eta},$$

wobei $\Pi_{\Psi^\perp} = (\mathbf{I} - \Psi \cdot (\Psi^T \Psi)^{-1} \Psi^T)$ die Projektionsmatrix auf das senkrechte Komplement zum Spaltenraum von Ψ darstellt.

Mit $\tilde{\eta} = \eta + n_\eta = \Psi \theta + n_\eta$ und $\Pi_{\Psi^\perp} \Psi = \mathbf{0}$ folgt

$$\hat{n}_\eta = \Pi_{\Psi^\perp} \cdot (\Psi \theta + n_\eta) = \Pi_{\Psi^\perp} n_\eta.$$

Die Projektionsmatrix Π_{Ψ^\perp} ist symmetrisch, so dass $\Pi_{\Psi^\perp}^T \Pi_{\Psi^\perp} = \Pi_{\Psi^\perp} \Pi_{\Psi^\perp} = \Pi_{\Psi^\perp}$ gilt. Damit ergibt sich

$$\hat{n}_\eta^T \hat{n}_\eta = n_\eta^T \Pi_{\Psi^\perp}^T \Pi_{\Psi^\perp} n_\eta = n_\eta^T \Pi_{\Psi^\perp} n_\eta.$$

Mit der Spur („tr“ für trace), d. h. die Summe der Diagonalelemente einer Matrix, gilt¹⁷

$$\hat{\mathbf{n}}_{\eta}^T \hat{\mathbf{n}}_{\eta} = \text{tr}(\Pi_{\Psi \perp} \mathbf{n}_{\eta} \mathbf{n}_{\eta}^T),$$

und damit erhält man für den Erwartungswert

$$E\{\hat{\mathbf{n}}_{\eta}^T \hat{\mathbf{n}}_{\eta}\}_r = \text{tr}(\Pi_{\Psi \perp} \cdot E\{\mathbf{n}_{\eta} \mathbf{n}_{\eta}^T\}_r) = \sigma_{\eta}^2 \cdot \text{tr}(\Pi_{\Psi \perp}).$$

Für die Spur von $\Pi_{\Psi \perp}$ gilt

$$\text{tr}(\Pi_{\Psi \perp}) = \text{tr}(\mathbf{I}_k) - \text{tr}(\Psi \cdot (\Psi^T \Psi)^{-1} \cdot \Psi^T)$$

und mit der Umformung¹⁸ $\text{tr}(\Psi \cdot (\Psi^T \Psi)^{-1} \cdot \Psi^T) = \text{tr}((\Psi^T \Psi)^{-1} \cdot \Psi^T \Psi) = \text{tr}(\mathbf{I}_p)$ ergibt sich letztlich

$$E\{\hat{\mathbf{n}}_{\eta}^T \hat{\mathbf{n}}_{\eta}\}_r = \sigma_{\eta}^2 \cdot (k - p).$$

Damit ist

$$\hat{\sigma}_{\eta}^2 = \frac{1}{k - p} \cdot \hat{\mathbf{n}}_{\eta}^T \hat{\mathbf{n}}_{\eta} = \frac{1}{k - p} \cdot \tilde{\eta}^T \cdot (\mathbf{I} - \Psi \cdot (\Psi^T \Psi)^{-1} \cdot \Psi^T) \cdot \tilde{\eta} \quad (8.8)$$

ein biasfreier Schätzer für die Varianz des Messrauschen.

Ein Sonderfall dieser Gleichung wurde schon im Kapitel 4 auf Seite 78 betrachtet. Dort wurde die Varianz des Messrauschen geschätzt, wenn mehrere Messungen $\tilde{x}_{k'}$ einer konstanten Größe x durchgeführt sind. Die Schätzung einer konstanten Größe x entspricht dem LS-Problem mit $\theta = [x]$ und $\Psi_k^T = [1]$. Dabei ist $p = 1$, und man erhält aus Gl. (8.8) den Zusammenhang

$$\hat{\sigma}_{x,k}^2 = \frac{1}{k-1} \cdot \sum_{k'} (\tilde{x}_{k'} - \hat{x}_k)^2.$$

(Die nötigen Umformungen sind nicht unbedingt offensichtlich. Aber der Faktor $\frac{1}{k-1}$ ist erkennbar.)

8.4 Varianz der Schätzwerte

Die wesentliche Eigenschaft der Schätzwerte ist deren Genauigkeit, die durch deren Varianz beschrieben werden kann.

In diesem Abschnitt wird zunächst gezeigt, dass die Varianz der Schätzwerte monoton mit der Anzahl der Messwerte abnimmt. D. h. es kann niemals passieren, dass ein Messwert die Schätzung – im Sinne der Varianz – verschlechtert. Im Anschluss werden allgemein Einflussfaktoren auf die Varianz anhand von Beispielen diskutiert. Zuletzt wird die Konsistenz des Schätzers betrachtet, d. h. es wird diskutiert, unter welchen Voraussetzungen die Varianz für $k \rightarrow \infty$ gegen null geht.

¹⁷ $\mathbf{x}^T \mathbf{A} \mathbf{x} = \sum_m \sum_l x_m a_{ml} x_l = \sum_m \sum_l a_{ml} x_l x_m = \sum_m \sum_l a_{ml} \cdot (\mathbf{x} \mathbf{x}^T)_{lm} = \sum_m (\mathbf{A} \cdot \mathbf{x} \mathbf{x}^T)_{mm} = \text{tr}(\mathbf{A} \cdot \mathbf{x} \mathbf{x}^T)$

¹⁸ Für $\mathbf{A} \in \mathbb{R}^{n \times m}$ und $\mathbf{B} \in \mathbb{R}^{m \times n}$ gilt $\text{tr}(\mathbf{AB}) = \sum_{k=1}^n (\mathbf{AB})_{kk} = \sum_{k=1}^n \sum_{i=1}^m a_{ki} b_{ik} = \sum_{i=1}^m \sum_{k=1}^n b_{ik} a_{ki} = \text{tr}(\mathbf{BA})$.

8.4.1 Monotonie der Varianz der Schätzwerte

Es soll gezeigt werden, dass die Varianz der Schätzwerte mit wachsender Anzahl an Messpunkten monoton abnimmt, d. h. es soll gezeigt werden, dass

$$\text{Var}\{\hat{\theta}_{i,k+1}\}_r \leq \text{Var}\{\hat{\theta}_{i,k}\}_r, \quad i = 1, \dots, p \quad (8.9)$$

gilt, wobei p die Anzahl der zu schätzenden Parameter ist.

Das Gl. (8.9) erfüllt ist, kann man zum Einen mit der Eigenschaft der Effizienz begründen. Dazu betrachtet man zwei Schätzer für $k + 1$ Messungen. Den normalen LS-Schätzer, und einen modifizierten LS-Schätzer, der nur die ersten k Messwerte verwendet. Da der normale LS-Schätzer effizient ist, kann es keinen besseren Schätzer bei $k + 1$ Messungen geben, insbesondere kann der modifizierte LS-Schätzer nicht besser sein, d. h. seine Schätzwerte können keine geringere Varianz haben als die des normalen LS-Schätzers. Demnach kann die Varianz bei dem normalen LS-Schätzer nicht steigen, wenn die Anzahl der Messpunkte erhöht wird.

Alternativ kann man Gl. (8.9) auch wie folgt beweisen, ohne dass ein Rückgriff auf die Eigenschaft der Effizienz erfolgen muss.

Unter Annahme einer konstanten Varianz σ_η^2 des Messrauschen gilt

$$\text{Var}\{\hat{\theta}_{i,k}\}_r = \sigma_\eta^2 \cdot (\mathbf{L}_k^{-1})_{ii}, \quad i = 1, \dots, p,$$

und damit ist

$$(\mathbf{L}_{k+1}^{-1})_{ii} \leq (\mathbf{L}_k^{-1})_{ii}, \quad i = 1, \dots, p,$$

zu zeigen.

Für die Informationsmatrix \mathbf{L}_{k+1} zur Messung $k + 1$ gilt

$$\mathbf{L}_{k+1} = \mathbf{L}_k + \boldsymbol{\psi}_{k+1} \boldsymbol{\psi}_{k+1}^T.$$

Daraus kann die Inverse nach der Sherman-Morrison-Formel ([Zurmühl und Falk 1984, S. 310]) angegeben werden,

$$\mathbf{L}_{k+1}^{-1} = (\mathbf{L}_k + \boldsymbol{\psi}_{k+1} \boldsymbol{\psi}_{k+1}^T)^{-1} = \mathbf{L}_k^{-1} - \frac{\mathbf{L}_k^{-1} \boldsymbol{\psi}_{k+1} \boldsymbol{\psi}_{k+1}^T \mathbf{L}_k^{-1}}{1 + \boldsymbol{\psi}_{k+1}^T \mathbf{L}_k^{-1} \boldsymbol{\psi}_{k+1}}.$$

Diese Gleichung ist auch Bestandteil des rekursiven LS-Verfahrens.

Man erhält die interessierenden Diagonaleinträge durch die Multiplikation von \mathbf{L}_{k+1}^{-1} mit den entsprechenden Einheitsvektoren von links und von rechts,

$$(\mathbf{L}_{k+1}^{-1})_{ii} = \mathbf{e}_i^T \mathbf{L}_{k+1}^{-1} \mathbf{e}_i.$$

Die i -te Spalte von \mathbf{L}_k^{-1} wird mit $\mathbf{p}_{k,i}$ bezeichnet, d. h. $\mathbf{L}_k^{-1} \mathbf{e}_i = \mathbf{p}_{k,i}$. Aufgrund der Symmetrie von \mathbf{L}_k^{-1} gilt damit auch $\mathbf{e}_i^T \mathbf{L}_k^{-1} = \mathbf{p}_{k,i}^T$, d. h. die i -te Zeile ist die Transponierte der i -ten Spalte. Damit ergibt sich für die Diagonalelemente der Ausdruck

$$(\mathbf{L}_{k+1}^{-1})_{ii} = (\mathbf{L}_k^{-1})_{ii} - \frac{\mathbf{p}_{k,i}^T \boldsymbol{\psi}_{k+1} \boldsymbol{\psi}_{k+1}^T \mathbf{p}_{k,i}}{1 + \boldsymbol{\psi}_{k+1}^T \mathbf{L}_k^{-1} \boldsymbol{\psi}_{k+1}}, \quad i = 1, \dots, p.$$

Da \mathbf{L}_k^{-1} positiv definit ist, ist der Nenner des Bruchs immer positiv. Der Zähler $\mathbf{p}_{k,i}^T \boldsymbol{\psi}_{k+1} \boldsymbol{\psi}_{k+1}^T \mathbf{p}_{k,i}$ ist ebenfalls immer größer oder gleich Null, da $\boldsymbol{\psi}_{k+1} \boldsymbol{\psi}_{k+1}^T$ positiv semidefinit ist. Damit ergibt sich der gesuchte Zusammenhang

$$0 < (\mathbf{L}_{k+1}^{-1})_{ii} \leq (\mathbf{L}_k^{-1})_{ii}, \quad i = 1, \dots, p,$$

wobei die zusätzlich angegebene Relation „ $0 <$ “ daher kommt, dass \mathbf{L}_{k+1} und damit auch \mathbf{L}_{k+1}^{-1} immer positiv definit ist und damit nie negative Diagonalelemente auftreten können.

8.4.2 Einflussfaktoren auf die Varianz

Die Kovarianzmatrix $\mathbf{P} = \text{Cov}(\{\hat{\theta}\}_r, \{\hat{\theta}\}_r)$ lautet $\mathbf{P} = \sigma_\eta^2 \cdot \mathbf{L}^{-1}$ (Gl. (8.5)). Die Varianzen $\text{Var}\{\hat{\theta}_j\}_r$ der Schätzwerte sind damit die Einträge auf der Diagonalen von \mathbf{P} , $\text{Var}\{\hat{\theta}_j\}_r = (\mathbf{P})_{jj}$.

Direkt ersichtlich und anschaulich ist, dass die Varianzen der Schätzwerte umso größer sind, je größer die Varianz σ_η^2 des Rauschens ist.

Die Summe der Varianzen entspricht der Spur der Kovarianzmatrix und damit, bis auf einen Faktor, der Spur der Inversen von \mathbf{L}_k ,

$$\sum_j \text{Var}\{\hat{\theta}_{j,k}\}_r = \text{tr}(\mathbf{P}_k) = \sigma_\eta^2 \cdot \text{tr}(\mathbf{L}_k^{-1}).$$

Die Spur einer quadratischen Matrix entspricht der Summe der Eigenwerte λ_j , d. h. es gilt

$$\sum_j \text{Var}\{\hat{\theta}_{j,k}\}_r = \sigma_\eta^2 \cdot \sum_j \lambda_{\mathbf{L}^{-1},j} = \sigma_\eta^2 \cdot \sum_j \frac{1}{\lambda_{\mathbf{L},j}}.$$

Im letzten Schritt wurde verwendet, dass die Eigenwerte der Inversen einer invertierbaren Matrix gerade die Kehrwerte der Eigenwerte der ursprünglichen Matrix sind.¹⁹

Hieraus lässt sich erkennen, dass die Eigenwerte von \mathbf{L} möglichst (gleichmäßig) groß sein sollten. (Der kleinste Eigenwert dominiert die Summe). Um dies genauer zu quantifizieren müssten noch die Größenordnungen der zu schätzenden Parameter berücksichtigt werden. Hier geht es aber um die prinzipiellen Zusammenhänge.

Die Eigenwerte $\lambda_{\mathbf{L},j}$ von $\mathbf{L}_k = \Psi_k^T \Psi_k$ wiederum sind mit den Singulärwerten²⁰ $\sigma_{\Psi,j}$ von Ψ_k verbunden,

$$\sigma_{\Psi,j} = \sqrt{\lambda_{\mathbf{L},j}}.$$

(Die Singulärwerte sind definitionsgemäß der Größe nach zu sortieren, $\sigma_{\Psi,1} \geq \sigma_{\Psi,2} \geq \dots$)

Kein Singulärwert von Ψ_k bzw. kein Eigenwert von \mathbf{L}_k darf null sein. Dies ist aber schon durch den vorausgesetzten vollen Rang von Ψ_k gegeben. (Dies würde einer unendlich hohen Varianz mindestens eines Schätzwertes entsprechen.)

Liegt etwas Erfahrung bezüglich der Singulärwerte einer Matrix als die „Verstärkung“ der durch Ψ_k gegebenen Abbildung vor, so kann man daraus auf Kriterien zum idealen Aufbau der einzelnen Zeilen Ψ_k^T , also der Wahl der Eingangsgrößen, kommen.

Im Folgenden soll der prinzipielle Einfluss der Eingangsgrößen auf die Varianz der Schätzwerte jedoch auf anschaulichem Wege anhand von Beispielen erläutert werden.

Beispiel mit zwei Parametern

Es wird zunächst das Beispiel der Identifikation der Coulomb'schen und viskosen Reibung betrachtet. In Abbildung 8.8a sind dazu zwei Messungen bei $M = 0,15 \text{ Nm}$ und $0,175 \text{ Nm}$ vorgenommen.

Die blaue durchgezogene Linie zeigt den wahren Verlauf. Die roten Kreuze markieren die Messpunkte auf der wahren Gerade. Aufgrund des Messrauschen wird man aber nicht die Geschwindigkeitswerte

¹⁹ Für invertierbare \mathbf{A} gilt: „ λ ist Eigenwert von \mathbf{A} “ $\Leftrightarrow \mathbf{Ax} = \lambda \mathbf{x} \Leftrightarrow \lambda^{-1} \mathbf{x} = \mathbf{A}^{-1} \mathbf{x} \Leftrightarrow \lambda^{-1}$ ist Eigenwert von \mathbf{A}^{-1} “.

²⁰ Wenn der Buchstabe σ für Singulärwerte verwendet ist, dann ist dies immer unmittelbar explizit bemerkt. In allen anderen Fällen steht σ für die Standardabweichung.

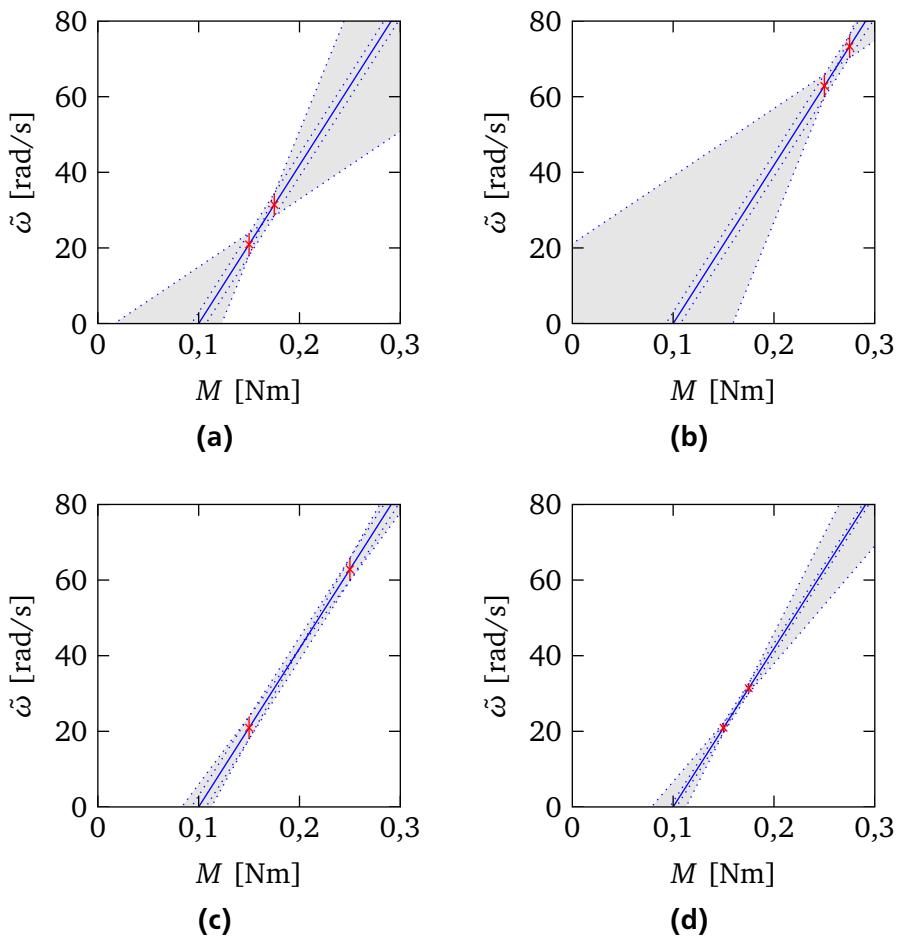


Abbildung 8.8

messen, die diesen Kreuzen entsprechen, sondern eine etwas zu hohe oder zu niedrige Geschwindigkeit. Dies ist durch die kurzen roten senkrechten Striche an den Messpunkten angedeutet. Geht man davon aus, dass die tatsächlichen Messwerte jeweils irgendwo auf den kurzen roten Strichen liegen, dann ergibt sich der in grau eingezeichnete Bereich, indem die Gerade, die sich aus den beiden Messpunkten errechnen lässt, liegen kann.

Misst man bei $M = 0,25 \text{ Nm}$ und $0,275 \text{ Nm}$ (anstelle von $M = 0,15 \text{ Nm}$ und $0,175 \text{ Nm}$), ergibt sich das in Abbildung 8.8b Gezeigte. Die Unsicherheit der Steigung der Geraden ist dieselbe wie in Abbildung 8.8a. (Die Unsicherheit der Geschwindigkeit bei $M = 0$ nimmt deutlich dazu. Hierzu wird später noch etwas gesagt.)

Legt man die Messpunkte weiter auseinander auf $M = 0,15 \text{ Nm}$ und $0,25 \text{ Nm}$, so ergibt sich eine deutlich kleinere Unsicherheit für die identifizierte Gerade, siehe Abbildung 8.8c.

Werden bei $M = 0,15 \text{ Nm}$ und $0,175 \text{ Nm}$ jeweils fünf Messungen durchgeführt, so macht es keinen Unterschied, ob diese als Einzelpunkte dem LS-Verfahren übergeben werden, oder ob man zunächst die Messpunkte für jeden der beiden Werte für M mittelt, und dann das LS-Verfahren auf die beiden Mittelwerte anwendet. Durch die Mittelung beträgt die Varianz je gemittelten Messpunkt bei einem Fünftel der Varianz des Messrauchens. Die Standardabweichung der gemittelten Werte beträgt damit $1/\sqrt{5} = 0,44$ der Standardabweichung des Messrauchens. Entsprechend sind die roten Striche in Abbildung 8.8d nur 0,44 so lang wie die in den Abbildungen zuvor. Damit nimmt die Unsicherheit verglichen mit dem Fall nur einer Messung je Moment (Abbildung 8.8a) deutlich ab.

Liegen Messungen an mehr als zwei Werten für M vor, ist es nicht mehr so anschaulich wie in Abbildung 8.8 darzustellen. Jedoch sollte es einsichtig sein, dass die Varianz der Schätzwerte bei den Messpunkten nach Abbildung 8.9a größer ist als bei den Messpunkten nach Abbildung 8.9b. (Es gilt $\sigma_{\hat{d}}/d' = 2,6\%$ für die Eingangsdaten nach Abbildung 8.9a und $\sigma_{\hat{d}}/d' = 1,8\%$ für Abbildung 8.9b.)

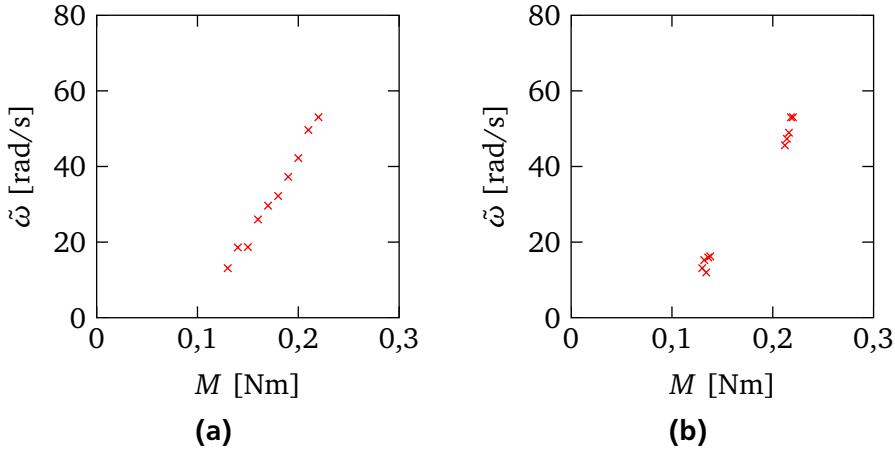


Abbildung 8.9

Aus diesen Beispielen lässt sich ableiten, dass die Eingangsgröße ein möglichst breites Gebiet „aufspannen“ sollte. Aus Sicht der Varianz sollten die Messpunkte dabei möglichst an den Rändern dieses Gebietes liegen. Zumindest bezüglich der Steigung sind die Absolutwerte der Eingangsgröße nicht relevant, es kommt nur auf deren Abstand ab.

Wenn der zu identifizierende Zusammenhang keine Konstante beinhaltet, so gilt dies nicht mehr in dieser Form. Wäre bekannt, dass die Coulomb'sche Reibung vernachlässigbar ist, und soll demnach nur die viskose Reibung nach dem Zusammenhang

$$\omega_{k'} = d' \cdot M_{k'}$$

identifiziert werden, so ergeben sich die in Abbildung 8.10 dargestellten Zusammenhänge. In diesem Fall ist es günstig, möglichst große Werte für M zu wählen. Führt man mehrere Messungen durch, so sollten diese alle bei möglichst hohen Werten für M vorliegen. Die Streuung der Eingangswerte untereinander spielt in diesem Fall keine Rolle mehr. (Man kann es sich auch so vorstellen, dass bei $M = 0$ quasi unendlich viele Messungen vorliegen, die alle ein $\tilde{\omega} = 0$ geben.)

Beispiel mit drei Parametern

Für zwei Eingangsgrößen ψ_2 und ψ_3 sind in Tabelle 8.2 bzw. Abbildung 8.11 vier Beispiele für mögliche Eingangsdaten gegeben. (Die erste Eingangsgröße gehört zu einer Konstanten, $\psi_1 \equiv 1$.) Diese unterscheiden sich neben der Anzahl der Messpunkte vor Allem im Grad der (empirischen) Korrelation zwischen den beiden Eingangsgrößen ψ_2 und ψ_3 .

Ideale Korrelation

Die Daten in Abbildung 8.11a sind so konstruiert, dass ψ_3 um drei zunimmt, wenn ψ_2 um vier zunimmt. Die beiden Folgen sind daher ideal korreliert. Damit kann z. B. $\psi_{3,k'}$ über

$$\psi_{3,k'} = 12 + \frac{3}{4} \cdot (\psi_{2,k'} - 35)$$

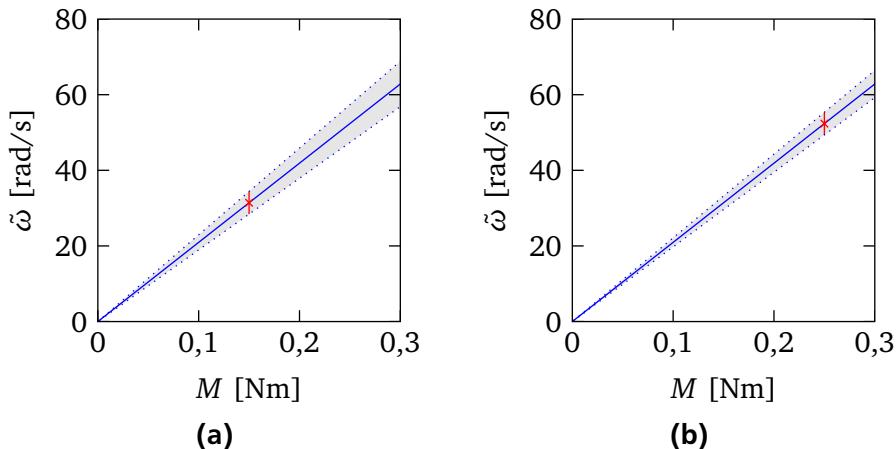


Abbildung 8.10: Zur Unsicherheit der Streuung

Tabelle 8.2: Eingangsdaten zu Abbildung 8.11

ψ_2	31	35	39
ψ_3	9	12	15

(a) Ideale Korrelation
(Abbildung 8.11a)

ψ_2	33	31	35	33	34	36	37
ψ_3	9	9	12	12	12	15	15

(b) Starke Korrelation (Abbildung 8.11b)

ψ_2	37	31	37
ψ_3	9	9	15

(c) Schwache Korrelation
(Abbildung 8.11c)

ψ_2	24	31	37
ψ_3	9	9	15

(d) Mittlere Korrelation
(Abbildung 8.11d)

durch $\psi_{2,k'}$ ausgedrückt werden, so dass neben der Konstanten Eins nur eine Eingangsgröße vorliegt. Damit können aber keine drei Parameter geschätzt werden und es ist anschaulich klar, dass in diesem Fall die Matrix $\Psi^T \Psi$ singulär ist, also nicht invertiert werden kann. (Bei idealer Korrelation könnten auch beliebig viele Messpunkte vorliegen, $\Psi^T \Psi$ bleibt immer singulär.)

Tabelle 8.3: Varianzen zu den Eingangsdaten aus Abbildung 8.11

Eingangsdaten	$\sigma_{\theta_2}^2 / \sigma_\eta^2$	$\sigma_{\theta_3}^2 / \sigma_\eta^2$
8.11a		$\Psi^T \Psi$ singulär
8.11b	0,217	0,150
8.11c	0,056	0,056
8.11d	0,041	0,144

Starke Korrelation

In Abbildung 8.11b sind sieben Datenpunkte dargestellt, die noch relativ gut auf einer Geraden in der ψ_2 - ψ_3 -Ebene liegen. Die sich damit ergebenden Varianzen sind in Tabelle 8.3 aufgeführt.

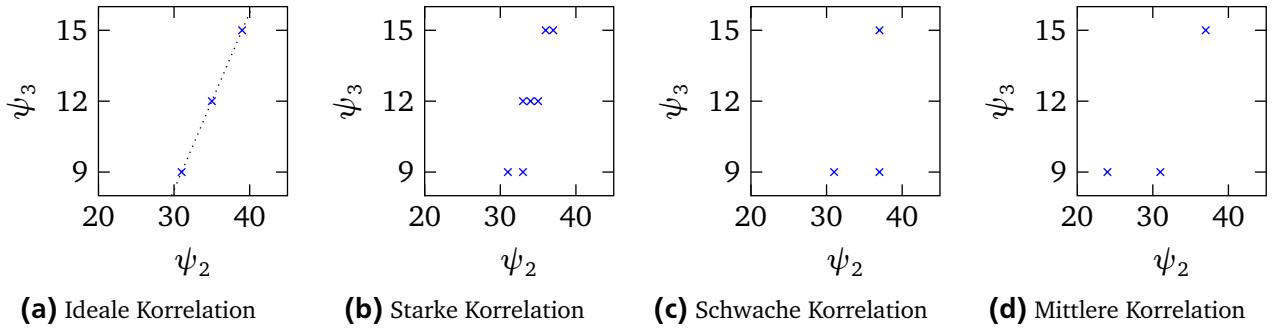


Abbildung 8.11: Beispiele für Eingangsdaten bei zwei Eingangsgrößen ψ_2 und ψ_3

Schwache Korrelation

In Abbildung 8.11c sind nur drei Datenpunkte vorhanden. Diese sind aber so gelegt, dass diese quasi das einhüllende Rechteck der Punkte aus Abbildung 8.11b bilden. Vergleicht man die sich ergebenden Varianzen in Tabelle 8.3 mit denen, die sich bei der starken Korrelation ergeben, so zeigt sich, dass die Varianzen im Fall der schwachen Korrelation deutlich geringer sind, obwohl nur $k = 3$ anstatt $k = 7$ Datenpunkte vorliegen.

Mittlere Korrelation

In Abbildung 8.11d sind ebenfalls nur drei Datenpunkte vorhanden. Im Vergleich zu Abbildung 8.11c ist der rechte Punkt bei $\psi_3 = 9$ nach links gerückt. Die Korrelation zwischen ψ_2 und ψ_3 nimmt damit wieder zu, anschaulich lassen sich diese Punkte wieder besser durch eine Gerade verbinden. Tabelle 8.3 zeigt zwar, dass die Schätzung von θ_2 besser wird, was hier auf die Verdoppelung des Wertebereichs von ψ_2 zurückzuführen ist. Allerdings wird θ_3 deutlich schlechter geschätzt.

Fazit

Bezüglich der Streuung der Werte einer *einzelnen* Eingangsgröße ist zu unterscheiden, ob eine Konstante mitgeschätzt wird oder nicht. Bei der Schätzung statischer Zusammenhänge ist das Vorhandensein einer Konstante meist der Fall. Dann kommt es im Wesentlichen auf die Breite der Streuung der Eingangswerte an. Der Absolutwert ist nicht wesentlich.

Wenn der Prozess keine Konstante beinhaltet, was später bei der Identifikation dynamischer Zusammenhänge der Fall ist, dann sind die Absolutwerte relevant. Diese sollten möglichst hoch sein.

In beiden Fällen müssen die Eingangswerte der *verschiedenen* Eingangsgrößen möglichst gut streuen, d. h. es sollte eine möglichst schlechte (empirische) Korrelation dieser Werte auftreten.

Anmerkungen

Identifikation bezüglich des Nennpunkts

Wenn Zusammenhänge mit einem konstanten Anteil identifiziert werden sollen ($\psi_{1,k'} \equiv 1$), ist eine direkte Interpretation der Varianz des Parameters, der dem konstanten Anteil entspricht nur dann sinnvoll möglich, wenn die Identifikation bezüglich des Nennpunktes durchgeführt wird.

Dies meint, dass nicht von

$$\eta_{k'} = \theta_1 + \theta_2 \cdot \psi_{2,k'} + \theta_3 \cdot \psi_{3,k'} + \dots + \theta_p \cdot \psi_{p,k'}$$

ausgegangen werden sollte, sondern von

$$\eta_{k'} = \theta'_1 + \theta_2 \cdot (\psi_{2,k'} - \psi_{1,N}) + \theta_3 \cdot (\psi_{3,k'} - \psi_{3,N}) + \cdots + \theta_p \cdot (\psi_{p,k'} - \psi_{p,N}),$$

wobei $\psi_{j,N}$ der Nennpunkt von der j -ten Eingangsgröße ψ_j , d. h. die Größe, um diese die j -te Eingangsgröße in der Regel Werte annimmt, ist. Dabei ändert sich nur der Wert des Parameters θ_1 zu θ'_1 , alle anderen Parameter bleiben unverändert.

Dazu ist in Abbildung 8.12 für den Fall $p = 2$ (inklusive der Konstanten) die Schätzung der Geraden mit zwei Messungen bei $\psi_2 = 40$ und 60 dargestellt. Die durchgezogene Linie entspricht dem tatsächlichen Zusammenhang. Durch das Messrauschen entstehen bei den beiden Eingangswerten von 40 und 60 Unsicherheiten, die durch die kleinen senkrechten Striche dargestellt sind. Der Bereich, in dem die damit geschätzte Gerade liegt, ist durch den grauen Bereich angedeutet. Es ist zu erkennen, dass dieser Bereich bei $\psi_2 = 50$ deutlich kleiner ist als bei $\psi_2 = 0$. D. h. die Varianz des Schätzwertes $\{\hat{\theta}_1\}_r$ wäre deutlich größer als die des Schätzwertes $\{\hat{\theta}'_1\}_r$. Da man sich in der Anwendung immer um den Nennpunkt aufhalten wird, ist die Varianz von $\{\hat{\theta}'_1\}_r$ aussagekräftiger. (Vgl. dazu auch die Bemerkung von [Freund, Wilson und Sa 2006, S. 42, 50] bezüglich der Interpretierbarkeit von $\hat{\theta}'_1$.)

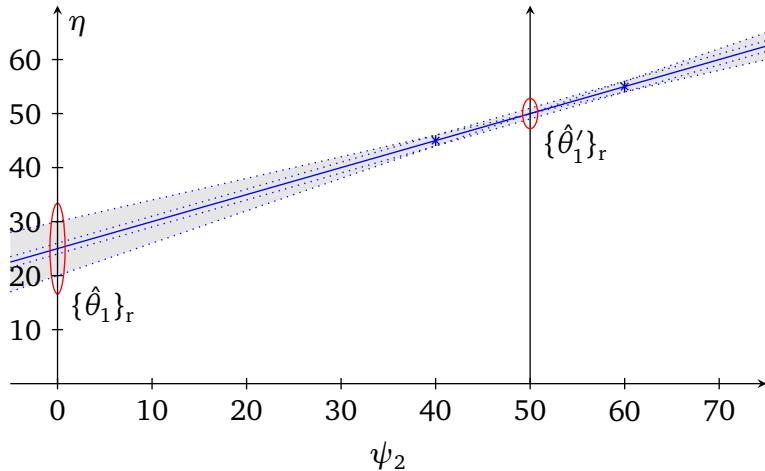


Abbildung 8.12: Vergleich von $\hat{\theta}_1$ und $\hat{\theta}'_1$

Die geschätzte Gerade ist aber in beiden Fällen dieselbe, und auch die Varianz, die sich mit den geschätzten Parametern für den gleichen Wert von ψ_2 ergibt ist in beiden Fällen identisch.

Letztlich ist dies nur eine andere Definition der Eingänge ψ_j , so dass dies im Weiteren nicht explizit ausgeschrieben ist.

Skalierung der Eingangsgrößen

Um Einträge gleicher Größenordnung in der L-Matrix und damit günstigere Bedingungen für die numerische Berechnung zu erhalten, sollten die Eingangsgrößen normiert werden. Dazu wird die Gleichung

$$\eta_{k'} = \theta_1 \cdot \psi_{1,k'} + \theta_2 \cdot \psi_{2,k'} + \cdots + \theta_p \cdot \psi_{p,k'}$$

in

$$\eta_{k'} = \psi_{1,k'} \cdot s_1 \cdot \frac{\theta_1}{s_1} + \psi_{2,k'} \cdot s_2 \cdot \frac{\theta_2}{s_2} + \cdots + \psi_{p,k'} \cdot s_p \cdot \frac{\theta_p}{s_p}$$

mit geeigneten Skalierungsfaktoren s_k umgeschrieben. Nun werden die Größen $\frac{\theta_j}{s_j}$ als neue Parameter θ'_j aufgefasst. Damit wird der Parametervektor

$$\boldsymbol{\theta}' = [\theta'_1 \quad \cdots \quad \theta'_p]^T$$

gesucht, wobei der Eingangsvektor

$$\psi_{k'}^T = [\psi_{1,k'} \cdot s_1 \quad \psi_{2,k'} \cdot s_2 \quad \cdots \quad \psi_{p,k'} \cdot s_p]$$

lautet.

8.4.3 Konsistenz

Im vorherigen Abschnitt wurde erläutert, wie die Eingangsgrößen gewählt werden müssen, um alle Parameter identifizieren zu können, und dabei möglichst geringe Varianzen der Schätzwerte zu erhalten. Jetzt soll der Frage nachgegangen werden, unter welchen Umständen die Varianzen für $k \rightarrow \infty$ gegen Null gehen, also die Schätzung konsistent ist.

Dazu wird wieder die Konfiguration der Eingangsgrößen wie in Abbildung 8.8a betrachtet. D.h. es werden nur Messungen bei zwei Werten der Eingangsgröße durchgeführt. In dem Beispiel ist dies bei $M_a = 0,15 \text{ Nm}$ und $M_b = 0,175 \text{ Nm}$.

Erhöht man nun die Anzahl der Messungen, indem man immer weitere Messungen bei M_b durchführt, d.h. man führt eine Messung bei $M = M_a$ und $k - 1$ Messungen bei $M = M_b$ durch, so ergibt sich

$$\mathbf{L}_k = \begin{bmatrix} 1 & M_a \\ M_a & M_a^2 \end{bmatrix} + (k-1) \cdot \begin{bmatrix} 1 & M_b \\ M_b & M_b^2 \end{bmatrix} = \begin{bmatrix} k & M_a + (k-1) \cdot M_b \\ M_a + (k-1) \cdot M_b & M_a^2 + (k-1) \cdot M_b^2 \end{bmatrix}.$$

Die Determinante dieser Matrix lautet

$$\det(\mathbf{L}_k) = k \cdot M_a^2 + k \cdot (k-1) \cdot M_b^2 - (M_a + (k-1) \cdot M_b)^2$$

und nach ein paar Zwischenschritten erhält man

$$\det(\mathbf{L}_k) = (k-1) \cdot (M_1 - M_2)^2.$$

Damit ergibt sich für die Inverse

$$\mathbf{L}_k^{-1} = \frac{1}{(k-1) \cdot (M_a - M_b)^2} \cdot \begin{bmatrix} M_a^2 + (k-1) \cdot M_b^2 & -M_a - (k-1) \cdot M_b \\ -M_a - (k-1) \cdot M_b & k \end{bmatrix}.$$

Bildet man den Grenzübergang für $k \rightarrow \infty$, so erhält man

$$\lim_{k \rightarrow \infty} \mathbf{L}_k^{-1} = \begin{bmatrix} \frac{M_b^2}{(M_a - M_b)^2} & -\frac{M_b}{(M_a - M_b)^2} \\ -\frac{M_a}{(M_a - M_b)^2} & \frac{1}{(M_a - M_b)^2} \end{bmatrix}.$$

Die Einträge von \mathbf{L}_k^{-1} und damit auch die Varianzen in $\mathbf{P}_k = \sigma_\eta^2 \mathbf{L}_k^{-1}$ gehen für $k \rightarrow \infty$ demnach nicht gegen Null, der Schätzer ist mit diesen Eingangsgrößen also nicht konsistent.

Anschaulich kann man sich das auch damit erklären, dass die Gerade zwar an der Abszisse M_b für $k \rightarrow \infty$ immer fester fixiert wird, aber deren Steigung durch die Unsicherheit der einzigen Messung bei M_a nie genau bekannt sein wird. Dies ist in Abbildung 8.13a dargestellt. Der „sichere“ Punkte bei M_b ist durch den roten Kreis markiert.

Für den Schätzwert, der sich mit den geschätzten Parametern zum Punkt M_b ergibt,

$$\hat{\eta}_b = \hat{\theta}_1 + \hat{\theta}_2 \cdot M_b,$$

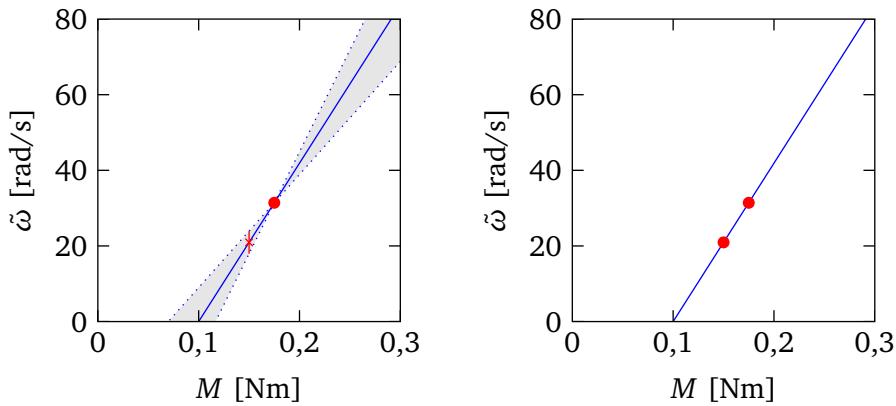


Abbildung 8.13: Verhalten der Unsicherheiten für $k \rightarrow \infty$

ergibt sich die Varianz von

$$\begin{aligned}\text{Var}\{\hat{\eta}_b\}_r &= \text{Var}\{\hat{\theta}_1 + \hat{\theta}_2 \cdot M_b\}_r \\ &= \text{Var}\{\hat{\theta}_1\}_r + M_b^2 \cdot \text{Var}\{\hat{\theta}_2\}_r + 2 \cdot M_b \cdot \text{Cov}(\{\hat{\theta}_1\}_r, \{\hat{\theta}_2\}_r) \\ &= 0.\end{aligned}$$

An diesem einen Abszissenwert ist der Ordinatenwert nach $k \rightarrow \infty$ Messungen damit erwartungsgemäß exakt bekannt. Für $M = M_a$ ergibt sich dagegen

$$\begin{aligned}\text{Var}\{\hat{\eta}_a\}_r &= \text{Var}\{\hat{\theta}_1 + \hat{\theta}_2 \cdot M_a\}_r \\ &= \text{Var}\{\hat{\theta}_1\}_r + M_a^2 \cdot \text{Var}\{\hat{\theta}_2\}_r + 2 \cdot M_a \cdot \text{Cov}(\{\hat{\theta}_1\}_r, \{\hat{\theta}_2\}_r) \\ &= \sigma_\eta^2.\end{aligned}$$

Hier entspricht die Varianz gerade der einer einzigen Messung.

Man kann auch sagen, dass in der „Eingangsrichtung“ $\Psi^T = [1 \ M_b]$ unendlich viele Informationen vorliegen, und damit die Varianz für diese Richtung gegen Null geht. Um die Varianzen für alle möglichen Eingangsgrößen gegen Null gehen zu lassen, müssen aber in zwei Richtungen unendlich viele Informationen vorliegen.

Führt man die k Messungen dagegen so durch, dass immer (abwechselnd) bei $M = M_a$ und $M = M_b$ gemessen wird, so ergibt sich

$$\mathbf{L}_k = \frac{k}{2} \cdot \begin{bmatrix} 1 & M_a \\ M_a & M_a^2 \end{bmatrix} + \frac{k}{2} \cdot \begin{bmatrix} 1 & M_b \\ M_b & M_b^2 \end{bmatrix} = \frac{k}{2} \cdot \begin{bmatrix} 2 & M_a + M_b \\ M_a + M_b & M_a^2 + M_b^2 \end{bmatrix}.$$

Dessen Determinante lautet

$$\det(\mathbf{L}_k) = \frac{k^2}{4} \cdot 2 \cdot (M_a^2 + M_b^2) - (-M_a - M_b)^2 = (M_a - M_b)^2$$

und damit kann die Inverse

$$\mathbf{L}_k^{-1} = \frac{2}{k} \cdot \frac{1}{(M_a - M_b)^2} \cdot \begin{bmatrix} M_a^2 + M_b^2 & -M_a - M_b \\ -M_a - M_b & 2 \end{bmatrix}$$

angegeben werden.

Bildet man in diesem Fall den Grenzwert für $k \rightarrow \infty$, so erhält man

$$\lim_{k \rightarrow \infty} \mathbf{L}_k^{-1} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Die Varianzen gehen gegen Null, und der Schätzer ist mit diesen Eingangsgrößen konsistent!

Die Forderung

$$\lim_{k \rightarrow \infty} \mathbf{L}_k^{-1} \stackrel{!}{=} \mathbf{0}$$

ist auch in der (etwas stärkeren) Forderung

$$\lim_{k \rightarrow \infty} k \cdot \mathbf{L}_k^{-1} \stackrel{!}{=} \mathbf{X} \quad \text{mit} \quad \det(\mathbf{X}) \neq 0$$

enthalten. Dies wiederum bedeutet, dass

$$\lim_{k \rightarrow \infty} \frac{1}{k} \cdot \mathbf{L}_k \stackrel{!}{=} \mathbf{X}^{-1}$$

gelten muss.

D. h. hinreichend für die Konsistenz des LS-Schätzers ist (neben den Bedingungen A1 bis A4), dass die Matrix $\frac{1}{k} \cdot \mathbf{L}_k$ für $k \rightarrow \infty$ gegen eine invertierbare Matrix konvergiert,

$$\det\left(\lim_{k \rightarrow \infty} \frac{1}{k} \cdot \mathbf{L}_k\right) \stackrel{!}{\neq} 0.$$

Für das erste Beispiel ergibt sich mit

$$\lim_{k \rightarrow \infty} \frac{1}{k} \cdot \mathbf{L}_k = \begin{bmatrix} 1 & M_b \\ M_b & M_b^2 \end{bmatrix}$$

eine nicht-invertierbare Matrix. Dagegen erhält man für das zweite Beispiel die invertierbare Matrix

$$\lim_{k \rightarrow \infty} \frac{1}{k} \cdot \mathbf{L}_k = \frac{1}{2} \cdot \begin{bmatrix} 2 & M_a + M_b \\ M_a + M_b & M_a^2 + M_b^2 \end{bmatrix}.$$

Am einfachsten ist dies sichergestellt, wenn eine Folge von Messwerten, die zu einer invertierbaren Matrix \mathbf{L}' führen, ständig wiederholt wird. Führt man K Wiederholungen einer solchen Folge von Messwerten durch, so gilt

$$\mathbf{L}^{-1} = (K \cdot \mathbf{L}')^{-1} = \frac{1}{K} \cdot \mathbf{L}'^{-1},$$

und für $K \rightarrow \infty$ gehen damit alle Einträge von \mathbf{L}^{-1} gegen Null.

8.5 Erweiterungen

8.5.1 Verfahren der gewichteten kleinsten Quadrate (Weighted Least Squares)

Bei dem normalen Least-Squares-Verfahren werden die k Schätzfehler \hat{e}_k im Gütemaß gleich gewichtet. Es können aber auch Gewichte $w_k > 0$ (bzw. $w_k \geq 0$, siehe unten) eingeführt werden, mit denen eine unterschiedliche Gewichtung der einzelnen Schätzfehler erhalten wird. So lautet V in diesem Fall

$$V = w_1 \cdot \hat{e}_1^2 + w_2 \cdot \hat{e}_2^2 + \cdots + w_k \cdot \hat{e}_k^2 = \hat{\mathbf{e}}^T \mathbf{W} \hat{\mathbf{e}},$$

mit der diagonalen Gewichtungsmatrix

$$\mathbf{W} = \text{diag}([w_1 \ w_2 \ \dots \ w_k]).$$

Aus diesem Gütemaß ergibt sich bei gleichem Vorgehen wie oben die Schätzgleichung

$$\hat{\boldsymbol{\theta}} = (\Psi^T \mathbf{W} \Psi)^{-1} \Psi^T \mathbf{W} \tilde{\boldsymbol{\eta}}.$$

Dabei können allgemein auch nicht-diagonale Gewichtungsmatrizen \mathbf{W} zugelassen werden, solange $\Psi^T \mathbf{W} \Psi$ invertierbar ist.²¹

Setzt man $\tilde{\boldsymbol{\eta}} = \boldsymbol{\eta} + \mathbf{n}_{\boldsymbol{\eta}}$ ein, und verwendet man $\boldsymbol{\eta} = \Psi \boldsymbol{\theta}$, so erhält man

$$\begin{aligned}\hat{\boldsymbol{\theta}} &= (\Psi^T \mathbf{W} \Psi)^{-1} \Psi^T \mathbf{W} \boldsymbol{\eta} + (\Psi^T \mathbf{W} \Psi)^{-1} \Psi^T \mathbf{W} \mathbf{n}_{\boldsymbol{\eta}} \\ &= \boldsymbol{\theta} + (\Psi^T \mathbf{W} \Psi)^{-1} \Psi^T \mathbf{W} \mathbf{n}_{\boldsymbol{\eta}}.\end{aligned}$$

Daraus folgt, wie beim normalen LS-Verfahren, dass (mit $E\{\mathbf{n}_{\boldsymbol{\eta}}|\Psi\}_r = \mathbf{0}$ unkorreliert)

$$E\{\hat{\boldsymbol{\theta}}\}_r = \boldsymbol{\theta}$$

gilt. Auch das gewichtete LS-Verfahren liefert unter den klassischen Annahmen also biasfreie Schätzwerte.

Schreibt man analog zum normalem LS-Verfahren die Kovarianzmatrix auf, erhält man

$$\mathbf{P}_W = (\Psi^T \mathbf{W} \Psi)^{-1} \cdot \Psi^T \mathbf{W} \cdot E\{\mathbf{n}_{\boldsymbol{\eta}} \mathbf{n}_{\boldsymbol{\eta}}^T\}_r \cdot \mathbf{W}^T \Psi \cdot (\Psi^T \mathbf{W} \Psi)^{-1}.$$

Ist die Kovarianzmatrix $E\{\mathbf{n}_{\boldsymbol{\eta}} \mathbf{n}_{\boldsymbol{\eta}}^T\}_r$ des Messrauschens bekannt, so kann

$$\mathbf{W} = \left(E\{\mathbf{n}_{\boldsymbol{\eta}} \mathbf{n}_{\boldsymbol{\eta}}^T\}_r \right)^{-1}.$$

gewählt werden, um die Varianz der Schätzwerte $\{\hat{\boldsymbol{\theta}}\}_r$ zu minimieren.

Um diese Behauptung zu beweisen, wird die Kovarianzmatrix des Messrauschens mit

$$\mathbf{R} = E\{\mathbf{n}_{\boldsymbol{\eta}} \mathbf{n}_{\boldsymbol{\eta}}^T\}_r$$

²¹ \mathbf{W} muss dazu nicht positiv definit sein, sondern die positive Semi-Definitheit ist ausreichend. Die Bedingung, dass das Produkt $\Psi^T \mathbf{W} \Psi$ invertierbar sein muss, bedeutet anschaulich, dass z. B. bei drei zu schätzenden Parametern auch mindestens drei Messungen ein Gewicht größer als Null erhalten müssen. (Und die Eingangsvektoren zu diesen drei Messungen dürfen nicht linear abhängig sein.)

abgekürzt. Mit dem Gewicht $\mathbf{W} = \mathbf{R}^{-1}$ ergibt sich

$$\mathbf{P}_{\mathbf{R}^{-1}} = (\Psi^T \mathbf{R}^{-1} \Psi)^{-1}$$

für die Kovarianzmatrix der Schätzwerte.

Ein Ansatz wäre jetzt zu beweisen, dass die Differenz zwischen der Kovarianzmatrix \mathbf{P}_W , die sich bei einer beliebigen Wahl von \mathbf{W} ergibt (die Regularität $\Psi^T \mathbf{W} \Psi$ vorausgesetzt), und $\mathbf{P}_{\mathbf{R}^{-1}}$ positiv semidefinit ist. Daraus kann dann wieder wie beim Beweis der BLUE-Eigenschaft bei in sich unkorreliertem Messrauschen konstanter Varianz darauf geschlossen werden, dass die Varianzen der Schätzwerte, die sich mit $\mathbf{W} = \mathbf{R}^{-1}$ ergeben, maximal so groß sind wie bei allen anderen Gewichtsmatrizen \mathbf{W} .

Jedoch ist es sogar etwas übersichtlicher (der Ansatz und das Vorgehen ist im Prinzip identisch), direkt die BLUE-Eigenschaft bezüglich des Rauschens $\mathbf{R} = E\{\mathbf{n}_\eta \mathbf{n}_\eta^T\}_r$ zu zeigen.

Für die Kovarianzmatrix eines allgemeinen linearen Schätzers gilt

$$\mathbf{P}_M = \mathbf{M} \cdot E\{\mathbf{n}_\eta \mathbf{n}_\eta^T\}_r \cdot \mathbf{M}^T = \mathbf{M} \mathbf{R} \mathbf{M}^T.$$

Die Differenz zwischen einem allgemeinen linearen Schätzers und dem mit $\mathbf{W} = \mathbf{R}^{-1}$ gewichteten LS-Verfahren ist damit

$$\mathbf{P}_M - \mathbf{P}_{\mathbf{R}^{-1}} = \mathbf{M} \mathbf{R} \mathbf{M}^T - (\Psi^T \mathbf{R}^{-1} \Psi)^{-1}.$$

Für biasfreie lineare Schätzer gilt immer $\mathbf{M} \Psi = \mathbf{I}$, und damit kann auch

$$\mathbf{P}_M - \mathbf{P}_{\mathbf{R}^{-1}} = \mathbf{M} \mathbf{R} \mathbf{M}^T - \mathbf{M} \Psi \cdot (\Psi^T \mathbf{R}^{-1} \Psi)^{-1} \cdot \Psi^T \mathbf{M}^T.$$

geschrieben werden. Es kann davon ausgegangen werden, dass die Kovarianzmatrix des Messrauschen positiv definit ist.²² Damit existiert die Cholesky-Zerlegung $\mathbf{R} = \mathbf{S}^T \mathbf{S}$ mit der invertierbaren Matrix \mathbf{S} , und man erhält

$$\begin{aligned} \mathbf{P}_M - \mathbf{P}_{\mathbf{R}^{-1}} &= \mathbf{M} \mathbf{S}^T \mathbf{S} \mathbf{M}^T - \mathbf{M} \Psi \cdot (\Psi^T \mathbf{S}^{-1} \mathbf{S}^T \Psi)^{-1} \cdot \Psi^T \mathbf{M}^T \\ &= \mathbf{M} \mathbf{S}^T \cdot \underbrace{(\mathbf{I} - \mathbf{S}^{-T} \Psi \cdot (\Psi^T \mathbf{S}^{-1} \mathbf{S}^T \Psi)^{-1} \cdot \Psi^T \mathbf{S}^{-1})}_{\Pi} \cdot \mathbf{S} \mathbf{M}^T. \end{aligned}$$

Die Matrix Π stellt wieder eine orthogonale Projektion dar, für die $\Pi = \Pi \Pi = \Pi \Pi^T$ gilt. Dies wird mit der Abkürzung $\mathbf{T} = \mathbf{S}^{-T} \Psi$ (und damit auch $\mathbf{T}^T = \Psi^T \mathbf{S}^{-1}$) etwas deutlicher:

$$\Pi = \mathbf{I} - \mathbf{T} \cdot (\mathbf{T}^T \mathbf{T})^{-1} \cdot \mathbf{T}^T.$$

Π und damit auch $\mathbf{P}_M - \mathbf{P}_{\mathbf{R}^{-1}}$ sind demnach positiv-semidefinit.

Mit der Wahl

$$\mathbf{W} = \mathbf{R}^{-1} = \left(E\{\mathbf{n}_\eta \mathbf{n}_\eta^T\}_r \right)^{-1}$$

ist das gewichtete LS-Verfahren also der beste lineare, biasfreie Schätzer für den Zusammenhang $\eta = \Psi^T \theta$.

Ist das Messrauschen unkorreliert, besitzt aber für jede Messung eine andere Varianz $\sigma_{\eta,k}^2$, so ist

$$\mathbf{R} = \text{diag}([\sigma_{\eta,1}^2 \ \sigma_{\eta,2}^2 \ \dots \ \sigma_{\eta,k}^2])$$

und das optimale Gewicht demnach

$$\mathbf{W} = \text{diag}\left[\frac{1}{\sigma_{\eta,1}^2} \ \frac{1}{\sigma_{\eta,2}^2} \ \dots \ \frac{1}{\sigma_{\eta,k}^2}\right].$$

Die Messungen sollten also umgekehrt zu ihrer Unsicherheit gewichtet werden.

²² Sie muss schon von der Definition her mindestens positiv semidefinit sein. Wäre sie aber nicht auch positiv definit, dann könnte man das Gewicht $\mathbf{W} = \mathbf{R}^{-1}$ nicht angeben.

8.5.2 Exponentielles Vergessen

Ein Sonderfall des gewichteten LS-Verfahrens ist das LS-Verfahren mit exponentiellem Vergessen. Bei diesem hat die Gewichtungsmatrix die Form

$$\mathbf{W}_k = \text{diag}([\lambda^k, \lambda^{k-1}, \dots, \lambda, 1]) ,$$

wobei $0 < \lambda < 1$ ist. Damit werden die Schätzfehler der weiter zurückliegenden Messungen geringer gewichtet als die der neueren Messungen.

In Abbildung 8.14 sind die Gewichtungsfaktoren für verschiedene λ für die letzten 50 zurückliegenden Datenpunkte dargestellt.

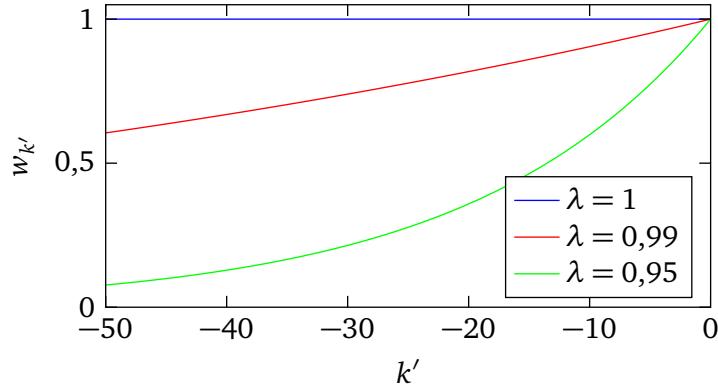


Abbildung 8.14: Gewichtsfaktoren bei exponentiellem Vergessen (aktueller Messpunkt bei $k' = 0$)

Der Vergessensfaktor kann im Allgemeinen auch variabel sein. Dies stellt sich dann am einfachsten über die rekursive Regel

$$\mathbf{W}_{k+1} = \begin{bmatrix} \lambda_{k+1} \mathbf{W}_k & \\ & 1 \end{bmatrix} \quad (8.10)$$

zur Bestimmung von \mathbf{W}_{k+1} dar.

8.6 Iterative und rekursive Berechnung

8.6.1 Iterative Berechnung

Wenn die Schätzung online erfolgen soll, dann kommt jeden Abtastschritt ein neuer Datensatz hinzu. Es müssen aber nicht in jedem Abtastschritt Produkte immer größer werdender Matrizen berechnet werden, sondern man kann die Berechnung iterativ ausführen.

Für die Matrix $\mathbf{L}_{k+1} = \Psi_{k+1}^T \Psi_{k+1}$ gilt

$$\Psi_{k+1}^T \Psi_{k+1} = [\Psi_k^T \quad \Psi_{k+1}] \begin{bmatrix} \Psi_k \\ \Psi_{k+1}^T \end{bmatrix} = \Psi_k^T \Psi_k + \Psi_{k+1} \Psi_{k+1}^T$$

und damit

$$\mathbf{L}_{k+1} = \mathbf{L}_k + \Psi_{k+1} \Psi_{k+1}^T .$$

Analog ergibt sich

$$\mathbf{r}_{k+1} = \Psi_{k+1}^T \tilde{\eta}_{k+1} = [\Psi_k^T \quad \Psi_{k+1}] \begin{bmatrix} \tilde{\eta}_k \\ \tilde{\eta}_{k+1} \end{bmatrix} = \Psi_k^T \tilde{\eta}_k + \Psi_{k+1} \tilde{\eta}_{k+1} = \mathbf{r}_k + \Psi_{k+1} \tilde{\eta}_{k+1}.$$

Zusammengefasst kann das LS-Verfahren damit in der Form

$$\hat{\theta}_k = \mathbf{L}_k^{-1} \mathbf{r}_k \quad (8.11)$$

mit

$$\mathbf{L}_k = \begin{cases} \mathbf{0} & k = 0 \\ \mathbf{L}_{k-1} + \Psi_k \Psi_k^T & k > 0 \end{cases} \quad (8.12)$$

und

$$\mathbf{r}_k = \begin{cases} \mathbf{0} & k = 0 \\ \mathbf{r}_{k-1} + \tilde{\eta}_k \Psi_k & k > 0 \end{cases} \quad (8.13)$$

geschrieben werden.

Soll neben den Parameterwerten auch nach Gl. (8.8),

$$\hat{\sigma}_{\eta,k}^2 = \frac{1}{k-p} \cdot \tilde{\eta}_k^T \cdot (\mathbf{I} - \Psi_k \cdot (\Psi_k^T \Psi_k)^{-1} \cdot \Psi_k^T) \cdot \tilde{\eta}_k,$$

die Varianz des Messrauschens bestimmt werden, so kann dies ebenfalls in iterativer Form geschrieben werden.

Durch Ausmultiplizieren erhält man

$$\tilde{\eta}_k^T \cdot (\mathbf{I} - \Psi_k \cdot (\Psi_k^T \Psi_k)^{-1} \cdot \Psi_k^T) \cdot \tilde{\eta}_k = \tilde{\eta}_k^T \tilde{\eta}_k - \underbrace{\tilde{\eta}_k^T \Psi_k}_{\mathbf{r}_k^T} \cdot \underbrace{(\Psi_k^T \Psi_k)^{-1}}_{\mathbf{L}_k^{-1}} \cdot \underbrace{\Psi_k^T \tilde{\eta}_k}_{\mathbf{r}_k}.$$

Die Faktoren des rechten Summanden können wie angegeben über die iterativ bestimmten Größen \mathbf{L}_k und \mathbf{r}_k ausgedrückt werden. Der erste Summand

$$H_k = \tilde{\eta}_k^T \tilde{\eta}_k$$

kann ebenfalls nach

$$H_k = \begin{cases} 0 & k = 0 \\ H_{k-1} + \tilde{\eta}_k^2 & k > 0 \end{cases}$$

iterativ berechnet werden. Somit erhält man

$$\tilde{\eta}_k^T \cdot (\mathbf{I} - \Psi_k \cdot (\Psi_k^T \Psi_k)^{-1} \cdot \Psi_k^T) \cdot \tilde{\eta}_k = H_k - \mathbf{r}_k^T \mathbf{L}_k^{-1} \mathbf{r}_k = H_k - \mathbf{r}_k^T \hat{\theta}_k$$

und damit

$$\hat{\sigma}_{\eta,k}^2 = \frac{1}{k-p} \cdot (H_k - \mathbf{r}_k^T \mathbf{L}_k^{-1} \mathbf{r}_k). \quad (8.14)$$

Zur Bestimmung des Schätzwertes für die Varianz muss also nur ein weiterer Skalar gespeichert werden.

Gewichtetes LS-Verfahren mit diagonaler Gewichtsmatrix

Die Schätzgleichung des gewichteten LS-Verfahrens lautet

$$\hat{\theta}_k = (\Psi_k^T \mathbf{W}_k \Psi_k)^{-1} \cdot \Psi_k^T \mathbf{W}_k \tilde{\eta}_k,$$

wobei im Weiteren von einer diagonalen Gewichtmatrix

$$\mathbf{W}_k = \begin{bmatrix} w_1 & & \\ & \ddots & \\ & & w_k \end{bmatrix}$$

ausgegangen wird. Es werden wie im Fall des normalen LS-Verfahrens die entsprechenden Abkürzungen $\mathbf{L}_k = \Psi_k^T \mathbf{W}_k \Psi_k$ und $\mathbf{r}_k = \Psi_k^T \mathbf{W}_k \tilde{\eta}_k$ verwendet, so dass auch hier die Schätzgleichung als $\hat{\theta}_k = \mathbf{L}_k^{-1} \mathbf{r}_k$ geschrieben werden kann. (Auf eine Kennzeichnung, dass \mathbf{L}_k und \mathbf{r}_k hier eine Gewichtung beinhalten, wird verzichtet.)

Damit ergibt sich

$$\begin{aligned} \mathbf{L}_{k+1} &= \Psi_{k+1}^T \mathbf{W}_{k+1} \Psi_{k+1} = [\Psi_k^T \quad \Psi_{k+1}] \begin{bmatrix} \mathbf{W}_k & \\ & w_{k+1} \end{bmatrix} \begin{bmatrix} \Psi_k \\ \Psi_{k+1}^T \end{bmatrix} = \Psi_k^T \mathbf{W}_k \Psi_k + \Psi_{k+1} w_{k+1} \Psi_{k+1}^T \\ &= \mathbf{L}_k + w_{k+1} \cdot \Psi_{k+1} \Psi_{k+1}^T \end{aligned}$$

und

$$\begin{aligned} \mathbf{r}_{k+1} &= \Psi_{k+1}^T \mathbf{W}_{k+1} \tilde{\eta}_{k+1} = [\Psi_k^T \quad \Psi_{k+1}] \begin{bmatrix} \mathbf{W}_k & \\ & w_{k+1} \end{bmatrix} \begin{bmatrix} \tilde{\eta}_k \\ \tilde{\eta}_{k+1} \end{bmatrix} = \Psi_k^T \mathbf{W}_k \tilde{\eta}_k + \Psi_{k+1} w_{k+1} \tilde{\eta}_{k+1} \\ &= \mathbf{r}_k + w_{k+1} \cdot \Psi_{k+1} \tilde{\eta}_{k+1} \end{aligned}$$

als iterative Berechnungsvorschrift.

LS mit exponentiellem Vergessen

Das LS-Verfahren mit exponentiellem Vergessen stellt einen Sonderfall des gewichteten LS-Verfahrens mit der Schätzgleichung

$$\hat{\theta}_k = (\Psi_k^T \mathbf{W}_k \Psi_k)^{-1} \cdot \Psi_k^T \mathbf{W}_k \tilde{\eta}_k$$

dar, bei dem die Gewichtsmatrix

$$\mathbf{W}_{k+1} = \begin{bmatrix} \lambda_{k+1} \mathbf{W}_k & \\ & 1 \end{bmatrix}$$

lautet. Es werden mit jeder neuen Messung die Gewichte der zurückliegenden Messungen verringert, während beim „normalen“ gewichteten LS-Verfahren die alten Gewichte unverändert bleiben.

Mit dieser Gewichtsmatrix ergeben sich die iterativen Berechnungsvorschriften

$$\begin{aligned} \mathbf{L}_{k+1} &= \Psi_{k+1}^T \mathbf{W}_{k+1} \Psi_{k+1} = [\Psi_k^T \quad \Psi_{k+1}] \begin{bmatrix} \lambda_{k+1} \mathbf{W}_k & \\ & 1 \end{bmatrix} \begin{bmatrix} \Psi_k \\ \Psi_{k+1}^T \end{bmatrix} = \lambda_{k+1} \Psi_k^T \mathbf{W}_k \Psi_k + \Psi_{k+1} \Psi_{k+1}^T \\ &= \lambda_{k+1} \mathbf{L}_k + \Psi_{k+1} \Psi_{k+1}^T \end{aligned}$$

und

$$\begin{aligned} \mathbf{r}_{k+1} &= \Psi_{k+1}^T \mathbf{W}_{k+1} \tilde{\eta}_{k+1} = [\Psi_k^T \quad \Psi_{k+1}] \begin{bmatrix} \lambda_{k+1} \mathbf{W}_k & \\ & 1 \end{bmatrix} \begin{bmatrix} \tilde{\eta}_k \\ \tilde{\eta}_{k+1} \end{bmatrix} = \lambda_{k+1} \Psi_k^T \mathbf{W}_k \tilde{\eta}_k + \Psi_{k+1} \tilde{\eta}_{k+1} \\ &= \lambda_{k+1} \mathbf{r}_k + \Psi_{k+1} \tilde{\eta}_{k+1}. \end{aligned}$$

8.6.2 Rekursives Least-Squares-Verfahren

Durch die iterative Berechnung wird das Problem umgangen, dass der Speicherbedarf mit wachsendem k unbeschränkt anwächst. Dennoch ist in jedem Schritt (wenn der Schätzwert zu jedem Schritt bestimmt werden soll) eine Inversion einer Matrix der Dimension $p \times p$ (bzw. eigentlich das Lösen eines quadratischen Gleichungssystems mit p Unbekannten) notwendig. Dies kann durch eine rekursive Formulierung umgangen werden.

Dazu wird

$$\hat{\theta}_{k+1} = \mathbf{L}_{k+1}^{-1} \mathbf{r}_{k+1}$$

zunächst als

$$\hat{\theta}_{k+1} = (\mathbf{L}_k + \boldsymbol{\psi}_{k+1} \boldsymbol{\psi}_{k+1}^T)^{-1} \cdot (\mathbf{r}_k + \boldsymbol{\psi}_{k+1} \tilde{\eta}_{k+1})$$

geschrieben.

Für die Inverse kann nach der Sherman-Morrison-Formel ([Zurmühl und Falk 1984, S. 310])

$$\mathbf{L}_{k+1}^{-1} = (\mathbf{L}_k + \boldsymbol{\psi}_{k+1} \boldsymbol{\psi}_{k+1}^T)^{-1} = \mathbf{L}_k^{-1} - \frac{\mathbf{L}_k^{-1} \boldsymbol{\psi}_{k+1} \boldsymbol{\psi}_{k+1}^T \mathbf{L}_k^{-1}}{1 + \boldsymbol{\psi}_{k+1}^T \mathbf{L}_k^{-1} \boldsymbol{\psi}_{k+1}}. \quad (8.15)$$

angegeben werden. Damit erhält man

$$\begin{aligned} \hat{\theta}_{k+1} &= \left(\mathbf{L}_k^{-1} - \frac{\mathbf{L}_k^{-1} \boldsymbol{\psi}_{k+1} \boldsymbol{\psi}_{k+1}^T \mathbf{L}_k^{-1}}{1 + \boldsymbol{\psi}_{k+1}^T \mathbf{L}_k^{-1} \boldsymbol{\psi}_{k+1}} \right) \cdot (\mathbf{r}_k + \boldsymbol{\psi}_{k+1} \tilde{\eta}_{k+1}) \\ &= \underbrace{\mathbf{L}_k^{-1} \mathbf{r}_k + \mathbf{L}_k^{-1} \boldsymbol{\psi}_{k+1} \tilde{\eta}_{k+1}}_{\hat{\theta}_k} - \frac{\mathbf{L}_k^{-1} \boldsymbol{\psi}_{k+1}}{1 + \boldsymbol{\psi}_{k+1}^T \mathbf{L}_k^{-1} \boldsymbol{\psi}_{k+1}} \boldsymbol{\psi}_{k+1}^T \underbrace{\mathbf{L}_k^{-1} \mathbf{r}_k}_{\hat{\theta}_k} - \frac{\mathbf{L}_k^{-1} \boldsymbol{\psi}_{k+1}}{1 + \boldsymbol{\psi}_{k+1}^T \mathbf{L}_k^{-1} \boldsymbol{\psi}_{k+1}} \boldsymbol{\psi}_{k+1}^T \mathbf{L}_k^{-1} \boldsymbol{\psi}_{k+1} \tilde{\eta}_{k+1}. \end{aligned}$$

Setzt man $\mathbf{P}'_k = \mathbf{L}_k^{-1}$ und führt man die Abkürzung

$$\gamma_k = \frac{\mathbf{P}'_k \boldsymbol{\psi}_{k+1}}{\boldsymbol{\psi}_{k+1}^T \mathbf{P}'_k \boldsymbol{\psi}_{k+1} + 1}$$

ein, so erhält man

$$\hat{\theta}_{k+1} = \hat{\theta}_k + \mathbf{P}'_k \boldsymbol{\psi}_{k+1} \tilde{\eta}_{k+1} - \gamma_k \boldsymbol{\psi}_{k+1}^T \hat{\theta}_k - \frac{\mathbf{P}'_k \boldsymbol{\psi}_{k+1}}{1 + \boldsymbol{\psi}_{k+1}^T \mathbf{P}'_k \boldsymbol{\psi}_{k+1}} \boldsymbol{\psi}_{k+1}^T \mathbf{P}'_k \boldsymbol{\psi}_{k+1} \tilde{\eta}_{k+1}.$$

Betrachtet man nur die Terme mit $\tilde{\eta}_{k+1}$, so erhält man

$$\begin{aligned} &\mathbf{P}'_k \boldsymbol{\psi}_{k+1} \tilde{\eta}_{k+1} - \frac{\mathbf{P}'_k \boldsymbol{\psi}_{k+1}}{1 + \boldsymbol{\psi}_{k+1}^T \mathbf{P}'_k \boldsymbol{\psi}_{k+1}} \boldsymbol{\psi}_{k+1}^T \mathbf{P}'_k \boldsymbol{\psi}_{k+1} \tilde{\eta}_{k+1} \\ &= \mathbf{P}'_k \boldsymbol{\psi}_{k+1} \cdot \left(1 - \frac{\boldsymbol{\psi}_{k+1}^T \mathbf{P}'_k \boldsymbol{\psi}_{k+1}}{1 + \boldsymbol{\psi}_{k+1}^T \mathbf{P}'_k \boldsymbol{\psi}_{k+1}} \right) \cdot \tilde{\eta}_{k+1} \\ &= \mathbf{P}'_k \boldsymbol{\psi}_{k+1} \cdot \left(\frac{1}{1 + \boldsymbol{\psi}_{k+1}^T \mathbf{P}'_k \boldsymbol{\psi}_{k+1}} \right) \cdot \tilde{\eta}_{k+1} \\ &= \gamma_k \cdot \tilde{\eta}_{k+1}, \end{aligned}$$

und damit

$$\begin{aligned}\hat{\theta}_{k+1} &= \hat{\theta}_k - \gamma_k \Psi_{k+1}^T \hat{\theta}_k + \gamma_k \cdot \tilde{\eta}_{k+1} \\ &= \hat{\theta}_k + \gamma_k \cdot (\tilde{\eta}_{k+1} - \Psi_{k+1}^T \hat{\theta}_k).\end{aligned}$$

Aus Gl. (8.15) erhält man

$$\begin{aligned}\mathbf{P}'_{k+1} &= \mathbf{P}'_k + \gamma_k \Psi_{k+1}^T \mathbf{P}'_k \\ &= (\mathbf{I} + \gamma_k \Psi_{k+1}^T) \cdot \mathbf{P}'_k.\end{aligned}$$

Nochmals zusammengefasst lautet der rekursive LS-Algorithmus

$$\begin{aligned}\gamma_k &= \frac{\mathbf{P}'_k \Psi_{k+1}}{\Psi_{k+1}^T \mathbf{P}'_k \Psi_{k+1} + 1} \\ \hat{\theta}_{k+1} &= \hat{\theta}_k + \gamma_k \cdot (\tilde{\eta}_{k+1} - \Psi_{k+1}^T \hat{\theta}_k) \\ \mathbf{P}'_{k+1} &= (\mathbf{I} - \gamma_k \Psi_{k+1}^T) \cdot \mathbf{P}'_k.\end{aligned}$$

Hier fällt auch die Matrixinversion weg. Die Matrix \mathbf{P}'_k ist die durch σ_η^2 geteilte Kovarianzmatrix der geschätzten Parameter ($\mathbf{P}_k = \sigma_\eta^2 \mathbf{P}'_k$) und muss für den ersten Schritt vorgegeben werden. In Abbildung 8.15 ist die Struktur des Algorithmus dargestellt.

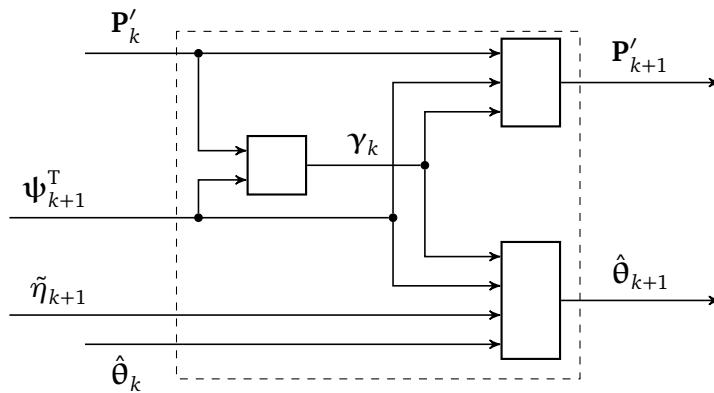


Abbildung 8.15: Rekursives LS-Verfahren

Der rekursive LS-Algorithmus kann auch mit der unskalierten Kovarianzmatrix angegeben werden, indem $\mathbf{P}'_k = \frac{1}{\sigma_\eta^2} \mathbf{P}_k$ eingesetzt wird. Man erhält

$$\begin{aligned}\gamma_k &= \frac{\mathbf{P}_k \Psi_{k+1}}{\Psi_{k+1}^T \mathbf{P}_k \Psi_{k+1} + \sigma_\eta^2} \\ \hat{\theta}_{k+1} &= \hat{\theta}_k + \gamma_k \cdot (\tilde{\eta}_{k+1} - \Psi_{k+1}^T \hat{\theta}_k) \\ \mathbf{P}_{k+1} &= (\mathbf{I} - \gamma_k \Psi_{k+1}^T) \cdot \mathbf{P}_k.\end{aligned}$$

Betrachtet man die Korrekturgleichung für den Schätzwert, $\hat{\theta}_{k+1} = \hat{\theta}_k + \gamma_k \cdot (\tilde{\eta}_{k+1} - \Psi_{k+1}^T \hat{\theta}_k)$, so erkennt man, dass der neue Schätzwert dem alten Schätzwert entspricht, der über die mit einem Faktor gewichtete Differenz $\tilde{\eta}_{k+1} - \Psi_{k+1}^T \hat{\theta}_k$ korrigiert wird. Der Term $\Psi_{k+1}^T \hat{\theta}_k$ ist der Ausgangswert, der sich mit den bis k vorliegenden Daten geschätzten Parametervektor $\hat{\theta}_k$ und den aktuellen Eingangswerten Ψ_{k+1}^T berechnet. Es handelt sich also um die Vorhersage des $k+1$ -ten Ausgangswertes aufgrund der Messungen bis k und wird auch als

$$\hat{\eta}_{k+1|k} = \Psi_{k+1}^T \hat{\theta}_k$$

geschrieben. Damit kann die Korrekturgleichung als $\hat{\theta}_{k+1} = \hat{\theta}_k + \gamma_k \cdot (\tilde{\eta}_{k+1} - \hat{\eta}_{k+1|k})$ geschrieben werden.

Wahl der Anfangswerte

Startet man das rekursive Verfahren mit dem normalen LS-Verfahren, d. h. führt man so viele Messungen durch, bis alle Eingänge (gut) angeregt sind und damit die Schätzwerte und die Kovarianzmatrix (numerisch sicher) bestimmt werden können, und verwendet man die so bestimmten Werte für $\hat{\theta}$ und $P = \sigma_{\eta}^2 L^{-1}$ als Startwerte für das rekursive LS-Verfahren, so entsprechen die Werte für alle k (bis auf numerische Abweichungen) exakt den Werten des normalen LS-Verfahrens.

Wenn das rekursive Verfahren jedoch gewählt wird, dann besteht häufig (rechnertechnisch) keine Möglichkeit, die Gleichung des normalen LS-Verfahrens zu lösen, womit diese Möglichkeit wegfällt. In diesem Fall können die Startwerte entweder aus Vorwissen, welches man über den Prozess und dessen Parameter besitzt, wählen, oder man gibt für $\hat{\theta}_0$ den besten Tipp vor, den man hat (im Zweifel $\hat{\theta}_0 = 0$) und setzt die Kovarianzmatrix in der Form

$$P_0 = \alpha I$$

mit einem sehr großen Wert für α vor. Für $k \rightarrow \infty$ ergeben dann das rekursive und das normale LS-Verfahren die gleichen Schätzwerte. (Das wird später im Zusammenhang mit dem Kalman-Filter nochmals aufgegriffen und begründet.)

Gewichtetes LS-Verfahren

Für das mit einer diagonalen Gewichtsmatrix gewichtete LS-Verfahren erhält man (Herleitung in Anhang D.1.1) das rekursive Verfahren

$$\begin{aligned}\gamma_k &= \frac{P'_k \Psi_{k+1}}{\Psi_{k+1}^T P'_k \Psi_{k+1} + \frac{1}{w_{k+1}}} \\ \hat{\theta}_{k+1} &= \hat{\theta}_k + \gamma_k \cdot (\tilde{\eta}_{k+1} - \Psi_{k+1}^T \hat{\theta}_k) \\ P'_{k+1} &= (I - \gamma_k \Psi_{k+1}^T) \cdot P'_k.\end{aligned}$$

Eine Interpretation von P'_k als Kovarianzmatrix von $\{\hat{\theta}\}_r$ kann nur dann erfolgen, wenn $w_{k'} = \frac{1}{\sigma_{\eta, k'}^2}$, also der Kehrwert der Varianz des Messrauschens ist.

LS-Verfahren mit exponentiellem Vergessen

Es ergibt sich (Anhang D.1.2)

$$\begin{aligned}\gamma_k &= \frac{P'_k \Psi_{k+1}}{\Psi_{k+1}^T P'_k \Psi_{k+1} + \lambda_{k+1}} \\ \hat{\theta}_{k+1} &= \hat{\theta}_k + \gamma_k \cdot (\tilde{\eta}_{k+1} - \Psi_{k+1}^T \hat{\theta}_k) \\ P'_{k+1} &= (I - \gamma_k \Psi_{k+1}^T) \cdot P'_k \cdot \frac{1}{\lambda_{k+1}}.\end{aligned}$$

Zur Interpretation ist hier insbesondere der Faktor $P'_k \cdot \frac{1}{\lambda_{k+1}}$ in der Berechnung der neuen „Kovarianzmatrix“ interessant. Dieser kann so interpretiert werden, dass die „Kovarianz“ P'_k der zurückliegenden Schätzung künstlich durch den Faktor $\frac{1}{\lambda_{k+1}}$ erhöht und damit der alte Schätzwert entsprechend „unsicherer“ gemacht wird, um damit die Mittelung zugunsten der neuen Messung zu verschieben.

8.7 Numerische Berechnung

In diesem Abschnitt wird die numerische Berechnung der Lösung des LS-Problems behandelt. Dieser basiert zu wesentlichen Teilen (abschnittsweise auch wörtlich) auf den „Ergänzungen zur Vorlesung

Identifikation dynamischer Systeme – Numerische Verfahren zur Identifikation linearer Systeme“ von Michael Vogt. Alle dabei angesprochenen Berechnungsverfahren sind gängige numerische Methoden und beispielsweise in [Golub und Van Loan 2013] zu finden.

Geht man von den Normalgleichungen

$$(\Psi^T \Psi) \cdot \hat{\theta} = \Psi^T \tilde{\eta}$$

des LS-Problems aus, so liegt es vielleicht nahe, die Lösung auch numerisch über

$$\hat{\theta} = (\Psi^T \Psi)^{-1} \cdot \Psi^T \tilde{\eta},$$

d. h. die explizite Inversion der Matrix $L = \Psi^T \Psi$ zu bestimmen. In Matlab wird dies z. B. nach Listing 8.1 erreicht. Von diesem Verfahren ist jedoch in der Praxis abzuraten, wie unten noch begründet wird.

Listing 8.1: Berechnung von $\hat{\theta}$ über explizite Inversion

```
1 theta = inv(Psi.' * Psi) * (Psi.' * eta); % SCHLECHT!
```

8.7.1 Berechnung des Schätzwertes bei der Methode der Normalgleichungen

Gauß-Verfahren

Das Gauß-Verfahren zerlegt eine Matrix A in ein Produkt aus einer unteren Dreiecksmatrix L („lower“²³) und oberen Dreiecksmatrix U („upper“),

$$A = L \cdot U.$$

Daher wird dieser Algorithmus auch als LU-Zerlegung bezeichnet. Ist die LU-Zerlegung von A bekannt, dann kann ein Gleichungssystem $Ax = b$ durch zweimaliges Rückeinsetzen gelöst werden, da L und U Dreiecksmatrizen sind. Diese Methode kann wie folgt auf die Normalgleichungen angewendet werden:

1. Berechne die Zerlegung $\Psi^T \Psi = L \cdot U$ mit dem Gauß-Algorithmus.
2. Löse $L \cdot z = \Psi^T \tilde{\eta}$ nach z durch Rückeinsetzen.
3. Löse $U \cdot \hat{\theta} = z$ nach $\hat{\theta}$ durch Rückeinsetzen.

In Matlab kann dies durch Listing 8.2 ausgedrückt werden. Der „\“-Operator von Matlab erkennt die Dreiecksstruktur der Matrizen U und L und führt die Rückeinsetzung durch.

Listing 8.2: Berechnung von $\hat{\theta}$ über LU-Zerlegung

```
[L, U] = lu(Psi.' * Psi);
theta = U \ ( L \ (Psi.' * eta) );
```

Implementiert man die LU-Zerlegung als Gleichungslöser, dann wird die transformierte „rechte Seite“ z gleich mitberechnet, so dass der zweite Schritt entfällt. Diese Variante benötigt weniger Rechenzeit und verursacht geringere Fehler als die explizite Inversion der Informationsmatrix.

²³ Ausschließlich in diesem Abschnitt bezeichnet L die untere Dreiecksmatrix beim Gauß-Verfahren. Ansonsten ist mit L immer die Matrix $\Psi^T \Psi$ gemeint.

Cholesky-Verfahren

Während das Gauß-Verfahren auf beliebige quadratische, eindeutig lösbarre lineare Gleichungssysteme angewendet werden kann, löst das dem Gauß-Verfahren sehr ähnliche Cholesky-Verfahren ausschließlich Systeme mit einer symmetrischen und positiv definiten Matrix, womit es auch für die Normalgleichungen des LS-Problems geeignet ist. Als „Ausgleich“ für den eingeschränkten Anwendungsbereich des Cholesky-Verfahrens besitzt dieses aber wesentlich bessere numerische Eigenschaften.

Das Cholesky-Verfahren zerlegt eine Matrix \mathbf{A} in ein Produkt

$$\mathbf{A} = \mathbf{R}^T \mathbf{R},$$

wobei \mathbf{R} eine obere Dreiecksmatrix ist. Verglichen mit dem Gauß-Verfahren gilt also $\mathbf{L} = \mathbf{R}^T$ und $\mathbf{U} = \mathbf{R}$. Die Matrix \mathbf{R} wird auch als Cholesky-Faktor von \mathbf{A} bezeichnet.

Mit der Cholesky-Verfahren ergibt sich damit das im Prinzip gleiche Vorgehen wie mit dem Gauß-Algorithmus (siehe auch Listing 8.3 für Beispielimplementierung in Matlab):

1. Berechne die Zerlegung $\Psi^T \Psi = \mathbf{R}^T \cdot \mathbf{R}$ mit dem Gauß-Algorithmus.
2. Löse $\mathbf{R}^T \cdot \mathbf{z} = \Psi^T \tilde{\eta}$ nach \mathbf{z} durch Rückeinsetzen.
3. Löse $\mathbf{R} \cdot \hat{\theta} = \mathbf{z}$ nach $\hat{\theta}$ durch Rückeinsetzen.

Listing 8.3: Berechnung von $\hat{\theta}$ über Cholesky-Zerlegung

```
R = chol(Psi.' * Psi);
theta = R \ (R.' \ (Psi.' * eta));
```

Auch das Cholesky-Verfahren kann direkt als Gleichungslöser implementiert werden, so dass die erste Rückeinsetzung entfällt.

Berechnung in Matlab

Wie in den Beispiellistings schon ersichtlich, stellt Matlab alle notwendigen Funktionen zur Bestimmung der LU- und Cholesky-Zerlegung sowie dem Rückeinsetzen zur Verfügung. Für die praktische Anwendung wird man aber zur Lösung der Normalgleichungen den „\“-Operator verwenden, der oben schon für die Ausführung der Rückeinsetzungen verwendet wurde. Mit diesem Operator kann mit $\mathbf{x} = \mathbf{A} \backslash \mathbf{b}$ die Lösung des Gleichungssystems

$$\mathbf{Ax} = \mathbf{b}$$

bestimmt werden, wobei die Struktur der Matrix \mathbf{A} automatisch erkannt wird und ein entsprechend geeignetes Verfahren gewählt wird.

Damit kann die Lösung der Normalgleichungen nach Listing 8.4 berechnet werden. Hier erkennt der „\“-Operator, dass der linke Operand symmetrisch und positiv definit ist und wählt daher die Cholesky-Zerlegung als Lösungsmethode.²⁴

Listing 8.4: Berechnung von $\hat{\theta}$ über „\“-Operator

```
theta = (Psi.' * Psi) \ (Psi.' * eta);
```

²⁴ Genau genommen prüft Matlab nur, ob die Matrix symmetrisch ist und positive Diagonaleinträge besitzt. Ist dies der Fall, versucht es die Cholesky-Zerlegung. Schlägt diese fehl, d. h. ist die Matrix – zumindest im Rahmen der Auflösung der Zahlendarstellung – nicht positiv definit, dann wird ein anderes Zerlegungsverfahren für symmetrische Matrizen angewandt.

8.7.2 Orthogonalisierungsverfahren

Orthogonalisierungsverfahren stellen eine Alternative zu dem im letzten Abschnitt behandelten LS-Verfahren (d. h. Lösung der Normalgleichungen) dar. Sie zeichnen sich durch erheblich bessere numerische Eigenschaften aus, weshalb sie in der Praxis häufig bei schlecht konditionierten Problemen zur Anwendung kommen.

Orthogonale Matrizen

Eine quadratische Matrix

$$\mathbf{Q} = [\mathbf{q}_1 \ \mathbf{q}_2 \ \cdots \ \mathbf{q}_n]$$

wird als „orthogonal“ bezeichnet, wenn die Spalten \mathbf{q}_i orthonormal zueinander sind, also

$$\mathbf{q}_i^T \mathbf{q}_j = \begin{cases} 1 & \text{für } i = j \\ 0 & \text{für } i \neq j \end{cases}$$

gilt. (Leider hat sich in der Literatur dafür der Begriff „orthogonal“ eingebürgert, auch wenn „orthonormal“ eigentlich passender wäre.)

Aus der Definition folgt

$$\mathbf{Q}^T \mathbf{Q} = \mathbf{I}$$

und damit gilt auch $\mathbf{Q}^{-1} = \mathbf{Q}^T$.

Eine wichtige Eigenschaft orthogonaler Matrizen ist, dass ein Vektor \mathbf{x} durch Multiplikation mit einer orthogonalen Matrix \mathbf{Q} seine euklidische Norm (2-Norm) nicht ändert, was durch

$$\|\mathbf{Q}\mathbf{x}\|_2^2 = \mathbf{x}^T \mathbf{Q}^T \mathbf{Q} \mathbf{x} = \mathbf{x}^T \mathbf{x} = \|\mathbf{x}\|_2^2$$

einfach zu sehen ist.

QR-Zerlegung

Mit der QR-Zerlegung wird eine $(k \times p)$ -Matrix Ψ in ein Produkt aus einer orthogonalen $(k \times k)$ -Matrix \mathbf{Q}^T und einer $(k \times p)$ -Matrix der Form $\begin{bmatrix} \mathbf{R} \\ \mathbf{0} \end{bmatrix}$, wobei \mathbf{R} eine obere Dreiecksmatrix ist, geschrieben,²⁵

$$\Psi = \mathbf{Q}^T \cdot \begin{bmatrix} \mathbf{R} \\ \mathbf{0} \end{bmatrix} \cdot \begin{array}{l} \} p \text{ Zeilen} \\ \} k-p \text{ Zeilen} \end{array}$$

Damit gilt

$$\mathbf{Q}\Psi = \begin{bmatrix} \mathbf{R} \\ \mathbf{0} \end{bmatrix}.$$

Diese Zerlegung kann beispielsweise mit der Householder-Transformation bestimmt werden, siehe Anhang D.2.

²⁵ Daneben ist auch die Notation verbreitet, dass \mathbf{R} die gesamte Matrix auf der rechten Seite ist, d. h. $\Psi = \mathbf{Q}^T \mathbf{R}$. Dabei wird dann auch \mathbf{R} als (nicht-quadratische) obere Dreiecksmatrix bezeichnet.

Anwendung auf LS-Problem

Es wird wieder die Aufgabe betrachtet, das $\hat{\theta}$ zu finden, welches die Summe der quadrierten Fehler minimiert. Mit der 2-Norm kann dies wegen $\sum \hat{e}_{k'} = \hat{\mathbf{e}}^T \hat{\mathbf{e}} = \|\hat{\mathbf{e}}\|_2$ auch als

$$\hat{\theta} = \arg \min_{\hat{\theta}} \|\hat{\mathbf{e}}(\hat{\theta})\|_2$$

geschrieben werden. Aufgrund der Eigenschaften einer orthogonalen Matrix gilt für den Vektor $\hat{\mathbf{e}}$ des Fehlers

$$\|\hat{\mathbf{e}}\|_2 = \|\mathbf{Q}\hat{\mathbf{e}}\|_2.$$

Damit erhält man

$$\|\hat{\mathbf{e}}\|_2 = \|\mathbf{Q}\hat{\mathbf{e}}\|_2 = \|\mathbf{Q} \cdot (\Psi\hat{\theta} - \tilde{\eta})\|_2 = \left\| \begin{bmatrix} \mathbf{R}\hat{\theta} \\ \mathbf{0} \end{bmatrix} - \mathbf{Q}\tilde{\eta} \right\|_2.$$

Das Produkt $\mathbf{Q}\tilde{\eta}$ wird entsprechend in $\begin{bmatrix} \mathbf{c} \\ \tilde{\mathbf{c}} \end{bmatrix}$ partitioniert, so dass sich weiter

$$\|\hat{\mathbf{e}}\|_2 = \left\| \begin{bmatrix} \mathbf{R}\hat{\theta} \\ \mathbf{0} \end{bmatrix} - \begin{bmatrix} \mathbf{c} \\ \tilde{\mathbf{c}} \end{bmatrix} \right\|_2 = \left\| \begin{bmatrix} \mathbf{R}\hat{\theta} - \mathbf{c} \\ -\tilde{\mathbf{c}} \end{bmatrix} \right\|_2$$

ergibt.

Anstelle der Minimierung der Norm $\|\hat{\mathbf{e}}\|_2$ kann äquivalent auch deren Quadrat $\|\hat{\mathbf{e}}\|_2^2$ minimiert werden. Für die 2-Norm, die der euklidischen Länge entspricht, kann damit

$$\|\hat{\mathbf{e}}\|_2^2 = \|\mathbf{R}\hat{\theta} - \mathbf{c}\|_2^2 + \|\tilde{\mathbf{c}}\|_2^2$$

geschrieben werden. Da $\tilde{\mathbf{c}}$ nur von den Eingangs- und Messdaten, nicht aber den Schätzwerten $\hat{\theta}$ abhängt ist $\|\tilde{\mathbf{c}}\|_2^2$ aus Sicht der Minimierungsaufgabe eine Konstante. $\hat{\theta}$ ist so zu wählen, dass $\|\mathbf{R}\hat{\theta} - \mathbf{c}\|_2^2$ (und damit $\|\mathbf{e}\|_2^2$) minimal wird.

Wird davon ausgegangen, dass Ψ vollen Spaltenrang hat, dann ist \mathbf{R} eine invertierbare Matrix und es kann immer $\mathbf{R}\hat{\theta} - \mathbf{c} = \mathbf{0}$ erreicht werden, was mit Sicherheit den minimalen Wert darstellt. $\hat{\theta}$ muss damit die Gleichung

$$\mathbf{R}\hat{\theta} = \mathbf{c}$$

erfüllen. Formal ergibt sich der Schätzwert dann als

$$\hat{\theta} = \mathbf{R}^{-1}\mathbf{c}.$$

Da \mathbf{R} aber eine obere Dreiecksmatrix ist, kann $\hat{\theta}$ einfach aus $\mathbf{R}\hat{\theta} = \mathbf{c}$ durch Rückeinsetzen bestimmt werden.

Für die praktische Berechnung wendet man (ähnlich wie beim Gauß- bzw. Cholesky-Verfahren) die QR-Zerlegung meist direkt auf die Matrix $\begin{bmatrix} \Psi & \tilde{\eta} \end{bmatrix}$ an, so dass das Ergebnis

$$\mathbf{Q} \cdot \begin{bmatrix} \Psi & \tilde{\eta} \end{bmatrix} = \begin{bmatrix} \mathbf{R} & \mathbf{c} \\ \mathbf{0} & \tilde{\mathbf{c}} \end{bmatrix} \quad \begin{array}{l} \} p \text{ Zeilen} \\ \} k-p \text{ Zeilen} \end{array}$$

alle benötigten Größen erhält. Hierbei ist zu beachten, dass das verwendete \mathbf{Q} weiterhin die Matrix ist, die sich aus der QR-Zerlegung von Ψ alleine (und nicht $[\Psi \quad \tilde{\eta}]$) ergibt. Bei der Berechnung wird aber die gesamte neue rechte Seite „mitgenommen“. Dabei kann der Algorithmus so implementiert werden, dass die Matrix \mathbf{Q} nicht explizit berechnet wird. Diese wird nicht mehr benötigt, da \mathbf{c} durch die Erweiterung der rechten Seite direkt mitberechnet wird. Dies bedeutet eine erhebliche Reduktion des Speicherbedarfs.

Führt man einen weiteren Schritt der Householder-Transformation durch, d. h. wendet man die QR-Zerlegung auf die komplette Matrix $[\Psi \quad \tilde{\eta}]$ an, so kann man direkt den Wert $\|\tilde{\mathbf{c}}\|_2^2$ des Residuums ablesen.

$$\mathbf{Q}' \cdot [\Psi \quad \tilde{\eta}] = \begin{bmatrix} \mathbf{R} & \mathbf{c} \\ \mathbf{0} & \|\tilde{\mathbf{c}}\|_2^2 \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \quad \begin{array}{l} \} p \text{ Zeilen} \\ \} 1 \text{ Zeile} \\ \} k-p-1 \text{ Zeilen} \end{array}$$

Allerdings bringt dies rechenzeittechnisch keinen Vorteil mehr, da $\|\tilde{\mathbf{c}}\|_2^2$ auch bei der Householder-Transformation bestimmt werden muss. Da auch \mathbf{Q}' in der Regel nicht gebraucht wird, macht dies nur Sinn, wenn mit vorhandenen Funktionen knapper Code erstellt werden soll.

In Listing 8.5 und 8.6 ist für beide Varianten der Matlab-Code gegeben. Aber auch hier wird man in der Praxis den „\“-Operator verwenden. Denn wendet man diesen auf eine nicht-quadratische Matrix an, so wird automatisch die Minimum-Norm-Lösung bestimmt. Somit wird man in Matlab in der Regel den Code nach Listing 8.7 benutzen.

Listing 8.5: Berechnung von $\hat{\theta}$ über QR-Zerlegung (Variante 1)

```

4 [Q, R] = qr(Psi);
  c = Q' * eta;
  theta = R(1:p,:) \ c(1:n);
  %res = norm(c(p+1:k))^2;

```

Listing 8.6: Berechnung von $\hat{\theta}$ über QR-Zerlegung (Variante 2)

```

1 S = qr([Psi, eta]);
  S = triu(S);
  theta = S(1:p,1:p) \ S(1:p,p+1);
  %res = S(p+1,p+1)

```

Listing 8.7: Berechnung von $\hat{\theta}$ über „\“-Operator

```

1 theta = Psi \ eta;
  %res = norm(Psi * theta - eta)^2;

```

8.7.3 Vergleich der Methoden

Sensitivitätsanalyse

Fehlerfortpflanzung in Gleichungssystemen

Es wird zunächst die Fehlerfortpflanzung in Gleichungssystemen untersucht (Sensitivitätsanalyse). Man geht dazu von einem linearen Gleichungssystem $\mathbf{Ax} = \mathbf{b}$ aus und untersucht, welchen Fehler Δx eine Störung $\Delta \mathbf{b}$ der rechten Seite verursacht.

Das neue Gleichungssystem lautet also

$$\mathbf{A} \cdot (\mathbf{x} + \Delta \mathbf{x}) = \mathbf{b} + \Delta \mathbf{b},$$

was sich wieder in

$$\mathbf{Ax} = \mathbf{b}$$

und

$$\mathbf{A}\Delta \mathbf{x} = \Delta \mathbf{b}$$

aufteilen lässt.

Unter Verwendung verträglicher Normen²⁶ folgt aus der Lösung des „gestörten Anteils“,

$$\Delta \mathbf{x} = \mathbf{A}^{-1} \Delta \mathbf{b},$$

und den Eigenschaften der Normen, dass

$$\|\Delta \mathbf{x}\| = \|\mathbf{A}^{-1} \Delta \mathbf{b}\| \leq \|\mathbf{A}^{-1}\| \cdot \|\Delta \mathbf{b}\|$$

gilt.

Für den ungestörten Teil gilt

$$\|\mathbf{b}\| = \|\mathbf{Ax}\| \leq \|\mathbf{A}\| \cdot \|\mathbf{x}\|$$

und damit

$$\frac{1}{\|\mathbf{x}\|} \leq \frac{\|\mathbf{A}\|}{\|\mathbf{x}\|}$$

Damit kann nun der relative Fehler

$$\begin{aligned} \frac{\|\Delta \mathbf{x}\|}{\|\mathbf{x}\|} &\leq \frac{\|\mathbf{A}^{-1}\| \cdot \|\Delta \mathbf{b}\|}{\|\mathbf{x}\|} \\ &\leq \|\mathbf{A}\| \cdot \|\mathbf{A}^{-1}\| \cdot \frac{\|\Delta \mathbf{b}\|}{\|\mathbf{b}\|} \end{aligned}$$

abgeschätzt werden.

Den Ausdruck

$$\text{cond}(\mathbf{A}) = \|\mathbf{A}\| \cdot \|\mathbf{A}^{-1}\|$$

nennt man Konditionszahl von \mathbf{A} . Die Berechnung der Konditionszahl ist in der Praxis sehr schwierig, da die dazu nötigen Rechenoperationen selbst fehlerbehaftet sind. Im Fall der Spektralnorm (2-Norm) kann man die Konditionszahl über die Singulärwertzerlegung nach

$$\text{cond}_2(\mathbf{A}) = \frac{\sigma_{\max}}{\sigma_{\min}}$$

berechnen, wobei σ_{\max} der größte und σ_{\min} der kleinste Singulärwert von \mathbf{A} ist. Als Singulärwerte bezeichnet man die Wurzeln der Eigenwerte von $\mathbf{A}^T \mathbf{A}$. In dieser Definition kann die Konditionszahl auch für nicht-quadratische Matrizen angegeben werden.

Für den Fehler in der Lösung gilt also

$$\frac{\|\Delta \mathbf{x}\|}{\|\mathbf{x}\|} \leq \text{cond}(\mathbf{A}) \cdot \frac{\|\Delta \mathbf{b}\|}{\|\mathbf{b}\|}.$$

In gleicher Weise kann man annehmen, dass zusätzlich zu den Fehlern $\Delta \mathbf{b}$ in \mathbf{b} auch noch Fehler $\Delta \mathbf{A}$ in \mathbf{A} vorhanden sind. Es zeigt sich, dass diese im Wesentlichen ähnlich eingehen.

²⁶ Die Matrixnorm $\|\mathbf{A}\|$ ist verträglich zur Vektornorm $\|\mathbf{x}\|$, wenn $\|\mathbf{Ax}\| \leq \|\mathbf{A}\| \cdot \|\mathbf{x}\|$ gilt. Beispielsweise ist die Spektralnorm $\|\mathbf{A}\|_2$ (größter Singulärwert von \mathbf{A}) mit $\|\mathbf{x}\|_2$ (2-Norm bzw. euklidischen Norm) verträglich.

Anwendung auf LS-Problem

Wendet man das Ergebnis der Sensitivitätsanalyse auf den numerisch kritischen Schritt des Weges über die Normalgleichungen an (die Berechnung der Lösung $\hat{\theta}$ des linearen Gleichungssystems $(\Psi^T \Psi) \cdot \hat{\theta} = (\Psi^T \tilde{\eta})$), erhält man

$$\frac{\|\Delta \hat{\theta}\|}{\|\hat{\theta}\|} \leq \text{cond}(\Psi^T \Psi) \cdot \frac{\|\Delta(\Psi^T \tilde{\eta})\|}{\|\Psi^T \tilde{\eta}\|} = \text{cond}^2(\Psi) \cdot \frac{\|\Delta(\Psi^T \tilde{\eta})\|}{\|\Psi^T \tilde{\eta}\|}.$$

(Rundungs)Fehler des Zwischenergebnisses $\Psi^T \tilde{\eta}$ schlagen sich also mit einem Faktor $\text{cond}^2(\Psi)$ im Ergebnis nieder.

Im Fall der QR-Zerlegung ist der numerisch kritische Schritt die Lösung des Gleichungssystems $R\hat{\theta} = c$ und es folgt

$$\frac{\|\Delta \hat{\theta}\|}{\|\hat{\theta}\|} \leq \text{cond}(R) \cdot \frac{\|\Delta c\|}{\|c\|}.$$

Weiter gilt

$$\text{cond}(R) = \text{cond}(\Psi),$$

da eine Multiplikation mit Q normerhaltend ist. D.h. bei der QR-Zerlegungen wirken sich Rundungsfehler im Zwischenergebnis c nur mit dem Faktor $\text{cond}(\Psi)$ aus. Damit ist die Sensitivität bezüglich Rundungsfehlern deutlich geringer als bei der Lösung über die Normalgleichungen.

Darüber hinaus lässt sich auch zeigen, dass

$$\frac{\|\Delta \hat{\theta}\|}{\|\hat{\theta}\|} \leq \text{cond}(\Psi) \cdot \frac{\|\Delta \tilde{\eta}\|}{\|\tilde{\eta}\|}$$

gilt. Dies bedeutet, dass Abweichungen im Vektor $\tilde{\eta}$ der Messwerte mit $\text{cond}(\Psi)$ eingehen. Dies ist bei beiden Lösungsmethoden der Fall. (Die Sensitivitätsanalyse, so wie hier vorgestellt, berücksichtigt an sich keine numerischen Rundungsfehler. Sie gibt lediglich an, wie sich eine Änderung der rechten Seite auf die Lösung des Gleichungssystems auswirkt. Und da beide Verfahren letztlich das gleiche zugrundeliegende Problem lösen, hat der gleiche Fehler $\Delta \tilde{\eta}$ auch die gleiche Auswirkung auf $\hat{\theta}$.)

Das Ergebnis der Sensitivitätsanalyse kann daher zum einen verwendet werden, um den Einfluss von Rundungsfehlern bei der Berechnung abzuschätzen. Dies funktioniert auch sehr gut, obwohl die genauere Abschätzung von Rundungsfehlern während der Rechnung deutlich komplizierter ist. Es kann hier aber auch eine Aussage darüber getroffen werden, wie sich das (relative) Messrauschen (bei ansonsten idealer Berechnung) auf den geschätzten Parameterwert auswirkt.

Beispiel

Zur Diskussion wird der Prozess

$$y_{k'} = \theta_1 + \theta_2 u_{k'} + \theta_3 u_{k'}^2 + \theta_4 u_{k'}^3 + \theta_5 u_{k'}^4 + \theta_6 u_{k'}^5$$

betrachtet, wessen Parameter θ_j (die hier alle als zwei gewählt sind) aus Messungen $\tilde{y}_{k'}$ identifiziert werden sollen.

In Standardform lauten die Größen des LS-Problems

$$\begin{aligned}\eta_{k'} &= [y_{k'}] \\ \Psi_{k'}^T &= [1 \ u_{k'} \ u_{k'}^2 \ u_{k'}^3 \ u_{k'}^4 \ u_{k'}^5] \\ \Theta &= [\theta_1 \ \theta_2 \ \theta_3 \ \theta_4 \ \theta_5 \ \theta_6]^T \\ &= [2 \ 2 \ 2 \ 2 \ 2 \ 2]^T.\end{aligned}$$

Die Parameter werden zunächst aus ungestörten Messwerten bestimmt (d. h. die Werte $\tilde{\eta}_{k'}$ enthalten nur die Rundungsfehler aufgrund der endlichen Rechengenauigkeit). Es werden zur Identifikation 100 gleichmäßig über ein Intervall verteilte Werte für $u_{k'}$ verwendet. Dabei wird einmal das Intervall $[-1, 1]$ und einmal das Intervall $[10, 12]$ verwendet. Der Verlauf der gesuchten Funktion in den jeweiligen Intervallen ist in Abbildung 8.16a bzw. (b) gezeigt.

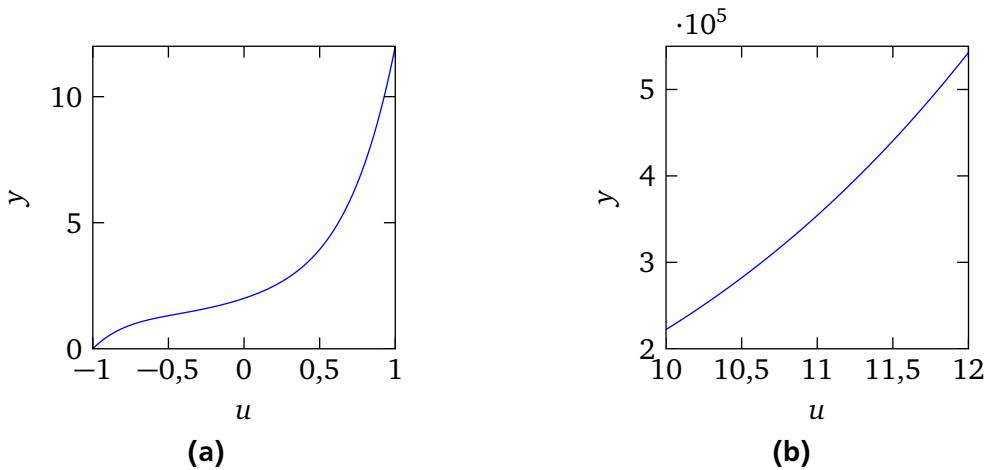


Abbildung 8.16

Intervall $[-1, 1]$

Mit den 100 gleichmäßig im Intervall $[-1, 1]$ verteilten $u_{k'}$ -Werten ergibt sich für die Eingangsmatrix eine Konditionszahl von $\text{cond}_2(\Psi) = 41,92$.

Werden ungestörte Messdaten verwendet, d. h. beinhaltet $\tilde{\eta}$ nur die Rundungsfehler aufgrund der Maschinengenauigkeit (die relativ, d. h. auf den jeweiligen Zahlenwert bezogen bei den verwendeten double-Werten im Bereich von $1 \cdot 10^{-16}$ liegt), ergeben sich die in Tabelle 8.4 aufgeführten Ergebnisse.

Neben den geschätzten Parameterwerten (die hier die Fehler $\Delta\hat{\theta}_i$ beinhalten) ist noch das Verhältnis $\frac{\|\Delta\hat{\theta}\|}{\|\hat{\theta}\|} / \frac{\|\Delta\tilde{\eta}\|}{\|\tilde{\eta}\|}$, der relative Fehler der Schätzwerte sowie das Residuum, d. h. die Abweichung der geschätzten Werte $\hat{\eta} = \Psi\hat{\theta}$ von den Messwerten $\tilde{\eta}$ angegeben.

Mit dem Verhältnis $\frac{\|\Delta\hat{\theta}\|}{\|\hat{\theta}\|} / \frac{\|\Delta\tilde{\eta}\|}{\|\tilde{\eta}\|}$ lässt sich die Abschätzung über die Konditionszahl von Ψ überprüfen. Für die Lösung über das Orthogonalisierungsverfahren gilt

$$\frac{\|\Delta\hat{\theta}\|}{\|\hat{\theta}\|} \leq \text{cond}(\Psi) \cdot \frac{\|\Delta c\|}{\|c\|}$$

und für die Lösung über die Normalgleichungen ist

$$\frac{\|\Delta\hat{\theta}\|}{\|\hat{\theta}\|} \leq \text{cond}^2(\Psi) \cdot \frac{\|\Delta(\Psi^T\tilde{\eta})\|}{\|\Psi^T\tilde{\eta}\|}.$$

Tabelle 8.4: Lösung des gut konditionierten LS-Problems (ohne Rauschen)

	Lsg. über Normalgl. mit Inversion	Lsg. über Normalgl. mit Cholesky-Zerlegung	Lsg. über QR-Zerlegung
$\hat{\theta}$	2,0000000000000008	1,999999999999999	2,000000000000000
	1,999999999999997	2,000000000000008	1,999999999999998
	2,0000000000000020	2,000000000000008	2,000000000000000
	1,999999999999947	1,999999999999967	2,000000000000008
	1,999999999999995	1,999999999999991	2,000000000000001
	2,0000000000000050	2,0000000000000027	1,999999999999991
$\frac{\ \Delta\hat{\theta}\ }{\ \hat{\theta}\ } / \frac{\ \Delta\tilde{\eta}\ }{\ \tilde{\eta}\ }$	186	110	29
$\frac{\ \Delta\hat{\theta}\ }{\ \hat{\theta}\ }$	$1,56 \cdot 10^{-14}$	$9,22 \cdot 10^{-15}$	$2,42 \cdot 10^{-15}$
$\ \Psi\hat{\theta} - \tilde{\eta}\ $	$1,67 \cdot 10^{-13}$	$1,49 \cdot 10^{-14}$	$1,01 \cdot 10^{-14}$

die numerisch relevante Abschätzung. Im Fall ohne Rauschen ergibt sich der Wert $\frac{\|\Delta\tilde{\eta}\|}{\|\tilde{\eta}\|}$ aus der endlichen Rechengenauigkeit der verwendeten Gleitkommadarstellung. Um diesen Quotienten abzuschätzen, ist jeder Eintrag von $\Delta\tilde{\eta}$ auf die Hälfte der Zahlenauflösung des jeweiligen Eintrags in $\tilde{\eta}$ gesetzt. Der Wert von $\frac{\|\Delta\tilde{\eta}\|}{\|\tilde{\eta}\|}$ liegt hier immer im Bereich von $8 \cdot 10^{-17}$. Die Quotienten $\frac{\|\Delta\epsilon\|}{\|\epsilon\|}$ und $\frac{\|\Delta(\Psi^T\tilde{\eta})\|}{\|\Psi^T\tilde{\eta}\|}$ besitzen ungefähr denselben Wert, so dass in den Tabellen der besseren Lesbarkeit halber auf diese Unterscheidung verzichtet ist.

Bei dem gut konditionierten Problem ist die Berechnung unkritisch, wie aus den gezeigten Ergebnissen hervorgeht. Auch wird die Abschätzung des relativen Fehlers von $\hat{\theta}$ über die Konditionszahlen in allen Fällen eingehalten.

Intervall [10, 12]

Verwendet man Messwerte aus dem Intervall [10, 12] zur Identifikation, so ist

$$\text{cond}_2(\Psi) = 7,53 \cdot 10^{11}$$

$$\text{bzw. } \text{cond}_2^2(\Psi) = 5,68 \cdot 10^{23}.$$

Diese sehr große Konditionszahl schlägt sich deutlich in den numerischen Ergebnissen, die in Tabelle 8.5 aufgelistet sind, nieder.

Die über die Normalgleichungen geschätzten Werte sind unbrauchbar, wobei die Berechnung über die explizite Bestimmung der Inverse völlig Phantasiewerte sind, wohingegen die über die Cholesky-Zerlegung²⁷ bestimmten Werte zumindest noch ein halbwegs kleines Residuum ergeben. D. h. die gefundenen Werte haben zwar wenig mit den eigentlichen Parameterwerten zu tun, lösen das zugrundeliegende Problem, die Minimierung des quadratischen Fehlers, aber noch einigermaßen gut. Da es bei der Systemidentifikation aber letztlich nicht darauf ankommt, die vorliegenden Messwerte gut abzubilden, sondern auch neue, unbekannte Messwerte erklären zu können, ist die Güte der geschätzten Parameter das einzig zählende Kriterium. (Hier wird der Parameter θ_6 , der zu u_k^5 gehört, noch relativ gut geschätzt. Letztlich bedeutet das hier, dass die anderen Parameter im Intervall [10, 12] kaum Einfluss auf den Prozess haben, was auch einsichtig ist. Würde man aber versuchen, mit den über die Normalgleichungen und Cholesky-Zerlegung geschätzten Parameterwerten den Wert bei z. B. $u_{k'} = 2$ zu bestimmen, würde sich ein sehr großer Fehler ergeben.)

²⁷ Genauer gesagt, die Lösung der Normalgleichungen über den „\“-Operator in Matlab.

Tabelle 8.5: Lösung des schlecht konditionierten LS-Problems (ohne Rauschen)

	Lsg. über Normalgl. mit Inversion	Lsg. über Normalgl. mit Cholesky-Zerlegung	Lsg. über QR-Zerlegung
$\hat{\theta}$	$1,0130 \cdot 10^4$	$-2,0522 \cdot 10^4$	$1,999822978737063$
	$1,0674 \cdot 10^{14}$	$9,3650 \cdot 10^3$	$2,000079952762307$
	$6,0025 \cdot 10^{13}$	$-1,7049 \cdot 10^3$	$1,999985567823533$
	$5,6431 \cdot 10^{12}$	$1,5744 \cdot 10^2$	$2,000001301464708$
	$8,8668 \cdot 10^{10}$	$-5,0709$	$1,999999941367653$
	2,0025	2,1285	$2,000000001055695$
$\frac{\ \Delta\hat{\theta}\ }{\ \hat{\theta}\ } / \frac{\ \Delta\tilde{\eta}\ }{\ \tilde{\eta}\ }$	$3,2 \cdot 10^{29}$	$5,9 \cdot 10^{19}$	$5,1 \cdot 10^{11}$
$\frac{\ \Delta\hat{\theta}\ }{\ \theta\ }$	$2,50 \cdot 10^{13}$	4618	$23,98 \cdot 10^{-5}$
$\ \Psi\hat{\theta} - \tilde{\eta}\ $	$1,75 \cdot 10^{17}$	0,0516	$1,61 \cdot 10^{-9}$

Mit der QR-Zerlegung ergeben sich hier noch sehr gute Ergebnisse, was die beschriebene Theorie bestätigt.

Die Abschätzung des relativen Fehlers von $\hat{\theta}$ über die Konditionszahlen wird bei der Lösung über die QR-Zerlegung sowie die Lösung über die Cholesky-Zerlegung eingehalten. Es zeigt sich auch, dass die sich ergebenden Werte für $\frac{\|\Delta\hat{\theta}\|}{\|\hat{\theta}\|} / \frac{\|\Delta\tilde{\eta}\|}{\|\tilde{\eta}\|}$ (bzw. eigentlich $\frac{\|\Delta\hat{\theta}\|}{\|\hat{\theta}\|} / \frac{\|\Delta c\|}{\|c\|}$ und $\frac{\|\Delta\hat{\theta}\|}{\|\hat{\theta}\|} / \frac{\|\Delta(\Psi^T\tilde{\eta})\|}{\|\Psi^T\tilde{\eta}\|}$) ziemlich nahe an den durch $\text{cond}(\Psi)$ bzw. $\text{cond}^2(\Psi)$ gegebenen Grenzen liegen, so dass diese Abschätzungen durchaus realistisch sind, auch wenn diese die numerischen Fehler während der Berechnung gar nicht explizit berücksichtigen.

Fazit

Dieses Beispiel zeigt, dass die numerische Berechnung über die QR-Zerlegung die besten Ergebnisse, hier sogar die einzige brauchbaren Ergebnisse, liefert. Damit sollte wenn immer möglich dieses Verfahren (oder ein anderes Orthogonalisierungsverfahren wie die Singulärwertzerlegung, auf die hier nicht eingegangen wurde) zur Bestimmung der Lösung des LS-Problems verwendet werden.

Zur Lösung über die QR-Zerlegung existiert mit dem DFSI-Verfahren (Discrete Square Root Filtering in Information Form, siehe Anhang D.3) auch eine rekursive Formulierung. Dieser sollte entsprechend aus numerischer Sicht dem in Abschnitt 8.6.2 vorgestellten Verfahren den Vorzug gegeben werden.

Vorteile der Berechnung über die Normalgleichungen sind eine schnellere Berechnung und ein geringerer Speicherbedarf. Für die meisten Anwendungen sollte dies jedoch eine untergeordnete Rolle spielen. Werden die Normalgleichungen zur Lösung verwendet, ist auf gar keinen Fall die Inverse $\Psi^T\Psi$ explizit zu berechnen, sondern das Gleichungssystem $(\Psi^T\Psi) \cdot \hat{\theta} = \Psi^T\tilde{\eta}$ ist direkt nach $\hat{\theta}$ zu lösen, z. B. wie hier vorgestellt über die Cholesky-Zerlegung.

Für eine Implementierung mit Festkommazahlen ist keines der betrachteten Verfahren geeignet. In diesem Fall muss auf Approximationsverfahren wie z. B. der NLMS-Algorithmus (Normalized Least Mean Squares) zurückgegriffen werden, [Isermann und Münchhof 2011, S. 306ff, S. 310ff].

8.8 Diskussion von Anregung, Kondition und Varianz

Abschließend soll noch der Zusammenhang zwischen der Wahl der Anregung des Prozesses, der Kondition der Matrix Ψ und der Varianz der geschätzten Parameter diskutiert werden.

Dazu wird nochmals das letzte, schlecht konditionierte Beispiel (Eingangsdaten im Intervall [10, 12]) zur numerischen Berechnung betrachtet, wobei jetzt noch ein Messrauschen mit einer Standardabweichung von 0,000001 (ein Millionstel) des Maximalwerts der Funktion in dem betrachteten Intervall auf die Messdaten aufaddiert ist. Dies ist ein äußerst geringer Wert, und in der Praxis ist prinzipiell mit einem höheren Rauschen zu rechnen.

Das Ergebnis ist in Tabelle 8.6 gezeigt.

Tabelle 8.6: Lösung des gut konditionierten LS-Problems (mit Rauschen)

	Lsg. über Normalgl. mit Inversion	Lsg. über Normalgl. mit Cholesky-Zerlegung	Lsg. über QR-Zerlegung
$\hat{\theta}$	$1,9161 \cdot 10^5$	$1,5781 \cdot 10^5$	$1,5446 \cdot 10^5$
	$1,0674 \cdot 10^{14}$	$-7,0026 \cdot 10^4$	$-6,8498 \cdot 10^4$
	$6,0025 \cdot 10^{13}$	$1,2406 \cdot 10^4$	$1,2127 \cdot 10^4$
	$5,6431 \cdot 10^{12}$	$-1,0942 \cdot 10^3$	$-1,0688 \cdot 10^3$
	$8,8668 \cdot 10^{10}$	50,334	49,180
	1,0166	1,1494	1,1704
$\frac{\ \Delta\hat{\theta}\ }{\ \hat{\theta}\ } / \frac{\ \Delta\tilde{\eta}\ }{\ \tilde{\eta}\ }$	$1,42 \cdot 10^{19}$	$2,01 \cdot 10^{10}$	$1,97 \cdot 10^{10}$
$\frac{\ \Delta\hat{\theta}\ }{\ \theta\ }$	$2,50 \cdot 10^{13}$	35334	34580
$\ \Psi\hat{\theta} - \tilde{\eta}\ $	$1,75 \cdot 10^{17}$	6,0129	6,0129

Man sieht, dass jetzt auch das QR-Verfahren zu völlig falschen Parameterwerten führt. Die Ergebnisse sind vergleichbar mit der Lösung über die Normalgleichungen und Cholesky-Zerlegung. Die Lösung über die explizite Inversenbildung ist auch hier völliger Unsinn, was auch von daher zu erwarten war, da bei der Inversenbildung, die schon im Beispiel zuvor nicht geglückt ist, die Messdaten noch gar keine Rolle spielen.

Das Problem, welches sich hier zeigt, ist dass auch die „Kondition“ des analytisch gelösten LS-Problems bezüglich Fehler in $\tilde{\eta}$ den Wert $\text{cond}(\Psi)$ hat. Ist diese Zahl hoch, dann können schon kleine Änderungen im Messvektor aufgrund des unvermeidlichen Messrauschen zu großen Abweichungen in den geschätzten Parametern führen.

Die Konditionszahl bezüglich der Spektralnorm, $\text{cond}_2(\Psi)$, gibt das Verhältnis von größtem zu kleinstem Singulärwert an. Daneben ist noch das absolute Niveau des kleinsten Singulärwertes relevant. Damit lassen sich prinzipiell vier Fälle unterscheiden, die in der Praxis natürlich stufenlos ineinander übergehen.

Wenn die Konditionszahl klein ist, dann bedeutet das, dass alle Eingangsrichtungen gleichmäßig angeregt werden, und damit die Varianzen aller Schätzwerte ähnlich sind. Ist der kleinste Singulärwert relativ groß, so sind alle Varianzen klein und eine gute Schätzung der Parameter ist möglich. Dies ist dazu noch numerisch unkritisch und damit der anzustrebende Fall. Umso mehr, desto geringer die Auflösung der verwendeten Zahlendarstellung ist, und je geringer die Möglichkeiten zur Implementierung numerisch stabiler Lösungsverfahren sind. Die Inverse von $\mathbf{L} = \Psi^T \Psi$ kann (zumindest bei moderater Parameteranzahl) noch sicher bestimmt werden, d. h. die Varianzen der Schätzwerte können angegeben werden.

Wenn die Konditionszahl und der kleinste Singulärwert klein ist, dann werden alle Eingangsrichtungen nur schlecht angeregt. Man kann zwar ohne numerische Probleme die Lösung bestimmen, die erhaltenen Schätzwerte sind aber unbrauchbar. Zumindest kann dies noch leicht an den noch sicher berechenbaren Varianzen erkannt werden.

Ist die Konditionszahl groß, dann bedeutet dies, dass die Eingangsrichtungen sehr unterschiedlich gut angeregt werden. Ist der kleinste Singulärwert dabei groß, dann bedeutet das, dass manche Eingangs-

richtungen gut und manche sehr gut angeregt werden. Damit ist eine sinnvolle Identifikation möglich, allerdings ist die Berechnung numerisch kritisch. Hierbei helfen dann numerisch stabile Verfahren wie die QR-Zerlegung. Störend ist in diesem Fall auch, dass die Varianzen der Schätzwerte nicht sicher bestimmt werden können.

Bei großer Konditionszahl aber (zu) niedrigem kleinsten Singulärwert wird mindestens eine Eingangsrichtung (zu) schlecht angeregt. Die Berechnung der Lösung bedarf numerisch stabiler Verfahren, die Lösung ist aber dennoch nicht brauchbar. Da die Varianzen nicht sicher berechnet werden können, ist dies nicht direkt zu erkennen. Wenn möglich, sollten die Parameterwerte mit weiteren Messdaten, die nicht zur Identifikation verwendet wurden und den gesamten Bereich abdecken, die das Modell beschreiben soll, getestet (validiert) werden.

Kann man das Eingangssignal zur Identifikation frei vorgeben, dann sollte man also für eine gute Anregung aller Eingangsrichtungen sorgen. Wenn man auf Messwerte angewiesen ist, die während des normalen Betriebes auftreten, können große Konditionszahlen dadurch entstehen, dass neben Phasen guter Anregung (in der mit dem Prozess „was passiert“) auch lange Phasen existieren, in denen sich der Prozess in einem Ruhezustand befindet. Die in den Ruhephasen gesammelten Daten sorgen für immer mehr Einträge in Ψ , die allerdings nur zu einer Eingangsrichtung gehören. Bei den numerischen Berechnungen können diese dann, aufgrund der endlichen Rechengenauigkeit, die anderen Anregungen überdecken. D. h. während analytisch gesehen jeder weitere Messpunkt die Schätzung zumindest nicht schlechter machen kann, gilt dies aus numerischer Sicht nicht mehr. (Die Diagonaleinträge von $\mathbf{L}_k^{-1} = (\Psi_k^T \Psi_k)^{-1}$ sinken monoton, während die Konditionszahl $\text{cond}(\Psi_k)$ mit k durchaus stark steigen kann.)

Um dennoch die Messdaten nutzen zu können, ist in diesem Fall eine stabile numerische Methode wie die QR-Zerlegung hilfreich. Alternativ könnten Messdaten, die in den ruhigen Phasen aufgenommen wurden, einfach verworfen werden. (Oder im Fall des rekursiven Verfahrens nicht verwendet werden. Dies gilt insbesondere bei der Identifikation mit einem Vergessensfaktor, da hierbei die Varianzen tatsächlich schlechter werden können!)

Fazit

Es sollte ein gute Anregung des Prozesses angestrebt werden. Numerisch stabile Verfahren wie allen voran die QR-Zerlegung können bei der Berechnung der Schätzwerte aus numerischen Gründen hilfreich bis nötig sein, allerdings kann kein numerisches Verfahren eine ungenügende Anregung kompensieren. (Solange kein wesentlicher Grund dagegen spricht, sollte aber unabhängig von der Konditionierung des Problems ein Orthogonalisierungsverfahren zur Lösung verwendet werden.)

9 Parametrische Identifikation dynamischer Systeme

9.1 Schätzung dynamischer Zusammenhänge aus Zeitbereichsdaten

9.1.1 Zeitdiskrete Übertragungsfunktionen

Es sollen die Parameter a_i und b_i einer zeitdiskreten Übertragungsfunktion

$$G(z) = \frac{b_n + b_{n-1}z^{-1} + \cdots + b_0z^{-n}}{1 + a_{n-1}z^{-1} + \cdots + a_0z^{-n}} \cdot z^{-d}, \quad d \geq 0,$$

geschätzt werden, die aus einem dynamischen Anteil n -ter Ordnung sowie ggf. einer Totzeit von d Abtastschritten besteht. Meist hat der dynamische Anteil keinen Durchgriff, so dass bekannt ist, dass $b_n = 0$ ist, und die Schätzung von

$$G(z) = \frac{b_{n-1}z^{-1} + \cdots + b_0z^{-n}}{1 + a_{n-1}z^{-1} + \cdots + a_0z^{-n}} \cdot z^{-d} = \frac{B(z^{-1})}{A(z^{-1})} \cdot z^{-d}, \quad d \geq 0,$$

ausgehen kann. Davon wird im Weiteren auch ausgegangen. Prinzipiell wird auch angenommen, dass die Systemordnung n und die Totzeit d bekannt sind.

Konkret für den Drei-Massen-Schwinger gilt $n = 5$, $d = 0$ und $b_5 = 0$ (kein Durchgriff), d. h.

$$G(z) = \frac{b_4z^{-1} + b_3z^{-2} + b_2z^{-3} + b_1z^{-5} + b_0z^{-5}}{1 + a_4z^{-1} + a_3z^{-2} + a_2z^{-3} + a_1z^{-4} + a_0z^{-5}}.$$

In Kapitel 3 wurde schon die Optimierung des Ausgangsfehlers betrachtet. Diese hat den Nachteil, dass in aller Regel Optimierungsaufgaben entstehen, die nicht geschlossen gelöst werden können.

Eine Alternative ist die Optimierung des sogenannten Gleichungsfehlers. Der dabei betrachtete Fehler

$$\hat{e}_{k'} = \tilde{y}_{k'} + \hat{a}_{n-1}\tilde{y}_{k'-1} + \cdots + \hat{a}_0\tilde{y}_{k'-n} - \hat{b}_{n-1}u_{k'-1-d} - \hat{b}_{n-2}u_{k'-2-d} - \cdots - \hat{b}_0u_{k'-n-d}$$

bestimmt sich, wie in Abbildung 9.1 dargestellt, aus den (als exakt bekannt vorausgesetzten) zurückliegenden Werten für $u_{k'}$ und den gemessenen Ausgangswerten $\tilde{y}_{k'}$.

Deutlicher wird der Zusammenhang mit dem bisher behandelten statischen LS-Problem, wenn man

$$\hat{y}_{k'} = -\hat{a}_{n-1}\tilde{y}_{k'-1} - \cdots - \hat{a}_0\tilde{y}_{k'-n} + \hat{b}_{n-1}u_{k'-1-d} + \hat{b}_{n-2}u_{k'-2-d} + \cdots + \hat{b}_0u_{k'-n-d}$$

als den Schätzwert auffasst, der sich anhand der Parameterschätzwerte \hat{a}_i und \hat{b}_i sowie den zurückliegenden Messwerten $\tilde{y}_{k'-1}, \dots, \tilde{y}_{k'-n}$ ergibt. Damit lautet der Schätzfehler

$$\hat{e}_{k'} = \tilde{y}_{k'} - \hat{y}_{k'},$$

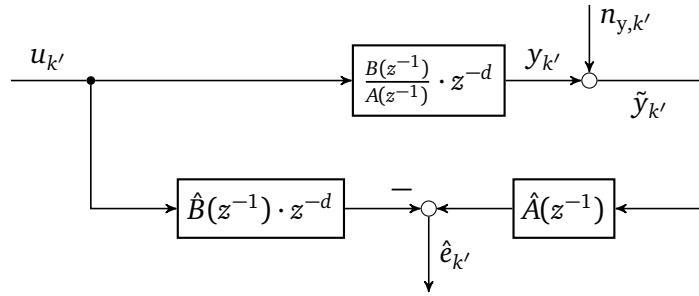


Abbildung 9.1: Gleichungsfehler

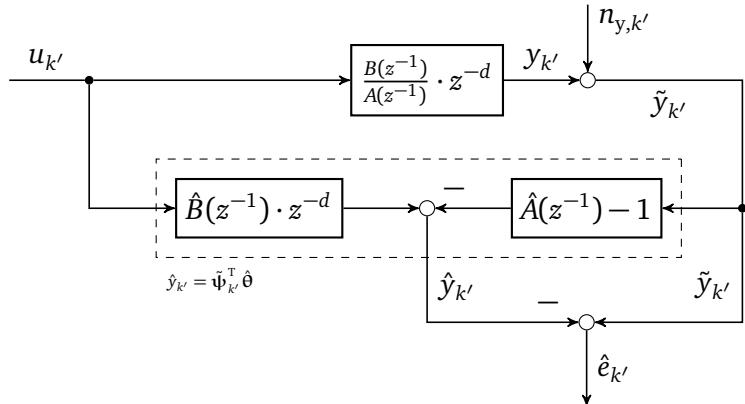


Abbildung 9.2: Alternative Darstellung des Gleichungsfehlers

wobei $\hat{y}_{k'}$ linear von den Parameterschätzwerten abhängt. Diese Struktur ist in Abbildung 9.2 verdeutlicht.

Damit kann durch die einfachen Definitionen

$$\tilde{\eta}_{k'} = \tilde{y}_{k'},$$

$$\tilde{\Psi}_{k'}^T = [-\tilde{y}_{k'-1} \ \cdots \ -\tilde{y}_{k'-n} \ u_{k'-1-d} \ \cdots \ u_{k'-n-d}]$$

sowie

$$\Theta = [a_{n-1} \ \cdots \ a_0 \ b_{n-1} \ \cdots \ b_0]^T$$

die Standardformulierung des LS-Problems (mit verrauschten Eingangsdaten) erhalten werden. Es gilt dann wieder

$$\hat{\eta}_{k'} = \tilde{\Psi}_{k'}^T \hat{\Theta},$$

und es ist das Ziel, den Fehler

$$\hat{e}_{k'} = \tilde{\eta}_{k'} - \hat{\eta}_{k'}$$

über k Auswertungen im quadratischen Sinne zu minimieren.

Die k Sätze von Mess- und Eingangswerten werden wieder zu dem Vektor bzw. Matrix

$$\tilde{\eta}_k = \begin{bmatrix} \tilde{y}_{k'_1} \\ \tilde{y}_{k'_2} \\ \vdots \\ \tilde{y}_{k'_k} \end{bmatrix} \quad \text{und} \quad \tilde{\Psi}_k = \begin{bmatrix} \tilde{\psi}_{k'_1}^T \\ \tilde{\psi}_{k'_2}^T \\ \vdots \\ \tilde{\psi}_{k'_k}^T \end{bmatrix} = \begin{bmatrix} -\tilde{y}_{k'_1-1} & \cdots & -\tilde{y}_{k'_1-n} & u_{k'_1-1-d} & \cdots & u_{k'_1-n-d} \\ -\tilde{y}_{k'_2-1} & \cdots & -\tilde{y}_{k'_2-n} & u_{k'_2-1-d} & \cdots & u_{k'_2-n-d} \\ \vdots & & \vdots & \vdots & & \vdots \\ -\tilde{y}_{k'_k-1} & \cdots & -\tilde{y}_{k'_k-n} & u_{k'_k-1-d} & \cdots & u_{k'_k-n-d} \end{bmatrix}$$

zusammengefasst. Die etwas sperrige Schreibweise mit k'_1, k'_2, \dots, k'_k bedeutet, dass die einzelnen Gleichungen nicht unbedingt zu zeitlich direkt hintereinander liegenden Zeitpunkten gehören müssen. (Meist wird man dies jedoch machen.) Für den Drei-Massen-Schwinger mit $n = 5$ und $d = 0$ ergibt sich z. B. (wenn das erste verfügbare Messwertpaar (u_0, \tilde{y}_0) ist)

$$\tilde{\eta}_k = \begin{bmatrix} \tilde{y}_5 \\ \tilde{y}_6 \\ \vdots \\ \tilde{y}_{5+k-1} \end{bmatrix} \quad \text{und} \quad \tilde{\Psi}_k = \begin{bmatrix} -\tilde{y}_4 & \cdots & -\tilde{y}_0 & u_4 & \cdots & u_0 \\ -\tilde{y}_5 & \cdots & -\tilde{y}_1 & u_5 & \cdots & u_1 \\ \vdots & & \vdots & \vdots & & \vdots \\ -\tilde{y}_{5+k-2} & \cdots & -\tilde{y}_{5+k-6} & u_{5+k-2} & \cdots & u_{5+k-6} \end{bmatrix}.$$

Der Parametervektor dazu lautet

$$\Theta = [a_4 \ a_3 \ a_2 \ a_1 \ a_0 \ b_4 \ b_3 \ b_2 \ b_1 \ b_0]^T.$$

Damit lautet die Lösung (formal)

$$\hat{\Theta}_k = (\tilde{\Psi}_k^T \tilde{\Psi}_k)^{-1} \tilde{\Psi}_k^T \tilde{\eta}_k.$$

Dieser Schätzwert garantiert den minimalen quadratischen Fehler bei den vorliegenden Messdaten, da diese Eigenschaft der LS-Lösung unabhängig von allen stochastischen Eigenschaften der involvierten Größen ist. Allerdings muss man sich bewusst machen, dass der Gleichungsfehler und nicht der Ausgangsfehler minimiert wird.

Insbesondere deswegen sind stochastische Aussagen wie Bias und Konsistenz der Schätzwerte wesentlich. Dies wird im Vergleich zu den bisher vorgestellten Aussagen etwas aufwendiger, da nicht mehr alle Annahmen erfüllt sind. Im Wesentlichen tritt das Problem auf, dass das auf die Ausgangswerte wirkende Rauschen mit den Eingangswerten¹ korreliert ist, sowie dass (je nach Sichtweise) der Eingangsvektor nicht exakt gemessen werden kann bzw. die Annahme A1 (das Zutreffen des LS-Prozessmodells) nicht erfüllt ist.

Zwar ist für jedes k' der Eingangsvektor $\tilde{\psi}_{k'}^T$ mit dem Messwert $\tilde{y}_{k'}$ unkorreliert, solange $\tilde{y}_{k'}$ (d. h. das Rauschen $n_{y,k'}$) in sich unkorreliert ist, da der aktuelle Messwert im Eingangsvektor nicht vorkommt. Allerdings ist die in A2 getroffene Annahme, $E\{\mathbf{n}_y | \Psi\}_r = \mathbf{0}$, weitreichender, da gefordert wird, dass jeder Rauschwert mit jedem Eingangsvektor unkorreliert ist. Also auch das Rauschen zum Zeitpunkt k' mit dem Eingangsvektor zu einem späteren Zeitpunkt $k' + x$.

Die Eingangsmatrix $\tilde{\Psi}$ ist hier mit einer Tilde markiert. Dazu ist anzumerken, dass aus Sicht des formulierten Gleichungsfehlers und damit des aufgestellten LS-Problems diese aber nicht unbedingt als gestört bzw. verrauscht anzusehen ist. Dies hängt davon ab, wie man den Gleichungsfehler auffasst.

Definiert man den Gleichungsfehler als den Unterschied zwischen dem gemessenen (und mit Rauschen gestörten) Ausgangswert $\tilde{y}_{k'}$ und dem Wert den man erhält, wenn man $y_{k'}$ mit den zurückliegenden

¹ Aus- und Eingangswerte des LS-Problems. Wenn von Ein- bzw. Ausgangswerten gesprochen wird, ist es aus dem Kontext zu erschließen, ob es sich um die Eingangs- bzw. Ausgangswerte des betrachteten dynamischen Systems ($u_{k'}$ und $y_{k'}$) oder des LS-Problems ($\psi_{k'}^T$ und $\eta_{k'}$) handelt.

Eingangswerten $u_{k'}$ und *gemessenen* (und nicht wahren) Ausgangswerten $\tilde{y}_{k'}$ bestimmt, dann ist der Eingangsvektor (per Definition) unverrauscht. Dafür ist die Annahme A1 nicht mehr erfüllt, da auch mit dem wahren Parametervektor $\Theta_{k'}$ der Zusammenhang $\eta_{k'} = y_{k'} = \tilde{\Psi}_{k'}^T \Theta$ im Allgemeinen nicht gelten wird. (Man kann nicht erwarten, dass – auch mit den richtigen Parametern – aus verrauschten zurückliegenden Daten der nächste, unverrauschte Ausgangswert korrekt bestimmt werden kann.)

Fasst man den Gleichungsfehler als (idealerweise) mit den wahren Ausgangswerten gebildet auf, so sind die entsprechenden Eingangswerte als verrauscht anzusehen. In diesem Fall liegt ein Error-In-Variables-Problem vor. Dafür ist die Annahme A1 erfüllt, da der Zusammenhang $\eta_{k'} = y_{k'} = \Psi_{k'}^T \Theta$ exakt gilt, wobei $\Psi_{k'}^T$ der mit den wahren Ausgangsdaten gebildete Eingangsvektor ist.

Letztlich macht es natürlich keinen Unterschied, welcher Interpretation man an dieser Stelle folgen möchte. Auf jeden Fall wird die Eingangsmatrix, die mit den verrauschten Ausgangswerten gebildet wird, hier mit einer Tilde geschrieben, um später Ψ ohne Tilde in naheliegender Weise für die Eingangsmatrix zu verwenden, die sich mit den exakten Ausgangswerten ergeben würde.

9.1.2 Stochastische Eigenschaften

Die Untersuchung des Bias geht von der Schätzgleichung

$$\hat{\theta}_k = (\tilde{\Psi}_k^T \tilde{\Psi}_k)^{-1} \cdot \tilde{\Psi}_k^T \tilde{\eta}_k$$

aus. Für den Messvektor $\tilde{\eta}_k$ kann auch

$$\tilde{\eta}_k = \Psi_k \Theta + \mathbf{n}_{\eta,k} \quad (9.1)$$

geschrieben werden, wobei Ψ_k die ungestörten Eingangswerte enthält (und damit $\Psi_k \Theta = \eta_k$ ist) und der Vektor

$$\mathbf{n}_{\eta,k} = [n_{y,k'_1} \quad \cdots \quad n_{y,k'_k}]^T$$

die Werte des Ausgangsrauschens der ersten k Messwerte² zusammenfasst.

Der Zusammenhang zwischen den gestörten und ungestörten Eingangswerten ist über

$$\tilde{\Psi}_k = \Psi_k + \mathbf{N}_{\psi,k} \quad (9.2)$$

gegeben, wobei

$$\mathbf{N}_{\psi,k} = \begin{bmatrix} -n_{y,k'_1-1} & \cdots & -n_{y,k'_1-n} & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ -n_{y,k'_k-1} & \cdots & -n_{y,k'_k-n} & 0 & \cdots & 0 \end{bmatrix}$$

die Rauschwerte zusammenfasst. Die Nullspalten gehören zu den Einträgen von $\tilde{\Psi}$, die aus den als ideal messbar angenommenen Eingangswerten $u_{k'}$ bestehen. In den Spalten, in denen in $\tilde{\Psi}$ die verschobenen, negierten, gemessenen Ausgangswerte stehen, enthalten hier die entsprechend verschobenen und negierten Werte des Ausgangsrauschens.

Mit Gl. (9.2) kann Gl. (9.1) auch als

$$\tilde{\eta}_k = (\tilde{\Psi}_k - \mathbf{N}_{\psi,k}) \cdot \Theta + \mathbf{n}_{\eta,k}$$

² Eigentlich der ersten k gemessenen Zusammenhänge. Entsprechend der angenommenen Systemordnung müssen mehr als k Messungen durchgeführt werden.

geschrieben werden. Setzt man diesen Ausdruck für $\tilde{\eta}$ in die Schätzgleichung ein, so erhält man

$$\begin{aligned}\hat{\theta}_k &= (\tilde{\Psi}_k^T \tilde{\Psi}_k)^{-1} \cdot \tilde{\Psi}_k^T \tilde{\eta}_k \\ &= (\tilde{\Psi}_k^T \tilde{\Psi}_k)^{-1} \cdot \tilde{\Psi}_k^T \cdot ((\tilde{\Psi}_k - N_{\psi,k}) \cdot \theta + n_{\eta,k}) \\ &= (\tilde{\Psi}_k^T \tilde{\Psi}_k)^{-1} \cdot \tilde{\Psi}_k^T \tilde{\Psi}_k \theta + (\tilde{\Psi}_k^T \tilde{\Psi}_k)^{-1} \cdot \tilde{\Psi}_k^T \cdot (n_{\eta,k} - N_{\psi,k} \theta).\end{aligned}$$

$(\tilde{\Psi}^T \tilde{\Psi})^{-1} \cdot \tilde{\Psi}^T \tilde{\Psi}$ ist für jede Realisierung die Einheitsmatrix, so dass der erste Summand gerade dem tatsächlichen Parametervektor θ entspricht. Damit ergibt sich für den Schätzfehler

$$\Delta \hat{\theta}_k = \hat{\theta}_k - \theta = (\tilde{\Psi}_k^T \tilde{\Psi}_k)^{-1} \cdot \tilde{\Psi}_k^T \cdot (n_{\eta,k} - N_{\psi,k} \theta).$$

Aufgrund der Korrelation zwischen $\tilde{\Psi}_k$ und $n_{\eta,k}$ sowie $N_{\psi,k}$ kann der Erwartungswert von $\{\Delta \hat{\theta}_k\}_r$ nicht bestimmt werden. Im Weiteren wird sich daher mit einer Analyse des asymptotischen Verhaltens zufrieden gegeben.

Dazu wird der Schätzfehler als

$$\Delta \hat{\theta}_k = \left(\frac{1}{k} \cdot \tilde{\Psi}_k^T \tilde{\Psi}_k \right)^{-1} \cdot \frac{1}{k} \cdot \tilde{\Psi}_k^T \cdot (n_{\eta,k} - N_{\psi,k} \theta) \quad (9.3)$$

geschrieben. Unter der Annahme, dass sowohl

$$\left\{ \left(\frac{1}{k} \cdot \tilde{\Psi}_k^T \tilde{\Psi}_k \right)^{-1} \right\}_r$$

als auch

$$\left\{ \frac{1}{k} \cdot \tilde{\Psi}_k^T \cdot (n_{\eta,k} - N_{\psi,k} \theta) \right\}_r$$

mindestens im schwachen Sinne konvergieren, kann nach dem Satz vom Slutsky der Wert bestimmt werden, gegen den $\{\Delta \hat{\theta}_k\}_r$ schwach konvergiert,

$$\operatorname{plim}_{k \rightarrow \infty} \{\Delta \hat{\theta}_k\}_r = \operatorname{plim}_{k \rightarrow \infty} \left\{ \left(\frac{1}{k} \cdot \tilde{\Psi}_k^T \tilde{\Psi}_k \right)^{-1} \right\}_r \cdot \operatorname{plim}_{k \rightarrow \infty} \left\{ \frac{1}{k} \cdot \tilde{\Psi}_k^T \cdot (n_{\eta,k} - N_{\psi,k} \theta) \right\}_r.$$

Dass die Annahme, dass $\frac{1}{k} \cdot \tilde{\Psi}_k^T \tilde{\Psi}_k$ für $k \rightarrow \infty$ gegen eine invertierbare Matrix konvergiert, sinnvoll ist und auch als Voraussetzung verwendet sollte, wurde schon im vorherigen Kapitel diskutiert.³ Somit wird jetzt

$$\frac{1}{k} \cdot \tilde{\Psi}_k^T \cdot (n_{\eta,k} - N_{\psi,k} \theta) \quad (9.4)$$

betrachtet. Die Differenz in der Klammer lautet ausgeschrieben

$$\begin{aligned}n_{\eta,k} - N_{\psi,k} \theta &= \begin{bmatrix} n_{y,k'_1} \\ \vdots \\ n_{y,k'_k} \end{bmatrix} - \begin{bmatrix} -n_{y,k'_1-1} & \cdots & -n_{y,k'_1-n} & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ -n_{y,k'_k-1} & \cdots & -n_{y,k'_k-n} & 0 & \cdots & 0 \end{bmatrix} \cdot \begin{bmatrix} a_{n-1} \\ \vdots \\ a_0 \\ b_{n-1} \\ \vdots \\ b_0 \end{bmatrix} \\ &= \begin{bmatrix} n_{y,k'_1} + a_{n-1}n_{y,k'_1-1} + \cdots + a_0n_{y,k'_1-n} \\ \vdots \\ n_{y,k'_k} + a_{n-1}n_{y,k'_k-1} + \cdots + a_0n_{y,k'_k-n} \end{bmatrix}.\end{aligned}$$

³ Welche Voraussetzungen sich damit für das Anregungssignal $u_{k'}$ dynamischer Prozesse ergeben, wird später behandelt.

Die Summen entsprechen gerade einer Filterung des Signales n_{y,k'_1} durch ein FIR-Filter $A(z^{-1})$. Wird der Filterausgang als $v_{k'}$ bezeichnet (siehe Abbildung 9.3a), dann entspricht die erste Summe gerade dem Wert dieses Ausgangssignals zum Zeitpunkt k'_1 ,

$$v_{k'_1} = n_{y,k'_1} + a_{n-1}n_{y,k'_1-1} + \cdots + a_0n_{y,k'_1-n} .$$

Entsprechend gibt die letzte Summe den Wert des Signals $v_{k'}$ zum Zeitpunkt k'_k an,

$$v_{k'_k} = n_{y,k'_k} + a_{n-1}n_{y,k'_k-1} + \cdots + a_0n_{y,k'_k-n} .$$

Dies kann man jedoch, was für später auch sinnvoller ist, andersherum auffassen. Und zwar, dass das Signal $n_{y,k'}$ durch ein über ein System $\frac{1}{A(z^{-1})}$ gefiltertes Signal $v_{k'}$ entsteht. Dies ist in Abbildung 9.3b dargestellt. Auch dann ergibt die erste Summe gerade $v_{k'_1}$ und die letzte $v_{k'_k}$.



Abbildung 9.3: Zusammenhang zwischen $n_{y,k'}$ und $v_{k'}$

Somit kann man

$$\mathbf{n}_{\eta,k} - \mathbf{N}_{\psi,k}\boldsymbol{\theta} = \begin{bmatrix} v_{k'_1} \\ \vdots \\ v_{k'_k} \end{bmatrix} = \mathbf{v}_k$$

schreiben. Welche Forderungen an den Zufallsprozess $\{v_{k'}\}_r$ sinnvoll sind, wird sich gleich ergeben.

Der zu untersuchende Term (9.4) kann damit auch als

$$\frac{1}{k} \cdot \tilde{\Psi}_k^T \cdot (\mathbf{n}_{\eta,k} - \mathbf{N}_{\psi,k}\boldsymbol{\theta}) = \frac{1}{k} \cdot \tilde{\Psi}_k^T \cdot \mathbf{v}_k = \frac{1}{k} \cdot \Psi_k^T \cdot \mathbf{v}_k + \frac{1}{k} \cdot \mathbf{N}_{\psi,k}^T \cdot \mathbf{v}_k \quad (9.5)$$

geschrieben werden.

Der ersten Faktor im ersten Summanden $\Psi_k^T \cdot \mathbf{v}_k$ beinhaltet die *ungestörten* Eingangs- und Ausgangsgrößen des Prozesses. Unabhängig davon, ob man diese nun als deterministische Größen oder als Zufallsprozess auffasst, ist die Forderung, dass $\{\mathbf{v}_k\}_r$ nicht mit diesen korreliert ist, sinnvoll und nicht einschränkend.⁴ Setzt man weiter voraus, dass $\{\mathbf{v}_k\}_r$ mittelwertfrei ist, also $E\{\mathbf{v}_k\}_r = \mathbf{0}$ gilt, so verschwindet der Erwartungswert des ersten Summanden,

$$E\{\Psi_k^T \cdot \mathbf{v}_k\}_r = \Psi_k^T \cdot E\{\mathbf{v}_k\}_r = \mathbf{0} .$$

(Dies auch für alle k , es ist kein Grenzübergang $k \rightarrow \infty$ nötig.)

Der zweite Summand der rechten Seite von Gl. (9.5) lautet ausgeschrieben

$$\frac{1}{k} \cdot \mathbf{N}_{\psi,k}^T \cdot \mathbf{v}_k = \frac{1}{k} \cdot \begin{bmatrix} -n_{y,k'_1-1} & \cdots & -n_{y,k'_k-1} \\ \vdots & & \vdots \\ -n_{y,k'_k-n} & \cdots & -n_{y,k'_k-n} \\ 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{bmatrix} \cdot \begin{bmatrix} v_{k'_1} \\ \vdots \\ v_{k'_k} \end{bmatrix} = -\frac{1}{k} \cdot \begin{bmatrix} \sum_{l=1}^k n_{y,k'_l-1} v_{k'_l} \\ \vdots \\ \sum_{l=1}^k n_{y,k'_l-n} v_{k'_l} \\ 0 \\ \vdots \\ 0 \end{bmatrix} .$$

⁴ Solange man nicht im geschlossenen Regelkreis identifizieren will.

In den auftretenden Summen treten immer nur Produkte von $n_{y,k'}$ und $v_{k'}$ auf, bei denen der Zeitpunkt von $n_{y,k'}$ mindestens um einen Schritt weiter zurückliegt als der Zeitpunkt des multiplizierten $v_{k'}$. Geht man jetzt davon aus, dass $\{v_{k'}\}_r$ in sich unkorreliert ist (z. B. weißes Rauschen), so sind damit auch die Faktoren in allen auftretenden Produkten unkorreliert, und man erhält z. B. für die erste Summe

$$E \left\{ \sum_{l=1}^k n_{y,k'_l-1} v_{k'_l} \right\}_r = \sum_{l=1}^k E\{n_{y,k'_l-1}\}_r \cdot E\{v_{k'_l}\}_r = 0.$$

Dies gilt für alle Summen, und somit ergibt sich

$$E\{\mathbf{N}_{\psi,k}^T \cdot \mathbf{v}_k\}_r = \mathbf{0}.$$

Damit gilt auch

$$E \left\{ \frac{1}{k} \cdot \tilde{\Psi}_k^T \cdot (\mathbf{n}_{\eta,k} - \mathbf{N}_{\psi,k} \theta) \right\}_r = \mathbf{0}. \quad (9.6)$$

Jetzt müsste noch überprüft werden, dass $\left\{ \frac{1}{k} \cdot \tilde{\Psi}_k^T \cdot (\mathbf{n}_{\eta,k} - \mathbf{N}_{\psi,k} \theta) \right\}_r = \left\{ \frac{1}{k} \cdot \tilde{\Psi}_k^T \cdot \mathbf{v}_k \right\}_r$ auch wenigstens schwach konvergiert. Dies könnte man z. B. machen, indem man die Kovarianzmatrix aufstellt und zeigt, dass diese für $k \rightarrow \infty$ gegen $\mathbf{0}$ geht. Damit wäre die Konvergenz im quadratischen Sinne nachgewiesen, welche die schwache Konvergenz einschließt. Darauf wird an dieser Stelle jedoch verzichtet. (Letztlich werden in $\tilde{\Psi}_k^T \cdot \mathbf{v}_k$ Summen von „harmlosen“ Zufallszahlen gebildet und dann durch k geteilt, also einfache Mittelwertbildungen vorgenommen. Die Varianz eines Mittelwertes sollte für $k \rightarrow \infty$ gegen null gehen.) Unter den genannten Voraussetzungen an $\{v_{k'}\}_r$ und damit $\{n_{y,k'}\}_r$ gilt also

$$\operatorname{plim}_{k \rightarrow \infty} \left\{ \frac{1}{k} \cdot \tilde{\Psi}_k^T \cdot (\mathbf{n}_{\eta,k} - \mathbf{N}_{\psi,k} \theta) \right\}_r = \mathbf{0} \quad (9.7)$$

und damit

$$\operatorname{plim}_{k \rightarrow \infty} \{\hat{\Delta\theta}_k\}_r = \mathbf{0}.$$

Der Schätzer ist also konsistent („praktisch“ asymptotisch biasfrei und die Varianz geht gegen Null), wenn $\{n_{y,k'}\}_r$ aus einem durch $\frac{1}{A(z^{-1})}$ gefilterten weißen Rauschen $\{v_{k'}\}_r$ entspricht.

Dieses spezielle Rauschmodell wird mit dem deterministischen Systemanteil zusammen als ARX-Modell bezeichnet, und ist in Abbildung 9.4 markiert. Dabei steht das „AR“ für „auto-regressive“ und bezieht sich auf den stochastischen Systemteil. Das „X“ steht für „eXogenous input“ und meint den deterministischen Eingang. (Im Kontext stochastischer Signale würde unter einem AR-Modell (ohne X) ein rein stochastisches Signalmodell wie z. B. in Abbildung 9.3b verstanden werden.)

Auch wenn Gl. (9.6) für alle k , und nicht nur für $k \rightarrow \infty$ gilt, so lässt sich hieraus nicht ableiten, dass der Schätzer biasfrei (und nicht nur asymptotisch biasfrei) ist. Dies liegt daran, dass für die Biasfreiheit nicht nur der zweite Faktor in Gl. (9.3) für sich einen Erwartungswert von null besitzen muss, sondern das der „komplette“ Erwartungswert von $\hat{\Delta\theta}_k$ (Gl. (9.3)) verschwinden muss. Und aufgrund der Korrelation der Faktoren kann der Erwartungswert nicht faktorweise bestimmt werden.

Dass die Biasfreiheit auf dem hier beschriebenen Wege nicht nachzuweisen ist, bedeutet jedoch nicht, dass der Schätzer (für endliche k) biasbehaftet sein muss. So sind in der Literatur Aussagen zu finden, dass bei einem vorliegenden ARX-Modell die Biasfreiheit gegeben ist. [Isermann und Münchhof 2011, S. 235]

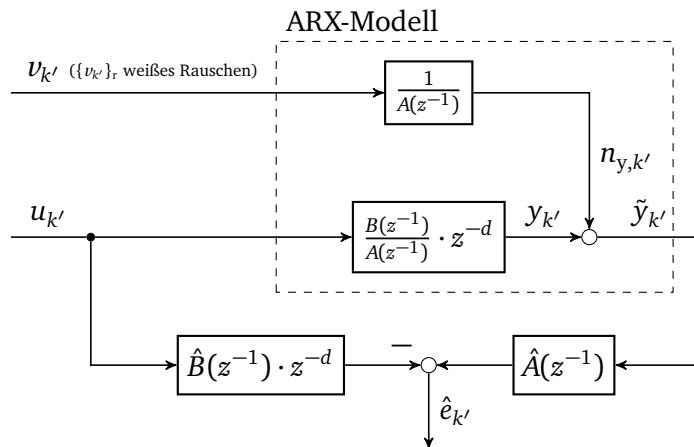


Abbildung 9.4: Gleichungsfehler mit ARX-Modell

Beispiel

Als Beispiel wird wieder das Modell des Drei-Massen-Schwingers betrachtet. Die Abtastzeit zur Identifikation wurde zu $T_s = 0,05$ s gewählt. Auf die – bei dieser Methode sehr wichtige – Wahl der Abtastzeit sowie zur Auswahl des Anregungssignals wird später noch ausführlicher eingegangen.

Mit der genannten Abtastzeit ergibt sich das zeitdiskrete Modell

$$G(z^{-1}) = \frac{0,0117z^{-1} + 0,2528z^{-2} + 0,5813z^{-3} + 0,2342z^{-4} + 0,01002z^{-5}}{1 - 2,429z^{-1} + 3,279z^{-2} - 3,175z^{-3} + 2,186z^{-4} - 0,7924z^{-5}} \quad (9.8)$$

für den Drei-Massen-Schwinger.

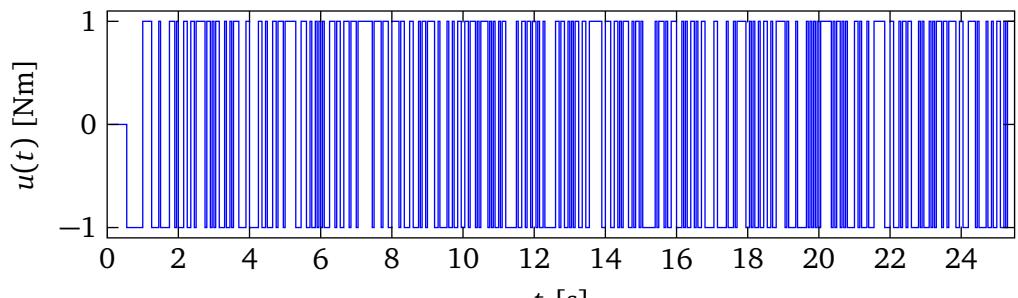
Das Eingangssignal für die folgenden Beispiele ist ein PRBS-Signal über ca. 25 s und in Abbildung 9.5a dargestellt.

Das Ausgangssignal, welches hier simulativ über die angegebene Übertragungsfunktion erzeugt ist, ist in Abbildung 9.5b gezeigt. Im gleichen Plot ist auch als ein Beispiel ein (diskretes) weißes Rauschen mit der Standardabweichung von $\sigma_y = 0,25$ rad/s gezeigt. Das (ungestörte) Ausgangssignal ist für die folgende Beispiele gleich, das Rauschen wird dagegen variiert.

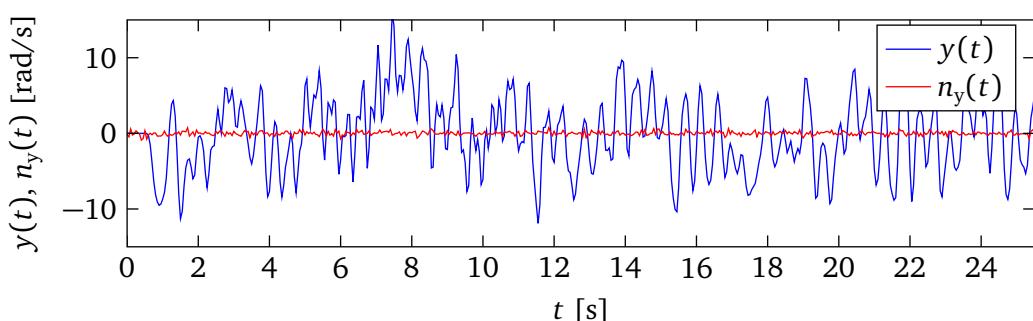
In Abbildung 9.6 sind die Ergebnisse gezeigt, wenn die Identifikation mit dem hier besprochenen Verfahren durchgeführt wird. Dabei sind in Abbildung 9.6a und (b) die geschätzten Koeffizienten der Übertragungsfunktion über den Zeitverlauf dargestellt. Die geschätzten Koeffizienten sind dabei jeweils auf die tatsächlichen Werte bezogen, so dass idealerweise ein Wert von \hat{a}_j/a_j bzw. \hat{b}_j/b_j von eins zu erreichen ist.

Prinzipiell zeigt sich zunächst ein Einschwingvorgang des Schätzers, der ungefähr die ersten 10 s (200 Schritte) benötigt. Es zeigen sich sowohl bei den geschätzten Nenner- als auch Zählerkoeffizienten auch für späte Zeitpunkte große Abweichungen von dem Idealwert von eins. Die Nennerkoeffizienten werden alle betragsmäßig unterschätzt. Bei den Zählerkoeffizienten werden b_3 und b_2 relativ gut geschätzt, b_1 wird betragsmäßig deutlich überschätzt. Die beiden Koeffizienten b_4 und b_0 werden extrem schlecht geschätzt und liegen meist außerhalb des dargestellten Bereichs. (Letzteres ist aber nicht zu überbewerten, da ein Blick auf die tatsächliche Übertragungsfunktion (9.8) zeigt, dass diese deutlich kleiner als die übrigen Zählerkoeffizienten sind. Damit haben sie für das Verhalten deutlich weniger Einfluss und sind leichter relativ falsch zu schätzen.)

Der Bias der Schätzwerte, den die Theorie bei einem weißen Rauschen vorhersagt, ist also deutlich zu erkennen.



(a) Eingangsgrößenverlauf



(b) Ausgangsgrößenverlauf und Messrauschen

Abbildung 9.5: Ein- und Ausgangsdaten zur Identifikation mit LS, weißes Rauschen, $\sigma_y = 0,25 \text{ rad/s}$

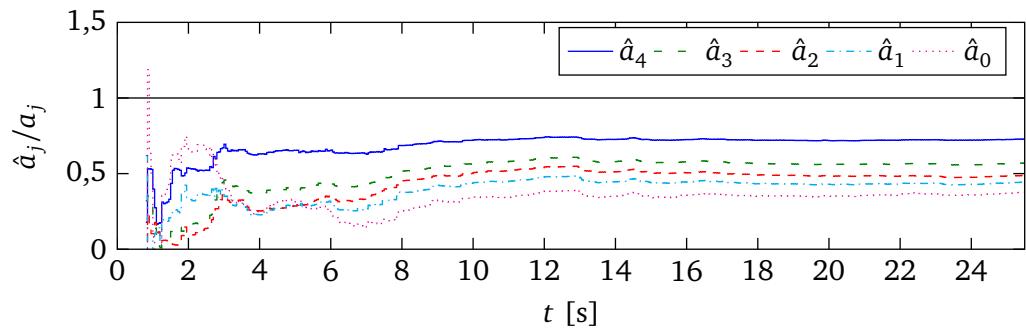
Aussagekräftiger als die reinen Koeffizientenwerte ist das Verhalten der damit beschriebenen Übertragungsfunktion. In Abbildung 9.6c ist der Amplitudengang der geschätzten Übertragungsfunktion $\hat{G}(z)$ mit dem der tatsächlichen Übertragungsfunktion $G(z)$ verglichen. Die jeweiligen Sprungantworten sind in Abbildung 9.6d gezeigt.

Es zeigt sich, dass das Systemverhalten mit den geschätzten Parametern nur sehr schlecht wiedergegeben wird.

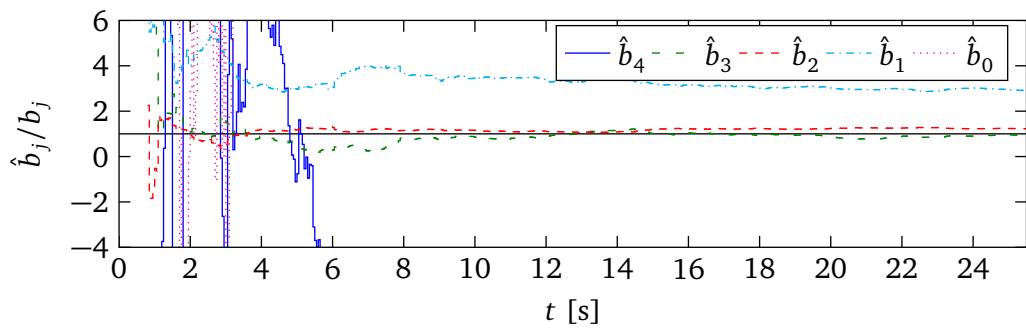
Man sollte sich an dieser Stelle auch bewusst machen, dass der Blick auf Abbildung 9.6d und der Vergleich der Sprungantworten, den man dabei automatisch vornimmt, einer Bewertung des Ausgangsfehlers entspricht. Das gleiche ist es, wenn man anhand der geschätzten Übertragungsfunktion und dem Eingangsgrößenverlauf u_k' die komplette Zeitlösung berechnet, und diese mit dem gemessenen Ausgangsgrößenverlauf \tilde{y}_k' vergleicht. Auch hierbei handelt es sich um die Bewertung des Ausgangsfehlers. Der Ausgangsfehler spielt für die Identifikation nach der hier besprochenen Methode aber keine Rolle, da der Gleichungsfehler optimiert wird.

Der Gleichungsfehler „schaut“ sich für jeden Zeitpunkt k' – im Fall des Drei-Massen-Schwingers – nur die letzten fünf Eingangswerte und die letzten fünf gemessenen Ausgangswerte an, und versucht die Parameter so einzustellen, dass (über alle k Auswertungen gesehen) der jeweils aktuelle Messwert damit möglichst gut erklärt werden kann. Daraus ist zumindest einsichtig, dass der Gleichungsfehler viel empfindlicher auf Rauschen reagiert als der Ausgangsfehler. Die Optimierung des Ausgangsfehlers bedeutet im Grunde, die Parameter so zu wählen, dass der (ohne Rückgriff auf gemessene Ausgangswerte) berechnete Ausgangsgrößenverlauf \hat{y}_k' möglichst „mittig“ durch die verrauschten Messdaten \tilde{y}_k' verläuft.

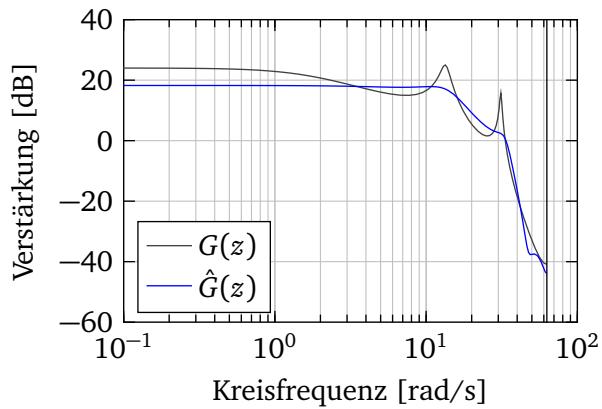
In Abbildung 9.7 sind die gleichen Plots wie in Abbildung 9.6 für den Fall dargestellt, dass die Standardabweichung des Rauschens halbiert ist.



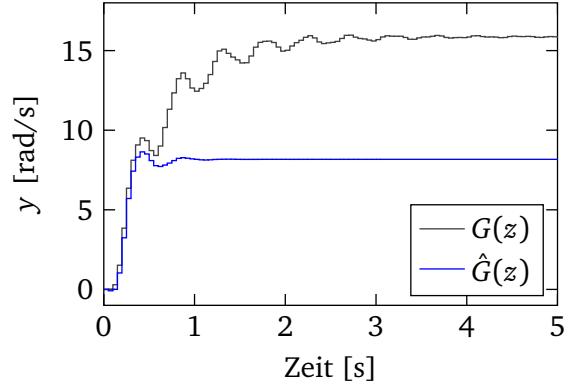
(a) Geschätzte Nennerkoeffizienten



(b) Geschätzte Zählerkoeffizienten

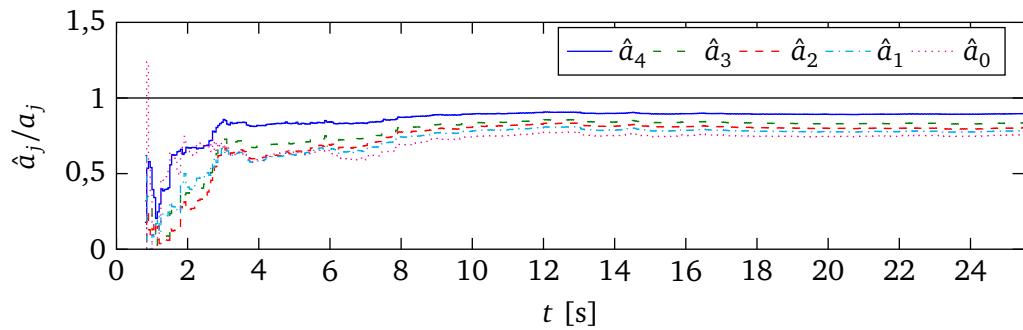


(c) Amplitudengang

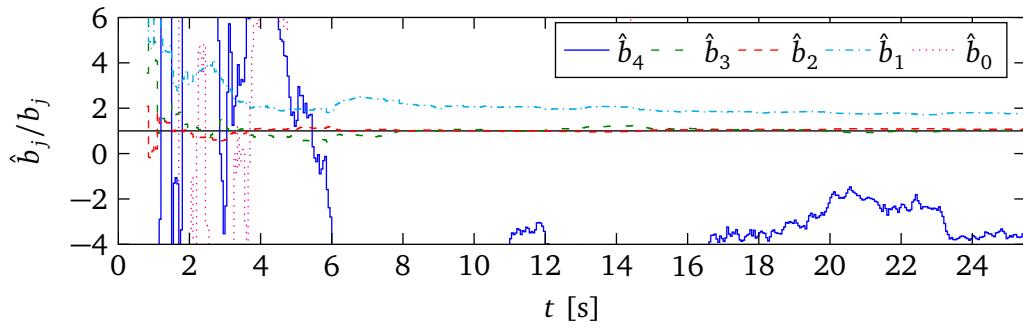


(d) Sprungantwort

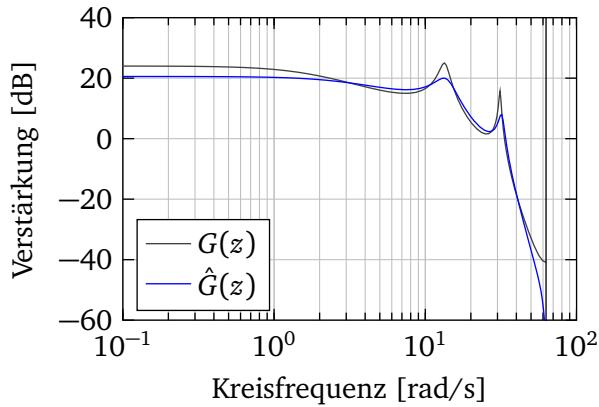
Abbildung 9.6: Identifikation mit LS, Störung weißes Rauschen, $\sigma_y = 0,25 \text{ rad/s}$



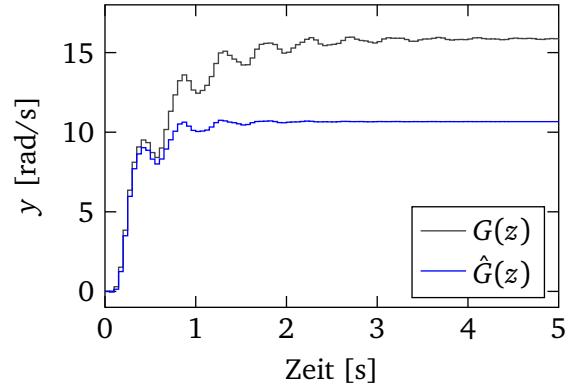
(a) Geschätzte Nennerkoeffizienten



(b) Geschätzte Zählerkoeffizienten



(c) Amplitudengang



(d) Sprungantwort

Abbildung 9.7: Identifikation mit LS, Störung weißes Rauschen, $\sigma_y = 0,125 \text{ rad/s}$

Der Bias in der Parameterschätzung ist noch deutlich erkennbar, jedoch im Vergleich zu dem stärkeren Rauschen reduziert. Der Amplitudengang sowie die Sprungantwort zeigen, dass auch das Systemverhalten jetzt zumindest in grober Näherung abgebildet wird.

Das dieses schlechte Schätzverhalten sich tatsächlich auf das Rauschen zurückführen lässt, ist in Abbildung 9.8 zu erkennen. Dort sind die geschätzten Parameterwerte gezeigt, die sich ergeben, wenn das Messrauschen $n_{y,k'}$ die Forderung erfüllt, durch weißes Rauschen und ein Formfilter $\frac{1}{A(z^{-1})}$ beschreibbar zu sein (ARX-Modell). Dabei ist die Standardabweichung auf $\sigma_y = 1,25 \text{ rad/s}$ gesetzt, d. h. das fünf-fache im Vergleich zu dem Rauschen welches in Abbildung 9.5b gezeigt ist und zu den Ergebnissen in Abbildung 9.6 geführt hat.

Die Parameterwerte schwingen sehr schnell auf die korrekten Werte ein. (Bis auf \hat{b}_4 und \hat{b}_0 , was aber aufgrund der schon genannten Gründe unkritisch ist.) Ein Vergleich der Systemantworten ist nicht dargestellt, es ist jedoch anhand der Parameterwerte schon erkennbar, dass das Systemverhalten fast perfekt abgebildet wird.

Wie extrem sich ein Rauschen, welches nicht dem ARX-Modell entspricht, auf die Schätzwerte auswirkt, ist nicht ohne Weiteres zu sagen. In Abbildung 9.9 ist der Fall dargestellt, dass das Rauschen durch ein weißes Rauschen und ein Filter $\frac{B(z^{-1})}{A(z^{-1})}$ erzeugt ist, wobei auch wieder die hohe Standardabweichung von $\sigma_y = 1,25 \text{ rad/s}$ verwendet ist. Dies entspricht einem weißen Eingangsrauschen des Systems. Es ist zu erkennen, dass dieses unkritisch wäre, obwohl es nicht der geforderten Form entspricht.

Alleine aus der Varianz oder dem „verrauscht sein“ der gemessenen Daten lässt sich also noch nicht sagen, wie gut oder schlecht die Ergebnisse der Identifikation sein werden.

Bei der Identifikation am realen Drei-Massen-Schwinger zeigt sich auch, dass die Identifikation mit dem LS-Verfahren dabei zu brauchbaren Ergebnissen führt. (Wobei die Validierung hierbei natürlich nicht anhand der bekannten, wahren Parameter erfolgen kann. Hierzu muss z. B. auf Bodediagramme, die mit Methoden der nicht-parametrischen Identifikation zu ermitteln sind, oder den Vergleich mit gemessenen Zeitantworten auf Testsignale⁵ zurückgegriffen werden.)

In Abbildung 9.10 ist noch der Verlauf der geschätzten Nennerkoeffizienten mit der rekursiven Formulierung der LS-Gleichung gezeigt, wobei einmal der Startwert $P'_0 = 1 \cdot I$ und im zweiten Fall der Startwert $P'_0 = 1000 \cdot I$ verwendet ist.

Der Einfluss des Faktors α in $P'_0 = \alpha \cdot I$ wird deutlich. Je größer α gewählt ist, desto schneller nähert sich die Lösung des rekursiven LS-Verfahrens dem normalen LS-Verfahren an.

9.1.3 Methode der Instrumentalvariablen

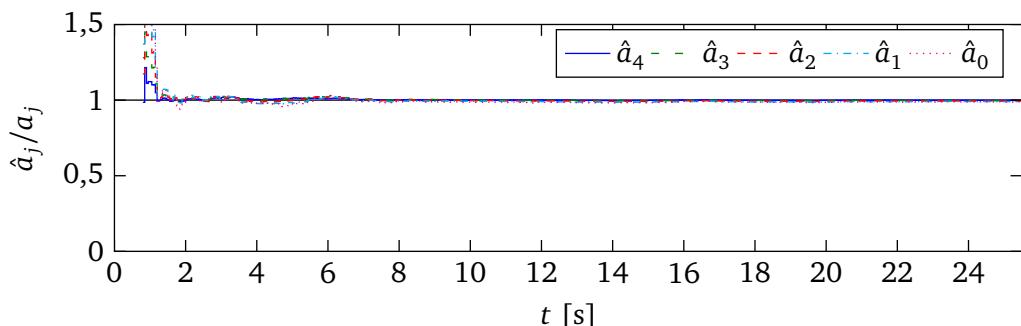
Die Einschränkung bezüglich des zulässigen Rauschens ist sehr groß. In kaum einen Fall wird ein Rauschen in genau dieser Form vorliegen. Daher sind in der Literatur eine Vielzahl von Methoden vorgeschlagen, die auch andere Rauschmodelle zulassen. Diese unterscheiden sich darin, welche Form das Rauschen besitzen darf. So können beispielsweise die Formfilter

$$\frac{1}{A(z^{-1}) \cdot F(z^{-1})}$$

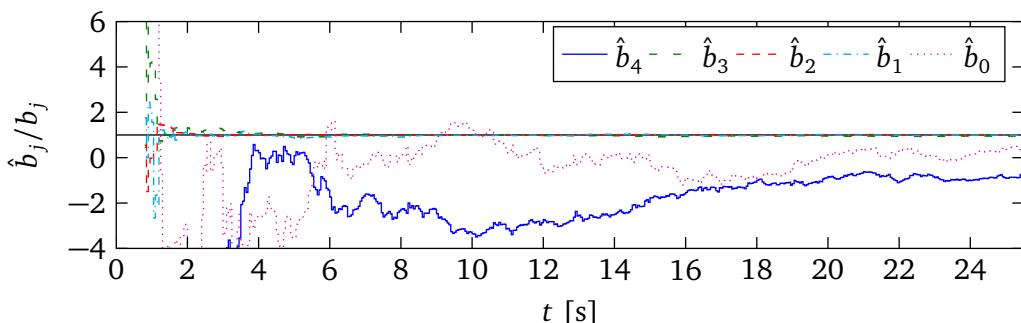
(Nenner muss Streckennennner enthalten, kein Zähler zulässig),

$$\frac{D(z^{-1})}{A(z^{-1})}$$

⁵ Wobei dann der Ausgangsfehler zugrundezulegen ist!

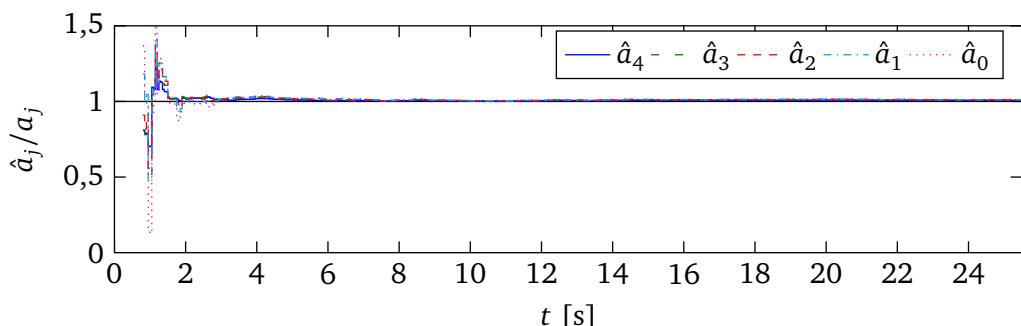


(a) Geschätzte Nennerkoeffizienten

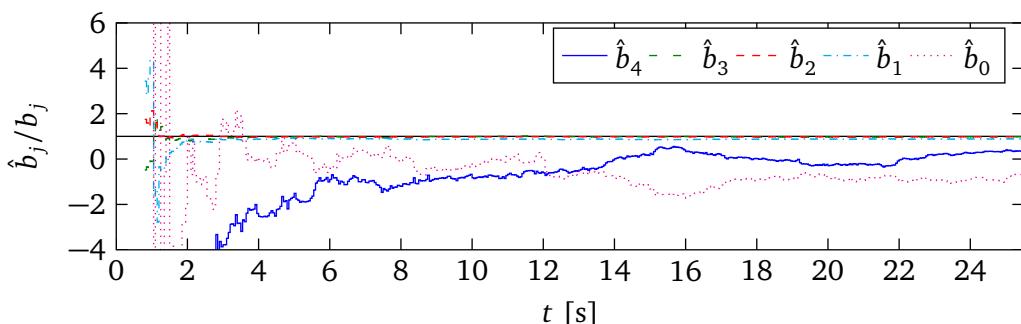


(b) Geschätzte Zählerkoeffizienten

Abbildung 9.8: Identifikation mit LS, ARX-Modell, $\sigma_y = 1,25 \text{ rad/s}$

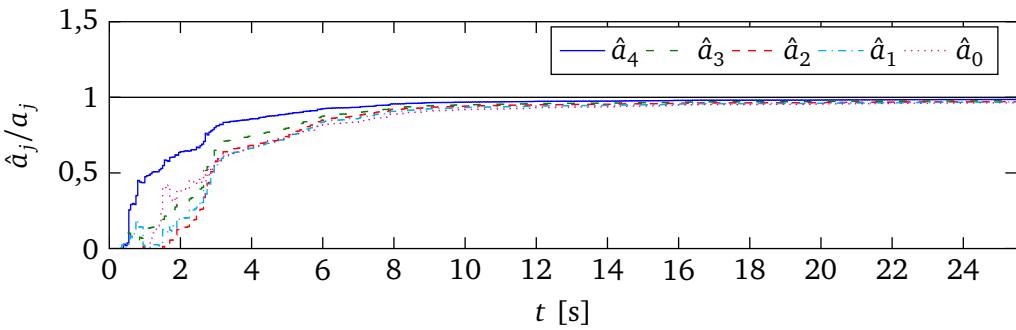


(a) Geschätzte Nennerkoeffizienten

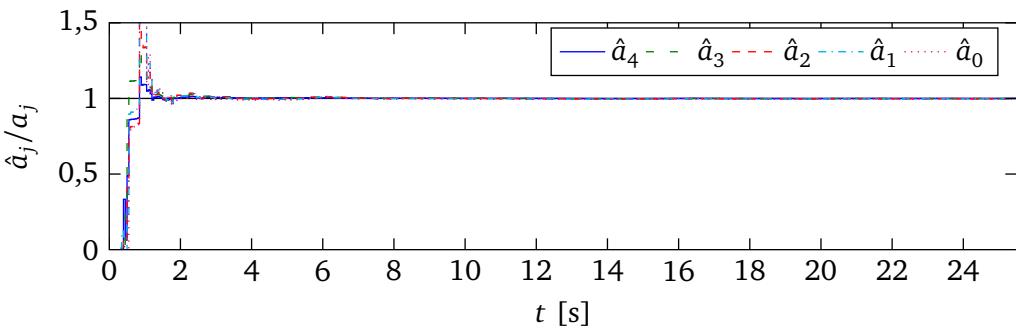


(b) Geschätzte Zählerkoeffizienten

Abbildung 9.9: Identifikation mit LS, Formfilter $B(z^{-1})/A(z^{-1})$, $\sigma_y = 1,25 \text{ rad/s}$



(a) Geschätzte Nennerkoeffizienten für $\alpha = 1$



(b) Geschätzte Nennerkoeffizienten für $\alpha = 1000$

Abbildung 9.10: Identifikation mit rekursivem LS, ARX-Modell, $\sigma_y = 1,25 \text{ rad/s}$

(Nenner muss Streckennenner entsprechen, Zähler aber beliebig) oder

$$\frac{D(z^{-1})}{C(z^{-1})}$$

(Rauschen unabhängig von Strecke) für das Rauschen zulässig sein. (Siehe [Isermann und Münchhof 2011, S. 581] für eine Übersicht. Die verschiedenen Verfahren sind z. B. in Kapitel 10 der genannten Quelle beschrieben.)

Hier soll im Weiteren mit der Methode der Instrumentalvariablen (im Deutschen zum Teil auch als „Hilfsvariablen“ bezeichnet) beispielhaft ein Verfahren vorgestellt werden. Dieses stellt auch keine besonderen Anforderungen an das Rauschen, d. h. sie lässt das allgemeinste Rauschmodell $\frac{D(z^{-1})}{C(z^{-1})}$ zu.

Einführung in die Methode

Die Grundüberlegung beginnt bei der ungestörten Gleichung

$$\eta_{k'} = \psi_{k'}^T \theta ,$$

die mit einem, im Grunde beliebigen, Vektor $\mathbf{z}_{k'}$ multipliziert wird. Der damit entstehende Zusammenhang

$$\mathbf{z}_{k'} \eta_{k'} = \mathbf{z}_{k'} \psi_{k'}^T \theta$$

ist natürlich exakt gültig.

Summiert man diesen Zusammenhang über k Paare $(\Psi_{k'}^T, \eta_{k'})$, wobei $\mathbf{z}_{k'}$ für jedes k' unabhängig gewählt werden kann, so erhält man

$$\underbrace{\frac{1}{k} \sum_{k'} \mathbf{z}_{k'} \eta_{k'}}_{\mathbf{r}'_k} = \underbrace{\left(\frac{1}{k} \sum_{k'} \mathbf{z}_{k'} \Psi_{k'}^T \right)}_{\mathbf{L}'_k} \cdot \boldsymbol{\theta}.$$

Die hier als \mathbf{L}'_k und \mathbf{r}'_k bezeichnete Matrizen enthalten nur ungestörte Größen und den Faktor $1/k$ und sind daher jeweils mit einem Strich gekennzeichnet.

Durch eine geeignete (bzw. eine nicht zu ungeschickte) Wahl von $\Psi_{k'}^T$ und $\mathbf{z}_{k'}$ ist zu erreichen, dass \mathbf{L}'_k invertierbar ist, und man kann den Zusammenhang nach

$$\boldsymbol{\theta} = \mathbf{L}'_k^{-1} \mathbf{r}'_k$$

umstellen.

Geht man von ideal bekannten $\Psi_{k'}^T$ aus, so besteht dennoch das Problem, dass nur die verrauschte Ausgangsgröße $\tilde{\eta}_{k'}$ messbar ist. Daher kann nur ein Schätzwert

$$\begin{aligned}\hat{\mathbf{r}}'_k &= \frac{1}{k} \sum_{k'} \mathbf{z}_{k'} \tilde{\eta}_{k'} \\ &= \frac{1}{k} \sum_{k'} \mathbf{z}_{k'} \eta_{k'} + \frac{1}{k} \sum_{k'} \mathbf{z}_{k'} n_{\eta, k'}\end{aligned}$$

für \mathbf{r}'_k bestimmt werden.

Werden die Vektoren $\mathbf{z}_{k'}$ jetzt so gewählt, dass

$$\operatorname{plim}_{k \rightarrow \infty} \left\{ \frac{1}{k} \sum_{k'} \mathbf{z}_{k'} n_{\eta, k'} \right\}_r = \mathbf{0}$$

gilt, so gilt auch

$$\operatorname{plim}_{k \rightarrow \infty} \{\hat{\mathbf{r}}'_k\}_r = \frac{1}{k} \sum_{k'} \mathbf{z}_{k'} \eta_{k'} = \mathbf{r}_k.$$

Man kann auch sagen, dass der Fehler $n_{\eta, k'}$ aus den Daten „wegkorreliert“ wird.⁶ Damit ist

$$\hat{\boldsymbol{\theta}}_k = \mathbf{L}'_k^{-1} \hat{\mathbf{r}}'_k$$

ein konsistenter, also sinnvoller Schätzer für die gesuchten Parameter. Schreibt man die Summen wieder in der Matrixschreibweise, wobei

$$\mathbf{Z}^T = [\mathbf{z}_1 \ \mathbf{z}_2 \ \cdots \ \mathbf{z}_k]$$

ist, so lautet die Schätzgleichung

$$\hat{\boldsymbol{\theta}}_k = (\mathbf{Z}_k^T \Psi_k)^{-1} \cdot \mathbf{Z}_k^T \tilde{\boldsymbol{\eta}}_k.$$

⁶ Vgl. [Ljung 1999, 562f].

(Wenn das Rauschen nicht mit den Eingangswerten korreliert ist und Z_k nicht mit dem Rauschen korreliert ist, also $E\{Z_k n_{\eta,k}\}_r = E\{Z_k\}_r \cdot E\{n_{\eta,k}\}_r$ gilt, dann ist dieser Schätzer biasfrei (und nicht nur konsistent), was sich wie beim normalen LS-Verfahren zeigen lässt.)

Wie sind nun die $z_{k'}$ zu wählen? Die Vektoren dürfen auf der einen Seite nicht mit dem Rauschen $\{n_{\eta,k'}\}_r$ korreliert sein. Zum anderen sollte die Korrelation mit den Eingangsvektoren $\Psi_{k'}^T$ möglichst hoch sein, um keine Nutzdaten „wegzukorrelieren“. Sind die Eingangsdaten und das Rauschen nicht korreliert, so liegt es damit nahe, für $z_{k'}$ gerade die Eingangsvektoren zu wählen,

$$z_{k'} = \Psi_{k'}.$$

Mit dieser Wahl erhält man dann wieder das normale LS-Verfahren. Andere Möglichkeiten $z_{k'}$ zu wählen sind aber dennoch zulässig! In diesem Fall führen sie aber zu schlechteren Varianzen der Schätzwerte, da das LS-Verfahren unter den gegebenen Voraussetzungen effizient bezüglich aller linearer biasfreier Schätzer ist.

Im Fall unkorrelierter Eingangsgrößen macht es daher keinen Sinn, etwas anderes als $Z_k = \Psi_k$ zu wählen. Anders sieht es aus, wenn die Eingangsgrößen mit dem Rauschen korreliert sind. Es kann dann ein Z_k gesucht werden, welches zwar mit den Eingangsdaten, aber nicht mit dem Rauschen korreliert ist. Genauer lauten die Bedingungen

$$\operatorname{plim}_{k \rightarrow \infty} \left\{ \frac{1}{k} \cdot Z_k^T n_{\eta,k} \right\}_r = \mathbf{0} \quad (9.9)$$

und

$$\det \left(\operatorname{plim}_{k \rightarrow \infty} \left\{ \frac{1}{k} \cdot Z_k^T \Psi_k \right\}_r \right) \neq 0. \quad (9.10)$$

Ein Z_k was diese Bedingungen erfüllt, wird als Instrumentalvariable bezeichnet.

Anwendung auf Identifikationsaufgabe

Bei der vorliegenden Identifikationsaufgabe liegt das Problem vor, dass die Eingangsdaten sowie das Rauschen korreliert sind. Darüber hinaus besteht auch das Problem, dass die Eingangsgrößen teilweise nur ebenfalls verrauscht gemessen sind. Letzteres wird hier aber durch den Ansatz der Instrumentalvariablen gleich mit erledigt, da das Messrauschen im Eingangsvektor sich aus dem gleichen Rauschprozess wie die Messung der Ausgangsgröße speist.

Da die Eingangsvektoren nur verrauscht gemessen werden können, also $\tilde{\Psi}_{k'}^T$ anstelle von $\Psi_{k'}^T$, kann auch L'_k nur geschätzt werden,

$$\hat{L}'_k = \frac{1}{k} \sum_{k'} z_{k'} \tilde{\Psi}_{k'}^T = \frac{1}{k} \sum_{k'} z_{k'} \Psi_{k'}^T + \frac{1}{k} \sum_{k'} z_{k'} n_{\psi,k'}^T.$$

Da die Einträge von $n_{\psi,k'}^T$ jedoch entweder aus Nullen oder aus Werten des Rauschprozesses $n_{\eta,k'}$ bestehen,

$$n_{\psi,k'}^T = [-n_{\eta,k'-1} \quad \cdots \quad -n_{\eta,k'-n} \quad 0 \quad \cdots \quad 0],$$

fallen auch diese durch die Korrelation weg, d. h. es gilt

$$\operatorname{plim}_{k \rightarrow \infty} \left\{ \frac{1}{k} \sum_{k'} z_{k'} n_{\psi,k'}^T \right\}_r = \mathbf{0}$$

und damit

$$\operatorname{plim}_{k \rightarrow \infty} \{\hat{\mathbf{L}}'_k\}_r = \frac{1}{k} \sum_{k'} \mathbf{z}_{k'} \boldsymbol{\Psi}_{k'}^T = \mathbf{L}'_k.$$

Findet man Instrumentalvariablen \mathbf{Z}_k , die die Bedingungen (9.9) und (9.10) erfüllen, so ist

$$\begin{aligned}\hat{\boldsymbol{\theta}}_k &= \hat{\mathbf{L}}'^{-1} \hat{\mathbf{r}}'_k \\ &= (\mathbf{Z}_k^T \tilde{\boldsymbol{\Psi}}_k)^{-1} \cdot \mathbf{Z}_k^T \tilde{\mathbf{n}}_k.\end{aligned}$$

ein konsistenter Schätzer für die gesuchten Parameter.

Wahl der Instrumentalvariablen – Mehrere Versuchsdurchführungen

Können mehrere Messungen mit gleichem Eingangsgrößenverlauf $u_{k'}$ (und gleichem Systemzustand zu Beginn) durchgeführt werden, dann bietet es sich an, die gemessene $\tilde{\boldsymbol{\Psi}}_k$ -Matrix eines anderen Versuchs als Instrumentalvariable zu verwenden. Der Nutzanteil ist dabei identisch, und das Rauschen nicht korreliert. [Pintelon und Schoukens 2001, S. 221] (Dort im Zusammenhang mit der Schätzung aus Frequenzbereichsdaten.)

Die Wirksamkeit dieser Methode ist an den in Abbildung 9.11 dargestellten Ergebnissen zu erkennen.

Der in Abbildung 9.5a gezeigte Eingangsgrößenverlauf ist dazu zweimal auf das sich jeweils in Ruhe befindliche System gegeben. Zu den Ausgangswerten ist ein weißes Rauschen mit der Standardabweichung von $\sigma_y = 0,25 \text{ rad/s}$ addiert.

Die Daten aus der zweiten Versuchsdurchführung sind als Instrumentalvariablen bei der Identifikation mit den Daten der ersten Durchführung verwendet. Die Ergebnisse sind mit Abbildung 9.6 zu vergleichen, da dort dasselbe Rauschen simuliert ist.

Es zeigt sich, dass die Koeffizienten zu Beginn schlechter geschätzt werden und sie langsamer konvergieren. Bis zum Ende der ungefähr 25 s (500 Auswertungen) konvergieren sie jedoch zu den tatsächlichen Werten. (Wieder mit Ausnahme von \hat{b}_4 und \hat{b}_0 .) Das Systemverhalten wird sehr gut wiedergegeben, was in Abbildung 9.11c und (d) zu erkennen ist.

Dieses Ergebnis bestätigt die vorhergesagte Konsistenz der Schätzung mit Instrumentalvariablen, zeigt aber auch, dass dafür die Varianz (die zwar gegen Null geht) für endliche k höher ist als bei dem normalen LS-Verfahren.

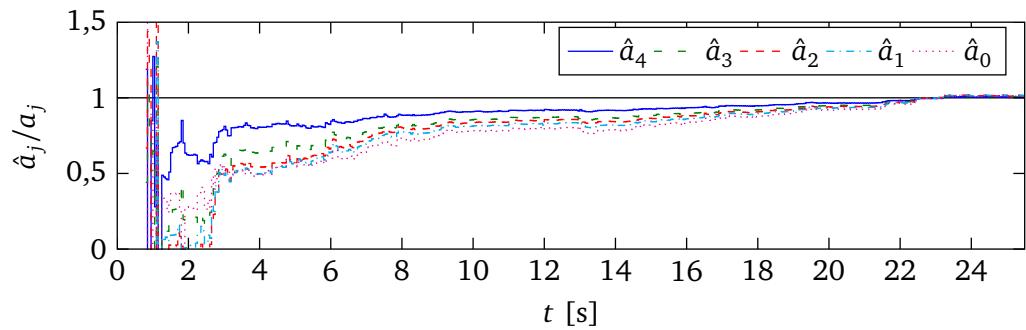
Wahl der Instrumentalvariablen – Eine Versuchsdurchführung

Ist es nicht möglich, mehrere gleiche Versuche durchzuführen, so wäre eine Möglichkeit zur Wahl der Instrumentalvariablen, anstelle der gemessenen Ausgangsgrößen $\tilde{y}_{k'}$ die ungestörten Eingänge $u_{k'}$ zu verwenden, da diese mit den ungestörten Ausgangsgrößen, nicht aber dem Rauschen korreliert sind. Dabei müssen diese um δ Schritte verschoben werden, da die Matrix \mathbf{Z}_k der Instrumentalvariablen vollen Spaltenrang besitzen muss. Es dürfen sich daher keine Spalten wiederholen. Damit erhält man

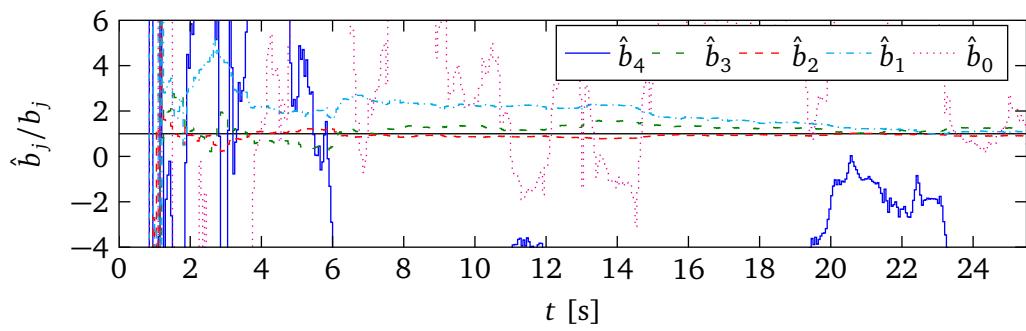
$$\mathbf{z}_{k'}^T = [u_{k'-1-\delta} \quad \cdots \quad u_{k'-n-\delta} \quad u_{k'-1-d} \quad \cdots \quad u_{k'-n-d}] \quad (9.11)$$

als Ansatz, wobei δ so gewählt werden sollte, eine möglichst gute Korrelation (möglichst große Einträge in $\mathbf{Z}^T \boldsymbol{\Psi}$) zu erhalten, ohne die Invertierbarkeit von $\mathbf{Z}^T \boldsymbol{\Psi}$ zu „zerstören“.

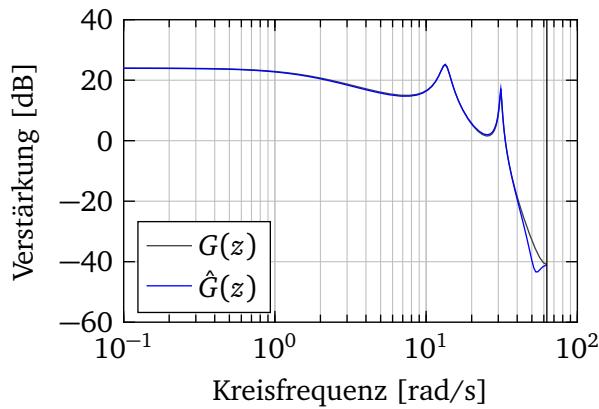
Weiter zurückliegende Eingangswerte sind aber natürlich weniger gut mit den Ausgangsgrößen korreliert als die Eingangsgrößen $u_{k'-d}$ bis $u_{k'-d-m}$. Eine Möglichkeit, Instrumentalvariablen mit einer besseren Korrelation zu erzeugen, ist es, Schätzwerte für die ungestörte Ausgangsgröße zu verwenden.



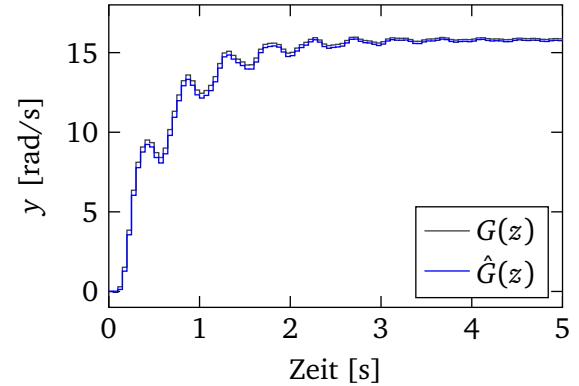
(a) Geschätzte Nennerkoeffizienten



(b) Geschätzte Zählerkoeffizienten



(c) Amplitudengang



(d) Sprungantwort

Abbildung 9.11: Identifikation mit Instrumentalvariablen, Störung weißes Rauschen, $\sigma_y = 0,25 \text{ rad/s}$

Dies bedeutet, dass mit einem Schätzwert $\hat{\theta}'_k$ für die Parameter, die über die Verwendung der Instrumentalvariablen nach Gl. (9.11) oder auch über ein einfaches LS-Verfahren bestimmt sind, Schätzwerte

$$\hat{y}'_{k'} = \Psi_{k'}^T \hat{\theta}'_k$$

für die Ausgangsgröße berechnet werden. Diese werden dann in

$$\mathbf{z}'_{k'} = [-\hat{y}'_{k'-1} \quad \cdots \quad -\hat{y}'_{k'-n} \quad u_{k'-1-d} \quad \cdots \quad u_{k'-n-d}]$$

als Instrumentalvariablen für einen weiteren Identifikationsvorgang (mit denselben Daten) verwendet. Dies kann dann iterativ wiederholt werden, bis sich die berechneten Schätzwerte nicht mehr (stark) ändern.

Basierend darauf gibt es auch eine rekursive Variante.

[Isermann und Münchhof 2011, 302ff]

9.1.4 Wahl der Abtastzeit und der Anregungsgröße

Im Vergleich zur in den früheren Kapiteln behandelte nichtparametrische Frequenzgangsschätzung oder die Optimierung des Ausgangsfehlers ergeben sich hier andere Anforderungen, die bei der Wahl der Abtastzeit und der Anregungsgröße zu berücksichtigen sind.

Bei der Frequenzgangsschätzung und bei der Optimierung des Ausgangsfehlers ist es prinzipiell besser bzw. schadet es nicht, eine möglichst kleine Abtastzeit zu wählen. (Wenn man nicht direkt einen zeitdiskreten Frequenzgang bzw. eine zeitdiskrete Übertragungsfunktion mit gegebener Abtastzeit erhalten will. Dann sollte diese gewählt werden.) Aus theoretischer Sicht spricht nichts gegen eine möglichst kleine Abtastzeit. Praktische Erwägungen, die die Abtastzeit letztlich begrenzen, sind die speicherbare Datenmenge, den Rechenaufwand und die Genauigkeit der Zahlendarstellung sowie möglicherweise die Zunahme des Rauschens des Messgerätes bei zunehmender Abtastfrequenz.

Bei der Identifikation über den Gleichungsfehler hingegen ist die richtige Abtastzeit wesentlich für eine gute Schätzung.

Bezüglich des Eingangssignals ist bei der Frequenzgangsschätzung sowohl zu Messbeginn als auch zu Messende zu beachten, dass bestimmte Bedingungen erfüllt sind. So muss sich das System (bezogen auf den Arbeitspunkt, um den identifiziert wird) entweder in Ruhe befinden, oder es ist bei periodischer Anregung darauf zu achten, dass das System sicher eingeschwungen ist und dass möglichst genau eine ganze Anzahl an Perioden gemessen wird.

Bei der Identifikation über die Optimierung des Ausgangsfehlers im Zeitbereich ist nur der Messbeginn zu beachten. Der Anfangszustand muss definiert und bekannt sein, damit die Simulation von dem gleichen Zustand ausgehen kann. Aufgrund der Kausalität kann die Messung aber zu jedem beliebigen Punkt beendet werden. Bei dem hier vorgestellten Gleichungsfehler enthält jeder Eingangsvektor aufgrund der zurückliegenden Ein- und Ausgangswerte alle Informationen über den Systemzustand um den Systemausgang des aktuellen Zeitpunktes zu schätzen. Die weitere Vorgesichte des Systems spielt keine Rolle mehr. Damit ist man – aus dieser Sicht – völlig frei und muss keine besonderen Bedingungen an den Start oder das Ende der Messung einhalten.

Bei der Frequenzgangsschätzung ist noch darauf zu achten, dass das Eingangssignal alle Frequenzen, für den man den Frequenzgang bestimmen möchte, auch ausreichend anregt. Dies ist bei der Optimierung des Ausgangsfehlers oder des Gleichungsfehlers nicht wesentlich. Wie noch gezeigt wird, ist es zwar wichtig, das System in den Frequenzbereichen anzuregen, in denen dessen Dynamik stattfindet, jedoch reichen dazu einzelne angeregte Frequenzen aus. D. h. das Anregungsspektrum muss nicht „glatt“ aussehen.

Abtastzeit

Allgemeine Überlegungen zur Wahl der Abtastzeit

Die weiter unten folgenden Überlegungen behandeln die Wahl der Abtastzeit nur aus Sicht des eigentlichen Identifikationsschritt. Auf andere Randbedingungen bei der Wahl wird nicht eingegangen, diese werden aber hier kurz erwähnt.

Zum einen legt das zur Verfügung stehende Stellglied die Abtastzeit fest. Man kann keine Systemdarstellung bezüglich einer Abtastzeit identifizieren, die kleiner als die des Stellgliedes ist. Langsamere Abtastzeiten als die des Stellgliedes sind jedoch möglich, sofern diese ein ganzes Vielfaches der Abtastzeit des Stellgliedes ist.

Zum anderen gibt die spätere Verwendung eine obere Grenze für die Abtastzeit vor. Soll beispielsweise ein Regler entworfen werden, so ergeben sich anhand der zu erreichenden Regelgüte implizit auch Anforderungen an die Abtastzeit, die nicht zu langsam sein darf. Man kann zwar kleinere Abtastzeiten wählen, die ein ganzer Teiler der Wunschabtastzeit sind, aber keine größeren.

Idealerweise würde man bezüglich der Systemdynamik eine möglichst schnelle Abtastzeit wählen. Dies führt auf ein diskretes Modell mit entsprechend kleiner Abtastzeit, aus dem man Modelle mit langsamerer Abtastzeit kleiner Auflösung bestimmen kann. Oder man geht von einer „quasikontinuierlichen“ Abtastung aus und erhält ein kontinuierliches Modell, welches wieder für beliebige Abtastzeiten diskretisiert werden kann.

Sollte zwischen diesen und den im Folgenden aufgezeigten Bedingungen an die Abtastzeit Widersprüche auftreten, so muss ein anderer Weg der Identifikation, z. B. über die nichtparametrische Frequenzgangsschätzung, gewählt werden.

Beispiel PT1

Zur Wahl der Abtastzeit wird als Beispiel die Identifikation eines PT_1 -Gliedes mit einer Zeitkonstante von $T = 1\text{ s}$ betrachtet. Zur Identifikation wird dieses mit einem Sprung angeregt, d. h.

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

Die ersten zwölf Sekunden der Sprungantwort werden zur Identifikation aufgezeichnet. Das Messsignal ist dabei mit einem Rauschen überlagert. Im Folgenden werden die drei Abtastzeiten $0,5\text{ s}$, $0,05\text{ s}$ und 5 s betrachtet.

In Abbildung 9.12a sind die mit $T_s = 0,5\text{ s}$ abgetasteten Daten gezeigt. Im Hintergrund ist das unverrauschte und verrauschte Ausgangssignal gezeigt. Für die abgetasteten Eingangswerte gilt $u_k \equiv 1$.

Mit diesen Daten werden die Koeffizienten der Übertragungsfunktion

$$\hat{G}(z) = \frac{\hat{b}_0 z^{-1}}{1 + \hat{a}_0 z^{-1}}$$

geschätzt.

In Abbildung 9.12b ist das Ergebnis der Schätzung anhand der Sprungantwort des sich ergebenden Systems gezeigt. Zum Vergleich ist noch die Sprungantwort des ursprünglichen Systems gezeigt. Das Systemverhalten wird relativ gut wiedergegeben.

In Abbildung 9.12c und (c) sind die entsprechenden Plots für eine Abtastzeit von $0,05\text{ s}$ dargestellt. Die Zeitkonstante ist hier deutlich zu klein geschätzt und die stationäre Verstärkung ist ebenfalls zu klein. Beträgt man den in Abbildung 9.12c gezeigten Ausschnitt der abgetasteten Messdaten, so erkennt man

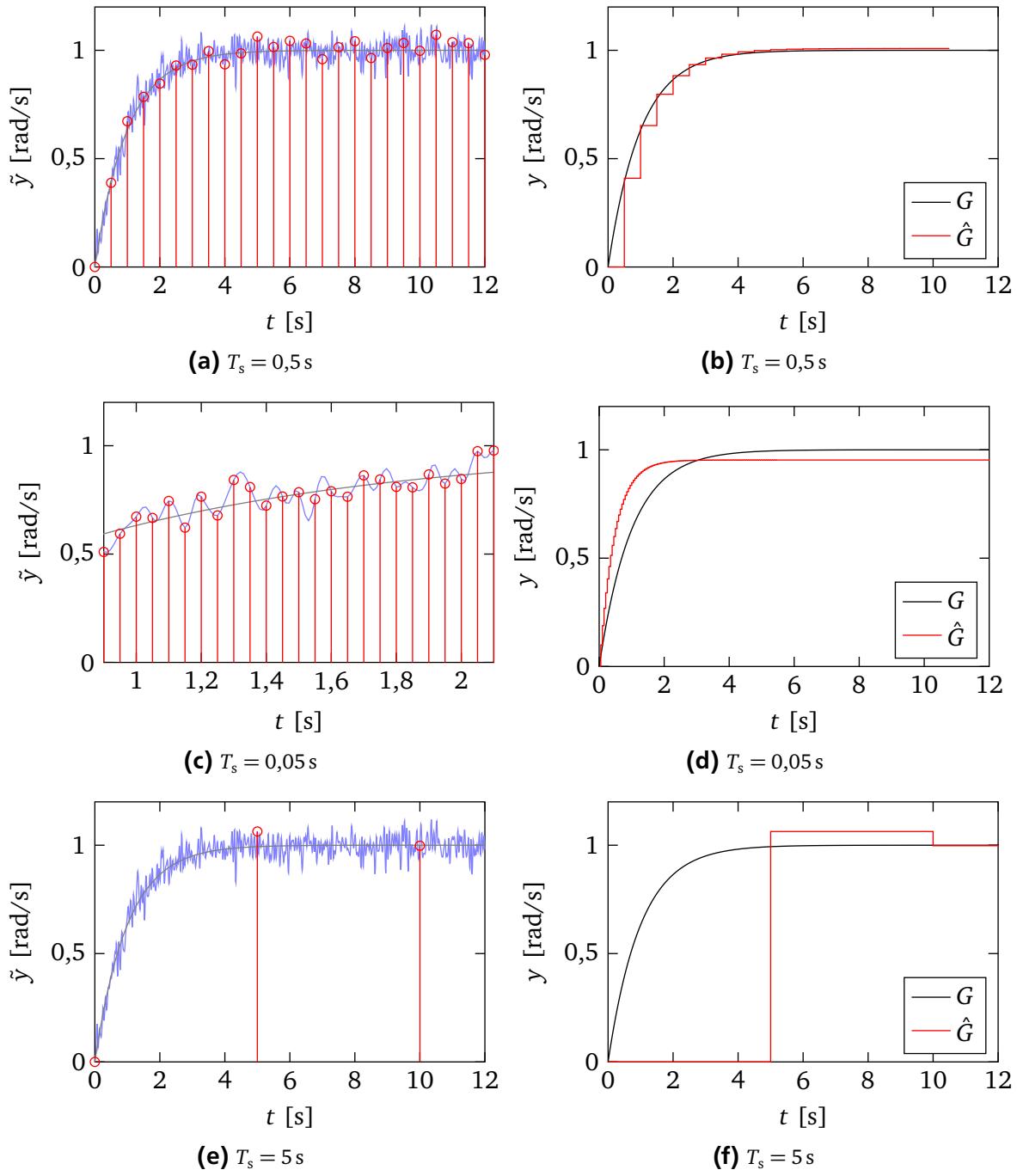


Abbildung 9.12: Einfluss verschiedener Abtastzeiten auf die Identifikation am Beispiel eines Systems mit PT_1 -Verhalten

die mit der kleinen Abtastzeit auftretende Problematik. Dabei muss man sich nochmals bewusst machen, was der Gleichungsfehler bewertet. Bei dem angesetzten System erster Ordnung schaut sich der Gleichungsfehler z. B. bei $t = 1,2\text{ s}$ den Messwert an, und bewertet den Fehler, den man (bei Kenntnis der Eingangsgröße bei $1,2\text{ s}$) dabei macht, den Ausgangswert bei $1,25\text{ s}$ vorherzusagen. In diesem Beispiel müsste für diese Einzelgleichung jedoch eine negative Systemverstärkung vorliegen. Auch müsste entweder die Zeitkonstante deutlich schneller als die wahre Systemzeitkonstante sein, oder die Systemverstärkung müsste betragsmäßig sehr hoch sein.

Solche Überlegungen lassen sich für jeden Zeitschritt anstellen, und meist wird man darauf kommen, dass für den jeweiligen Zeitschritt Parameterwerte optimal wären, die weit weg von den tatsächlichen liegen. Lediglich wenn das Rauschen die oben hergeleiteten Eigenschaften besitzt, wird sich dieser Effekt herausmitteln.

Man kann sich damit gut vorstellen, dass das Messrauschen umso mehr an Einfluss gewinnt, desto kürzer die Abtastzeit gewählt ist.

In Abbildung 9.12e und (e) ist noch der Fall mit einer Abtastzeit von 5 s gezeigt. Die Identifikation ergibt damit, dass das System im Wesentlichen eine Totzeit von einem Abtastschritt ist. Dies ist auch für die gewählte Abtastzeit in guter Näherung korrekt.

Es lässt sich also an diesem Beispiel festhalten, dass eine zu kurze Abtastzeit beim Gleichungsfehler schnell in Verbindung mit Rauschen (welches nicht der speziellen Bedingung genügt) zu völlig falschen Parameterwerten führen kann. Eine zu große Abtastzeit „verschluckt“ jedoch Dynamiken des Systems. Damit muss die Abtastzeit innerhalb eines gewissen Rahmens gefunden werden.

In diesem einfachen Beispiel ist dies noch einfach, da das System nur eine Zeitkonstante besitzt, dessen Größe sich aus der Betrachtung der Sprungantwort relativ gut abschätzen lässt. Bei etwas komplexeren Systemen, wie dem Drei-Massen-Schwinger ist dies aber nicht so offensichtlich. Die Zeitkonstanten des reellen Pols sowie die des langsameren konjugiert-komplexen Polpaars lassen sich in der Sprungantwort noch gut erkennen. Aber das schnellere Polpaar schon nicht mehr. Hat man schon einen Frequenzgang identifiziert, so kann man sich an diesem orientieren (oder aber gleich zur Identifikation des parametrischen Modells verwenden). Ansonsten bleibt zu prüfen, mit welchen Abtastzeiten man die Parameter schätzt, die auf Validierungsdaten angewandt das beste Ergebnis liefern.

Beispiel Drei-Massen-Schwinger

In Abbildung 9.13 ist der Einfluss der Abtastzeit bei der Identifikation des Drei-Massen-Schwingers an verschiedenen Identifikationsergebnissen gezeigt.

Die Taktzeit des PRBS ist für die ersten drei dargestellten Fälle $T_{ns} = 50\text{ ms}$. Bei der Abtastzeit $T_s = 100\text{ ms}$ ist die Taktzeit des PRBS ebenfalls (zwangsweise) auf 100 ms erhöht.

Es bestätigen sich die Ergebnisse, die aus dem Beispiel des PT_1 -Gliedes gewonnen wurden. Ausgehend von der Abtastzeit von 50 rad/s verschlechtert sich bei Verkleinerung der Abtastzeit das Identifikationsergebnis. Insbesondere bei der Abtastzeit von 10 ms hat das Ergebnis nicht mehr viel mit dem tatsächlichen System zu tun.

Eine Vergrößerung der Abtastzeit verbessert die Wiedergabe des tatsächlichen Systemverhalten im niedrigen Frequenzbereich, jedoch wird dabei nicht mehr die gesamte Systemdynamik erfasst.

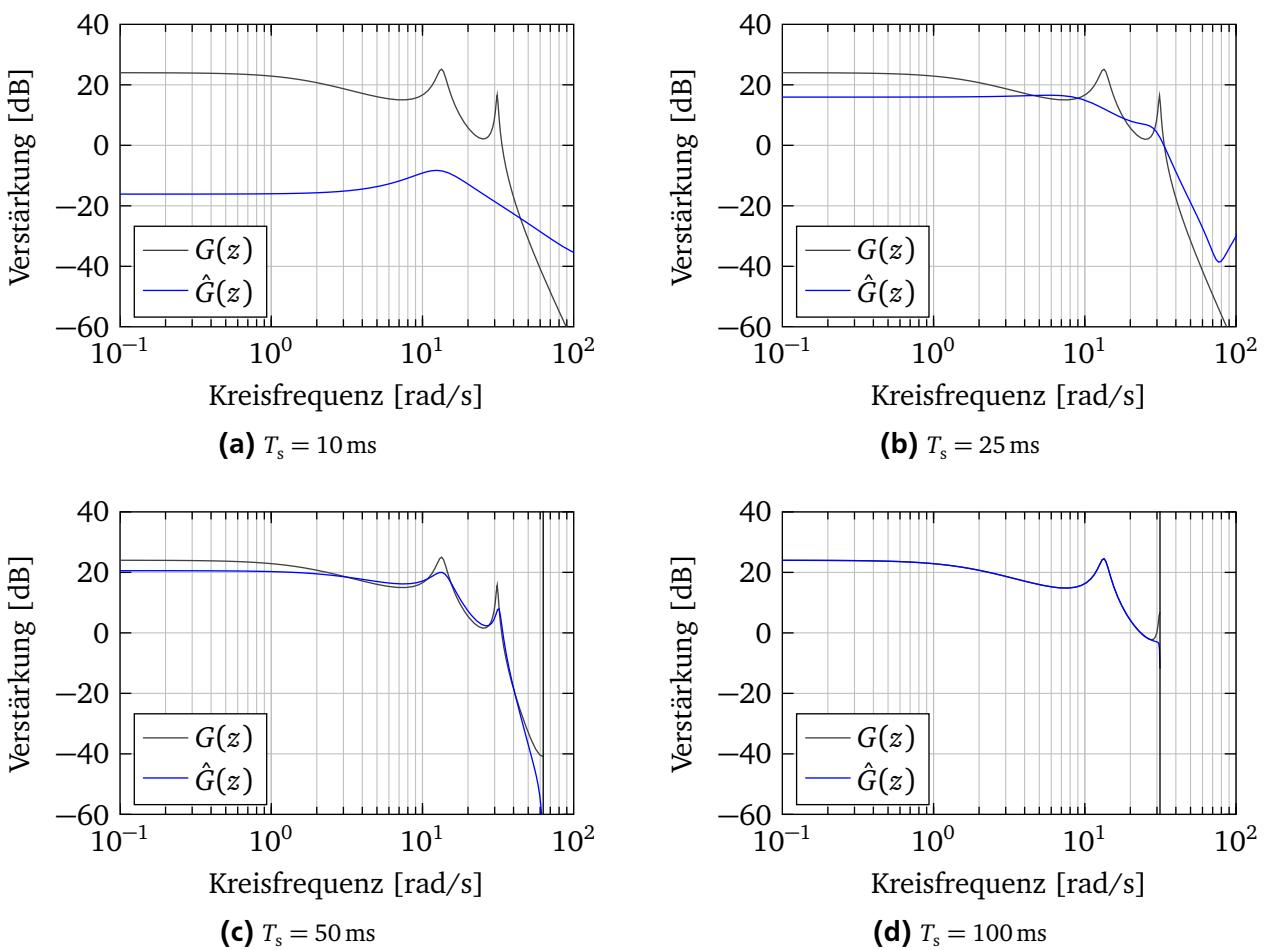


Abbildung 9.13: Wirkung verschiedener Abtastzeiten bei der Identifikation des Drei-Massen-Schwingers
(Identifikation mit LS, Störung weißes Rauschen, $\sigma_y = 0,125 \text{ rad/s}$)

Anregungsgröße

Es ist zur Identifikation sicherzustellen, dass die Matrix $\tilde{\Psi}$ den vollen Spaltenrang von p besitzt. Darüber hinaus soll sinnvollerweise

$$\operatorname{plim}_{k \rightarrow \infty} \left\{ \left(\frac{1}{k} \cdot \tilde{\Psi}_k^T \tilde{\Psi}_k \right)^{-1} \right\}_r$$

gelten, d. h. dass die Varianz der Schätzwerte für $k \rightarrow \infty$ gegen Null geht. (Aufgrund des Bias, der bei allgemeinem Rauschen entsteht, bedeutet dies hier nicht, dass der Schätzer konsistent ist.)

D. h. es ist ein Anregungssignal $u_{k'}$ zu verwenden, welches sicherstellt, dass alle Spalten von

$$\tilde{\Psi}_k = \begin{bmatrix} -\tilde{y}_{k'_1-1} & -\tilde{y}_{k'_1-2} & \cdots & -\tilde{y}_{k'_1-n} & u_{k'_1-1-d} & u_{k'_1-2-d} & \cdots & u_{k'_1-n-d} \\ -\tilde{y}_{k'_2-1} & -\tilde{y}_{k'_2-2} & \cdots & -\tilde{y}_{k'_2-n} & u_{k'_2-1-d} & u_{k'_2-2-d} & \cdots & u_{k'_2-n-d} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ -\tilde{y}_{k'_k-1} & -\tilde{y}_{k'_k-2} & \cdots & -\tilde{y}_{k'_k-n} & u_{k'_k-1-d} & u_{k'_k-2-d} & \cdots & u_{k'_k-n-d} \end{bmatrix} \quad (9.12)$$

linear unabhängig sind. Die Bedingung $\operatorname{plim}_{k \rightarrow \infty} \left\{ \left(\frac{1}{k} \cdot \tilde{\Psi}_k^T \tilde{\Psi}_k \right)^{-1} \right\}_r$ erfordert darüber hinaus, dass – anschaulich gesprochen – für $k \rightarrow \infty$ die Matrix $\tilde{\Psi}_k$ in eine mit k wachsende Anzahl an Teilblöcke (voller Breite, kleinerer Höhe) zerlegt werden kann, die alle für sich vollen Rang besitzen.

Eine Anregung $u_{k'}$, die dies sicherstellt (was neben dem Zeitverlauf von $u_{k'}$ auch von dem betrachteten System abhängt), wird als „ständige Anregung“ (persistent excitation) bezeichnet.

Darüber hinaus sollten die sich mit $\tilde{\Psi}$ ergebenden Varianzen der Schätzwerte natürlich möglichst gering sein.

Monofrequente, harmonische Anregung

Es soll als Anregungssignal eine abgetastete harmonische Schwingung

$$u_{k'} = u_0 \cdot \cos(\omega_u \cdot k \cdot T_s)$$

verwendet werden.

Damit stehen in dem zu den Eingangswerten gehörenden Teil von $\tilde{\Psi}$

$$\begin{bmatrix} u_{k'_1-1-d} & u_{k'_1-2-d} & \cdots & u_{k'_1-n-d} \\ u_{k'_2-1-d} & u_{k'_2-2-d} & \cdots & u_{k'_2-n-d} \\ \vdots & \vdots & & \vdots \\ u_{k'_k-1-d} & u_{k'_k-2-d} & \cdots & u_{k'_k-n-d} \end{bmatrix}.$$

Von oben nach unten gelesen steht in jeder Spalte eine (ggf. auch unregelmäßig abgetastete) harmonische Schwingung. Die Schwingungen der einzelnen Spalten besitzen dabei die gleiche Frequenz und Amplitude, aber eine unterschiedliche Phase.

Da aus zwei harmonischen Schwingungen gleicher Frequenz aber unterschiedlicher Phase (die Phasendifferenz darf kein Vielfaches von π sein) durch eine Verstärkung und Überlagerung jede beliebige harmonische Schwingung gleicher Frequenz erzeugt werden kann (d.h. es kann jede Amplitude und jede Phase erreicht werden), kann jede Spalte der angegebenen Teilmatrix aus zwei der anderen Spalten durch eine Linearkombination erzeugt werden. (Siehe Anhang E.3) Damit beträgt der Rang dieser Teilmatrix maximal zwei.

Dies gilt auch, wenn man im eingeschwungenen Zustand den Ausgang hinzunimmt, da dieser bei linearen Systemen auch eine harmonische Schwingung mit gleicher Frequenz (aber unterschiedlicher Amplitude und Phase) ist. D.h. auch die Matrix $\tilde{\Psi}$ (Gl. (9.12)) kann maximal einen Rang von zwei besitzen.

Damit können mit einer monofrequenten Anregung nur zwei Parameter, also ein PT_1 -Glied

$$G(z) = \frac{b_0 z^{-1}}{1 + a_0 z^{-1}}$$

identifiziert werden.

Dies ist auch von daher plausibel, da man mit einem Bodediagramm, welches nur einen einzigen Frequenzpunkt enthält, lediglich ein PT_1 -Glied konstruieren kann.

Breitbandige Anregung

Theoretisch benötigt man also für je zwei zu identifizierende Parameter eine anregende Frequenz. Praktisch wird man jedoch ohnehin Signale mit breitbandiger Anregung verwenden, die die oben angegebenen Bedingungen erfüllen. Letztlich wird schon durch einen einfachen Eingangssprung erreicht, dass $\tilde{\Psi}$ vollen Rang besitzt, bzw. eine Rechteckschwingung ergibt für $k \rightarrow \infty$ eine immer kleiner werdende Varianz. Allerdings sind diese Signale in der Regel nicht besonders günstig, was im Folgenden noch kurz diskutiert werden soll.

Dazu werden verschiedene Beispiele betrachtet. Zum einen das schon verwendete PRBS, wobei hier auch der Einfluss der Taktzeit T_{ns} betrachtet wird. Daneben werden der normale und „binäre“ Gleitsinus (letzteres meint einen Gleitsinus mit nachgeschalteter Signumfunktion) sowie eine einfache Rechteckschwingung betrachtet. Die Ergebnisse sind in Abbildung 9.14 bzw. Tabelle 9.1 gezeigt.

Bei den hier gezeigten Beispielen ist die Varianz (abgesehen von \hat{b}_0 und \hat{b}_4) mit der gewählten Messzeit kein sonderliches Problem. D. h. eine Verlängerung der Messzeit durch Wiederholung der Eingangssignale führt bei keinem der gezeigten Eingangssignale zu einer großen Änderung der geschätzten Parameter. Die Abweichungen der Parameter von den wahren Werten, die sich in den Bode-Plots zeigt, ist im Wesentlichen auf den Bias zurückzuführen. (Würde man die Bedingungen des ARX-Modells erfüllen, so ergäbe sich tatsächlich in allen Fällen der korrekte Verlauf, mit leichten Streuungen.) Das bedeutet aber auch, dass der Bias nicht nur vom Rauschen, sondern auch von dem gewählten Eingangssignal abhängt.

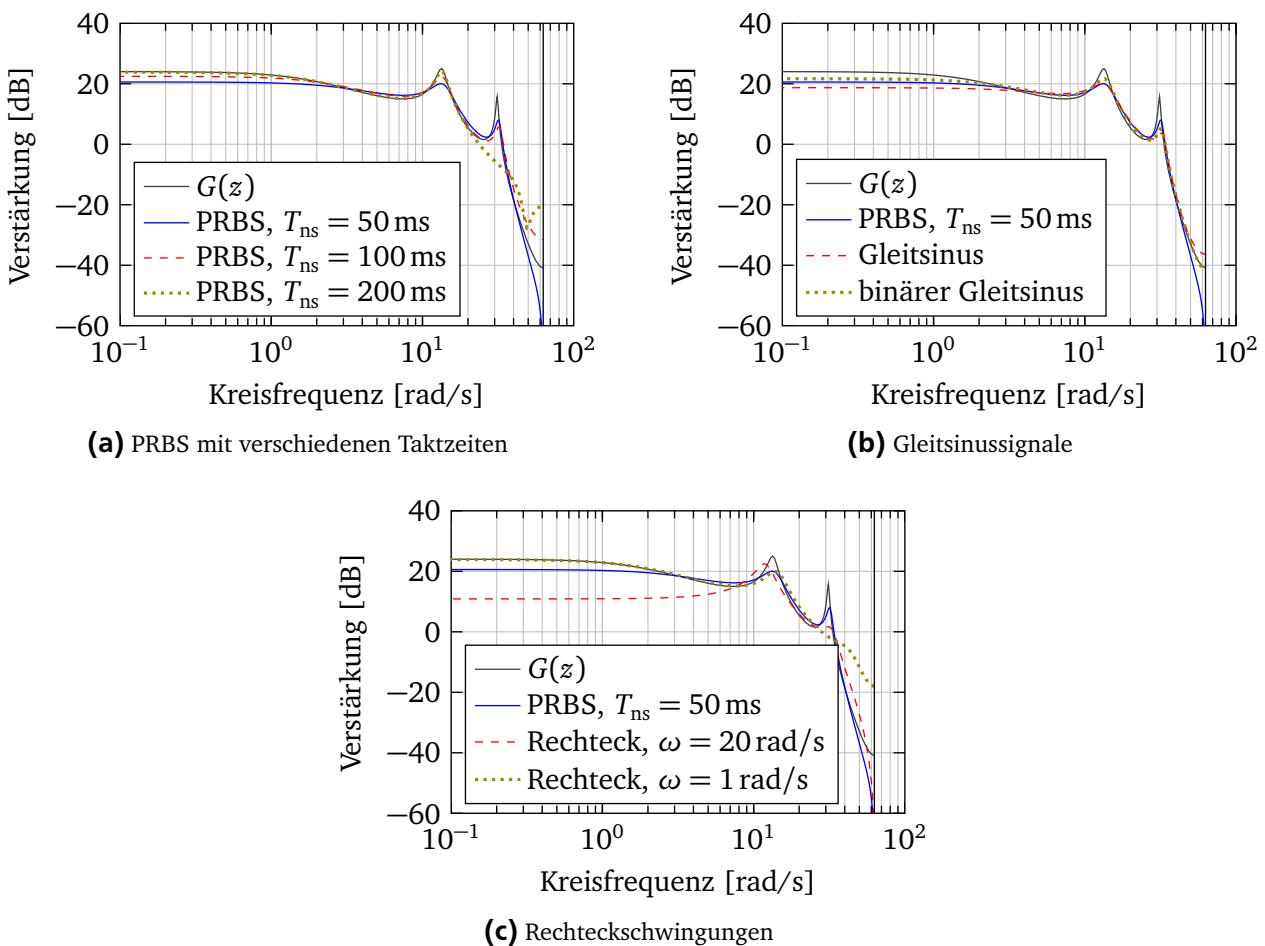


Abbildung 9.14: Wirkung verschiedener Anregungssignale bei der Identifikation des Drei-Massen-Schwingers
(Identifikation mit LS, Störung weißes Rauschen, $\sigma_y = 0,125 \text{ rad/s}$)

Bei binären Signalen werden die Eingänge immer nur kurz, dafür sehr stark angeregt. Beim Drei-Massen-Schwinger ist ein Sprung der Eingangsgröße innerhalb von fünf Abtastschritten „durch“. Die nächsten Schritte beinhalten nur noch Informationen über die stationäre Verstärkung und das Eigenverhalten des Systems, aber keine mehr über die Nullstellen des Systems.

Bei einem normalen Gleitsinus hingegen werden die Eingänge durchgehend angeregt, jedoch deutlich schwächer. Die Ergebnisse dieser Simulationen legen nahe, dass die starken, wenigen Sprünge dabei leichte Vorteile gegenüber der langsameren, aber ständigen Änderung der Eingangsgröße besitzen.

Tabelle 9.1: Relative Fehler

	\hat{a}_4	\hat{a}_3	\hat{a}_2	\hat{a}_1	\hat{a}_0	\hat{b}_3	\hat{b}_2	\hat{b}_1
PRBS, $T_{ns} = 50$ ms	-10 %	-16 %	-20 %	-22 %	-24 %	3 %	6 %	76 %
PRBS, $T_{ns} = 100$ ms	-8 %	-15 %	-20 %	-20 %	-18 %	-25 %	19 %	29 %
PRBS, $T_{ns} = 200$ ms	-35 %	-77 %	-98 %	-85 %	-55 %	84 %	19 %	60 %
Gleitsinus	-9 %	-17 %	-22 %	-22 %	-22 %	-4 %	18 %	93 %
Binärer Gleitsinus	-9 %	-17 %	-22 %	-22 %	-22 %	8 %	12 %	51 %
Rechteck, $\omega = 20$ rad/s	-38 %	-54 %	-65 %	-68 %	-92 %	-36 %	-16 %	225 %
Rechteck, $\omega = 1$ rad/s	-35 %	-73 %	-93 %	-85 %	-62 %	-18 %	45 %	210 %

9.1.5 Zeitvariante Systeme

In manchen Anwendungen ist es nötig, die Parameter eines zeitvarianten Systems im Zeitverlauf zu schätzen, z. B. um diese in adaptiven Regelungen zu verwenden oder um Fehler zu erkennen. Dabei bietet sich die (rekursive) LS-Schätzung mit einem Vergessensfaktor oder die Schätzung über ein Kalman-Filter an.

Durch einen Vergessensfaktor $\lambda < 1$ werden die weiter zurückliegende Messungen immer schwächer gewichtet und die Schätzung damit von den jüngeren Daten dominiert. Anhand der rekursiven Schreibweise lässt es sich auch so interpretieren, dass die Varianz durch den Vergessensfaktor künstlich vergrößert wird, womit die aktuellen Messdaten den Schätzwert stärker beeinflussen.

Die beiden Effekte, dass die andauernde Anregung die Varianz vermindert und der Vergessensfaktor diese erhöht, arbeiten dabei gegeneinander. Je nach Anregung und Vergessensfaktor stellt sich damit eine gewisse Streuung in den Schätzwerten ein. Problematisch wird es, wenn die Anregung ausbleibt, z. B. weil sich der Prozess länger in einem stationären Zustand befindet. Dadurch steigen die „Varianzen“ in \mathbf{P}' und kleines Rauschen auf den Eingängen wirkt sich sehr stark aus. Sollte dies auftreten können, so sollte nicht durchgehend mit dem gleichen λ geschätzt werden. So kann entweder λ manuell auf 1 gesetzt und das Vergessen für die Phasen unzureichender Anregung ausgeschaltet werden, oder man ignoriert Messungen, die in solche Phasen fallen, vollständig.

Ein Vorgehen, welches den Vergessensfaktor automatisch anpasst, beruht darauf, diese Anpassung anhand des „a posteriori“-Fehler vorzunehmen, d. h. die Differenz des aktuellen Messwerts von der mit *nach* dem Update vorhandenen Schätzwerten $\hat{\theta}_{k+1}$, also $\tilde{y}_{k+1} - \psi_{k+1}^T \hat{\theta}_{k+1}$. [Isermann und Münchhof 2011, S. 340f]

Wenn bekannt ist, zu welchen Zeiten eine Änderung der Parameter wahrscheinlich ist, dann kann auch die „Kovarianzmatrix“, \mathbf{P}' direkt manipuliert werden. Dadurch kann (auch gezielt einzelnen Parametern) eine höhere Varianz bzw. Unsicherheit zugesprochen werden, was eine stärkere Korrektur dieser Parameter zur Folge hat. Dieses Konzept wird im Zusammenhang mit dem Kalman-Filter nochmals genauer betrachtet.

Beispiel

Als Beispiel wird folgender Fall betrachtet. Ab 7,5 s entspricht das Modell wieder dem Standardmodell. Bis zur Zeit von 7,5 s ist die Federsteifigkeit der zweiten Drehfeder halbiert. Die Identifikation wird mit demselben Eingangssignal wie bisher durchgeführt. Das Rauschen ist durch $\frac{1}{A(s)}$ gefiltertes weißes Rauschen, wobei $A(s)$ der Nenner der Übertragungsfunktion des Standardmodells ist. D. h. ab 7,5 s liegt der ARX-Fall vor, vorher nicht.

In Abbildung 9.15 ist das Ergebnis gezeigt, wenn dennoch eine normale LS-Schätzung (hier in der rekursiven Formulierung mit $P'_0 = 1000 \cdot I$) über alle Daten durchgeführt wird.

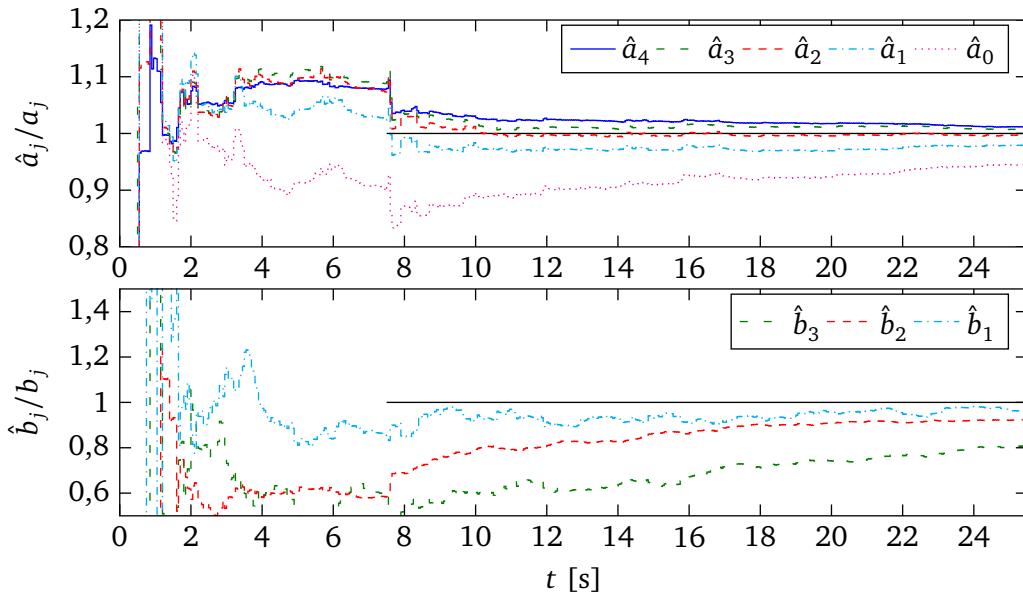


Abbildung 9.15: Schätzung mit $\lambda = 1$, Parameteränderung bei 7,5 s

In den Ergebnisplots sind die geschätzten Koeffizienten immer bezogen auf die Koeffizienten des Standardmodells angegeben. (Die Koeffizienten des diskretisierten Modells mit halber Federsteifigkeit weichen bis zu 25 % (Nenner) und 50 % (Zähler) von denen des Standardmodells ab.) D.h. in den ersten 7,5 s ist sich nicht an dem Referenzwert von 1 zu orientieren. Dazu kommt, dass die Schätzung innerhalb der ersten 7,5 s auch biasbehaftet ist.

Die Schätzwerte für b_4 und b_0 sind nicht dargestellt, da diese für den gezeigten Ausschnitt der Ordinatenachse zu sehr streuen.

Man erkennt in Abbildung 9.15, dass sich die Schätzwerte bei 7,5 s zunächst deutlich ändern, allerdings nähern sich die bezogenen Schätzwerte dann nur langsam dem Wert eins an. Dies liegt daran, dass die bis 7,5 s aufgezeichneten Werte immer voll berücksichtigt werden.

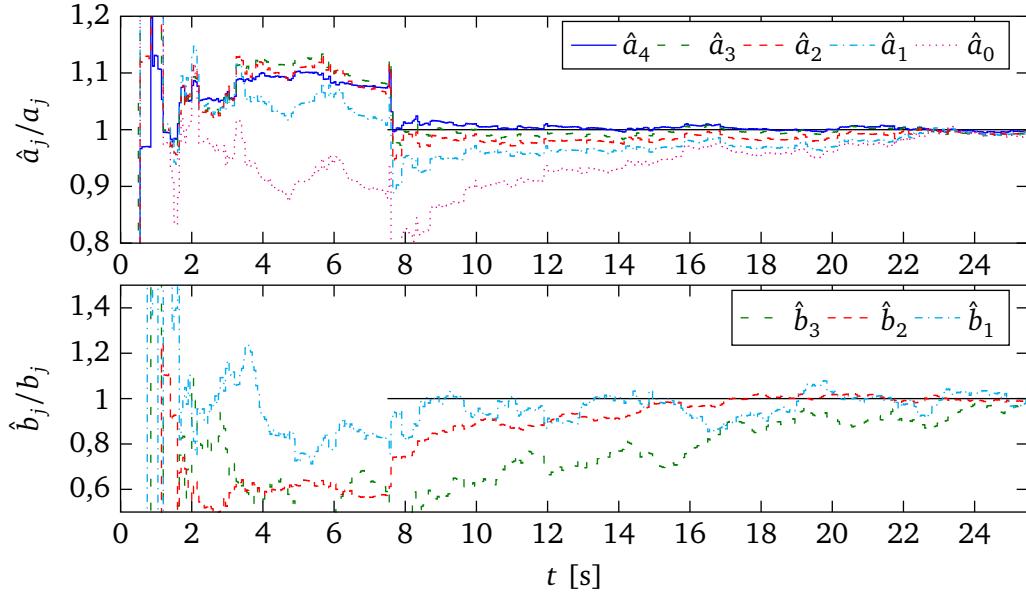
Je länger die Schätzung vor der Parameteränderung gelaufen wäre, desto langsamer würden die Schätzwerte nach der Parameteränderung gegen den neuen Endwert konvergieren.

In Abbildung 9.16a ist ein Vergessensfaktor von $\lambda = 0,99$ verwendet. Die Schätzwerte nähern sich nach der Parameteränderung deutlich schneller den wahren Koeffizienten an. Allerdings wird dies auch durch eine erhöhte Welligkeit (Varianz) der Schätzwerte erreicht.

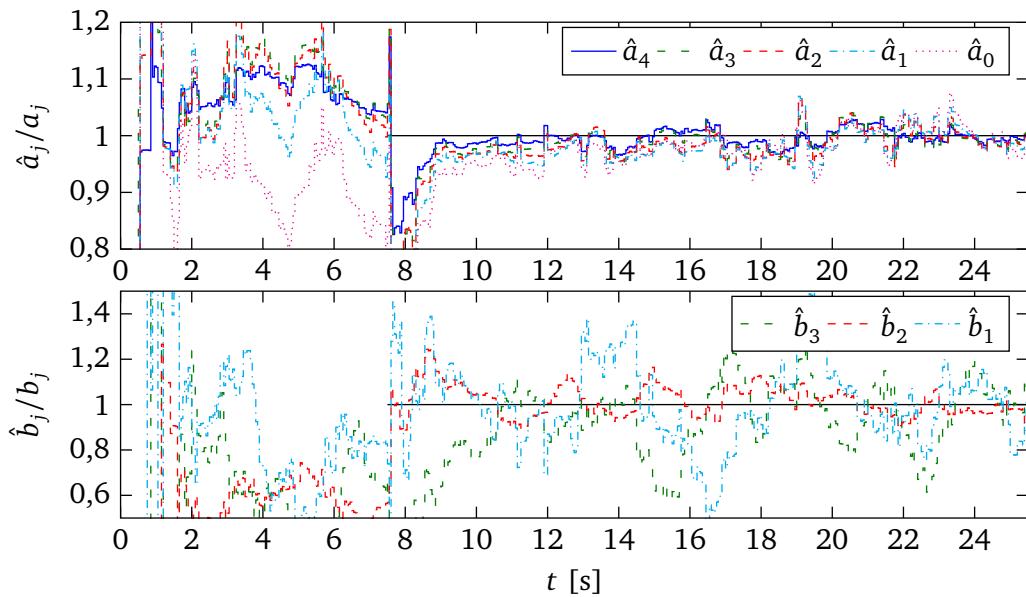
Dies wird mit der Wahl $\lambda = 0,95$ (Abbildung 9.16b) nochmals deutlicher. Hier werden die neuen Parameterwerte sehr schnell erreicht, jedoch treten starke Schwingungen auf, die auch für unendliche Messdauern nicht verschwinden.

9.1.6 Zeitkontinuierliche Übertragungsfunktionen

Zur Identifikation zeitkontinuierlicher Übertragungsfunktionen kann ein dem in Abschnitt 9.1.1 behandelten Verfahren für zeitdiskrete Systeme ähnliches Vorgehen angewandt werden. Praktisch ist dies aber kritischer als die Identifikation zeitdiskreter Übertragungsfunktionen aus Zeitbereichsdaten. Eine Alternative wäre hier der Weg über die Frequenzbereichsdarstellung, d. h. nichtparametrische Frequenzgänge. Darauf wird später noch eingegangen.



(a) $\lambda = 0,99$



(b) $\lambda = 0,95$

Abbildung 9.16: Schätzungen mit $\lambda = 0,99$ und $\lambda = 0,95$, Parameteränderung bei 7,5 s

Ausgangspunkt ist die Übertragungsfunktion

$$G(s) = \frac{b_m s^m + \dots + b_1 s + b_0}{a_n s^n + \dots + a_1 s + 1}$$

deren Parameter bestimmt werden sollen.

Aus dieser lässt sich die Differentialgleichung

$$y(t) = -a_n y^{(n)}(t) - \dots - a_2 \ddot{y}(t) - a_1 \dot{y}(t) + b_m u^{(m)}(t) + \dots + b_1 \dot{u}(t) + b_0 u(t)$$

ablesen.

Analog zum zeitdiskreten Fall wird der Gleichungsfehler betrachtet. Für diesen gilt

$$\hat{e}_{k'} = \tilde{y}(t_{k'}) - \hat{y}(t_{k'})$$

wobei

$$\hat{y}(t) = -\hat{a}_n \tilde{y}^{(n)}(t) - \dots - \hat{a}_2 \ddot{\tilde{y}}(t) - \hat{a}_1 \dot{\tilde{y}}(t) + \hat{b}_m u^{(m)}(t) + \dots + \hat{b}_1 \dot{u}(t) + \hat{b}_0 u(t)$$

wieder der Schätzwert ist, in dem neben dem bekannten Eingang auch der gemessene Ausgang einfließt.

Dieser Zusammenhang wird für k Zeitpunkte $t_{k'}$ gemessen. Mit

$$\tilde{\eta}_{k'} = \tilde{y}(t_{k'})$$

und

$$\tilde{\Psi}_{k'}^T = [-\tilde{y}^{(n)}(t_{k'}) \quad \dots \quad -\dot{\tilde{y}}(t_{k'}) \quad u^{(m)}(t_{k'}) \quad \dots \quad u(t_{k'})]$$

sowie dem Parametervektor

$$\Theta = [a_n \quad \dots \quad a_1 \quad b_m \quad \dots \quad b_0]^T$$

erhält man dann wieder ein LS-Problem.

Mit den Zusammenfassungen

$$\tilde{\eta}_k = \begin{bmatrix} \tilde{y}(t_{k'_1}) \\ \tilde{y}(t_{k'_2}) \\ \vdots \\ \tilde{y}(t_{k'_k}) \end{bmatrix} \quad \text{und} \quad \tilde{\Psi}_k = \begin{bmatrix} \tilde{\Psi}_{k'_1}^T \\ \tilde{\Psi}_{k'_2}^T \\ \vdots \\ \tilde{\Psi}_{k'_k}^T \end{bmatrix} = \begin{bmatrix} -\tilde{y}^{(n)}(t_{k'_1}) & \dots & -\dot{\tilde{y}}(t_{k'_1}) & u^{(m)}(t_{k'_1}) & \dots & u(t_{k'_1}) \\ -\tilde{y}^{(n)}(t_{k'_2}) & \dots & -\dot{\tilde{y}}(t_{k'_2}) & u^{(m)}(t_{k'_2}) & \dots & u(t_{k'_2}) \\ \vdots & & \vdots & \vdots & & \vdots \\ -\tilde{y}^{(n)}(t_{k'_k}) & \dots & -\dot{\tilde{y}}(t_{k'_k}) & u^{(m)}(t_{k'_k}) & \dots & u(t_{k'_k}) \end{bmatrix}$$

ergibt sich dessen Lösung zu

$$\hat{\Theta} = (\tilde{\Psi}_k^T \tilde{\Psi}_k)^{-1} \tilde{\Psi}_k^T \tilde{\eta}_k .$$

Auf die stochastischen Eigenschaften wird hier nicht weiter eingegangen. Kritischer als ein Bias ist in der praktischen Anwendung auch die notwendige Bestimmung der zeitlichen Ableitungen aus den abgetasteten Zeitdaten.

Bestimmung der zeitlichen Ableitungen

Numerisches Differenzieren

Wenn quasi kein Rauschen die Signale überlagert, können die Ableitungen durch numerisches Differenzieren, also z. B.

$$\hat{y}(t_{k'}) = \frac{y(t_k + T_s) - y(t_k - T_s)}{2T_s}$$

bestimmt werden. Ist jedoch mehr als nur unwesentliches Rauschen vorhanden oder werden auch höhere Ableitungen benötigt, so wird das Rauschen durch das Differenzieren stark verstärkt und dieses verdeckt leicht das eigentliche Nutzsignal in den Ableitungen. Daher ist ein Ansatz nötig, der eine gewisse Filterung vorsieht.

Zustandsvariablenfilter

Die Idee des Zustandsvariablenfilters beruht darauf, ein in Prinzip beliebiges zeitkontinuierliches Filter

$$G_F(s) = \frac{b_0}{a_0 + a_1 s + \cdots + a_{n_F-1} s^{n_F-1} + s^{n_F}}$$

(wobei in der Regel $b_0 = a_0$ ist, um eine stationäre Verstärkung von eins zu erhalten) in der Form

$$\begin{bmatrix} \dot{x}(t) \\ \ddot{x}(t) \\ \dddot{x}(t) \\ \vdots \\ x^{(n_F-1)}(t) \\ x^{(n_F)}(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & & & 0 \\ 0 & 0 & 1 & & & 0 \\ 0 & 0 & 0 & \ddots & & 0 \\ & & & & \ddots & \\ 0 & 0 & 0 & & & 1 \end{bmatrix} \begin{bmatrix} x(t) \\ \dot{x}(t) \\ \ddot{x}(t) \\ \vdots \\ x^{(n_F-2)}(t) \\ x^{(n_F-1)}(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ b_0 \end{bmatrix} \cdot y(t)$$

$$y_F(t) = [1 \ 0 \ 0 \ \dots \ \dots \ 0] \cdot \mathbf{x}(t)$$

darzustellen. Diese Form entspricht weitestgehend der Regelungsnormalform. (Bis darauf, dass der Koeffizient b_0 in der Eingangsmatrix anstelle der Ausgangsmatrix erscheint.)

In dieser Form entspricht der erste Filterzustand $x(t)$ gerade dem gefilterten Signal $y_F(t)$. Der zweite Filterzustand entspricht damit der ersten Ableitung $\dot{y}_F(t)$ des gefilterten Signals und so weiter. Werden die ersten beiden Ableitungen eines Signales benötigt, so ergibt sich die Ausgangsgleichung

$$\begin{bmatrix} y_F(t) \\ \dot{y}_F(t) \\ \ddot{y}_F(t) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \end{bmatrix} \cdot \mathbf{x}(t).$$

Das einfachste Filter ist ein PT_n -Glied mit n gleichen Zeitkonstanten T_F ,

$$G_F(s) = \frac{1}{(1 + T_F s)^n}.$$

Es können natürlich an dieser Stelle beliebige andere, bessere Filter verwendet werden. Das Butterworth-Filter ist in Anhang E.1 aufgeführt.

Benötigt man die ersten beiden Ableitungen eines gemessenen Signals, so könnte man als einfaches Beispiel das Filter vierter Ordnung

$$G_F(s) = \frac{1}{(1 + T_F s)^4} = \frac{\frac{1}{T_F^4}}{\left(\frac{1}{T_F} + s\right)^4} = \frac{\frac{1}{T_F^4}}{\frac{1}{T_F^4} + 4 \cdot \frac{1}{T_F^3} s + 6 \cdot \frac{1}{T_F^2} s^2 + 4 \cdot \frac{1}{T_F} s^3 + s^4}$$

ansetzen. Daraus lässt sich direkt das Zustandsvariablenfilter

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -\frac{1}{T_F^4} & -\frac{4}{T_F^3} & -\frac{6}{T_F^2} & -\frac{4}{T_F} \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 0 \\ 0 \\ 0 \\ \frac{1}{T_F^4} \end{bmatrix} \tilde{y}(t) \\ \begin{bmatrix} y_F(t) \\ \dot{y}_F(t) \\ \ddot{y}_F(t) \end{bmatrix} &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \mathbf{x}(t)\end{aligned}$$

ablesen.

Das Bodediagramm dieses Filters ist für $T_F = 1$ s in Abbildung 9.17a dargestellt. In Abbildung 9.17b ist zum Vergleich ein aus einem Butterworth-Filter vierter Ordnung konstruiertes Zustandsvariablenfilter gezeigt. Die Grenzfrequenz des Butterworth-Filters ist auf 1 rad/s gelegt. (Damit beträgt die Verstärkung beim PT_4 -Filter bei $\omega = 1$ rad/s gerade -12 dB und beim Butterworth-Filter -3 dB. Die Asymptoten schneiden sich in beiden Fällen bei $\omega = 1$ rad/s, so dass dies eine gute Vergleichbarkeit ergibt.)

Man erkennt, dass die Filter für niedrige Frequenzen wie ideale Differenzierer bezogen auf das ungefilterte Signal $y(t)$ (welches als waagrechte Linie bei 0 dB bzw. 0° eingezeichnet werden könnte) arbeiten und bei hohen Frequenzen auf alle Signale dämpfend wirken.

Bezogen auf das gefilterte Signal $y_F(t)$ sind die Ableitungen $\dot{y}_F(t), \ddot{y}_F(t), \dots$, die das Filter ausgibt, für alle Frequenzen exakt. Dieses ist die wesentliche Eigenschaft des Zustandsvariablenfilters und führt zu dem in Abbildung 9.18 gezeigten Blockschaltbild.

Es wird sowohl für das Ein- sowie das Ausgangssignal ein Zustandsvariablenfilter verwendet. Diese sind durch die gestrichelten Boxen markiert. Innerhalb der Zustandsvariablenfilter ist die Filterung und die Ableitungsbildung getrennt dargestellt. Dies entspricht zwar nicht der Berechnung der Größen im Filter, ist aber auf das Ein-Ausgangs-Verhalten bezogen korrekt.

Filtert man das Ein- und Ausgangssignal mit dem gleichen Filter $G_F(s)$, so wurde schon einführend zur nichtparametrischen Identifikation gezeigt, dass die Filterung sich dann in dem Sinne aufhebt, dass $\frac{Y(\omega)}{U(\omega)} = \frac{Y_F(\omega)}{U_F(\omega)} = H(\omega)$ gilt. ($Y_F(\omega)$ ist die Fouriertransformierte des ungestörten Anteils von $\tilde{y}_F(t)$, $\tilde{y}_F(t) = y_F(t) + n_{y,F}(t)$.)

Wichtig bei der Anwendung ist es, dass nicht nur die gefilterten Ableitungen, sondern auch das eigentliche Signal aus dem Filter verwendet wird, d. h. es ist $u_F(t)$ und $\tilde{y}_F(t)$ und nicht $u(t)$ und $\tilde{y}(t)$ zu verwenden. Ebenso muss auch die Eingangsgröße gefiltert werden, unabhängig davon, ob Ableitungen von $u(t)$ benötigt werden oder nicht. Der einzige Fehler entsteht dann durch die Diskretisierung des Filters.

Diskretisierung des Filters

Für die Implementierung ist dieses zeitkontinuierliche System mit den Methoden der digitalen Regelungstechnik in ein zeitdiskretes System umzuwandeln.

Dabei spielt es hier keine große Rolle, ob das diskretisierte Filter den Frequenzgang des kontinuierlichen Filters möglichst exakt wiedergibt. Wichtiger ist es zunächst, dass die Ableitungsbeziehungen der einzelnen Filterausgänge untereinander stimmen.

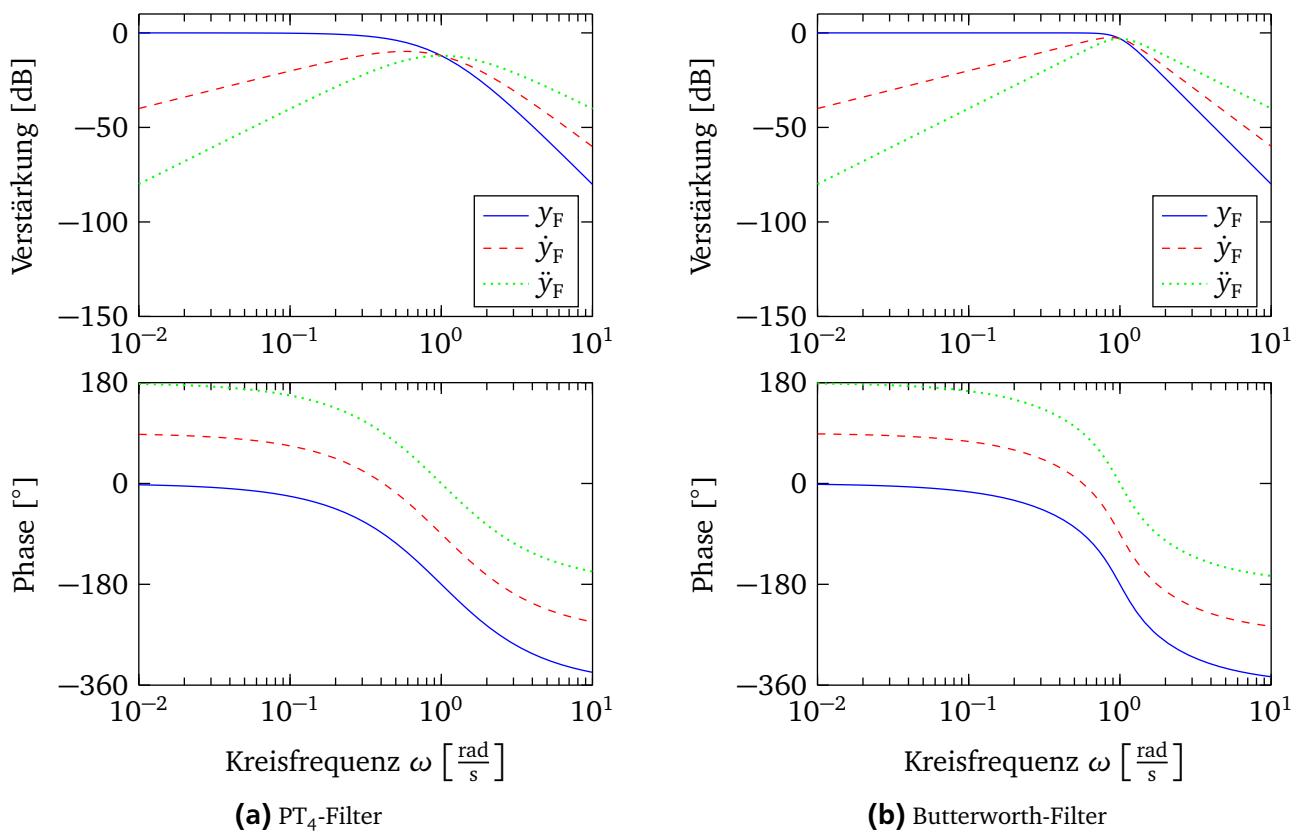


Abbildung 9.17: Bodediagramme von Zustandsvariablenfilter vierter Ordnung zur Erzeugung der ersten beiden Ableitungen

$H_{F,0}(\omega)$ soll hier den Frequenzgang des diskretisierten Filters ohne Ableitungsbildung bezeichnen. $H_{F,d}(\omega)$ den Frequenzgang auf den Ausgang der d -ten Ableitung.

Wird nun eine zeitkontinuierliche harmonische Schwingung der Frequenz ω gemessen, so erscheint diese nach der Filterung als abgetastete Schwingung die entsprechend $H_{F,0}(\omega)$ verstrtzt und phasenverschoben ist. Wesentlich ist, dass am Ausgang der ersten Ableitung eine (abgetastete) Schwingung gemessen wird, die mit dem Faktor ω verstrtzt ist und um 90° phasenverschoben ist, wie es sich aus den Ableitungsregeln einer harmonischen Schwingung ergibt. Fr die hheren Ableitungen gilt entsprechendes. Idealerweise gilt also

$$H_{F,d}(\omega) \stackrel{!}{=} (j\omega)^d \cdot H_{F,0}(\omega),$$

was aber sicher nicht für den gesamten Frequenzbereich zu erfüllen ist. (Dies entspricht der Differenziationsregel für zeitkontinuierliche Signale im Laplacebereich, „ s “.)

In Abbildung 9.19 ist diese Eigenschaft bezüglich der zweiten Ableitung eines Zustandsvariablenfilters vierter Ordnung (Butterworth) bei einer Grenzfrequenz von 1 rad/s und einer Abtastzeit von $T_s = 10 \text{ ms}$ für die Diskretisierung anhand der bilinearen Transformation sowie über ein Halteglied erster (FOH) und nullter Ordnung (ZOH) überprüft. Dazu ist jeweils der Amplitudengang des jeweiligen Filterfrequenzgangs bezüglich der zweiten Ableitung sowie den mit $(j\omega)^2$ multiplizierten Filterfrequenzgang bezüglich der nicht abgeleiteten Größe dargestellt. (Die Phasengänge sind jeweils deckungsgleich).

Es zeigt sich, dass im wesentlichen Arbeitsbereich alle Filter die Differentiation sehr gut wiedergeben. (Würde man die Abweichungen vergrößern, so würde sich die folgende Reihenfolge bezüglich steigender Abweichungen zweigen: FOH, ZOH, Bilinear.) Lediglich bei höheren Frequenzen ergeben sich hier sichtbare Abweichungen, wobei diese bei der bilinearen Transformation am größten sind. Dafür zeigt

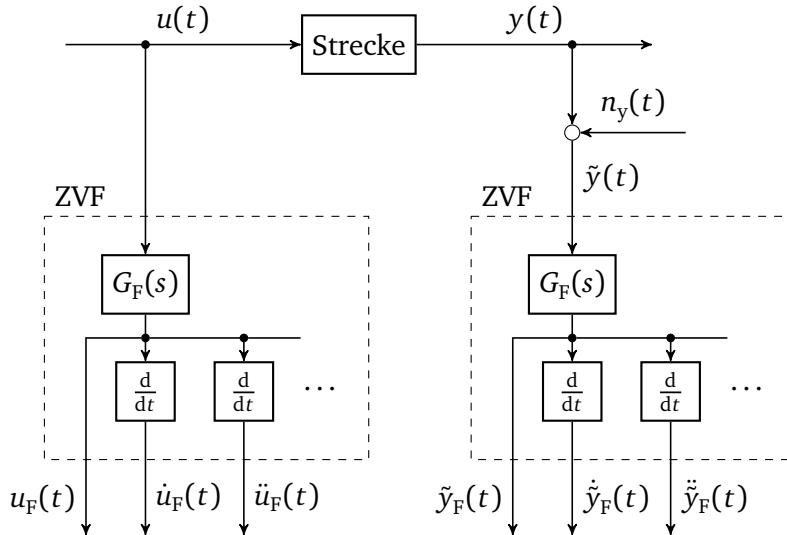


Abbildung 9.18: Zustandsvariablenfilter

sich aber auch, dass der Filterfrequenzgang des mit der bilinearen Transformation erzeugten Filters bei hohen Frequenzen eine höhere Dämpfung als bei den anderen Diskretisierungsmethoden aufweist. Bei der Nyquistfrequenz wäre der Amplitudengang sogar null. In den folgenden Beispielen zeigt sich, dass dies eine sehr nützliche Eigenschaft ist.

Beispiel

Um die Wirkung der verschiedenen Diskretierungsverfahren weiter zu untersuchen, wird in Abbildung 9.20 zunächst die Identifikation der zeitkontinuierlichen Übertragungsfunktion des Drei-Massen-Schwingers ohne Rauschen betrachtet.

Der Zustandsvariablenfilter ist im Zeitkontinuierlichen als Butterworth-Filter achter Ordnung ausgelegt. Es sind die drei Grenzfrequenzen 200 rad/s, 100 rad/s und 50 rad/s betrachtet. Die Beurteilung der Güte der Identifikation erfolgt anhand des Vergleichs der sich ergebenden Frequenzgänge mit dem tatsächlichen Frequenzgang.

Als Anregungssignal ist hier der (normale) Gleitsinus verwendet. Dies ermöglicht eine Abtastzeit von $T_s = 10 \text{ ms}$. (Wobei davon ausgegangen wird, dass der Gleitsinus quasikontinuierlich, also schneller abgetastet auf das System gegeben wird.) Würde man ein PRBS oder allgemeiner ein Eingangssignal verwenden, welches aus Sprüngen besteht, muss um Alias-Effekte zu vermeiden eine schnellere Abtastzeit gewählt werden. (Mit einer Abtastzeit von 1 ms ergeben sich hier mit dem PRBS ähnliche Ergebnisse wie bei den mit 10 ms abgetasteten Ausgangsdaten, wenn der Gleitsinus als Eingangsgröße verwendet wird.)

Es zeigt sich, dass prinzipiell die Identifikation besser wird, wenn die Grenzfrequenz zu niedrigeren Frequenzen hin verschoben wird. Dabei ist offenbar die Dämpfung bei den Frequenzen nahe der Nyquistfrequenz entscheidend, da der Filter der sich über die bilineare Transformation ergibt schon bei einer Grenzfrequenz von 200 rad/s (die Nyquistfrequenz liegt hier schon bei 314 rad/s) zu einer sehr guten Identifikation führt.

Da hier kein Rauschen simuliert ist, ergeben sich die Abweichungen ausschließlich aus (schwachen) Alias-Effekten, Fehlern bei der Ableitungsbestimmung bei hohen Frequenzen und numerischen Effekten.

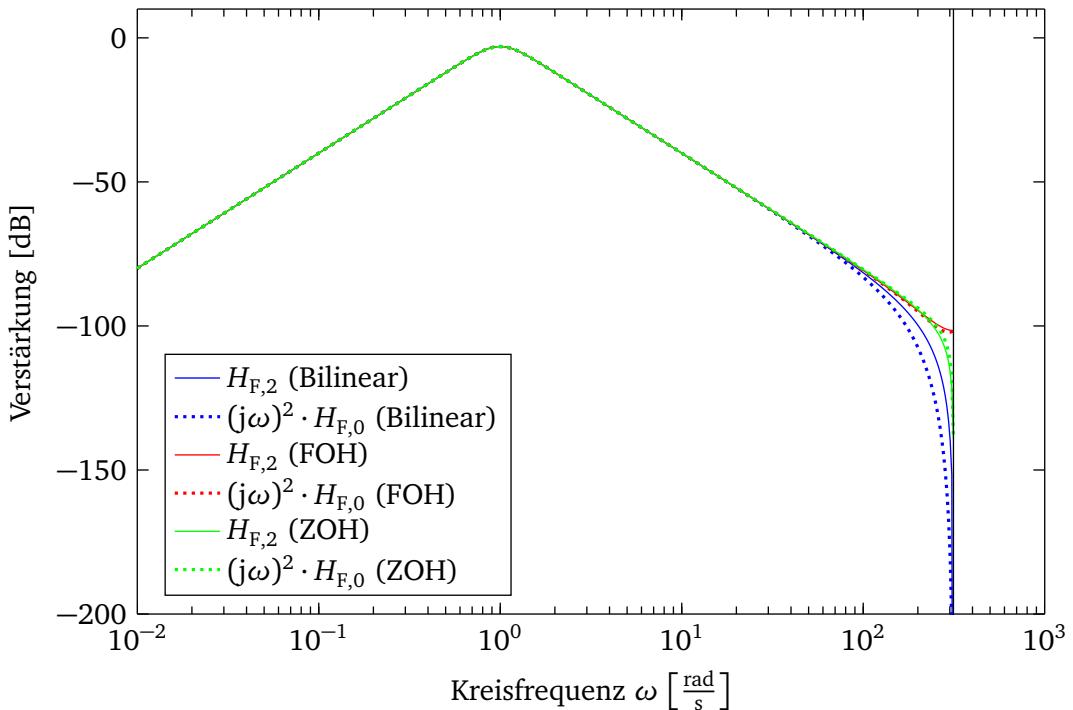


Abbildung 9.19: Zweite Ableitung bei verschiedenen Diskretisierungsmethoden

Da der mit der bilinearen Transformation diskretisierte Filter gewissermaßen der robusteste ist, wird im nächsten Beispiel nur dieser verwendet. In Abbildung 9.21 ist das Ergebnis gezeigt, wenn auf die abgetasteten gemessenen Signale (vor der Anwendung des Zustandvariablenfilters) noch ein weißes Rauschen mit der Standardabweichung von $\sigma_y = 0,25 \text{ rad/s}$ aufaddiert ist.

Es zeigt sich dabei, dass die Grenzfrequenz im Vergleich zum ungestörten Fall deutlich nach unten gesetzt werden muss. Mit einer Grenzfrequenz von 25 rad/s (diese liegt noch unterhalb der zweiten Resonanz des Systems!) ergeben sich sehr gute Ergebnisse. Bei 50 rad/s werden die Resonanzen zwar noch prinzipiell „erkannt“, aber der quantitative Verlauf des Frequenzgangs aus den identifizierten Parametern stimmt nicht mehr.

Neben der Ableitungsbestimmung ist also auch die Filterwirkung der Zustandsvariablenfilter ein wesentlicher Aspekt. (Und nicht nur ein Nebeneffekt, der nötig ist, die Ableitungen überhaupt bestimmen zu können.) Dass die hohen Frequenzanteile der Störungen einen starken Einfluss auf das Ergebnis der Identifikation haben, kann man sich wie folgt erklären. Wenn das Rauschen über den gesamten Frequenzbereich die gleiche Dichte aufweist, so haben die Schwingungen, in die es zerlegt werden kann, für alle Frequenzen die gleiche Amplitude. Damit haben die entsprechenden Ableitungen, bei denen die Amplituden mit der jeweiligen Frequenz multipliziert werden, entsprechend höhere Werte bei höheren Frequenzen. In dem Fall des Drei-Massen-Schwingers werden die ersten fünf Ableitungen benötigt, was diesen Effekt noch verstärkt. Die Messungen, d. h. die Zeilen in $\tilde{\Psi}_k$ mit den höchsten Einträgen dominieren das Ergebnis, da eine betragsmäßig gleiche Abweichung der Parameterwerte von dem jeweiligen „Zeilenoptimum“ einen größeren Fehler verursacht, wenn die Einträge in $\tilde{\Psi}^T$ höher sind.

Bleibt noch anzumerken, dass die Bestimmung einer zeitkontinuierlichen Übertragungsfunktion numerisch kritischer ist. Dies liegt an den schon angesprochenen großen Unterschieden in den Größenordnungen der Einträge von $\tilde{\Psi}_k$, welche zu sehr schlecht konditionierten Matrizen führen.

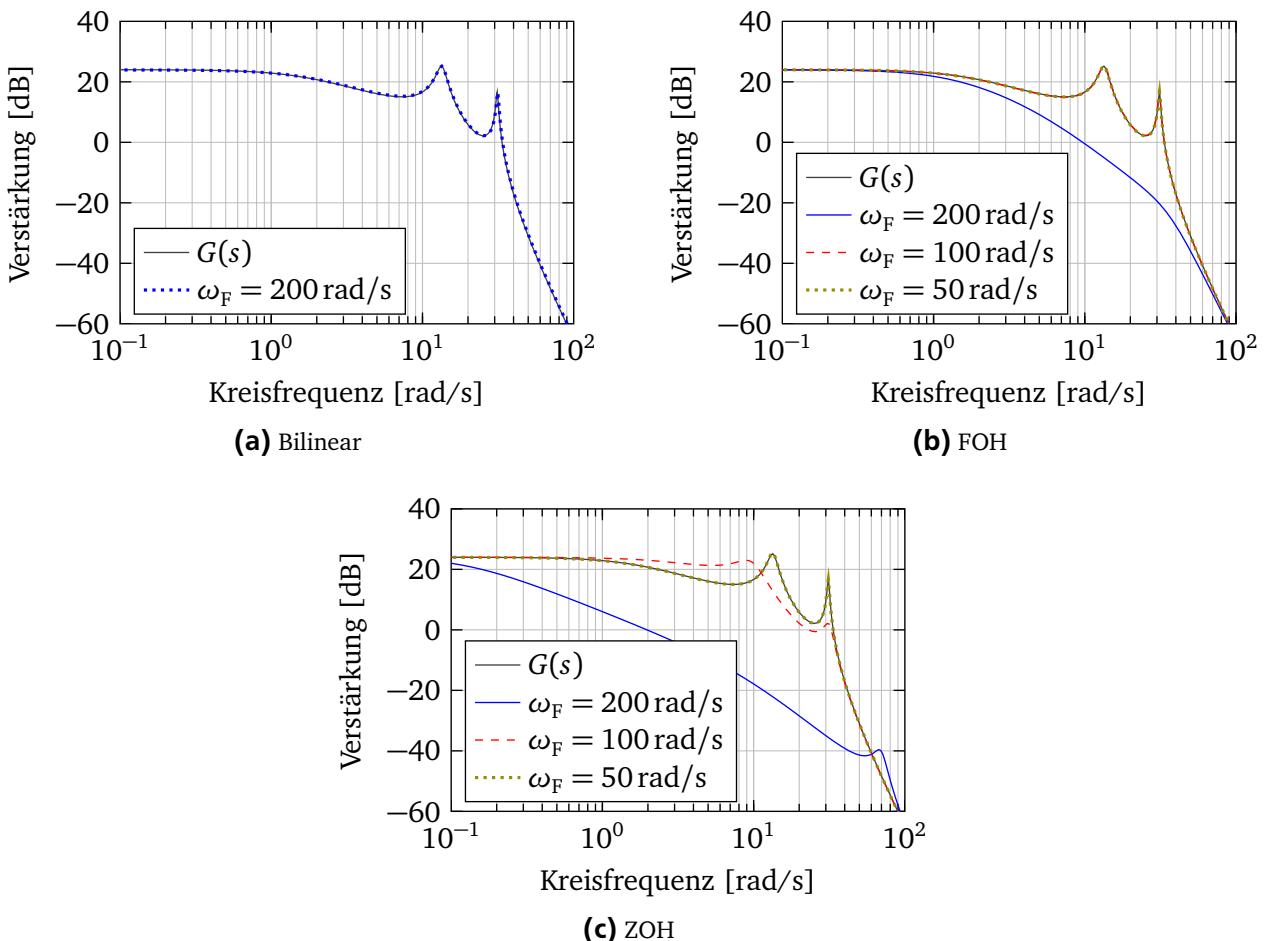


Abbildung 9.20: Wirkung verschiedener Filtergrenzfrequenzen und Diskretisierungsverfahren bei der Identifikation des Drei-Massen-Schwingers
($T_s = 10 \text{ ms}$, ohne Rauschen)

Filterung bei Frequenzbereichs- und Zeitbereichsmethoden

An dieser Stelle soll noch kurz ein Unterschied bezüglich der Filterung bei Frequenzbereichs- und Zeitbereichsmethoden angesprochen werden.

Wenn aus Zeitdaten eine Frequenzgangsschätzung erfolgen soll, so ist jede Filterung, die über eine Anti-Aliasing-Filterung hinausgeht, im Wesentlichen unnötig. Dies kann man sich darüber klar machen, dass bei der Frequenzgangsschätzung der Quotient aus gemessenen Ausgangs- und Eingangsspektrum gebildet wird, d. h. $\hat{H}(\omega) = \frac{\tilde{Y}(\omega)}{U(\omega)} = \frac{Y(\omega) + N_y(\omega)}{U(\omega)}$. Dämpft man die Ausgangsdaten (inklusive Rauschen) und Eingangsdaten mit demselben Filter mit dem Frequenzgang $H_F(\omega)$, so gilt $\hat{H}(\omega) = \frac{H_F(\omega) \cdot (Y(\omega) + N_y(\omega))}{H_F(\omega) \cdot U(\omega)}$. Es ändert sich also nichts an der Schätzung.

Kommt man nun auf die Idee, nur den Ausgang zu filtern, um starke Störungen zu dämpfen, so ist dies auch keine Lösung. Zwar kann bei geeigneter Wahl von $H_F(\omega)$ der Schätzwert $\hat{H}'(\omega) = \frac{H_F(\omega) \cdot (Y(\omega) + N_y(\omega))}{U(\omega)}$ näher an dem tatsächlichen Wert $H(\omega)$ liegen, aber letztlich bedeutet dies, dass man eine biasbehaftete Schätzung durchführt (was man einfach erkennt, wenn man die Störung wieder auf Null setzt), um damit im gestörten Fall zu besseren Ergebnissen zu kommen. Um das Filter $H_F(\omega)$ richtig einzustellen, müsste man aber Informationen über das System und die Störungen haben. Wenn man diese Informationen jedoch hätte, müsste man nicht identifizieren.

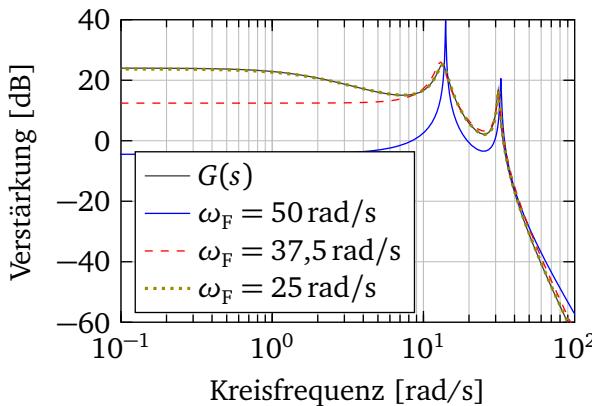


Abbildung 9.21: Wirkung verschiedener Filtergrenzfrequenzen bei der Identifikation des Drei-Massen-Schwingers
(Störung weißes Rauschen, $\sigma_y = 0,25 \text{ rad/s}$, ZVF: Butterworth achter Ordnung, Diskretisierung bilinear)

Es sollte also darauf geachtet werden, das Rauschen von vornehmerein so gut wie möglich zu vermeiden. Im Weiteren sollte dann das Hauptaugenmerk darauf liegen, die Spektren so exakt wie möglich zu berechnen und dabei wenig Fehler zu machen. Wenn man möglichst fehlerfrei im Frequenzbereich angekommen ist, ist die Filterung kein Problem mehr. Im Frequenzbereich können beliebige Bandstopfilter mit idealen Flanken realisiert werden: Man lässt die Frequenzpunkte, die durch Störungen unbrauchbar sind, einfach weg. Dabei werden weder die Amplitude noch die Phase der behaltenen Frequenzpunkte beeinflusst.

Mit den im Frequenzbereich gefilterten Daten kann man dann mit den unten in Abschnitt 9.2 vorgestellten Methoden ein parametrisches Modell identifizieren.

Bei der parametrischen Identifikation mit Zeitbereichsdaten ist im Gegensatz zur nichtparametrischen Frequenzgangsschätzung eine Filterung der Daten im Zeitbereich z. T. sinnvoll und nötig. Dies zeigt sich z. B. an dem gerade behandelten Beispiel. Hätte man zunächst einen Frequenzgang geschätzt und dann damit eine Übertragungsfunktion identifiziert, so würde man die hohen Frequenzen, in denen quasi nur Rauschen zu sehen ist, einfach im nächsten Identifikationsschritt weglassen. Geht man diesen Zwischenschritt nicht und will direkt von den Zeitbereichsdaten zu einer Übertragungsfunktion kommen, dann muss dass Rauschen schon im Zeitbereich möglichst entfernt werden.

9.2 Schätzung von Übertragungsfunktionen aus Frequenzbereichsdaten

Die Übertragungsfunktion ist in der Form

$$G(s) = \frac{b_m s^m + \dots + b_1 s + b_0}{a_n s^n + \dots + a_1 s + 1}$$

gegeben. Der Frequenzgang lautet damit

$$H(\omega) = \frac{b_m \cdot (j\omega)^m + \dots + b_1 \cdot (j\omega) + b_0}{a_n \cdot (j\omega)^n + \dots + a_1 \cdot (j\omega) + 1} \quad (9.13)$$

bzw.

$$H(\omega) = \frac{B(j\omega)}{A(j\omega)} .$$

Liegt ein gemessener Frequenzgang $\tilde{H}(\omega)$ (d. h. ein durch ein nichtparametrisches Verfahren geschätzter Frequenzgang) vor, dann liegt es wieder nahe, für k Frequenzpunkte $\omega_{k'}$ den Ausgangsfehler

$$\hat{e}_{a,k'} = \tilde{H}(\omega_{k'}) - \frac{\hat{B}(j\omega_{k'})}{\underbrace{\hat{A}(j\omega_{k'})}_{\hat{H}(\omega_{k'})}}$$

zu minimieren. In der Gütfunktion

$$V = \sum_{k'} w_{k'} \cdot |e_{a,k'}|^2$$

werden dabei in der Regel Gewichte vorgesehen. Als Gewicht bietet sich z. B. die (geschätzte) Varianz des geschätzten Frequenzganges am jeweiligen Frequenzstützpunkt⁷ an,

$$V = \sum_{k'} \frac{|\tilde{H}(\omega_{k'}) - \hat{H}(\omega_{k'})|}{\sigma_{\tilde{H},k'}^2}.$$

Der Ausgangsfehler hat jedoch auch hier den Nachteil, dass dieser auf ein Optimierungsproblem führt, welches nicht linear in den Parametern ist.

Minimierung des Gleichungsfehlers

Im Folgenden wird eine dem Gleichungsfehler bei der Schätzung aus Zeitbereichsdaten äquivalente Form aufgestellt.

Der Fehler wird wieder analog dem Gleichungsfehler gebildet, d. h.

$$\begin{aligned} \hat{e}_{k'} &= \hat{A}(j\omega_{k'}) \cdot \tilde{H}(\omega_{k'}) - \hat{B}(j\omega_{k'}) \\ &= \tilde{H}(\omega_{k'}) - \underbrace{\left((1 - \hat{A}(j\omega_{k'})) \cdot \tilde{H}(\omega_{k'}) + \hat{B}(j\omega_{k'}) \right)}_{\hat{H}'(\omega_{k'})}. \end{aligned} \quad (9.14)$$

Die zu minimierende Gütfunktion ist durch

$$V = \sum_{k'} w_{k'} \cdot |\hat{e}_{k'}|^2$$

gegeben. Ziel ist es wieder, die Summe der quadratischen Fehler zu minimieren, wobei hier auch ein Gewicht $w_{k'}$ für die einzelnen Fehler vorgesehen ist. Da der Schätzfehler $\hat{e}_{k'}$ eine komplexe Größe ist, ist in der Gütfunktion vor der Quadrierung der Betrag zu nehmen. Es gilt $|\hat{e}_{k'}|^2 = \hat{e}_{k'} \hat{e}_{k'}^*$, wobei $\hat{e}_{k'}^*$ das konjugiert-komplexe von $\hat{e}_{k'}$ bezeichnet. Damit erhält man weiter⁸ $\hat{e}_k = \hat{e}_{R,k'} + j\hat{e}_{I,k'}$ und kann die Gütfunktion auch als

$$\begin{aligned} V &= \sum_{k'} w_{k'} \cdot (\hat{e}_{R,k'}^2 + \hat{e}_{I,k'}^2) \\ &= \sum_{k'} w_{k'} \hat{e}_{R,k'}^2 + \sum_{k'} w_{k'} \hat{e}_{I,k'}^2 \end{aligned}$$

⁷ Frequenzstützpunkte werden hier in der Regel mit l' indiziert. Da die „Messungen“ zum Aufstellen eines LS-Problems jedoch mit k' indiziert werden, wird in diesem Abschnitt von dieser Konvention abgewichen.

⁸ In diesem Abschnitt steht X_R immer für den Realteil der Größe X und X_I entsprechend für den Imaginärteil, d. h. $X = X_R + j \cdot X_I$. Diese Notation wird nicht für jede Größe explizit eingeführt.

schreiben. In dieser Darstellung treten zum einen keine komplexen Größen mehr auf, und zum anderen erkennt man, das sich die Gütefunktion mit

$$\hat{\mathbf{e}}_k = [\hat{e}_{R,1} \quad \hat{e}_{I,1} \quad \hat{e}_{R,2} \quad \hat{e}_{I,2} \quad \cdots \quad \hat{e}_{R,k} \quad \hat{e}_{I,k}]^T$$

(der Vektor $\hat{\mathbf{e}}_k$ hat $2k$ Einträge) und

$$\mathbf{W}_k = \text{diag}([w_1 \quad w_1 \quad w_2 \quad w_2 \quad \cdots \quad w_k \quad w_k])$$

in der Form

$$V = \hat{\mathbf{e}}_k^T \mathbf{W}_k \hat{\mathbf{e}}_k$$

schreiben lässt. Man erhält $2k$ Gleichungen anstelle von k , dafür treten keine komplexen Größen mehr auf.

Aufstellen der Matrizen

Mit dem Frequenzgang (9.13) kann der Gleichungsfehler (9.14) als

$$\hat{e}_{k'} = \tilde{H}(\omega_{k'}) + \hat{a}_n \cdot \tilde{H}(\omega_{k'}) \cdot (j\omega_{k'})^n + \cdots + \hat{a}_1 \cdot \tilde{H}(\omega_{k'}) \cdot (j\omega_{k'}) - \hat{b}_m \cdot (j\omega_{k'})^m - \cdots - \hat{b}_1 \cdot (j\omega_{k'}) - \hat{b}_0 \quad (9.15)$$

angegeben werden.

Daraus können die – komplexen – Größen

$$\tilde{\eta}_{k'} = \tilde{H}(\omega_{k'})$$

und

$$\tilde{\Psi}_{k'}^T = [-\tilde{H}(\omega_{k'}) \cdot (j\omega_{k'})^n \quad \cdots \quad -\tilde{H}(\omega_{k'}) \cdot (j\omega_{k'}) \cdot (j\omega_{k'})^m \quad \cdots \quad (j\omega_{k'}) \quad 1]$$

sowie der (reelle) Parametervektor

$$\hat{\theta} = [a_n \quad a_{n-1} \quad \cdots \quad a_1 \quad b_m \quad b_{m-1} \quad \cdots \quad b_0]^T$$

des LS-Problems abgelesen werden.

Damit kann der Gleichungsfehler kurz als

$$\hat{e}_{k'} = \tilde{\eta}_{k'} - \tilde{\Psi}_{k'}^T \hat{\theta}$$

geschrieben werden. Der Fehler wird in Real- und Imaginärteil aufgeteilt,

$$\begin{aligned} \hat{e}_{R,k'} &= \text{Re}(\hat{e}_{k'}) = \text{Re}(\tilde{\eta}_{k'}) - \text{Re}(\tilde{\Psi}_{k'}^T) \cdot \hat{\theta} \quad \text{und} \\ \hat{e}_{I,k'} &= \text{Im}(\hat{e}_{k'}) = \text{Im}(\tilde{\eta}_{k'}) - \text{Im}(\tilde{\Psi}_{k'}^T) \cdot \hat{\theta}. \end{aligned}$$

Damit ergeben sich

$$\tilde{\eta}_k = \begin{bmatrix} \text{Re}(\tilde{\eta}_1) \\ \text{Im}(\tilde{\eta}_1) \\ \text{Re}(\tilde{\eta}_2) \\ \text{Im}(\tilde{\eta}_2) \\ \vdots \\ \text{Re}(\tilde{\eta}_k) \\ \text{Im}(\tilde{\eta}_k) \end{bmatrix} = \begin{bmatrix} \text{Re}(\tilde{H}(\omega_1)) \\ \text{Im}(\tilde{H}(\omega_1)) \\ \text{Re}(\tilde{H}(\omega_2)) \\ \text{Im}(\tilde{H}(\omega_2)) \\ \vdots \\ \text{Re}(\tilde{H}(\omega_k)) \\ \text{Im}(\tilde{H}(\omega_k)) \end{bmatrix} \quad \text{und} \quad \tilde{\Psi}_k = \begin{bmatrix} \text{Re}(\tilde{\Psi}_1^T) \\ \text{Im}(\tilde{\Psi}_1^T) \\ \text{Re}(\tilde{\Psi}_2^T) \\ \text{Im}(\tilde{\Psi}_2^T) \\ \vdots \\ \text{Re}(\tilde{\Psi}_k^T) \\ \text{Im}(\tilde{\Psi}_k^T) \end{bmatrix} \quad (9.16)$$

sowie die Gewichtsmatrix

$$\mathbf{W}_k = \begin{bmatrix} w_1 & & & \\ & w_1 & & \\ & & w_2 & \\ & & & w_2 \\ & & & \ddots \\ & & & & w_k \\ & & & & & w_k \end{bmatrix}.$$

Wenn die Programmiersprache mit der gearbeitet wird komplexe Zahlen direkt unterstützt (z. B. Matlab) ist dieser Aufbau zu empfehlen. Ist dies nicht der Fall, dann können $\tilde{\Psi}_k$ und $\tilde{\eta}_k$ auch wie in Anhang E.2 beschrieben aufgestellt werden. (Das Ergebnis ist letztlich dasselbe, lediglich treten in den Zwischenstufen keine komplexen Größen auf.)

Die Lösung des damit beschriebenen gewichteten LS-Problems

$$\hat{\theta} = \arg \min_{\hat{\theta}} \hat{\mathbf{e}}_k^T \mathbf{W}_k \hat{\mathbf{e}}_k = \arg \min_{\hat{\theta}} (\tilde{\eta}_k - \tilde{\Psi}_k \hat{\theta})^T \cdot \mathbf{W}_k \cdot (\tilde{\eta}_k - \tilde{\Psi}_k \hat{\theta})$$

lautet

$$\hat{\theta}_k = (\tilde{\Psi}_k^T \mathbf{W}_k \tilde{\Psi}_k)^{-1} \tilde{\Psi}_k^T \mathbf{W}_k \tilde{\eta}_k.$$

Dabei ist die Wahl der Gewichte wesentlich, wie das in Abbildung 9.22 gezeigte Beispiel deutlich macht. Das dort gezeigte Ergebnis beruht auf einer Schätzung ohne Gewichtung, d. h. $w_{k'} \equiv 1$. Der durch die grauen Punkte dargestellte gemessene Frequenzgang wurde aus zehn einzelnen Gleitsinusanregungen bestimmt. Offensichtlich wird dieser nur sehr schlecht durch die identifizierte Übertragungsfunktion abgebildet.

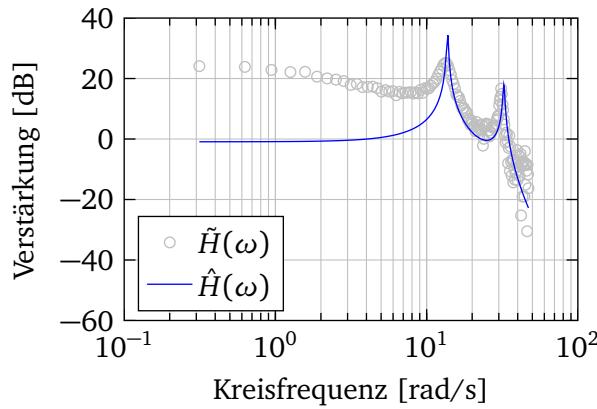


Abbildung 9.22: Schätzung mit $w_{k'} \equiv 1$

Um dies genauer zu betrachten, ist in Abbildung 9.23 oben das gleiche gezeigt, lediglich in linearer Achsenteilung. Hier werden die Abweichungen noch deutlicher.

Im unteren Plot von Abbildung 9.23 sind die Beiträge $|e_{k'}|^2$ der Einzelfehler zu der Gütfunktion dargestellt. Offenbar tragen die niedrigen Frequenzen hier am meisten zum Gesamtfehler bei. Auch überpräsentiert sind die hohen Frequenzen. Zwar ist der Beitrag $|e_{k'}|^2$ dort geringer als bei den niedrigen Frequenzen, aber dafür ist die absolute Abweichung im Frequenzgang deutlich geringer. Die Abweichungen bei den Resonanzstellen tragen kaum zu dem Gesamtfehler bei.

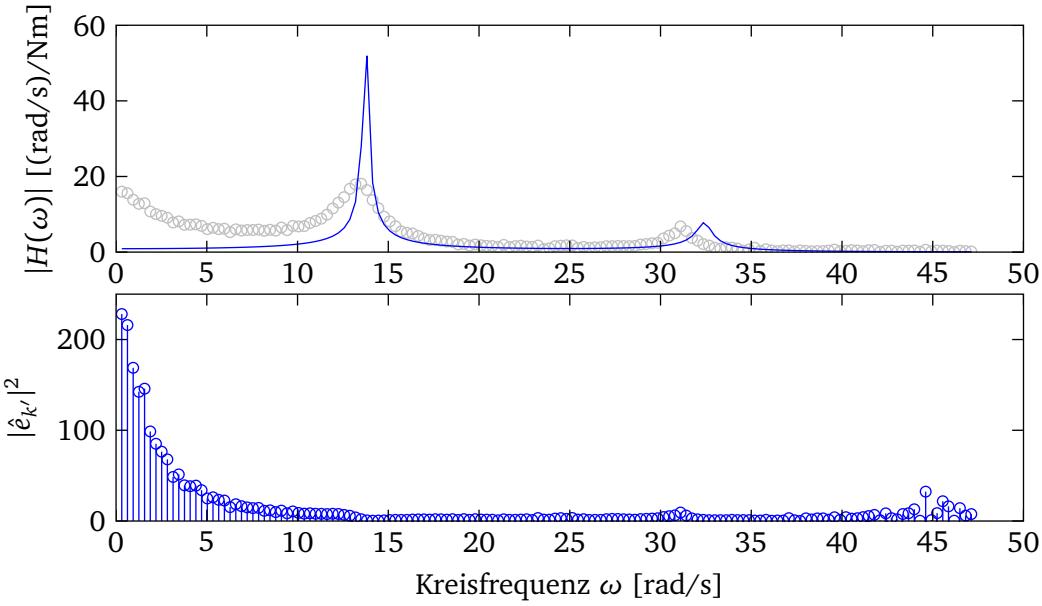


Abbildung 9.23: Schätzung mit $w_{k'} \equiv 1$, lineare Achenskalierung und Angabe der Einzelfehler

Es ist auch hier zu beachten, dass $\hat{e}_{k'}$ der Gleichungsfehler ist. Bei der Betrachtung der Abweichung im Frequenzgangsplot bewertet man aber automatisch den Ausgangsfehler. Daher „widerspricht“ sich die beobachtete Abweichung und der Fehlerbeitrag $|e_{k'}|^2$ bezüglich des Ausgangsfehlers bei derselben Frequenz.

Dies ist auch der Grund für das schlechte Ergebnis bezüglich des Ausgangsfehlers, d. h. die großen Abweichungen zwischen gemessenen Frequenzgang und der daraus geschätzten Übertragungsfunktion.

Wahl der Gewichte

Ein Kriterium zur Wahl der Gewichte ist die Unsicherheit in den gemessenen Frequenzgängen. So ist eine mögliche Wahl, den Kehrwert des quadrierten relativen Fehlers $\frac{|\Delta\tilde{H}(\omega_{k'})|}{|\tilde{H}(\omega_{k'})|}$ zu verwenden,

$$w_{k'} = \left(\frac{|\tilde{H}(\omega_{k'})|}{|\Delta\tilde{H}(\omega_{k'})|} \right)^2$$

Dabei ist die Unsicherheit $\Delta\tilde{H}(\omega_{k'})$ bzw. der relative Fehler $\frac{|\Delta\tilde{H}(\omega_{k'})|}{|\tilde{H}(\omega_{k'})|}$ aus der Streuung der Punkte im Frequenzgang abzuschätzen oder bei der Schätzung des Frequenzganges mitzuschätzen.

Da die relative Unsicherheit mit steigender Frequenz in der Regel zunimmt, wird damit auch dem Effekt entgegengewirkt, dass die zu höheren Frequenzen gehörenden Fehler aufgrund der Potenzierungen der Frequenz im Eingangsvektor deutlich stärker in den Gesamtfehler eingehen. (Allerdings ist dieses Gewicht bei Nullstellen problematisch, da in deren Umgebung die relative Unsicherheit in der Regel sehr groß ist. Damit werden die Frequenzpunkte um die Nullstellen kaum gewichtet, so dass die Optimierung die Freiheitsgrade der angesetzten Zählnullstellen dazu nutzen wird, die Abweichungen in Bereichen hoher Systemverstärkung zu vermindern. Hierbei können durchaus auch manuelle Anpassungen des Gewichts nötig sein.)

Diesen Effekt der stärkeren Gewichtung kann man auch wie folgt interpretieren. Der Ausgangsfehler lautet

$$\hat{e}_{a,k'} = \tilde{H}(\omega_{k'}) - \frac{\hat{B}(j\omega_{k'})}{\hat{A}(j\omega_{k'})}.$$

Wird dieser mit $\hat{A}(\omega_{k'})$ durchmultipliziert, erhält man

$$\hat{A}(\omega_{k'}) \cdot \hat{e}_{a,k'} = \hat{A}(\omega_{k'}) \cdot \tilde{H}(\omega_{k'}) - \hat{B}(\omega_{k'}) .$$

Auf der rechten Seite steht der hier verwendete Gleichungsfehler $\hat{e}_{k'}$. Man kann den Gleichungsfehler also als den mit $|\hat{A}(\omega_{k'})|^2$ gewichteten Ausgangsfehler interpretieren.

Daher liegt ein Ansatz darin, diese Gewichtung, die im Allgemeinen hohe Frequenzen zu stark gewichtet, wieder rückgängig zu machen. Daher können die Gewichte

$$w_{k'} = \frac{1}{|A(\omega_{k'})|^2}$$

angesetzt werden. (Das zu schätzende Polynom $\hat{A}(\omega_{k'})$ kann nicht verwendet werden, da dies ja wieder auf ein Optimierungsproblem, welches nichtlinear in den Parametern ist, führen würde. Nämlich genau auf die Optimierung des Ausgangsfehlers.) Das wahre Polynom $A(\omega_{k'})$ ist unbekannt, jedoch kann ein iterativer Ansatz angewandt werden. [Isermann und Münchhof 2011, S. 373f]

Letzterer Ansatz kann auch mit dem Ansatz der Gewichtung über die Varianz bzw. $\Delta\tilde{H}(\omega_{k'})$ gekoppelt werden, also z. B.

$$w_{k'} = \frac{1}{\sigma_{\tilde{H},k'}^2} \cdot \frac{1}{|A(\omega_{k'})|^2} .$$

(Hierbei ist nicht der relative Schätzfehler verwendet, da durch den Faktor $\frac{1}{|A(\omega_{k'})|^2}$ quasi der Ausgangsfehler bewertet wird. Und bei diesem ist die Bewertung mit dem absoluten Schätzfehler bzw. Varianz sinnvoll.)

Prinzipiell ist es einem jedoch auch freigestellt, bestimmte Frequenzbereiche nach eigenem Ermessen stärker oder schwächer zu gewichten.

Beispiele

Als erstes Beispiel wird die Gewichtung nach

$$w_{k'} = \left(\frac{|\tilde{H}(\omega_{k'})|}{|\Delta\tilde{H}(\omega_{k'})|} \right)^2$$

verwendet. Dabei wird $|\Delta\tilde{H}(\omega_{k'})|^2$ durch die geschätzte Varianz $\hat{\sigma}_{\tilde{H},k'}^2$ des gemessenen Frequenzgangs abgeschätzt, so dass sich

$$w_{k'} = \frac{|\tilde{H}(\omega_{k'})|^2}{\hat{\sigma}_{\tilde{H},k'}^2} \quad (9.17)$$

ergibt. (Ein gemeinsamer Faktor, also z. B. das zweifache oder dreifache der Varianz spielt keine Rolle für die Gewichtung.)

Die simulative „Messung“ der Frequenzgänge ist hier zehnmal mit gleichem, unverrauschem Eingangsgrößensignal erfolgt. Damit ist es möglich, ohne Nachteile in Kauf nehmen zu müssen, für jede der Einzelschätzungen einen Frequenzgang $\tilde{H}_m(\omega_l)$ zu ermitteln⁹ Die zehn Einzelmessungen sind zur Verdeutlichung der Streuung im oberen Plot von Abbildung 9.24 übereinander geplottet dargestellt.

⁹ Der mit den nicht-parametrischen Methoden geschätzte Frequenzgang wird für den hier behandelten Schritt der parametrischen Identifikation als „gemessen“ aufgefasst, und daher mit einer Tilde gekennzeichnet.

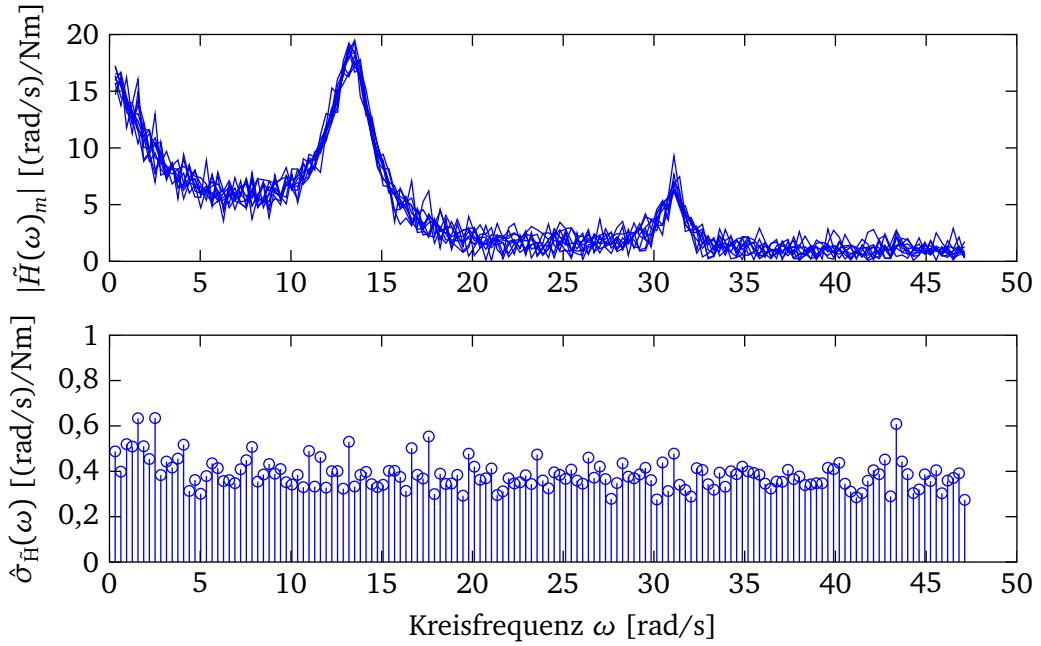


Abbildung 9.24: Schätzung der Varianz des gemessenen Frequenzgangs aus 10 Einzelmessung \tilde{H}_m

Daraus ergibt sich

$$\tilde{H}(\omega_{l'}) = \frac{1}{10} \sum_{m=1}^{10} \tilde{H}_m(\omega_{l'}) .$$

$\tilde{H}(\omega_{l'})$ ist beispielsweise in Abbildung 9.22 bzw. Abbildung 9.23 durch die grauen Punkte dargestellt.

Die Schätzwert $\hat{\sigma}_{\tilde{H}, l'}^2$ der Varianz $\text{Var}\{\tilde{H}(\omega_{l'})\}_r$ berechnet sich über

$$10 \cdot \hat{\sigma}_{\tilde{H}, l'}^2 = \frac{1}{9} \cdot \sum_{m=1}^{10} (\text{Re}(\tilde{H}_m(\omega_{l'})) - \text{Re}(\tilde{H}(\omega_{l'})))^2 + \frac{1}{9} \cdot \sum_{m=1}^{10} (\text{Im}(\tilde{H}_m(\omega_{l'})) - \text{Im}(\tilde{H}(\omega_{l'})))^2 .$$

(Der Faktor von 10 auf der linken Seite ergibt sich daher, dass auf der rechten Seite der Schätzwert der Varianz für eine Messung steht. Da zehn Messungen gemittelt sind, reduziert sich die Varianz für den Schätzwert $\tilde{H}(\omega_{l'})$ um den Faktor 10.) Die sich ergebenden Werte sind als Standardabweichung im unteren Plot von Abbildung 9.24 gezeigt.

Es zeigt sich, dass die Varianz hier relativ konstant ist. Aufgrund dessen, dass die Absolutwerte von $\tilde{H}(\omega)$ aber für hohe Frequenzen deutlich geringer sind, wird mit der angegebenen Gewichtung (9.17) deutlich weniger Gewicht auf die Frequenzbereiche gelegt, bei denen die Verstärkung des Systems gering ist.

In Abbildung 9.25 sind die Ergebnisse gezeigt, die sich mit diesem Gewicht ergeben. Es wird direkt deutlich, dass das Ergebnis besser ist, als im Fall von $w_{k'} \equiv 1$.

Im zweiten Plot von Abbildung 9.25 sind die ungewichteten, quadrierten Einzelfehler gezeigt, im dritten Plot die gewichteten. Der unterste Plot zeigt das Gewicht, wobei dieses in dB angegeben ist. Man erkennt, dass durch die Gewichtung der Schwerpunkt der Optimierung auf die niedrigen Frequenzen und die Resonanzstellen gelegt wird. Sieht man von dem einen Ausreißer bei ca. 45 rad/s ab, so ist der Beitrag der gewichteten Fehler für die niedrigen und hohen Frequenzen ungefähr gleich. Ungewichtet wäre bei dem gegebenen Verlauf von $\hat{H}(\omega)$ der Beitrag der Fehler bei den hohen Frequenzen deutlich dominant. D. h. die ungefähr gleichen absoluten Abweichungen vom gemessenen Frequenzgang würden bei den

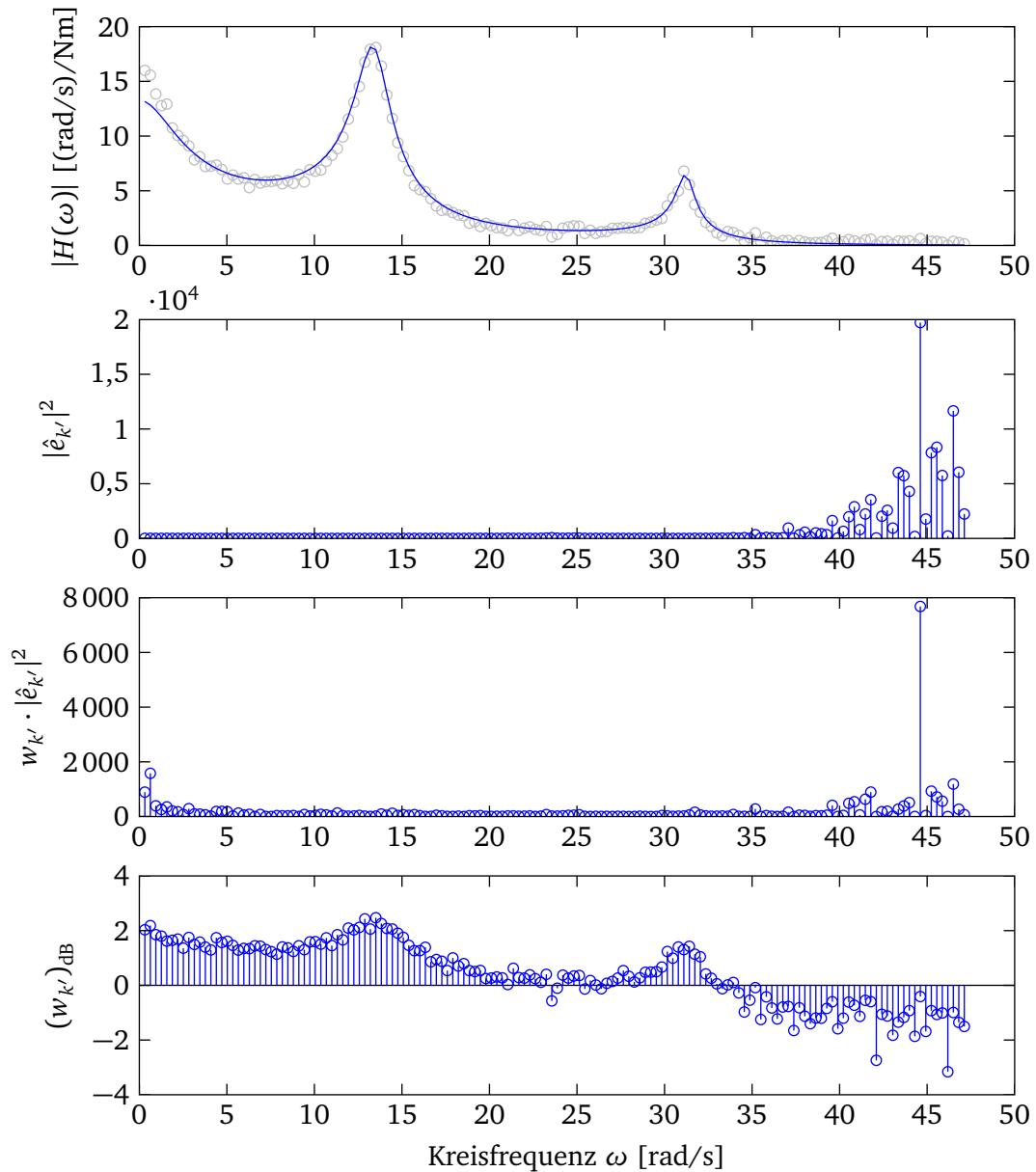


Abbildung 9.25: Schätzung mit $w_{k'} = \frac{|\tilde{H}(\omega_{k'})|^2}{\sigma_{\tilde{H}, k'}^2}$

hohen Frequenzen deutlich stärker gewertet. Dies führt dazu, dass bei der oben gezeigten ungewichteten Schätzung die Parameter so gefunden werden, dass die Abweichung bei den hohen Frequenzen möglichst gering wird, auch wenn dies sehr große Abweichungen bei den niedrigen Frequenzen hervorruft.

Die relativen Abweichungen der geschätzten Koeffizienten liegen im Bereich von 8 % bis 22 %.

In Abbildung 9.26 sind noch die Ergebnisse gezeigt, die sich über den Gewichtsansatz

$$w_{k'} = \frac{1}{|A(j\omega_{k'})|^2}$$

ergeben. Dabei ist mit dem Gewicht $w_{k'} \equiv 1$ begonnen, und dann für jede weitere der drei Iterationen der Faktor $A(j\omega_{k'})$ über das geschätzte Nennerpolynom des vorangegangenen Schrittes bestimmt.

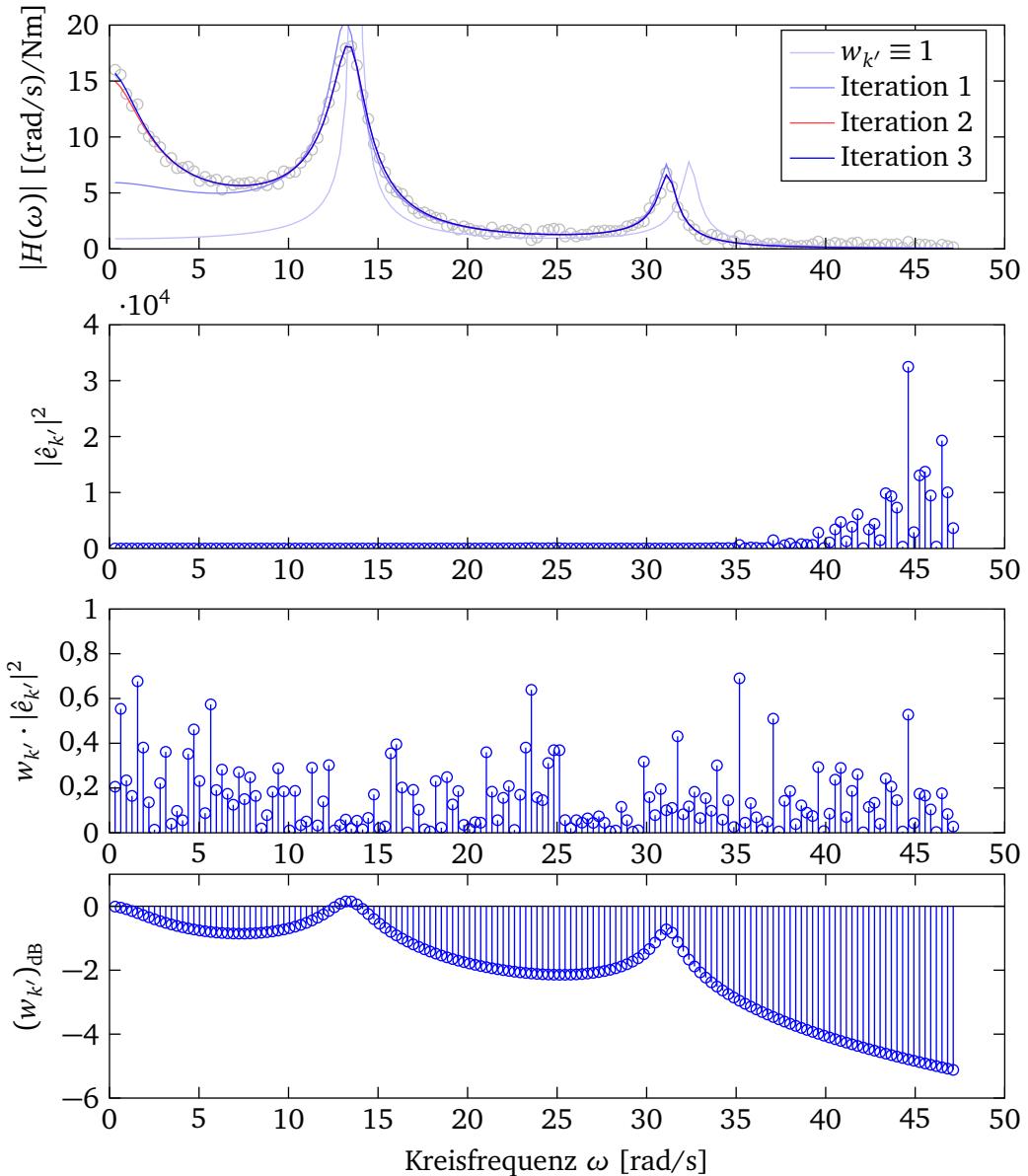


Abbildung 9.26: Schätzung mit $w_{k'} = \frac{1}{|A(j\omega_{k'})|^2}$
(iterativ, in den unteren Plots sind die Werte der letzten Iteration dargestellt)

Man erkennt, dass die Schätzung sehr schnell konvergiert. Die zweite Iteration zeigt lediglich noch kleine Abweichung bei sehr niedrigen Frequenzen, mit der dritten Iteration ist ein zumindest optisch einwand-

freies Ergebnis erzielt. Dies bestätigt auch die Güte der geschätzten Koeffizienten. Die relative Abweichung der geschätzten Koeffizienten liegen nach der zweiten Iteration zwischen 2 % und 5 %, und nach der dritten Iteration liegen alle unterhalb von 0,4 %.

Der Verlauf des Gewichtes über die Frequenz ist deutlich glatter als im Fall der Gewichtung über den geschätzten relativen Fehler, womit auch kein deutlicher Ausreißer mehr auftritt. Die gewichtete, quadrierte Einzelfehler haben über den gesamten Frequenzbereich gleichmäßige Werte.

Variation

Anstelle des gemessenen Frequenzgangs $\tilde{H}(\omega)$ kann auch ein Gleichungsfehler formuliert werden, der von den gemessenen Spektren $\tilde{Y}(\omega_{k'})$ und $\tilde{U}(\omega_{k'})$ ausgeht. Dies führt auf die Gütfunktion

$$V = \sum_{k'} |\hat{A}(j\omega_{k'}) \cdot \tilde{Y}(\omega_{k'}) - \hat{B}(j\omega_{k'}) \cdot \tilde{U}(\omega_{k'})|^2 ,$$

welche ein LS-Problem, welches linear in den Parametern ist, beschreibt.

Berechnung über QR-Methode mit Gewicht

Die Gütfunktion des LS-Verfahrens mit Gewichtung lautet

$$V = \hat{\mathbf{e}}^T \mathbf{W} \hat{\mathbf{e}} = (\tilde{\boldsymbol{\eta}}^T + \hat{\boldsymbol{\theta}}^T \boldsymbol{\Psi}^T) \cdot \mathbf{W} \cdot (\boldsymbol{\Psi} \hat{\boldsymbol{\theta}} + \tilde{\boldsymbol{\eta}}) .$$

Zerlegt man \mathbf{W} in die Form $\mathbf{W} = \mathbf{S}^T \mathbf{S}$, so erhält man¹⁰

$$\begin{aligned} V &= \hat{\mathbf{e}}^T \mathbf{W} \hat{\mathbf{e}} = (\tilde{\boldsymbol{\eta}}^T + \hat{\boldsymbol{\theta}}^T \boldsymbol{\Psi}^T) \cdot \mathbf{S}^T \mathbf{S} \cdot (\boldsymbol{\Psi} \hat{\boldsymbol{\theta}} + \tilde{\boldsymbol{\eta}}) \\ &= (\tilde{\boldsymbol{\eta}}^T \mathbf{S}^T + \hat{\boldsymbol{\theta}}^T \boldsymbol{\Psi}^T \mathbf{S}^T) \cdot (\mathbf{S} \boldsymbol{\Psi} \hat{\boldsymbol{\theta}} + \mathbf{S}^T \tilde{\boldsymbol{\eta}}) . \end{aligned}$$

Man kann anstelle der Lösung des gewichteten LS-Verfahrens also auch die Lösung des ungewichteten LS-Verfahrens mit

$$\begin{aligned} \boldsymbol{\Psi}_S &= \mathbf{S} \boldsymbol{\Psi} \\ \tilde{\boldsymbol{\eta}}_S &= \mathbf{S} \tilde{\boldsymbol{\eta}} \end{aligned}$$

bestimmen.

Bezüglich der Lösung über die Normalgleichungen macht dies keinen Unterschied, da ohnehin alles wieder ausmultipliziert wird. Aber in dieser Form kann die Lösung über die QR-Zerlegung bestimmt werden. Dies kann hier sehr schnell notwendig werden, da die Größenordnungen der Matrixeinträge durch die Potenzierung der Frequenzen extrem unterschiedlich sein können.

Schätzen von diskreten Übertragungsfunktionen

Für diskrete Übertragungsfunktionen kann das gleiche Vorgehen angewendet werden. Der Frequenzgang ergibt sich über

$$H(\omega) = G(z = e^{j\omega T_s}) = \frac{b_{n-1}z^{-1} + \dots + b_0z^{-n}}{1 + a_{n-1}z^{-1} + \dots + a_0z^{-n}} \cdot z^{-d} , \quad d \geq 0$$

¹⁰ Es kann beispielsweise die Quadratwurzel $\mathbf{S} = \mathbf{S}^T$ der positiv-semidefiniten Matrix \mathbf{W} verwendet werden, $\mathbf{W} = \mathbf{S} \mathbf{S}^T = \mathbf{S}^T \mathbf{S}$. Ist \mathbf{W} eine Diagonalmatrix (der Regelfall für diese Anwendung), $\mathbf{W} = \text{diag}([w_1 \ w_2 \ \dots \ w_k])$, so erhält man \mathbf{S} , indem man von den Diagonaleinträgen einzeln die Wurzeln bildet, $\mathbf{S} = \text{diag}([\sqrt{w_1} \ \sqrt{w_2} \ \dots \ \sqrt{w_k}])$. Bei positiv definitem \mathbf{W} kann die Aufteilung $\mathbf{W} = \mathbf{S}^T \mathbf{S}$ auch über die Cholesky-Zerlegung bestimmt werden. (In diesem Fall ist \mathbf{S} nicht die Quadratwurzel von \mathbf{W} .)

aus der diskreten Übertragungsfunktion. Es liegt wieder die Messung $\tilde{H}'(\omega_k')$ des Frequenzgangs an k Frequenzpunkten $\omega_{k'}$ vor. Dieser ist mit der Abtastzeit $T_{s,\text{FG}}$ ermittelt.

Wenn der Frequenzgang mit einer anderen Abtastzeit geschätzt wurde, als die Abtastzeit T_s für die die Übertragungsfunktion ermittelt werden soll, dann ist darauf zu achten, dass keine Frequenzen oberhalb der Nyquistfrequenz genommen werden.¹¹ Es sollte darüber hinaus ein gewisser Abstand zur Nyquistfrequenz eingehalten werden. Bei der bezüglich eines Haltegliedes nullter Ordnung am Eingang exakten Diskretisierung von zeitkontinuierlichen Systemen bzw. der Reduzierung der Abtastzeit von zeitdiskreten Systemen ist für Frequenzen nahe der Nyquistfrequenz meist ein Unterschied in den Frequenzgängen festzustellen. Daher sollte man die Optimierung nicht zwingen, diesen für die gewünschte „Zielabtastzeit“ falschen Verlauf zu berücksichtigen.

Dies ist in Abbildung 9.27 verdeutlicht. Es ist dort in schwarz der Frequenzgang des kontinuierlichen Systems dargestellt. Dazu sind die Frequenzgänge der mit Abtastzeiten zwischen 0,001 s und 0,1 s abgetasteten Systeme gezeigt. Die senkrechten schwarzen Striche markieren die Nyquistfrequenzen zu den drei langsameren Abtastzeiten.

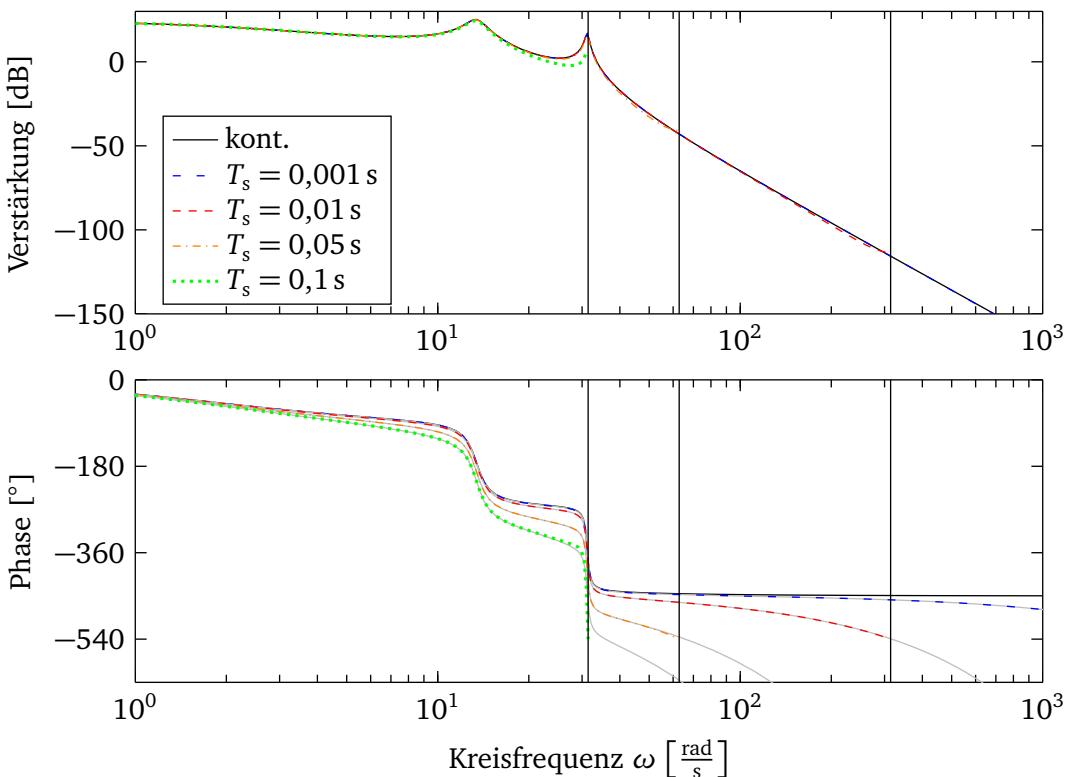


Abbildung 9.27

Für $T_s = 0,001\text{ s}$ zeigt sich in der Amplitude im gezeigten Frequenzbereich kein sichtbarer Unterschied zum kontinuierlichen System. Für die anderen Abtastzeiten ist im Bereich unterhalb der jeweiligen Nyquistfrequenz eine kleine, bei $T_s = 0,1\text{ s}$ auch größere Abweichung zu erkennen. (Wobei eine Abtastzeit von $0,1\text{ s}$ für das vorliegende System – wenn auch nicht verboten – schon sehr knapp gewählt ist.)

Im Phasengang zeigen sich für alle Abtastzeiten eine Auswirkung, d. h. ein Unterschied zum kontinuierlichen System, wobei bei $T_s = 0,001\text{ s}$ dieser erst bei Frequenzen auftritt, die praktisch sicherlich nicht zur Identifikation verwendet werden. Der zu sehende Phasenabfall hat seine Ursache in der Diskretisierung

¹¹ Wenn die gleiche Abtastzeit verwendet wird, dann spielt das keine Rolle, da dann auch der gemessene Frequenzgang bezüglich der Nyquistfrequenz symmetrisch ist.

mit einem Halteglied nullter Ordnung. Hierdurch entsteht bezogen auf das kontinuierliche System eine Totzeit T_{zoh} von der halben Abtastzeit,

$$T_{\text{zoh}} = \frac{1}{2} \cdot T_s . \quad (9.18)$$

Dies ist in Abbildung 9.27 im Phasengang gut zu erkennen, da in grau jeweils der Phasengang des kontinuierlichen Systems unter Beachtung der entsprechenden Totzeit nach (9.18) eingezeichnet ist.

Entsprechend beinhaltet der aus mit $T_{s,\text{FG}}$ erzeugten Eingangsdaten bestimmte Frequenzgang $\tilde{H}'(\omega_k)$ eine „Diskretisierungstotzeit“ von

$$T_{\text{zoh,FG}} = \frac{1}{2} \cdot T_{s,\text{FG}} .$$

(Hierbei ist die Abtastzeit der Generierung des Eingangssignals relevant. Sollte das Eingangssignal tatsächlich kontinuierlich gewesen sein, dann ist $T_{s,\text{FG}} = 0$, unabhängig von der Abtastung der Ausgangswerte!)

Die gemessene Übertragungsfunktion sollte mit der Differenz

$$\Delta T_{\text{zoh}} = T_{\text{zoh}} - T_{\text{zoh,FG}}$$

dieser beiden Totzeiten auf die Abtastzeit T_s angepasst werden, für die die Übertragungsfunktion identifiziert werden soll. D. h.

$$\tilde{H}(\omega_{k'}) = \tilde{H}'(\omega_{k'}) \cdot e^{-j\omega_{k'} \cdot \Delta T_{\text{zoh}}} .$$

Mit der Abkürzung

$$z_{k'} = e^{j\omega_{k'} T_s}$$

lautet der Gleichungsfehler

$$\hat{e}_{k'} = \tilde{H}(\omega_{k'}) + \hat{a}_{n-1} \cdot \tilde{H}(\omega_{k'}) \cdot z_{k'}^{-1} + \cdots + \hat{a}_0 \cdot \tilde{H}(\omega_{k'}) \cdot z_{k'}^{-n} - \hat{b}_{n-1} \cdot z_{k'}^{-1-d} - \cdots - \hat{b}_0 \cdot z_{k'}^{-n-d} .$$

Daraus können analog dem zeitkontinuierlichen Fall die – wieder komplexen – Größen

$$\tilde{\eta}_{k'} = \tilde{H}(\omega_{k'})$$

und

$$\tilde{\Psi}_{k'}^T = [-\tilde{H}(\omega_{k'}) \cdot z_{k'}^{-1} \quad \cdots \quad -\tilde{H}(\omega_{k'}) \cdot z_{k'}^{-n} \quad z_{k'}^{-1-d} \quad \cdots \quad z_{k'}^{-n-d}]$$

sowie der (reelle) Parametervektor

$$\theta = [a_{n-1} \quad \cdots \quad a_0 \quad b_{n-1} \quad \cdots \quad b_0]^T$$

des LS-Problems angegeben werden.

Zur Lösung wird wieder die Aufteilung in Real- und Imaginärteil vorgenommen. Die Anmerkungen bezüglich des Gewichts gelten hier auch.

Numerisch ist die Identifikation einer zeitdiskreten Übertragungsfunktion deutlich unkritischer, da der Betrag von $z_{k'}^x$ immer eins ist. Damit treten keine extremen Unterschiede in den Einträgen von Ψ auf.

9.3 Mehrgrößensysteme

9.3.1 Systemmodelle

Das Ein-Ausgangsverhalten eines allgemeinen linearen MIMO-Systems mit p' Eingängen und q Ausgängen kann über die Übertragungsmatrix

$$\mathbf{G}(s) = \begin{bmatrix} G_{11}(s) & \cdots & G_{1p'}(s) \\ \vdots & \ddots & \vdots \\ G_{q1}(s) & \cdots & G_{qp'}(s) \end{bmatrix}$$

dargestellt werden, wobei $G_{ij}(s)$ die Übertragungsfunktion vom j -ten Eingang auf den i -ten Ausgang bezeichnet. (Für den zeitdiskreten Fall ergibt sich die gleiche Struktur, nur ist die Unabhängigkeit der Übertragungsfunktionen natürlich z.) von jeweils zwei Ein- und Ausgängen entspricht dies der in Abbildung 9.28a dargestellten Struktur.

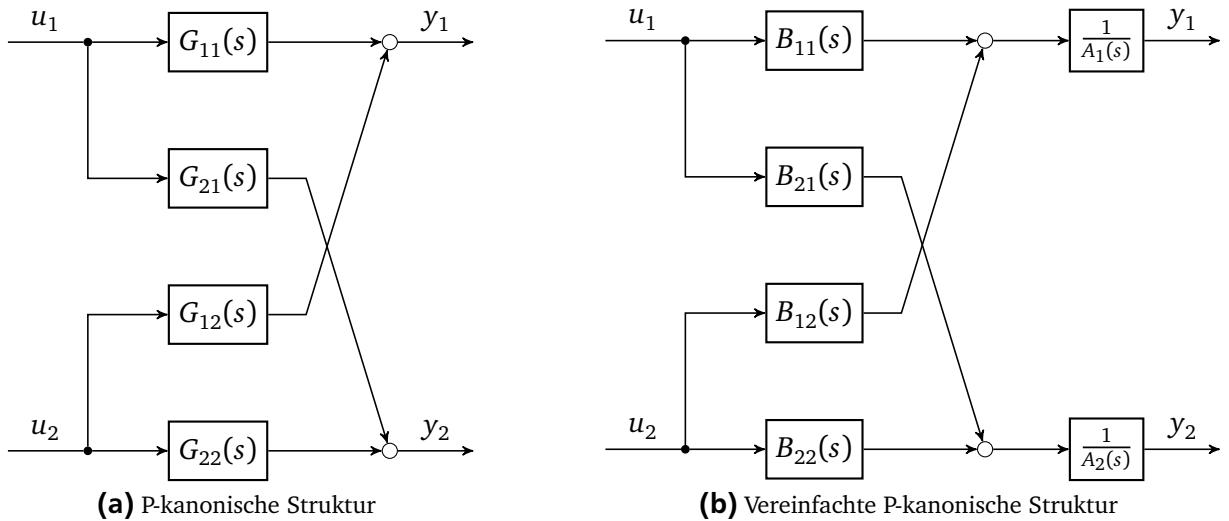


Abbildung 9.28

In dieser allgemeinen Darstellung sind die einzelnen Übertragungsfunktionen völlig unabhängig, werden also durch unabhängige Zähler- und Nennerpolynome beschrieben, $G_{ij}(s) = \frac{B_{ij}(s)}{A_{ij}(s)}$.

Häufig können aber Annahmen über die Struktur des MIMO-Systems gemacht werden. Eine Möglichkeit ist in Abbildung 9.28b dargestellt, welche auch als vereinfachte P-kanonische Struktur bezeichnet wird [Isermann und Münchhof 2011, S. 430f]. Dabei besitzt das System bezüglich jeder der beiden Ausgänge eine feste Eigendynamik, und lediglich der Einfluss der beiden Eingänge, d. h. die Verstärkung und Nullstellen, unterscheiden sich für jeden Ausgang. Als Übertragungsmatrix aufgestellt lautet dies (für den Fall zwei Ein- und zwei Ausgänge)

$$\mathbf{G}(s) = \begin{bmatrix} \frac{B_{11}(s)}{A_1(s)} & \frac{B_{12}(s)}{A_1(s)} \\ \frac{B_{21}(s)}{A_2(s)} & \frac{B_{22}(s)}{A_2(s)} \end{bmatrix}.$$

Eine solche Vereinfachung bzw. die Annahme einer solchen Struktur der Übertragungsmatrix kann jedoch nicht einfach „definiert“ werden, sondern muss sich aus dem tatsächlichen System heraus ergeben.

Bei mechanischen Systemen sind darüber hinaus noch weiterreichende Annahmen zutreffend. Betrachtet man als Beispiel den Drei-Massen-Schwinger und nimmt an, dass auch an der zweiten Trägheit ein externes Moment aufgebracht werden kann und dass auch die Geschwindigkeit der zweiten Trägheit gemessen wird, so erhält man ein MIMO-System mit zwei Eingängen

$$\mathbf{u}(t) = [M_M(t) \quad M_{M2}(t)]^T$$

und zwei Ausgängen

$$\mathbf{y}(t) = [\omega_2 \quad \omega_3]^T.$$

Der Aufbau des Systems ist in Abbildung 9.29 dargestellt.

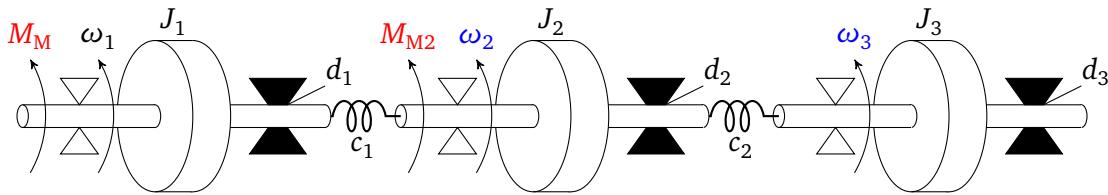


Abbildung 9.29: Drei-Massen-Schwinger als MIMO-System

Die Übertragungsmatrix wird also durch vier Übertragungsfunktionen gebildet. Dabei ist aber der Nenner $A(s)$ in allen vier Übertragungsfunktionen identisch, d. h.

$$\mathbf{G}(s) = \begin{bmatrix} \frac{B_{11}(s)}{A(s)} & \frac{B_{12}(s)}{A(s)} \\ \frac{B_{21}(s)}{A(s)} & \frac{B_{22}(s)}{A(s)} \end{bmatrix}.$$

Dies ergibt sich unmittelbar aus der mechanischen Modellbildung, da die Eigendynamik des mechanischen Systems unabhängig von den Ein- und Ausgängen immer identisch ist. (Es könnte lediglich zu Kürzungen mancher Pole durch entsprechende Nullstellen auf einzelnen Übertragungspfaden kommen.) Es ist dann auch anzuraten, dieser strukturellen Tatsache bei der Identifikation Rechnung zu tragen. Dies nicht nur aus pragmatischen Gründen wie der damit verbundenen Reduzierung von zu optimierenden Parametern (wobei ein solcher Ansatz andererseits eine Zerlegung in kleinere Teilidentifikationsprobleme unterbindet) sondern auch, um ein sinnvolles Modell zu erhalten. Es wäre ein schwer zu begründendes Systemverhalten, wenn die Resonanzen des Systems bei der Messung der Winkelgeschwindigkeit der zweiten Trägheit bei anderen Frequenzen liegen als bei der Messung der Winkelgeschwindigkeit der dritten Trägheit.

Geht man von einer Zustandsraumdarstellung der Ordnung n aus, so erhält man (vollständige Steuer- und Beobachtbarkeit vorausgesetzt) bei der Überführung in eine Übertragungsmatrix immer denselben Nenner $A(s)$ der Ordnung n , wobei durchaus Kürzungen durch Nullstellen auftreten können. Kennt man die Gesamtordnung n des Systems, so kann man dieses sicher darstellen, wenn man einen gleiches Nennerpolynom $A(s)$ für alle Übertragungsfunktionen ansetzt, und für jeden Übertragungspfad einen eigenen Zählerpolynom $B_{ij}(s)$ der Ordnung $n - 1$ (bei durchgriffsfreien Systemen). Ggf. kann durch Systemwissen die Ordnung der Zählerpolynome noch weiter eingeschränkt werden.

Wenn das System jedoch aus unabhängigen Teilsystemen besteht, oder aus Teilsystemen mit einer nur sehr schwachen Kopplung, so führt dieser Ansatz auf sehr viele zu optimierende Parameter. Es ist dann abzuwägen, ob – bei schwacher Kopplung – der Verlust systematischer Zusammenhänge gegenüber einer einfacheren Identifikation vertretbar ist.

Leider ist es für eine einfache Schätzung, die auf die lineare Abhängigkeit von den Parametern angewiesen ist, nicht möglich, für verschiedene Teilsysteme bestimmte Pole und Nullstellen gleich zu setzen.

Bei dem Drei-Massen-Schwinger sollte auf jeden Fall das gleiche Nennerpolynom angesetzt werden, so dass sich als allgemeines Systemmodell

$$\mathbf{G}(s) = \begin{bmatrix} \frac{b_{11,4}s^4 + b_{11,3}s^3 + b_{11,2}s^2 + b_{11,1}s + b_{11,0}}{a_5s^5 + a_4s^4 + a_3s^3 + a_2s^2 + a_1s + 1} & \frac{b_{12,4}s^4 + b_{12,3}s^3 + b_{12,2}s^2 + b_{12,1}s + b_{12,0}}{a_5s^5 + a_4s^4 + a_3s^3 + a_2s^2 + a_1s + 1} \\ \frac{b_{21,4}s^4 + b_{21,3}s^3 + b_{21,2}s^2 + b_{21,1}s + b_{21,0}}{a_5s^5 + a_4s^4 + a_3s^3 + a_2s^2 + a_1s + 1} & \frac{b_{22,4}s^4 + b_{22,3}s^3 + b_{22,2}s^2 + b_{22,1}s + b_{22,0}}{a_5s^5 + a_4s^4 + a_3s^3 + a_2s^2 + a_1s + 1} \end{bmatrix} \quad (9.19)$$

mit 25 Parametern ergibt. Darüber hinaus lässt sich auch einfach – mit etwas Mechanikkenntnis – die Anzahl der Nullstellen der einzelnen Übertragungspfade

$$\begin{bmatrix} M_M \rightarrow \omega_2 & M_{M2} \rightarrow \omega_2 \\ M_M \rightarrow \omega_3 & M_{M2} \rightarrow \omega_3 \end{bmatrix}$$

angeben.¹² Man erhält dann mit

$$\mathbf{G}(s) = \begin{bmatrix} \frac{b_{11,2}s^2 + b_{11,1}s + b_{11,0}}{a_5s^5 + a_4s^4 + a_3s^3 + a_2s^2 + a_1s + 1} & \frac{b_{12,4}s^4 + b_{12,3}s^3 + b_{12,2}s^2 + b_{12,1}s + b_{12,0}}{a_5s^5 + a_4s^4 + a_3s^3 + a_2s^2 + a_1s + 1} \\ \frac{b_{21,0}}{a_5s^5 + a_4s^4 + a_3s^3 + a_2s^2 + a_1s + 1} & \frac{b_{22,2}s^2 + b_{22,1}s + b_{22,0}}{a_5s^5 + a_4s^4 + a_3s^3 + a_2s^2 + a_1s + 1} \end{bmatrix} \quad (9.20)$$

eine Darstellung mit 17 Parametern.

9.3.2 Identifikation aus Zeitbereichsdaten

Eingangssignale

Die Eingänge können, wie im Fall der nichtparametrischen Schätzung, prinzipiell einzeln oder zusammen angeregt werden. Je nach den Abhängigkeiten (gemeinsame Nenner) der einzelnen Übertragungsfunktionen kann es jedoch sein, dass diese Messungen in einer gemeinsamen Schätzung verarbeitet werden.

Werden die Eingänge gemeinsam angeregt, dann ist darauf zu achten, dass über alle Messungen gesehen keine linearen Abhängigkeiten auftreten. Dazu kann wiederum dieselbe Technik wie bei der nichtparametrischen Schätzung verwendet werden. Dabei ist ein identisches Grundsignal mit verschiedenen Faktoren auf die einzelnen Eingänge geschaltet. Bei der Identifikation aus Zeitbereichsdaten muss dies nicht auf Einzelmessungen aufgeteilt werden, sondern die verschiedenen Variationen können direkt hintereinander auf das System gegeben werden.

Ein Signaltyp, der die lineare Unabhängigkeit in sich sicherstellt, ist das „amplitudenmodulierte PRBS“, siehe z. B. [Isermann und Münchhof 2011, S. 442f]. Dieses hat insbesondere für Verfahren die auf Korrelationsfunktionen beruhen (und hier nicht behandelt sind) die wesentliche Eigenschaft, dass die Kreuzkorrelationsfunktionen verschwinden und die Autokorrelationsfunktion von weißem Rauschen noch möglichst gut nachgebildet wird.

Identifikation

Liegt ein System in der Form

$$\mathbf{G}(s) = \begin{bmatrix} \frac{B_{11}(s)}{A_{11}(s)} & \frac{B_{12}(s)}{A_{12}(s)} \\ \frac{B_{21}(s)}{A_{21}(s)} & \frac{B_{22}(s)}{A_{22}(s)} \end{bmatrix},$$

¹² Man kann dazu entweder die Differenzordnung der einzelnen Pfade bestimmen, indem man die Speicher zählt, um vom jeweiligen Ein- zum Ausgang zu kommen. Oder man kann die Anzahl der möglichen Tilgungen für jeden Übertragungspfad zählen.

also in allgemeiner P-kanonischer Form vor, so ist es für das nachfolgend vorgestellte Vorgehen nötig, dass dann die Eingänge einzeln angeregt werden. Damit zerfällt dieses System für die Identifikation in vier SISO-Systeme.

Ist das System in vereinfachter P-kanonischer Form gegeben, z. B.

$$\mathbf{G}(s) = \begin{bmatrix} \frac{b_{11,1}s+b_{11,0}}{a_{1,2}s^2+a_{1,1}s+1} & \frac{b_{12,0}}{a_{1,2}s^2+a_{1,1}s+1} \\ \frac{b_{21,1}s+b_{21,0}}{a_{2,1}s^2+a_{2,1}s+1} & \frac{b_{22,0}}{a_{2,2}s^2+a_{2,1}s+1} \end{bmatrix},$$

so können beide Eingänge gemeinsam angeregt werden. Die Schätzung kann für die Übertragungspfade zum ersten und zum zweiten Ausgang getrennt durchgeführt werden, es sind also bei Systemen in vereinfachter P-kanonischer Form mehrere MISO-Identifikationen durchzuführen. (Man kann diese zwar auch in einer Schätzung zusammenfassen, nur ändert dies nichts am Ergebnis und bringt nur numerische Nachteile mit sich.)

Die Differentialgleichung für die Ausgangsgröße $y_1(t)$ lautet in dem Beispiel

$$y_1(t) = -a_{1,2}\ddot{y}_1(t) - a_{1,1}\dot{y}_1(t) + b_{11,1}\dot{u}_1(t) + b_{11,0}u_1(t) + b_{12,0}u_2(t).$$

Analog zu dem in Abschnitt 9.1.6 beschriebenen Vorgehen für den SISO-Fall wird der Gleichungsfehler

$$\hat{\epsilon}_{1,k'} = \tilde{y}_1(t_{k'}) - \hat{y}_1(t_{k'})$$

mit

$$\hat{y}_1(t) = -\hat{a}_{1,2}\ddot{\hat{y}}_1(t) - \hat{a}_{1,1}\dot{\hat{y}}_1(t) + \hat{b}_{11,1}\dot{\hat{u}}_1(t) + \hat{b}_{11,0}u_1(t) + \hat{b}_{12,0}u_2(t)$$

aufgestellt.

Daraus ergibt sich das LS-Problem mit den Größen

$$\tilde{\eta}_{k'} = \tilde{y}_1(t_{k'})$$

und

$$\tilde{\Psi}_{k'}^T = [-\ddot{\tilde{y}}_1(t_{k'}) \quad -\dot{\tilde{y}}_1(t_{k'}) \quad \dot{\hat{u}}_1(t_{k'}) \quad u_1(t_{k'}) \quad u_2(t_{k'})]$$

sowie dem Parametervektor

$$\Theta = [a_{1,2} \quad a_{1,1} \quad b_{11,1} \quad b_{11,0} \quad b_{12,0}]^T.$$

Bezüglich des zweiten Ausgangs ergibt sich für die Parameter $[a_{2,2} \quad a_{2,1} \quad b_{21,1} \quad b_{21,0} \quad b_{22,0}]^T$ ein analoges LS-Problem, welches unabhängig gelöst werden kann.

Bei dem Beispielsystem

$$\mathbf{G}(s) = \begin{bmatrix} \frac{b_{11,1}s+b_{11,0}}{a_{2}s^2+a_{1}s+1} & \frac{b_{12,0}}{a_{2}s^2+a_{1}s+1} \\ \frac{b_{21,1}s+b_{21,0}}{a_{2}s^2+a_{1}s+1} & \frac{b_{22,0}}{a_{2}s^2+a_{1}s+1} \end{bmatrix}$$

sind alle Nenner identisch, und damit müssen beide Ausgänge gemeinsam betrachtet werden.

Je zu berücksichtigendem Zeitpunkt $t_{k'}$ sind also zwei Fehler

$$\hat{e}_{1,k'} = \begin{bmatrix} \tilde{y}_1(t_{k'}) - \hat{y}_1(t_{k'}) \\ \tilde{y}_2(t_{k'}) - \hat{y}_2(t_{k'}) \end{bmatrix}$$

zu minimieren, die sich analog zu oben ergeben. ($\hat{e}_{1,k'}$ ist, genauso wie die gleich verwendete Größe $\tilde{\eta}_{k'}$, ein Vektor und $\tilde{\Psi}_{k'}^T$ ist eine Matrix. Die Bezeichnung erfolgt hier aber analog zum SISO-Fall.)

Damit ist das LS-Problem durch

$$\tilde{\eta}_{k'} = \begin{bmatrix} \tilde{y}_1(t_{k'}) \\ \tilde{y}_2(t_{k'}) \end{bmatrix}$$

und

$$\tilde{\Psi}_{k'}^T = \begin{bmatrix} -\ddot{\tilde{y}}_1(t_{k'}) & -\dot{\tilde{y}}_1(t_{k'}) & \dot{u}_1(t_{k'}) & u_1(t_{k'}) & u_2(t_{k'}) & 0 & 0 & 0 \\ -\ddot{\tilde{y}}_2(t_{k'}) & -\dot{\tilde{y}}_2(t_{k'}) & 0 & 0 & 0 & \dot{u}_1(t_{k'}) & u_1(t_{k'}) & u_2(t_{k'}) \end{bmatrix}$$

mit

$$\theta = [a_2 \ a_1 \ b_{11,1} \ b_{11,0} \ b_{12,0} \ b_{21,1} \ b_{21,0} \ b_{22,0}]^T$$

gegeben. Je k' ergeben sich also zwei Einträge in dem Gesamtvektor $\tilde{\eta}_k$ bzw. zwei Zeilen in der Gesamt-matrix $\tilde{\Psi}_k$.

9.3.3 Identifikation aus Frequenzbereichsdaten

Bei der Identifikation aus dem Frequenzbereich ist davon auszugehen, dass für jeden Übertragungspfad ein gemessener Frequenzgang vorliegt.

Wird für jede Übertragungsfunktion $G_{ij}(s)$ ein eigener Nenner angesetzt ergibt sich damit die Aufgabe entsprechend viele SISO-Systeme zu identifizieren. Sind die Übertragungsfunktionen gekoppelt, so muss die Identifikation entsprechend gruppenweise erfolgen.

Es wird wieder das Beispiel

$$G(s) = \begin{bmatrix} \frac{b_{11,1}s + b_{11,0}}{a_2s^2 + a_1s + 1} & \frac{b_{12,0}}{a_2s^2 + a_1s + 1} \\ \frac{b_{21,1}s + b_{21,0}}{a_2s^2 + a_1s + 1} & \frac{b_{22,0}}{a_2s^2 + a_1s + 1} \end{bmatrix}$$

verwendet. Da hier der Nenner für alle Übertragungspfade identisch ist, muss die Identifikation in einem Schritt erfolgen.

Für den Frequenzgang des Übertragungspfades $u_1 \rightarrow y_1$ gilt

$$H_{11}(\omega) = \frac{b_{11,1} \cdot (j\omega) + b_{11,0}}{a_2 \cdot (j\omega)^2 + a_1 \cdot (j\omega) + 1}.$$

Analog zu dem Vorgehen, welches auf Gl. (9.15) im SISO-Fall geführt hat, ergibt sich für den Schätzfehler

$$\hat{e}_{11,k'} = \tilde{H}_{11}(\omega_{k'}) + a_2 \cdot (\omega_{k'})^2 \cdot H_{11}(\omega_{k'}) + a_1 \cdot (\omega_{k'}) \cdot H_{11}(\omega_{k'}) - b_{11,1} \cdot (\omega_{k'}) - b_{11,0}$$

für den Frequenzgang $H_{11}(\omega)$ beim k' -ten Frequenzpunkt $\omega_{k'}$. Entsprechend ergibt sich für die anderen Frequenzgänge

$$\begin{aligned}\hat{e}_{12,k'} &= \tilde{H}_{12}(\omega_{k'}) + a_2 \cdot (\mathrm{j}\omega_{k'})^2 \cdot H_{12}(\omega_{k'}) + a_1 \cdot (\mathrm{j}\omega_{k'}) \cdot H_{12}(\omega_{k'}) - b_{12,0} \\ \hat{e}_{21,k'} &= \tilde{H}_{21}(\omega_{k'}) + a_2 \cdot (\mathrm{j}\omega_{k'})^2 \cdot H_{21}(\omega_{k'}) + a_1 \cdot (\mathrm{j}\omega_{k'}) \cdot H_{21}(\omega_{k'}) - b_{21,1} \cdot (\mathrm{j}\omega_{k'}) - b_{21,0} \\ \hat{e}_{22,k'} &= \tilde{H}_{22}(\omega_{k'}) + a_2 \cdot (\mathrm{j}\omega_{k'})^2 \cdot H_{22}(\omega_{k'}) + a_1 \cdot (\mathrm{j}\omega_{k'}) \cdot H_{22}(\omega_{k'}) - b_{22,0}.\end{aligned}$$

Fasst man die vier Fehler zusammen, so erhält man

$$\hat{e}_{k'} = \begin{bmatrix} \hat{e}_{11,k'} \\ \hat{e}_{12,k'} \\ \hat{e}_{21,k'} \\ \hat{e}_{22,k'} \end{bmatrix} = \begin{bmatrix} \tilde{H}_{11}(\omega_{k'}) \\ \tilde{H}_{12}(\omega_{k'}) \\ \tilde{H}_{21}(\omega_{k'}) \\ \tilde{H}_{22}(\omega_{k'}) \end{bmatrix} - \begin{bmatrix} \Psi_{11,k'}^T \\ \Psi_{12,k'}^T \\ \Psi_{21,k'}^T \\ \Psi_{22,k'}^T \end{bmatrix} \cdot \hat{\Theta} = \tilde{\eta}_{k'} - \Psi_{k'}^T \hat{\Theta},$$

wobei $\hat{e}_{k'}$ sowie $\tilde{\eta}_{k'}$ Vektoren sind und $\Psi_{k'}^T$ eine Matrix ist.

Es ergeben sich also je Frequenzpunkt k' vier Messwerte

$$\tilde{\eta}_{k'} = \begin{bmatrix} \tilde{H}_{11}(\omega_{k'}) \\ \tilde{H}_{12}(\omega_{k'}) \\ \tilde{H}_{21}(\omega_{k'}) \\ \tilde{H}_{22}(\omega_{k'}) \end{bmatrix}$$

und mit dem Parametervektor

$$\Theta_{k'} = [a_2 \quad a_1 \quad b_{11,1} \quad b_{11,0} \quad b_{12,0} \quad b_{21,1} \quad b_{21,0} \quad b_{22,0}]^T$$

die vier Eingangsvektoren

$$\Psi_{k'}^T = \begin{bmatrix} -(\mathrm{j}\omega_{k'})^2 \cdot H_{11}(\omega_{k'}) & -(\mathrm{j}\omega_{k'}) \cdot H_{11}(\omega_{k'}) & \mathrm{j}\omega_{k'} & 1 & 0 & 0 & 0 & 0 \\ -(\mathrm{j}\omega_{k'})^2 \cdot H_{12}(\omega_{k'}) & -(\mathrm{j}\omega_{k'}) \cdot H_{12}(\omega_{k'}) & 0 & 0 & 1 & 0 & 0 & 0 \\ -(\mathrm{j}\omega_{k'})^2 \cdot H_{21}(\omega_{k'}) & -(\mathrm{j}\omega_{k'}) \cdot H_{21}(\omega_{k'}) & 0 & 0 & 0 & \mathrm{j}\omega_{k'} & 1 & 0 \\ -(\mathrm{j}\omega_{k'})^2 \cdot H_{22}(\omega_{k'}) & -(\mathrm{j}\omega_{k'}) \cdot H_{22}(\omega_{k'}) & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Diese sind wieder nach Gl. (9.16) getrennt in Real- und Imaginärteil in $\tilde{\eta}_k$ und Ψ_k einzutragen, so dass in $\tilde{\eta}_k$ und Ψ_k je gemessener Frequenz $\omega_{k'}$ acht Zeilen stehen.

Bezüglich der Gewichtung gilt das, was schon zum SISO-Fall gesagt wurde. Wenn das Gewicht $w_{k'} = \frac{1}{|A(\omega_{k'})|}$ verwendet wird, so ist dieses für alle vier komplexen Fehler bzw. die acht reellen Gleichungen zu dem entsprechenden k' identisch. Berücksichtigt das Gewicht auch die (relative) Unsicherheit bei der Frequenzgangsmessung, ergibt sich für jeden Fehler $\hat{e}_{ij,k'}$ ein eigenes Gewicht. Auf jeden Fall sollte nicht der Einfachheit halber eine durchschnittliche relative Unsicherheit für alle Frequenzgänge je Frequenzpunkt verwendet werden. Dies führt beim Vorhandensein von Nullstellen in einzelnen Übertragungspfaden zu Problemen. Bei Nullstellen sind Gewichte, die auf der relativen Unsicherheit der gemessenen Frequenzgänge beruhen, ohnehin kritisch.

Daher eignet sich das Gewicht $w_{k'} = \frac{1}{|A(\omega_{k'})|}$ dann eher. Die relative Unsicherheit als Gewicht kann aber zur Generierung als Startwert dienen, wenn mit einem Initialgewicht von $w_{k'} \equiv 1$ keine guten Startwerte für eine Iteration gefunden werden, da – wenn auch nicht die Zähler – doch zumindest der Nenner einigermaßen gut geschätzt werden sollte.

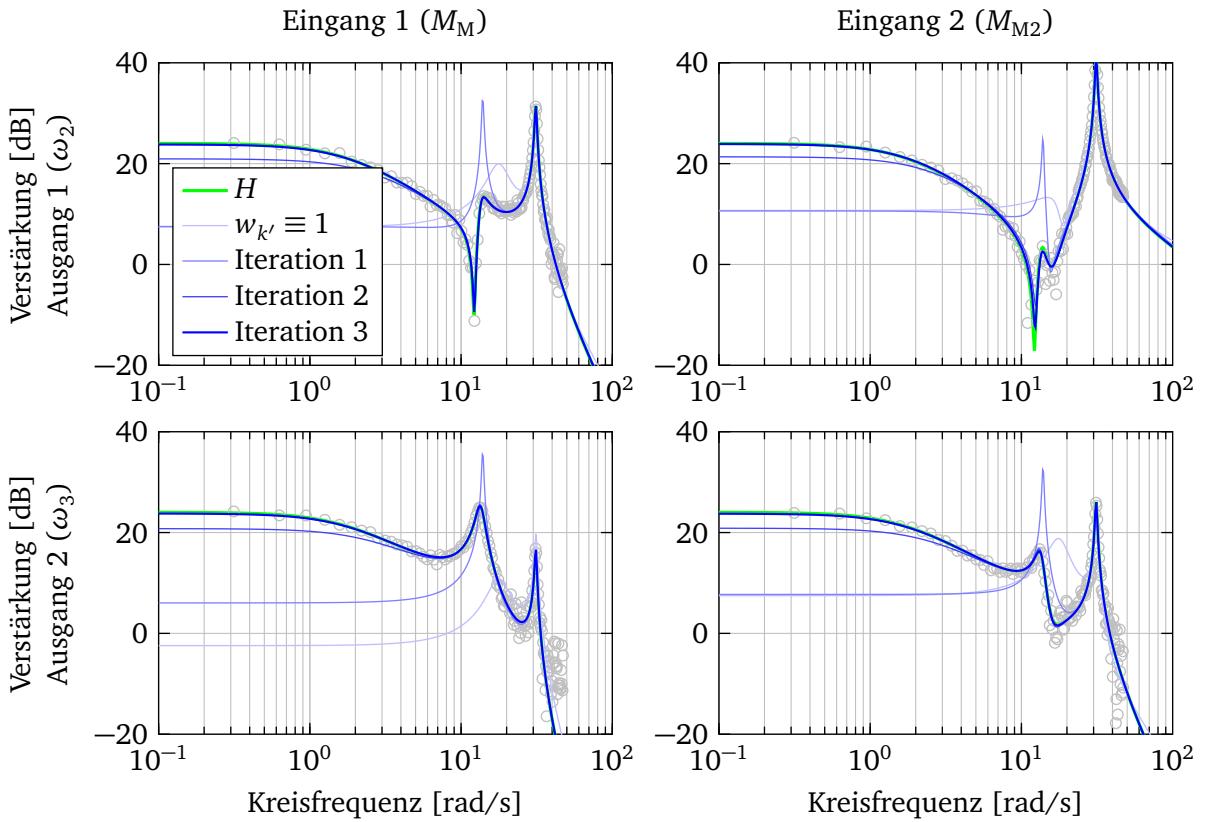


Abbildung 9.30: Identifikation eines MIMO-Modells aus gemessenen Frequenzgängen

Beispiel Drei-Massen-Schwinger

In Abbildung 9.30 ist das Ergebnis bei der Anwendung auf den Drei-Massen-Schwinger gezeigt. Der parametrischen Identifikation anhand des Ansatzes (9.20) mit 17 Parametern liegen dabei für jeden Übertragungspfad ein aus zehn einzelnen Gleitsinusanregungen geschätzter Frequenzgang zu Grunde, der jeweils als graue Punkte in den Plots eingezeichnet ist. Im Hintergrund ist als breite grüne Linie der tatsächliche Frequenzgang eingezeichnet.

Die erste Schätzung ist mit $w_{k'} \equiv 1$, also ungewichtet durchgeführt. Das Ergebnis ist als hellblaue Linie dargestellt und es zeigt sich, dass dieses sehr schlecht ist.

Anschließend sind noch drei Iterationen (in dunkler werdendem Blau) mit dem Gewicht $w_{k'} = \frac{1}{|\hat{A}(\omega_{k'})|}$ durchgeführt, wobei $\hat{A}(\omega_{k'})$ jeweils aus dem Ergebnis des vorhergehenden Schrittes bestimmt ist. Nach diesen drei Iterationen zeigt sich ein sehr gutes Schätzergebnis.

Eine Schätzung anhand Gl. (9.19), dem Modellansatz mit 25 Parametern ohne vorherige Berücksichtigung der Differenzenordnungen der einzelnen Übertragungspfade, führt im hier gezeigten Frequenzbereich zu einem sehr ähnlichen Ergebnis. Jedoch müsste man in diesem Fall, um zu einem guten Modell zu kommen, die „nicht-dominanten Nullstellen“ der einzelnen Übertragungspfade per Hand entfernen, insbesondere die nicht-minimalphasigen, die bei diesem System anschaulich nicht passen.

9.4 Identifikation im geschlossenen Regelkreis

Teilweise muss die Identifikation eines Systems G im geschlossenen Regelkreis erfolgen. Dies ist zwangsläufig der Fall, wenn das System instabil ist, aber auch wenn die Identifikation aus anderen Gründen

im normalen, geregelten Betrieb erfolgen soll. Damit ergibt sich in diesen Fällen eine Struktur wie in Abbildung 9.31 dargestellt, wobei G_R ein (dynamischer) Regler ist und w_R die Sollgröße ist.

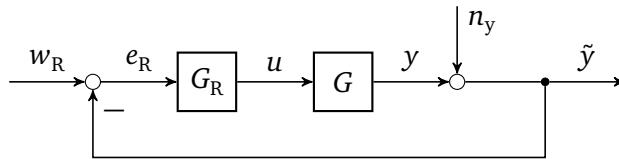


Abbildung 9.31: Geschlossener Regelkreis

Es gibt dazu noch Variationen, auf die hier aber nicht weiter eingegangen werden soll, siehe z. B. [Ljung 1999, Kap. 13.4, 13.5] und [Isermann und Münchhof 2011, Kap. 13]. Hier soll nur ein kurzer Überblick über die Möglichkeiten bzw. Problemfelder der Identifikation im geschlossenen Regelkreis gegeben werden.

Direkte und indirekte Identifikation

Indirekte Identifikation

Bei der indirekten Identifikation wird zunächst das Gesamtsystem des geschlossenen Regelkreises vom Eingang w_R zum Ausgang \tilde{y} identifiziert. Dies ist in Abbildung 9.32 (für einen linearen Regler und eine lineare Strecke) dargestellt. Es können dafür alle bisher behandelten Methoden verwendet werden, ggf. ist auf das angepasste Rauschmodell zu achten.

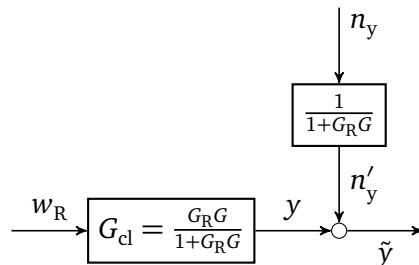


Abbildung 9.32: Indirekte Identifikation

Aus dem Modell G_{cl} für das Gesamtsystem muss dann das eigentlich interessierende Teilmодell extrahiert werden, wofür der Regler genau bekannt sein muss. Problematisch dabei ist auch, dass viele in der Praxis vorhandene, im Prinzip einfache Regler doch über viele Nichtlinearitäten wie Begrenzungen der Stellgröße und Änderungsrate der Stellgröße oder Anti-Windup-Maßnahmen verfügen [Ljung 1999, S. 435].

Direkte Identifikation

Bei der direkten Identifikation wird neben dem gestörten Systemausgang \tilde{y} noch die Eingangsgröße u gemessen. Dies hat den Vorteil, dass der Regler nicht bekannt sein muss. Dafür können nicht mehr alle Identifikationsverfahren einfach angewandt werden.

So verstärkt sich die Korrelation zwischen dem Rauschen n_y und dem LS-Eingangsvektor ψ , da dann auch die Eingangsgröße u des Systems mit n_y korreliert ist. Bei niedrigen Rauschlevel und hohen Sollgrößensprüngen (also einem hohen Nutzsignalanteil) kann dies noch ignoriert werden. Andernfalls ist dies aber zu berücksichtigen.

Anregung

Ausreichende Anregung

Für eine ausreichende Anregung ist es nicht unbedingt genügend, den Regler als Festwertregler um einen Arbeitspunkt zu betreiben. Dies zeigt schon das folgende, [Ljung 1999, S. 430f] entnommene Beispiel.

Es soll die Strecke

$$G(z) = \frac{b_0 z^{-1}}{1 + a_0 z^{-1}}$$

identifiziert werden, wobei während der Messung ein P-Regler

$$u_{k'} = -K_P \cdot (\tilde{y}_{k'} + w_{R,k'})$$

verwendet wird.

Ist die Sollgröße $w_{R,k'}$ während der gesamten Messung gleich 0, so gilt die Differenzengleichung

$$\hat{y}_{k'} = -a_0 \tilde{y}_{k'-1} + b_0 u_{k_1} = -a_0 \tilde{y}_{k'-1} - b_0 K_P \tilde{y}_{k'-1}$$

und damit erhält man, für $\Theta = [a_0 \quad b_0]^T$, bei der direkten Identifikation den Eingangsvektor

$$\tilde{\Psi}_{k'}^T = [-\tilde{y}_{k'-1} \quad -K_P \tilde{y}_{k'-1}] .$$

Die sich aus diesen Vektoren ergebenden Spalten der Matrix $\tilde{\Psi}_k$ sind linear abhängig und damit ist die Matrix $\mathbf{L}_k = \tilde{\Psi}_k^T \tilde{\Psi}_k$ nicht invertierbar.

Über Variationen der Sollgröße $w_{R,k'}$ ist es aber möglich, eine ausreichende Anregung sicherzustellen.

In [Ljung 1999, S. 432] wird dieses Thema noch genauer behandelt und gezeigt, dass auch Regler ausreichend hoher Ordnung, zeitvariante oder nichtlineare Regler eine ausreichende Anregung sicherstellen können, auch wenn $w_{R,k'} \equiv 0$ ist.

Allgemeiner Einfluss des Reglers

Ein weiterer von [Ljung 1999, S. 434, 440] herausgehobener Aspekt ist der, dass durch den Regler letztlich weniger Informationen über das System gemessen werden.

Ist beispielsweise ein Regler für den Drei-Massen-Schwinger als Kompensationsregler ausgelegt, der das konjugiert-komplexe Polpaar näherungsweise kompensiert, dann bedeutet das, dass der Regler diese Frequenzen gar nicht mehr auf das System gibt, und die Identifikation in diesem Frequenzbereich schlechter sein wird.

9.5 Anmerkungen zum Aufstellen von LS-Problemen

Zum Teil liegen zusätzliche Informationen über das System vor, welche in der Schätzung berücksichtigt werden können. Damit ergeben sich etwas andere LS-Probleme als die bisher hier behandelten. Das Aufstellen dieser Probleme ist aber relativ einfach (sofern möglich, d. h. die LS-Prozesse linear in den Parametern sind). Dies ist in diesem Abschnitt an einem einfachen Beispiel gezeigt. Dieses Vorgehen lässt sich auch allgemein zum Aufstellen von LS-Problemen verwenden, ob es sich um dynamische oder statische Anwendungsfälle handelt.

Insbesondere bei kleineren zeitkontinuierlichen Systemen ist es teilweise sinnvoll, den Ansatz nicht mit den allgemeinen Koeffizienten a_j und b_j zu schreiben, sondern direkt mit physikalischen Parametern. Dabei sind dann manche physikalischen Parameter teilweise relativ gut bekannt, wie z. B. Massen. Diese können dann auf die bekannten festen Werte gesetzt werden, um so die Schätzung der unsicheren Parameter wie z. B. Dämpfungen und Federsteifigkeiten zu verbessern.

Als einfaches Beispiel wird der Fall betrachtet, dass ein zeitdiskretes PT₁-Glied,

$$G(z) = \frac{b_0 z^{-1}}{1 + a_0 z^{-1}},$$

identifiziert werden soll. Es wird jetzt aber davon ausgegangen, dass dessen stationäre Verstärkung

$$K = \frac{b_0}{1 + a_0}$$

bekannt ist. Mit dieser Gleichung kann einer der beiden Parameter durch den anderen ausgedrückt werden, so z. B.

$$b_0 = K + K \cdot a_0.$$

Setzt man dies in die Übertragungsfunktion ein, erhält man

$$G(z) = \frac{(K + K a_0) \cdot z^{-1}}{1 + a_0 z^{-1}}$$

und daraus die Differenzengleichung

$$y_{k'} = -a_0 y_{k'-1} + (K + K a_0) \cdot u_{k'-1}.$$

Es ist ein Schätzer für den Parameter a_0 gesucht. Dies erfolgt schrittweise nach folgendem Schema.

1. Sinnvollen Fehler definieren:

Es ist ein Fehler festzulegen, dessen Minimierung sinnvoll ist. In diesem Fall wird sich an dem Gleichungsfehler orientiert.

Dazu wird der auf den zurückliegendem Messwert basierende Schätzwert

$$\hat{y}_{k'} = -\hat{a}_0 \tilde{y}_{k'-1} + (K + K \hat{a}_0) \cdot u_{k'-1}$$

aufgestellt und damit der Fehler

$$\begin{aligned}\hat{e}_{k'} &= \tilde{y}_{k'} - \hat{y}_{k'} \\ &= \tilde{y}_{k'} + \hat{a}_0 \tilde{y}_{k'-1} - (K + K \hat{a}_0) \cdot u_{k'-1}\end{aligned}$$

gebildet.

2. Aufteilen der einzelnen Bestandteile der Gleichung:

Alles, was nicht von den Parametern abhängt (abgesehen von dem Fehler) nach links, alles andere nach rechts.

$$\tilde{y}_{k'} + K u_{k'-1} = \hat{a}_0 \cdot (\tilde{y}_{k'-1} - K u_{k'-1}) + \hat{e}_{k'}$$

Auf der rechten Seite muss für jeden zu schätzenden Parameter genau ein Produktterm vorliegen. Ist eine solche Aufteilung nicht möglich, kann versucht werden, neue Parameter zu definieren. Dabei ist aber darauf zu achten, dass sich die alten Parameter eindeutig aus den neuen bestimmen lassen.

3. Ablesen der Größen des LS-Problems:

$$\underbrace{\tilde{y}_{k'} + Ku_{k'-1}}_{\tilde{\eta}_{k'}} = \underbrace{(\tilde{y}_{k'-1} - Ku_{k'-1}) \cdot \hat{a}_0}_{\Psi_k^T \hat{\theta}} + \hat{e}_{k'}$$

- Alles was nicht von den Parametern abhängt, ist die gemessene Ausgangsgröße des LS-Problems, hier also die linke Seite

$$\tilde{\eta}_{k'} = \tilde{y}_{k'} + Ku_{k'-1}.$$

- Der Parametervektor erklärt sich von selbst,

$$\theta = [a_0].$$

- Für jeden Parameter (hier nur a_0) ist alles, was mit diesem multipliziert wird, die entsprechende LS-Eingangsgröße,

$$\Psi_k^T = [\tilde{y}_{k'-1} - Ku_{k'-1}].$$

Dabei spielt es keine Rolle, ob Ein- oder Ausgangsgrößen des physikalischen Systems ($u_{k'}, \tilde{y}_{k'}$) nichtlinear oder in mehr als einem Eintrag in Ψ^T auftauchen. (Dies wäre ggf. bei einer Analyse der stochastischen Eigenschaften zu beachten, aber das LS-Problem kann man in jedem Fall aufstellen und lösen.)

Auf diese Weise kann insbesondere mit konstanten Offsets und bekannten Parameterwerten umgegangen werden. Letzteres meint, dass manche der angesetzten Parameter nicht identifiziert werden müssen, sondern schon bekannt sind.

Sind von einem schon aufgestellten LS-Problem ein oder mehrere Parameter bekannt bzw. auf einen festen Wert zu setzen, so kann dies auch wie folgt durchgeführt werden. Dazu wird der Parametervektor (ggf. nach umsortieren) in die Form

$$\theta = \begin{bmatrix} \theta_f \\ \theta_e \end{bmatrix}$$

gebracht, wobei θ_f die festen bzw. bekannten Parameter bezeichnet („fixed“), und θ_e die noch zu schätzenden Parameter („estimate“). Der Eingangsvektor teilt sich entsprechend auf,

$$\Psi_k^T = [\Psi_{f,k}^T \quad \Psi_{e,k}^T].$$

Diese Aufteilung wird auch für die in Ψ_k zusammengefassten Eingangsvektoren verwendet, $\Psi_k = [\Psi_{f,k} \quad \Psi_{e,k}]$.

Die LS-Prozessgleichung lautet dann

$$\eta_{k'} = \Psi^T \theta = \Psi_{f,k}^T \theta_f + \Psi_{e,k}^T \theta_e,$$

was zu

$$\eta'_{k'} = \eta_{k'} - \Psi_{f,k}^T \theta_f = \Psi_{e,k}^T \theta_e$$

umgestellt werden kann, wobei die neue Ausgangsgröße $\eta'_{k'} = \eta_{k'} - \Psi_{f,k}^T \theta_f$ eingeführt ist. Damit liegt wieder ein normaler „LS-Prozess“ vor.

Mit den neuen Messgrößen

$$\tilde{\eta}'_{k'} = \tilde{\eta}_{k'} - \Psi_{f,k'}^T \Theta_f$$

und der Schätzung

$$\hat{\eta}'_{k'} = \Psi_{e,k'}^T \hat{\Theta}_e$$

liegt mit

$$\hat{e}_{k'} = \tilde{\eta}'_{k'} - \hat{\eta}'_{k'}$$

ein normales LS-Problem vor, dessen Lösung durch

$$\begin{aligned}\hat{\Theta}_{e,k} &= (\Psi_{e,k}^T \Psi_{e,k})^{-1} \Psi_{e,k}^T \tilde{\eta}'_k \\ &= (\Psi_{e,k}^T \Psi_{e,k})^{-1} \Psi_{e,k}^T \cdot (\tilde{\eta}_k - \Psi_{f,k} \Theta_f)\end{aligned}$$

gegeben ist.

10 Kalmanfilter

10.1 Parameteridentifikation über Zustandsschätzung

Bei dem Kalmanfilter handelt es sich um einen Zustandsbeobachter. Wie Zustandsbeobachter zur Parameteridentifikation verwendet werden können, wird im Folgenden gezeigt.

Als einfaches (lineares) System wird eine Drehträgheit betrachtet die reibungsbehaftet gelagert ist und auf die ein Moment $M(t)$ aufgebracht werden kann. Die Reibung wird als lineare, viskose Reibung angenommen. Ausgangsgröße ist der gemessene Winkel $\varphi(t)$. Damit ergibt sich die Zustandsraumdarstellung

$$\begin{bmatrix} \dot{\varphi}(t) \\ \dot{\omega}(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & -\frac{d}{J} \end{bmatrix} \cdot \begin{bmatrix} \varphi(t) \\ \omega(t) \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{J} \end{bmatrix} \cdot M(t)$$
$$y(t) = [1 \ 0] \cdot \begin{bmatrix} \varphi(t) \\ \omega(t) \end{bmatrix}$$

Mit einem Beobachter, wie z. B. einem Kalmanfilter, lassen sich die Systemzustände φ und ω aus dem bekannten Eingangsgrößenverlauf $M(t)$ und den Messungen des Winkels $\varphi(t)$ bestimmen, wobei der Kalmanfilter auch Mess- und Systemrauschen explizit berücksichtigt.

Wenn nun auch die Dämpfung d nicht bekannt ist, so kann für diese ein weiterer Zustand eingefügt werden. Um dies deutlich zu machen, wird dieser nicht als d , sondern x_d bezeichnet. Die Zustandsgleichung für diesen Zustand lautet

$$\dot{x}_d(t) = 0,$$

was bedeutet, dass davon ausgegangen wird, dass die Dämpfung konstant ist und weder von den anderen Zuständen noch der Eingangsgröße abhängt.

Mit diesem neuen Zustand ergibt sich die Zustandsraumdarstellung

$$\begin{bmatrix} \dot{\varphi}(t) \\ \dot{\omega}(t) \\ \dot{x}_d(t) \end{bmatrix} = \begin{bmatrix} \omega(t) \\ \frac{-1}{J} \cdot x_d(t) \cdot \omega(t) + \frac{1}{J} \cdot M(t) \\ 0 \end{bmatrix}$$
$$y(t) = \varphi(t).$$

Diese ist nichtlinear, was typisch für den Fall ist, dass Parameter als Zustände geschrieben werden.

Verwendet man einen geeigneten Schätzer für dieses nichtlineare System, so erhält man neben den Werten der ursprünglichen Zustände φ und ω auch einen Schätzwert für die Dämpfung $x_d = d$.

Das Kalmanfilter ist ein Beobachter für lineare Systeme. Mit dem erweiterten Kalmanfilter (extended Kalmanfilter) existiert aber auch eine Anpassung auf nichtlineare Systeme.

Beobachtbarkeit

Bei nichtlinearen Systemen kann die Beobachtbarkeit auch von der Eingangsgröße bzw. deren Verlauf abhängen. Dies bedeutet für die Systemidentifikation, dass teilweise eine Anregung nötig ist, um die als Zustand geschriebenen Parameter beobachten (also identifizieren) zu können.

10.2 Hinführendes Beispiel

Als einführendes Beispiel wird ein Messbecher betrachtet, der mit Wasser gefüllt wird. Die Füllung erfolgt bis zu einem Eichstrich, so dass die Menge des Wassers mit einer entsprechenden Unsicherheit bekannt ist.

Ausgangspunkt (Startwert)

Die Masse an Wasser direkt nach der Füllung des Bechers wird mit x_0 bezeichnet. Der genaue Wert ist unbekannt, aber durch eine sorgfältige Füllung geht man davon aus, dass sich 1 kg Wasser im Becher befinden sollte. Die angenommene Masse, der Schätzwert, wird mit \hat{x}_k bezeichnet. Es gilt also $\hat{x}_0 = 1 \text{ kg}$. Die Unsicherheit dieser Angabe wird über eine Standardabweichung $\sigma_{\hat{x},0} = 0,01 \text{ kg}$ abgeschätzt.

Berücksichtigung einer Messung

Um den Schätzwert für die Wassermasse zu verbessern soll eine Messung vorgenommen werden. Der Messwert wird mit \tilde{y} bezeichnet und entspricht dem mit einem Messfehler gestörten Wert für x ,

$$\tilde{y}_k = x_k + n_{y,k} .$$

Der Messfehler $\{n_y\}_r$ sei mittelwertfrei und besitze die Varianz $\sigma_{y,k}^2$, die von Messung zu Messung verschieden sein kann.

Wird der Becher gewogen und die Masse des leeren Bechers vom Messwert abgezogen, so sollte dieser Wert relativ genau der Masse des Wassers entsprechen. Die Standardabweichung des Messfehlers der Waage betrage hier $\sigma_{y,w} = 0,001 \text{ kg}$. Wird dagegen eine Messung derart durchgeführt, dass die Füllhöhe gemessen wird und damit ein „Messwert“ für die Masse des Wassers berechnet wird, dann ist dieser Wert deutlich unsicherer. Dessen Unsicherheit liege im Bereich von $\sigma_{y,F} = 0,01 \text{ kg}$.

Im ersten Fall sollte es eher sinnvoll sein, den alten, initialen Schätzwert vollständig zu ersetzen. Im zweiten Fall sollte eine gleichberechtigte Mittelung günstiger sein.

Mathematisch ausgedrückt stellt sich die Frage, wie zwei Werte $\hat{x}_{1,-} = \hat{x}_0$ und \tilde{y}_1 mit den Varianzen $\sigma_{\hat{x},1,-}^2$ bzw. $\sigma_{y,1}^2$ optimal gemittelt werden sollen.¹ Da beide Werte per Annahme erwartungstreu sind, ist eine gewichtete Mittelung der Form

$$\hat{x}_{1,+} = (1 - \gamma) \cdot \hat{x}_{1,-} + \gamma \tilde{y}_1 \quad (10.1)$$

sinnvoll, die diese Eigenschaft auch für den gemittelten Wert erhält. Als weiteres Kriterium für die Berechnung kann die Varianz $\sigma_{\hat{x},1,+}^2$ des gemittelten Wertes dienen. Diese sollte minimal sein.

Diese Fragestellung wurde schon im Kapitel über die stochastischen Grundlagen behandelt, mit dem Ergebnis, dass das Optimum im genannten Sinne durch

$$\gamma = \frac{\sigma_{\hat{x},1,-}^2}{\sigma_{\hat{x},1,-}^2 + \sigma_{y,1}^2}$$

erreicht wird.

Um diese Mittelung optimal durchführen zu können, ist die Kenntnis der Varianz des Schätzwertes vor der Mittelung nötig. Um für weitere Messungen gewappnet zu sein, sollte also nicht nur der neue Schätzwert, sondern auch dessen Varianz

$$\sigma_{\hat{x},1,+}^2 = (1 - \gamma)^2 \cdot \sigma_{\hat{x},1,-}^2 + \gamma^2 \sigma_y^2$$

berechnet werden.

¹ $\hat{x}_{1,-}$ bezeichnet den Schätzwert vor der ersten Messung, und $\hat{x}_{1,+}$ ist dann der Wert nach der ersten Messung.

Berücksichtigung einer Messung unter Betrachtung des Schätzfehlers

Genau genommen möchte man nicht die Varianz des Schätzwertes minimieren, sondern die Varianz des Schätzfehlers

$$\Delta\hat{x}_{1,+} = x_1 - \hat{x}_{1,+} . \quad (10.2)$$

Solange der wahre Werte x_1 deterministisch ist, wie bisher angenommen, macht dies keinen Unterschied. Wenn aber zugelassen werden soll, dass sich auch der wahre Zustand zufällig ändern kann, so ist die Betrachtung des Schätzfehlers notwendig.

Den Schätzfehler nach der Mittelung erhält man, wenn man in Gl. (10.2) $\hat{x}_{1,+}$ durch Gl. (10.1) ersetzt. Mit ein paar Umformungen erhält man über

$$\begin{aligned}\Delta\hat{x}_{1,+} &= x_1 - \hat{x}_{1,+} = x_1 - (1 - \gamma) \cdot \hat{x}_{1,-} - \gamma \cdot (x_1 + n_{y,1}) \\ &= (1 - \gamma) \cdot (x_1 - \hat{x}_{1,-}) - \gamma n_{y,1}\end{aligned}$$

den Ausdruck

$$\Delta\hat{x}_{1,+} = (1 - \gamma) \cdot \Delta\hat{x}_{1,-} - \gamma n_{y,1} . \quad (10.3)$$

Dieser Ausdruck ist Gl. (10.1) sehr ähnlich.

Wenn die Varianz des Schätzfehlers vor der Mittelung mit $\sigma_{\Delta\hat{x},1,-}^2$ bezeichnet wird, so beträgt das optimale Gewicht

$$\gamma = \frac{\sigma_{\Delta\hat{x},1,-}^2}{\sigma_{\Delta\hat{x},1,-}^2 + \sigma_{y,1}^2} . \quad (10.4)$$

Die Mittelung selber wird mit diesem γ über Gl. (10.1) vorgenommen, d. h. das Gewicht zur Mittelung des Schätzwertes wird über die Betrachtung der Varianz des Schätzfehlers bestimmt.

Der Schätzfehler selber ist zu keinem Zeitpunkt bekannt. (Ansonsten würde man auch den genauen gesuchten Wert kennen.) Jedoch kann man die Varianz des Schätzfehlers angeben, solange die Varianz des Schätzfehlers zu Beginn bekannt ist. Dies ist auch wichtig, da diese Varianz in Gl. (10.4) zur Berechnung des optimalen Gewichts benötigt wird.

Aus Gl. (10.3) erhält man mit den Regeln zur Rechnung mit Varianzen und der Annahme, dass $\{\Delta\hat{x}_{1,-}\}_r$ nicht mit $\{n_{y,1}\}_r$ korreliert ist die Gleichung

$$\sigma_{\Delta\hat{x},1,+}^2 = (1 - \gamma)^2 \cdot \sigma_{\Delta\hat{x},1,-}^2 + \gamma^2 \sigma_{y,1}^2 .$$

Diese Annahme ist dann erfüllt, wenn der Prozess $\{n_{y,k}\}_r$ in sich unkorelliert ist.

Berücksichtigung einer weiteren Messung

Wird eine weitere Messung \tilde{y}_2 vorgenommen, so kann diese genauso berücksichtigt werden. D. h. es wird $\hat{x}_{2,-} = \hat{x}_{1,+}$ und entsprechend $\sigma_{\Delta\hat{x},2,-}^2 = \sigma_{\Delta\hat{x},1,+}^2$ gesetzt, und damit der neue Schätzwert $\hat{x}_{2,+}$ sowie die Varianz des Schätzfehlers $\sigma_{\Delta\hat{x},2,+}^2$ nach diesem Schritt bestimmt.

Zeitschritt

Die nächste Messung soll 24 Stunden später erfolgen. Innerhalb dieser Zeit wird etwas Wasser verdunstet sein, so dass der alte Schätzwert $\hat{x}_{2,+}$ nicht mehr korrekt ist. Würde man einfach nach den bisherigen

Vorgehen den Schätzwert $\hat{x}_{2,+}$ mit dem dritten Messwert \tilde{y}_3 mitteln, so macht man einen systematischen Fehler.

Wenn Wissen über die Systemdynamik, d. h. hier der Verdunstungsvorgang, existiert, so ist es sinnvoll dieses einfließen zu lassen.

Angenommen, man hätte das zeitdiskrete Modell (Abtastzeit 24 Stunden)

$$x_{k+1} = 0,9 \cdot x_k \quad (10.5)$$

für dieses System vorliegen, nach dem je Tag 10 % der Wassermasse verdunstet.²

Dann ist es sofort einsichtig, dass auch der Schätzwert entsprechend angepasst werden sollte, d. h.

$$\hat{x}_{3,-} = 0,9 \cdot \hat{x}_{2,+} . \quad (10.6)$$

Damit ändert sich aber auch die Varianz des Schätzwertes und damit die des Schätzfehlers. (Wenn $\hat{x}_{3,-}$ das 0,9-fache von $\hat{x}_{2,+}$ ist, so werden auch alle Streuungen von $\{\hat{x}_{2,+}\}_r$ mit 0,9 multipliziert und damit verringert.)

Um die Varianz des Schätzfehlers anzugeben, wird zunächst eine Gleichung für den Schätzfehler abgeleitet. Dazu wird in $\Delta\hat{x}_{3,-} = x_3 - \hat{x}_{3,-}$ der linke Summand x_3 durch die Zustandsgleichung (10.5) des Systems, und der rechte Summand $\hat{x}_{3,-}$ durch Gl. (10.6) ersetzt,

$$\begin{aligned} \Delta\hat{x}_{3,-} &= 0,9 \cdot x_2 - 0,9 \cdot \hat{x}_{2,+} \\ &= 0,9 \cdot \Delta\hat{x}_{2,+} . \end{aligned}$$

Hieraus ergibt sich direkt

$$\sigma_{\Delta\hat{x},3,-}^2 = 0,9^2 \cdot \sigma_{\Delta\hat{x},2,+}^2 .$$

Es ist mit $\hat{x}_{3,-}$ und $\sigma_{\Delta\hat{x},3,-}^2$ der Schätzwert für die Masse und die Varianz des Schätzfehlers nach dem ersten Zeitschritt von 24 Stunden bekannt. Damit kann die Mittelung des Schätzwertes $\hat{x}_{3,-}$ mit dem Messwert \tilde{y}_3 im Sinne der optimalen Varianz erfolgen.

Zeitschritt mit Rauschen

Oben ist ein sehr einfaches Modell für das System angenommen. Die Verdunstung wird z. B. auch von der relativen Luftfeuchte der Umgebung abhängen.

Befindet sich das System in einem klimatisierten Raum, so sind die Umgebungsbedingungen kontrolliert, es können aber ggf. kurzzeitige Störungen auftreten. Diese könnten durch einen entsprechenden Störterm als Systemrauschen berücksichtigt werden. Die Zustandsgleichung lautet dann

$$x_{k+1} = 0,9 \cdot x_k + n_{x,k} ,$$

wobei $\{n_{x,k}\}_r$ in der Regel als mittelwertfreies, in sich unkorreliertes Rauschen mit der Varianz $\sigma_{x,k}^2$ angenommen wird.

Die Mittelung mit neuen Messwerten wird von dieser Änderung der Systemgleichung nicht beeinflusst, jedoch die Berechnung, wie sich die Kovarianz des Schätzfehlers über die Zeit ändert.

² Wahrscheinlich hängt die Verdunstung weniger von der Masse an Flüssigkeit, sondern im Wesentlichen von der Querschnittsfläche der offenen Oberfläche ab. Man kann sich jetzt überlegen, wie das Gefäß aussehen müsste, damit die Oberfläche proportional zur Wassermasse ist. Das Beispiel dient hier aber mehr als Gedankenexperiment.

Für die Übertragung des Schätzwerts über den Zeitwert ändert sich nichts, d. h. in dem Beispiel würde man wieder

$$\hat{x}_{3,-} = 0,9 \cdot \hat{x}_{2,+}$$

verwenden, da bei einem mittelwertfreien Rauschen $\{n_{x,k}\}_r$ dies immer noch den erwartungstreuen Schätzwert $\{\hat{x}_{3,-}\}_r$ liefert. (Solange $\{\hat{x}_{2,+}\}_r$ erwartungstreu ist.)

Für den Schätzfehler ergibt sich jetzt aber

$$\begin{aligned}\Delta\hat{x}_{3,-} &= 0,9 \cdot x_{2,+} + n_{x,2} - 0,9 \cdot \hat{x}_{2,+} \\ &= 0,9 \cdot \Delta\hat{x}_{2,+} + n_{x,2},\end{aligned}$$

und damit erhält man

$$\sigma_{\Delta\hat{x},3,-}^2 = 0,9^2 \cdot \sigma_{\Delta\hat{x},2,+}^2 + \sigma_{x,2}^2.$$

Die Varianz des neuen Schätzwertes erhöht sich also um die Varianz des Systemrauschens.

Je größer die Varianz $\sigma_{x,k}^2$ des Systemrauschens ist, desto höher ist auch die Varianz des Schätzfehlers nach dem Zeitschritt, und desto stärker wird dieser durch neue Messdaten korrigiert.

Dieses Vorgehen ist sehr nützlich, wenn das Systemverhalten in einem gewissen Maß zufällig ist, jedoch sollte es nicht missbraucht werden, um eine zu ungenaue Modellierung des vorliegenden Systems zu kompensieren. So könnte man auch die Idee kommen,

$$x_{k+1} = n_{x,k}$$

anzusetzen. Dabei wird die Masse der Flüssigkeit prinzipiell als konstant angesehen, und jede – auch deterministische – Änderung der Masse wird durch das Rauschen erklärt.

Allerdings führt dies zu schlechteren Ergebnissen als der obige Ansatz. Zum einen sind die 90 % Masseverlust pro Tag nicht unsicher, sondern relativ sicher bekannt. Die Varianz des Schätzfehlers wird damit unnötig schlechter gerechnet als sie ist, und die Mittelung damit zugunsten der Messungen verschoben.

Auch ist im vorliegenden Fall die Annahme, dass das Systemrauschen mittelwertfrei und unkorreliert ist, nicht erfüllt. Das Rauschen wird hier immer negative Werte annehmen, und sich von Tag zu Tag auch nicht stark unterscheiden. Damit ist z. B. der sich ergebende Schätzwert nicht mittelwertfrei, und auch die Aussagen bezüglich der Optimalität des gesamten Schätzvorganges, auf die später eingegangen wird, gelten nicht mehr.

Zeitschritt mit Eingangsgröße

Wenn die relative Luftfeuchtigkeit gemessen werden kann, so kann diese auch als Eingangsgröße berücksichtigt werden. Linearisiert³ könnte man z. B. das System

$$x_{k+1} = 0,9 \cdot x_k + 0,05 \cdot u_k + n_{x,k}$$

erhalten.

Berücksichtigt man die Wirkung der Eingangsgröße auch bei der Prädiktion des Schätzwertes, so erfolgt diese nach der Gleichung

$$\hat{x}_{3,-} = 0,9 \cdot \hat{x}_{2,+} + 0,05 \cdot u_2.$$

³ Ohne den Zusammenhang zwischen Luftfeuchte und Verdunstung nachzuschlagen, kann man sicher davon ausgehen, dass dies eine sehr starke Vereinfachung darstellt. Aber für den Zweck dieses Beispiels spielt das keine Rolle.

Für den Schätzfehler ändert sich damit nichts, was

$$\begin{aligned}\Delta \hat{x}_{3,-} &= 0,9 \cdot x_{2,+} + 0,05 \cdot u_2 + n_{x,2} - 0,9 \cdot \hat{x}_{2,+} - 0,05 \cdot u_2 \\ &= 0,9 \cdot \Delta \hat{x}_{2,+} + n_{x,2}\end{aligned}$$

zeigt. Damit bleibt auch die Berechnung der Varianz des Schätzfehler unverändert.

Eingangsgrößen müssen demnach lediglich in der Prädiktion des Schätzwertes auf naheliegende Art berücksichtigt werden.

Zeitschritt mit genauerem Systemmodell

Wird die relative Luftfeuchte wie gerade beschrieben als Eingangsgröße betrachtet, so kann sich herausstellen, dass ein diskretes Modell mit einer Abtastzeit von einem Tag für den vorliegenden Fall etwas zu ungenau ist. Die Luftfeuchte ändert sich ja über den Tag, und mit einer einzigen Messung eine Vorhersage über das Verhalten über einen ganzen Tag zu machen ist etwas viel verlangt.

Es spricht dann aber nichts dagegen, die Abtastzeit auf ein sinnvolles Maß zu verkürzen. Das bedeutet dann aber auch nicht, dass deswegen mehr Messungen der Ausgangsgröße durchgeführt werden müssen. Wenn die Abtastzeit des Systemmodells eine Stunde beträgt, dann können auch 24 Prädiktionsschritte durchgeführt werden, bevor wieder eine Messung der Ausgangsgröße berücksichtigt wird.

Im Grenzfall kann auch ein zeitkontinuierliches Modell verwendet werden. Dabei ist dann auch die Varianz des Schätzfehlers zeitkontinuierlich zu berechnen.

Auch könnte ein nichtlineares Modell verwendet werden, wobei hier die Varianz des Schätzfehlers dann geeignet zu berechnen ist.

Letztlich geht es nur darum, dass System über den Zeitschritt möglichst genau zu simulieren, und dabei auch die Varianz des Schätzfehlers zu verfolgen.

Fazit

Eine optimale Berücksichtigung der Messwerte beruht auf der Berücksichtigung der Varianzen von Messwert und bisherigem Schätzfehler. Diese müssen daher bekannt sein. Die Varianz des Messwertes ergibt sich aus den Eigenschaften der Messung bzw. Annahmen darüber. Die Varianz des Schätzfehlers folgt aus der gesamten „Schätzgeschichte“ und muss ständig mitgeführt und aktualisiert werden, ausgehend von einer Annahme über die Varianz des Schätzfehlers zu Beginn.

Dieses Vorgehen, Größen im Zeitverlauf zu schätzen, lässt sich also in zwei Schritte unterteilen. Liegt eine neue Messung vor, so wird ein

- Korrekturschritt

durchgeführt. Zwischen den Messungen muss das System (inklusive dessen stochastischen Eigenschaften) simuliert werden, im Kontext des Kalmanfilters wird dies als

- Prädiktionsschritt

bezeichnet.

Dabei brauchen sich Korrekturschritt und Prädiktionsschritt nicht immer abzuwechseln. Solange Messungen vorliegen wird ein Korrekturschritt durchgeführt. Liegt nach einem Prädiktionsschritt keine Messung vor, so kann sich direkt ein weiterer Prädiktionsschritt anschließen.

10.3 Diskretes Kalmanfilter

Es wird das zeitdiskrete System

$$\begin{aligned}\mathbf{x}_{k+1} &= \mathbf{A}_k \mathbf{x}_k + \mathbf{B}_k \mathbf{u}_k + \mathbf{n}_{x,k} \\ \tilde{\mathbf{y}}_k &= \mathbf{C}_k \mathbf{x}_k + \mathbf{D}_k \mathbf{u}_k + \mathbf{n}_{y,k}\end{aligned}$$

betrachtet.

Das Systemrauschen $\mathbf{n}_{x,k}$ und das Messrauschen $\mathbf{n}_{y,k}$ sind jeweils mittelwertfrei, in sich unkorreliert und besitzen die Kovarianzmatrizen

$$\begin{aligned}\text{Cov}(\{\mathbf{n}_{x,k}\}_r, \{\mathbf{n}_{x,k}\}_r) &= \mathbf{P}_{x,k} \quad \text{und} \\ \text{Cov}(\{\mathbf{n}_{y,k}\}_r, \{\mathbf{n}_{y,k}\}_r) &= \mathbf{P}_{y,k}.\end{aligned}$$

Das Systemrauschen beschreibt tatsächliche Änderungen des Zustandsvektors. Idealerweise würde der Schätzwert diesem Rauschen daher folgen. Das Messrauschen beschreibt Störungen, die der Schätzwert idealerweise ignorieren würde.

10.3.1 Korrekturschritt

Liegen neue Messdaten vor, so besteht die Aufgabe darin, ausgehend von dem geschätzten Systemzustand $\hat{\mathbf{x}}_{k,-}$ mit der bekannten Kovarianz $\mathbf{P}_{\Delta,k,-}$ des Schätzfehlers und den Messwerten $\tilde{\mathbf{y}}_k$ mit der bekannten Kovarianz $\mathbf{P}_{y,k}$ einen neuen Schätzwert $\hat{\mathbf{x}}_{k,+}$ zu bestimmen, der biasfrei ist und dabei die minimale Varianz besitzt.

Da hier bei allen zeitabhängigen Größen nur der Zeitpunkt k auftritt, wird dieser in diesem Abschnitt im Index weggelassen.

Ansatz

Es wird von einem linearen Ansatz der Form⁴

$$\hat{\mathbf{x}}_+ = \Gamma_x \cdot \hat{\mathbf{x}}_- + \Gamma_y \cdot \tilde{\mathbf{y}} + \gamma_0$$

ausgegangen, wobei Γ_x , Γ_y und γ_0 noch zu bestimmen sind.

Die Korrektur soll erwartungstreue sein, d. h. unter der Voraussetzung, dass $E\{\hat{\mathbf{x}}_-\}_r = E\{\mathbf{x}\}_r$ gilt, soll auch $E\{\hat{\mathbf{x}}_+\}_r = E\{\mathbf{x}\}_r$ gelten, bzw. unter der Voraussetzung, dass $E\{\Delta\hat{\mathbf{x}}_-\}_r = \mathbf{0}$ ist, soll $E\{\Delta\hat{\mathbf{x}}_+\}_r = \mathbf{0}$ gelten. Nimmt man den Erwartungswert des allgemeinen Ansatzes,

$$E\{\hat{\mathbf{x}}_+\}_r = \Gamma_x \cdot E\{\hat{\mathbf{x}}_-\}_r + \Gamma_y \cdot E\{\tilde{\mathbf{y}}\}_r + \gamma_0,$$

so ergibt sich für ein System ohne Durchgriff ($\mathbf{D} = \mathbf{0}$)

$$E\{\hat{\mathbf{x}}_+\}_r = \Gamma_x \cdot E\{\mathbf{x}\}_r + \Gamma_y \mathbf{C} \cdot E\{\mathbf{x}\}_r + \gamma_0 \stackrel{!}{=} E\{\mathbf{x}\}_r.$$

Dieser Zusammenhang soll für alle $E\{\mathbf{x}\}_r$ erfüllt sein. Damit folgt $\gamma_0 = \mathbf{0}$ und weiter ergibt sich der Zusammenhang

$$\Gamma_x = \mathbf{I} - \Gamma_y \mathbf{C}$$

⁴ Eigentlich ein affiner Ansatz, aber in der Literatur wird in diesem Kontext dies immer als linearer Ansatz bezeichnet.

zwischen den Gewichten Γ_x und Γ_y .

Das Gewicht Γ_y wird im Weiteren als \mathbf{K} geschrieben, also $\Gamma_y = \mathbf{K}$. Damit ergibt sich die Korrekturgleichung

$$\begin{aligned}\hat{\mathbf{x}}_+ &= (\mathbf{I} - \mathbf{KC}) \cdot \hat{\mathbf{x}}_- + \mathbf{K}\tilde{\mathbf{y}} \\ &= \hat{\mathbf{x}}_- + \mathbf{K} \cdot (\tilde{\mathbf{y}} - \mathbf{C}\hat{\mathbf{x}}_-),\end{aligned}$$

in der \mathbf{K} noch bestimmt werden muss.

Sollte das System einen Durchgriff besitzen, so ergibt sich die Gleichung

$$\hat{\mathbf{x}}_+ = \hat{\mathbf{x}}_- + \mathbf{K} \cdot (\tilde{\mathbf{y}} - (\mathbf{C}\hat{\mathbf{x}}_- + \mathbf{Du})),$$

um die Erwartungstreue sicherzustellen ($\gamma_0 = \mathbf{KD}\mathbf{u}$).

Minimierung der Varianz

Durch den oben gegebenen Ansatz ist die Erwartungstreue des Korrekturschritts für beliebige \mathbf{K} gegeben. Dieses ist jetzt so zu wählen, dass die Varianz der Schätzwerte minimal wird.

Betrachtet man den Ansatz, dann erkennt man, dass jede Zeile \mathbf{k}_i^T von \mathbf{K} nur einen Einfluss auf die i -te Komponente von $\hat{\mathbf{x}}_+$ besitzt. Damit können die einzelnen Komponenten einzeln betrachtet werden,

$$\hat{x}_{i,+} = (\mathbf{e}_i^T - \mathbf{k}_i^T \mathbf{C}) \cdot \hat{\mathbf{x}}_- + \mathbf{k}_i^T \tilde{\mathbf{y}}.$$

Für den Schätzfehler $\Delta\hat{x}_{i,+}$ der i -ten Komponente gilt

$$\begin{aligned}\Delta\hat{x}_{i,+} &= x_i - \hat{x}_{i,+} \\ &= \mathbf{e}_i^T \mathbf{x} - (\mathbf{e}_i^T - \mathbf{k}_i^T \mathbf{C}) \cdot \hat{\mathbf{x}}_- - \mathbf{k}_i^T \tilde{\mathbf{y}} \\ &= \mathbf{e}_i^T \mathbf{x} - (\mathbf{e}_i^T - \mathbf{k}_i^T \mathbf{C}) \cdot \hat{\mathbf{x}}_- - \mathbf{k}_i^T \mathbf{Cx} - \mathbf{k}_i^T \mathbf{Cn}_y \\ &= (\mathbf{e}_i^T - \mathbf{k}_i^T \mathbf{C}) \cdot (\mathbf{x} - \hat{\mathbf{x}}_-) - \mathbf{k}_i^T \mathbf{Cn}_y,\end{aligned}$$

d. h.

$$\Delta\hat{x}_{i,+} = (\mathbf{e}_i^T - \mathbf{k}_i^T \mathbf{C}) \cdot \Delta\hat{\mathbf{x}}_- - \mathbf{k}_i^T \mathbf{Cn}_y.$$

(Bei Systemen mit Durchgriff ergibt sich der selbe Ausdruck für den Schätzfehler nach der Korrektur.)

Sinnvoll ist es, den mittleren quadrierten Fehler aller Komponenten $\{\Delta\hat{x}_{i,+}\}_r$ zu minimieren. Da der Schätzer per Konstruktion und Voraussetzung biasfrei ist, entspricht die Minimierung der mittleren quadrierten Fehler der Minimierung der Varianzen.

Aufgrund dessen, dass über die \mathbf{k}_i^T die Mittelung für die einzelnen Komponenten unabhängig stattfindet, kann

$$\text{Var}\{\Delta\hat{x}_{i,+}\}_r \rightarrow \min \quad \forall i$$

gefordert werden.⁵

Für die Varianz von $\{\Delta\hat{x}_{i,+}\}_r$ ergibt sich unter der Annahme, dass das Messrauschen in sich unkorreliert, und damit das aktuelle Messrauschen nicht mit dem aktuellen Schätzwert, in dem es noch nicht verarbeitet wurde, korreliert ist,

$$\text{Var}\{\Delta\hat{x}_{i,+}\}_r = (\mathbf{e}_i^T - \mathbf{k}_i^T \mathbf{C}) \cdot \mathbf{P}_{\Delta,-} \cdot (\mathbf{e}_i - \mathbf{C}^T \mathbf{k}_i) + \mathbf{k}_i^T \mathbf{P}_y \mathbf{k}_i, \quad (10.7)$$

⁵ Häufig wird die Minimierung von $\sum_i \text{Var}\{\Delta\hat{x}_{i,+}\}_r$ bzw. $\sum_i \text{MSE}\{\Delta\hat{x}_{i,+}\}_r$ gefordert. Dies ist äquivalent, jedoch kommt bei der Minimierung der Einzelvarianzen deutlicher hervor, dass dabei kein „Kompromiss“ eingegangen werden muss.

wobei $\mathbf{P}_{\Delta,-}$ für die Kovarianzmatrix des Schätzfehlers vor der Mittelung steht. Umstellen ergibt

$$\text{Var}\{\Delta \hat{x}_{i,+}\}_r = \mathbf{e}_i^T \mathbf{P}_{\Delta,-} \mathbf{e}_i - \mathbf{e}_i^T \mathbf{P}_{\Delta,-} \mathbf{C}^T \mathbf{k}_i - \mathbf{k}_i^T \mathbf{C} \mathbf{P}_{\Delta,-} \mathbf{e}_i + \mathbf{k}_i^T \cdot (\mathbf{C} \mathbf{P}_{\Delta,-} \mathbf{C}^T + \mathbf{P}_y) \cdot \mathbf{k}_i .$$

Der Vektor \mathbf{k}_i^T soll nun so gewählt werden, dass $\text{Var}\{\Delta \hat{x}_{i,+}\}_r$ minimal wird. Es wird daher die Ableitung nach \mathbf{k}_i gebildet,

$$\frac{d\text{Var}\{\Delta \hat{x}_{i,+}\}_r}{d\mathbf{k}_i} = -2\mathbf{e}_i^T \mathbf{P}_{\Delta,-} \mathbf{C}^T + 2 \cdot \mathbf{k}_i^T \cdot (\mathbf{C} \mathbf{P}_{\Delta,-} \mathbf{C}^T + \mathbf{P}_y) ,$$

und

$$\frac{d\text{Var}\{\Delta \hat{x}_{i,+}\}_r}{d\mathbf{k}_i} \stackrel{!}{=} 0$$

gefordert. Damit folgt

$$\mathbf{k}_i^T = \mathbf{e}_i^T \mathbf{P}_{\Delta,-} \mathbf{C}^T \cdot (\mathbf{C} \mathbf{P}_{\Delta,-} \mathbf{C}^T + \mathbf{P}_y)^{-1} .$$

Setzt man die einzelnen \mathbf{k}_i^T wieder zu \mathbf{K} zusammen, so erhält man

$$\mathbf{K} = \mathbf{P}_{\Delta,-} \mathbf{C}^T \cdot (\mathbf{C} \mathbf{P}_{\Delta,-} \mathbf{C}^T + \mathbf{P}_y)^{-1} . \quad (10.8)$$

Damit ist die Kalman-Verstärkung \mathbf{K} gegeben.

Um für weiter Schritte die Kovarianzmatrix verfolgen zu können, muss auch der Wert von $\mathbf{P}_{x,+}$ bestimmt werden. Mit den Rechenregeln zu Varianzen ergibt sich

$$\mathbf{P}_{\Delta,+} = (\mathbf{I} - \mathbf{K} \mathbf{C}) \cdot \mathbf{P}_{\Delta,-} \cdot (\mathbf{I} - \mathbf{C}^T \mathbf{K}^T) + \mathbf{K} \mathbf{P}_y \mathbf{K}^T . \quad (10.9)$$

Multipliziert man die rechte Klammer aus, erhält man

$$\begin{aligned} \mathbf{P}_{\Delta,+} &= (\mathbf{I} - \mathbf{K} \mathbf{C}) \cdot \mathbf{P}_{\Delta,-} - (\mathbf{I} - \mathbf{K} \mathbf{C}) \cdot \mathbf{P}_{\Delta,-} \cdot \mathbf{C}^T \mathbf{K}^T + \mathbf{K} \mathbf{P}_y \mathbf{K}^T \\ &= (\mathbf{I} - \mathbf{K} \mathbf{C}) \cdot \mathbf{P}_{\Delta,-} - \mathbf{P}_{\Delta,-} \cdot \mathbf{C}^T \mathbf{K}^T + \mathbf{K} \mathbf{C} \cdot \mathbf{P}_{\Delta,-} \cdot \mathbf{C}^T \mathbf{K}^T + \mathbf{K} \mathbf{P}_y \mathbf{K}^T \\ &= (\mathbf{I} - \mathbf{K} \mathbf{C}) \cdot \mathbf{P}_{\Delta,-} - \mathbf{P}_{\Delta,-} \cdot \mathbf{C}^T \mathbf{K}^T + \mathbf{K} \cdot (\mathbf{C} \cdot \mathbf{P}_{\Delta,-} \cdot \mathbf{C}^T + \mathbf{P}_y) \cdot \mathbf{K}^T . \end{aligned}$$

Setzt man im rechten Summanden für das linke \mathbf{K} die Kalman-Verstärkung (10.8) ein, so ergibt sich

$$\mathbf{P}_{\Delta,+} = (\mathbf{I} - \mathbf{K} \mathbf{C}) \cdot \mathbf{P}_{\Delta,-} . \quad (10.10)$$

Diese einfache Gleichung gilt aufgrund des letzten Schrittes nur, wenn \mathbf{K} die oben berechnete Kalman-Verstärkung ist. Gl. (10.9) dagegen ist für beliebige Verstärkungen \mathbf{K} gültig.

Zusammengefasst lauten die Gleichungen für den Korrekturschritt also

$$\begin{aligned} \mathbf{K} &= \mathbf{P}_{\Delta,-} \mathbf{C}^T \cdot (\mathbf{C} \mathbf{P}_{\Delta,-} \mathbf{C}^T + \mathbf{P}_y)^{-1} , \\ \hat{\mathbf{x}}_+ &= \hat{\mathbf{x}}_- + \mathbf{K} \cdot (\tilde{\mathbf{y}} - \underbrace{(\mathbf{C} \hat{\mathbf{x}}_- + \mathbf{D} \mathbf{u})}_{\hat{\mathbf{y}}_-}) \quad \text{und} \end{aligned}$$

$$\mathbf{P}_{\Delta,+} = (\mathbf{I} - \mathbf{K} \mathbf{C}) \cdot \mathbf{P}_{\Delta,-} .$$

10.3.2 Prädiktion

Setzt man auch für die Prädiktion einen linearen Schätzansatz

$$\hat{\mathbf{x}}_{k+1} = \Gamma_x \hat{\mathbf{x}}_k + \Gamma_u \mathbf{u}_k + \gamma_0$$

an (wobei Γ_x und γ_0 hier andere Größen als zuvor sind). Auch der Prädiktionsschritt soll erwartungstreue sein. D. h. unter der Voraussetzung, dass $E\{\hat{\mathbf{x}}_k\}_r = E\{\mathbf{x}_k\}_r$ ist, soll auch

$$E\{\hat{\mathbf{x}}_{k+1}\}_r = E\{\mathbf{x}_{k+1}\}_r$$

gelten.

Für die linke Seite gilt

$$E\{\hat{\mathbf{x}}_{k+1}\}_r = \Gamma_x \cdot E\{\mathbf{x}_k\}_r + \Gamma_u \mathbf{u}_k + \gamma_0$$

und für die rechte

$$E\{\mathbf{x}_{k+1}\}_r = \mathbf{A}_k \cdot E\{\mathbf{x}_k\}_r + \mathbf{B}_k \mathbf{u}_k .$$

Da $E\{\hat{\mathbf{x}}_{k+1}\}_r = E\{\mathbf{x}_{k+1}\}_r$ für alle Werte von $E\{\mathbf{x}_k\}_r$ und \mathbf{u}_k gelten sollen, folgt daraus

$$\hat{\mathbf{x}}_{k+1} = \mathbf{A}_k \hat{\mathbf{x}}_k + \mathbf{B}_k \mathbf{u}_k . \quad (10.11)$$

Damit stehen auch keine Freiheitsgrade mehr zur Verfügung.

Es bleibt, die Kovarianzmatrix $\mathbf{P}_{\Delta,k+1}$ des Schätzfehlers nach dem Zeitschritt zu bestimmen. Der Schätzfehler lautet

$$\begin{aligned} \Delta \hat{\mathbf{x}}_{k+1} &= \mathbf{x}_{k+1} - \hat{\mathbf{x}}_{k+1} \\ &= \mathbf{A}_k \mathbf{x}_k + \mathbf{B}_k \mathbf{u}_k + \mathbf{n}_{x,k} - (\mathbf{A}_k \hat{\mathbf{x}}_k + \mathbf{B}_k \mathbf{u}_k) \\ &= \mathbf{A}_k \cdot \Delta \hat{\mathbf{x}}_k + \mathbf{n}_{x,k} . \end{aligned}$$

Wenn das Systemrauschen $\{\mathbf{n}_{x,k}\}_r$ in sich unkorreliert ist, dann ist das Rauschen und der Schätzfehler zum selben Zeitpunkt ebenfalls unkorreliert. Damit gilt

$$\mathbf{P}_{\Delta,k+1} = \text{Var}\{\Delta \hat{\mathbf{x}}_{k+1}\}_r = \mathbf{A}_k \mathbf{P}_{\Delta,k} \mathbf{A}_k^T + \mathbf{P}_{x,k} . \quad (10.12)$$

Der erste Summand beschreibt den Einfluss des deterministischen Systemverhaltens auf den Schätzwert. Der zweite Einfluss berücksichtigt die Schwankungen des Systemzustandes. Diese erhöhen die Kovarianz des Schätzfehlers, da sich der wahre Wert vom Schätzwert entfernt.

Dagegen steht beim Korrekturschritt der zweite Summand in Gl. (10.9), $\mathbf{K} \mathbf{P}_y \mathbf{K}^T$, für die Erhöhung der Unsicherheit der Schätzung aufgrund des Messrauschens, d. h. der Schätzwert bewegt sich vom zu schätzenden Wert weg.

10.3.3 Zusammenfassung

Die Gleichungen des Kalmanfilters werden hier nochmal kompakt zusammengestellt. Dabei wird auch die Notation angepasst. Nach den Gleichungen folgen Erläuterungen.

- Systemgleichungen

$$\begin{aligned}\mathbf{x}_{k+1} &= \mathbf{A}_k \mathbf{x}_k + \mathbf{B}_k \mathbf{u}_k + \mathbf{n}_{x,k} \\ \tilde{\mathbf{y}}_k &= \mathbf{C}_k \mathbf{x}_k + \mathbf{D}_k \mathbf{u}_k + \mathbf{n}_{y,k}\end{aligned}$$

- Eigenschaften des Rauschens

$$\begin{array}{lll}\mathbf{Q}_k = E\{\mathbf{n}_{x,k} \mathbf{n}_{x,k}^T\}_r & E\{\mathbf{n}_{x,k}\}_r = \mathbf{0} & E\{\mathbf{n}_{x,k} \mathbf{n}_{x,l}^T\}_r = 0 \quad \forall k \neq l \\ \mathbf{R}_k = E\{\mathbf{n}_{y,k} \mathbf{n}_{y,k}^T\}_r & E\{\mathbf{n}_{y,k}\}_r = \mathbf{0} & E\{\mathbf{n}_{y,k} \mathbf{n}_{y,l}^T\}_r = 0 \quad \forall k \neq l\end{array}$$

- Initialschätzwert

$$E\{\hat{\mathbf{x}}_0\}_r = E\{\mathbf{x}_0\}_r \quad \text{Cov}(\{\mathbf{x}_0 - \hat{\mathbf{x}}_0\}_r, \{\mathbf{x}_0 - \hat{\mathbf{x}}_0\}_r) = \mathbf{P}_0$$

- Prädiktionsschritt

$$\begin{aligned}\hat{\mathbf{x}}_{k+1,-} &= \mathbf{A}_k \hat{\mathbf{x}}_{k,+} + \mathbf{B}_k \mathbf{u}_k \\ \mathbf{P}_{k+1,-} &= \mathbf{A}_k \mathbf{P}_{k,+} \mathbf{A}_k^T + \mathbf{Q}_k\end{aligned}$$

- Korrekturschritt

$$\begin{aligned}\mathbf{K}_k &= \mathbf{P}_{k,-} \mathbf{C}_k^T \cdot (\mathbf{C}_k \mathbf{P}_{k,-} \mathbf{C}_k^T + \mathbf{R}_k)^{-1} \\ \hat{\mathbf{x}}_{k,+} &= \hat{\mathbf{x}}_{k,-} + \mathbf{K}_k \cdot (\tilde{\mathbf{y}}_k - (\mathbf{C}_k \hat{\mathbf{x}}_{k,-} + \mathbf{D}_k \mathbf{u}_k)) \\ \mathbf{P}_{k,+} &= (\mathbf{I} - \mathbf{K}_k \mathbf{C}_k) \cdot \mathbf{P}_{k,-}\end{aligned}$$

Erläuterungen

- Das System kann zeitvariant sein und einen Durchgriff besitzen.
- Das Rauschen ist jeweils mittelwertfrei und in sich unkorreliert. Es ist jedoch keine Annahme über die Verteilung des Rauschens getroffen. Meist wird zwar weißes Rauschen bzw. weißes, gaußverteiltes Rauschen angesetzt, jedoch spielt das nur in weitergehenden Betrachtungen eine Rolle. Es existieren auch Erweiterungen des Kalmanfilters, um korrigiertes Rauschen zuzulassen.
Die Kovarianzmatrizen werden einer weitverbreiteten Notation folgend als \mathbf{Q} und \mathbf{R} bezeichnet.
- Die Biasfreiheit des Prädiktions- und des Korrekturschritts wurde für den Fall sichergestellt, dass der Eingangsschätzwert jeweils biasfrei ist. Daher muss auch der Anfangswert $\hat{\mathbf{x}}_0$ biasfrei sein. Es ist auch die Kovarianzmatrix des Schätzfehlers anzugeben. Diese wird in der Regel auch einfach als \mathbf{P}_k , ohne beschreibenden Index, bezeichnet.
- Die Nomenklatur in den zusammengefassten Gleichungen ist auf immer abwechselnde Prädiktions- und Korrekturschritte ausgelegt,

$$\dots \hat{\mathbf{x}}_{k,+} \xrightarrow{k \rightarrow k+1} \hat{\mathbf{x}}_{k+1,-} \xrightarrow{\tilde{\mathbf{y}}_{k+1}} \hat{\mathbf{x}}_{k+1,+} \xrightarrow{k+1 \rightarrow k+2} \hat{\mathbf{x}}_{k+2,-} \dots$$

Von diesem Schema kann aber abgewichen werden! Es wäre nur zu unübersichtlich, eine entsprechend allgemeine Indizierung zu verwenden. Das Kalmanfilter ist schematisch in Abbildung 10.1 dargestellt.

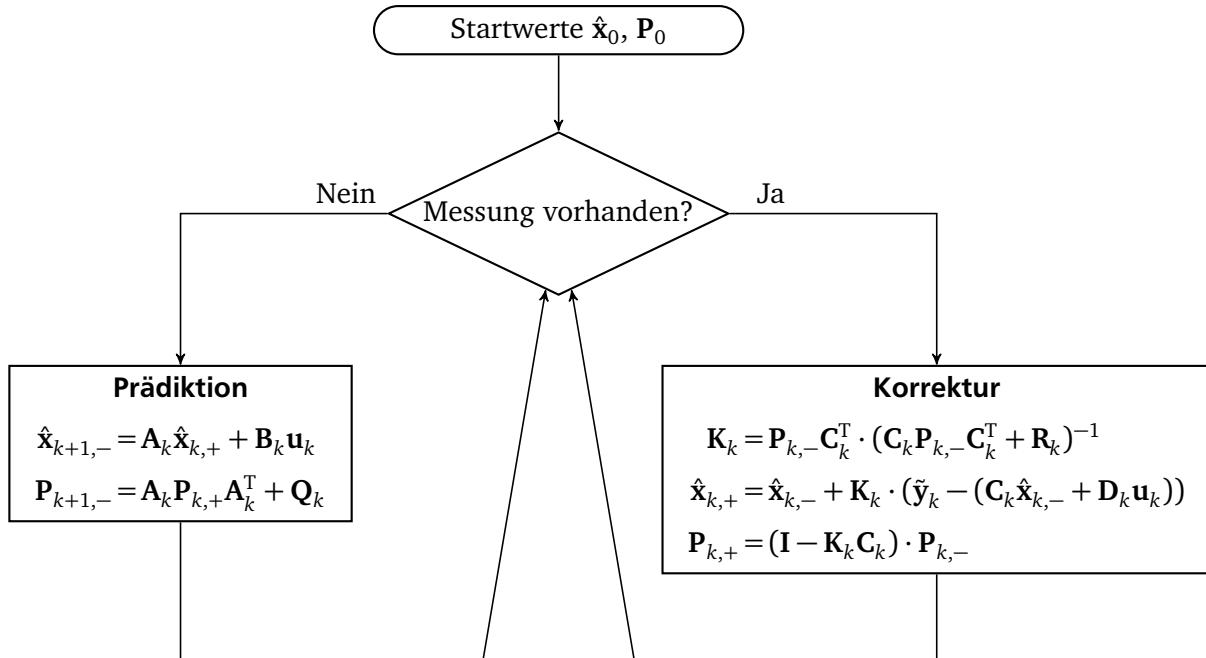


Abbildung 10.1: Kalmanfilter

Anmerkungen

- Es wurde hier jeder Schritt einzeln betrachtet. Damit ist – wenn es auch naheliegend scheint – noch nicht bewiesen, dass es sich hierbei auch um die insgesamt beste Lösung handelt. D. h. in einem ersten Schritt wird eine Kalman-Verstärkung K_1 derart bestimmt, dass der korrigierte Schätzwert die minimale Varianz besitzt. Im zweiten Schritt wird die Kalman-Verstärkung K_2 entsprechend so bestimmt, dass der im zweiten Schritt zu korrigierende Schätzwert, der schon über K_1 beeinflusst wurde, minimale Varianz besitzt. Es bleibt die Frage, ob ein anderes K_1 im Hinblick auf die Varianz nach der zweiten Korrektur vielleicht besser gewesen wäre.

Die Antwort auf diese Frage ist nein, siehe z. B. [Brammer und Siffling 1989]. Dort wird dies über das „Orthogonalitätsprinzip“ begründet. (Genauer werden die Schätzgleichungen über dieses Prinzip hergeleitet, womit diese es konstruktionsbedingt erfüllen.) Das Orthogonalitätsprinzip besagt, dass $\{\Delta \hat{x}_k\}_r$ genau dann die minimale Varianz (unter allen linearen Schätzern) hat, wenn der Schätzfehler zu dem Zeitpunkt k stochastisch orthogonal zu allen bisherigen Messwerten ist, d. h. es muss

$$E\{\Delta \hat{x}_k \cdot \tilde{y}_{k'}^T\}_r \stackrel{!}{=} \mathbf{0} \quad \forall k' \leq k$$

gelten.

- Es wurde, wie auch beispielsweise in [Brammer und Siffling 1989], die Erwartungstreue der Schätzwerte vorausgesetzt. Es kann aber auch direkt von der Minimierung der mittleren quadrierten Schätzfehler ausgegangen werden, wie z. B. in [Kamen und Su 1999].

Bei den gleichen Ansätzen für den Prädiktions- und Korrekturschritt ergeben sich damit auch die gleichen Ergebnisse, was in Anhang F.1 gezeigt ist. Lediglich entspricht \mathbf{P}_k dann nicht der Kovarianzmatrix

$$E\{(\Delta \hat{\mathbf{x}}_k - E\{\Delta \hat{\mathbf{x}}_k\}_r) \cdot (\Delta \hat{\mathbf{x}}_k - E\{\Delta \hat{\mathbf{x}}_k\}_r)^T\}_r$$

des Schätzfehlers⁶, da hier nicht die Biasfreiheit $E\{\Delta \hat{\mathbf{x}}_k\}_r = \mathbf{0}$ vorausgesetzt wird, sondern der „Korrelationsmatrix“

$$E\{\Delta \hat{\mathbf{x}}_k \cdot \Delta \hat{\mathbf{x}}_k^T\}_r,$$

auf dessen Diagonale die MSE-Werte der einzelnen geschätzten Zustände abgelesen werden können.

- Nimmt man noch die Annahme hinzu, dass die Rauschprozesse gaußverteilt sind, so ist das Kalmanfilter – ähnlich dem LS-Verfahren – nicht nur optimal bezüglich aller linearen Zustandsschätzer, sondern auch bezüglich aller nichtlinearer Schätzer.
- In einer „Standard-Beobachter-Anwendung“, d. h. zeitinvariantes System (feste Abtastzeit), immer abwechselnde Prädiktions- und Korrekturschritte und zeitinvariantes System- und Messrauschen, konvergiert die Kalmanverstärkung in der Regel relativ schnell gegen einen festen Wert

$$\mathbf{K}_\infty = \mathbf{P}_\infty \mathbf{C}^T \cdot (\mathbf{C} \mathbf{P}_\infty \mathbf{C}^T + \mathbf{R})^{-1},$$

wobei sich \mathbf{P}_∞ aus der Lösung der Matrix-Riccati-Gleichung

$$\mathbf{P}_\infty = \mathbf{A} \cdot (\mathbf{P}_\infty - \mathbf{P}_\infty \mathbf{C}^T \cdot (\mathbf{C} \mathbf{P}_\infty \mathbf{C}^T + \mathbf{R})^{-1} \cdot \mathbf{C} \mathbf{P}_\infty) \cdot \mathbf{A}^T + \mathbf{Q}$$

ergibt. Im Fall einer klassischen Beobachtungsaufgabe dürfte es in den meisten Fällen sinnvoll sein, mit dieser stationären Verstärkung eine Luenberger-Beobachter-Struktur zu verwenden.

Der Nutzen des Kalmanfilters wie hier beschrieben liegt darin, dass sich dieser auf nichtlineare Systeme erweitern lässt sowie dass auch unregelmäßige Messungen verarbeitet werden können.

- Setzt man $\mathbf{x} = \boldsymbol{\theta}$, $\mathbf{C}_k = \boldsymbol{\psi}_k^T$ und $\tilde{y}_k = \tilde{\eta}_k$ sowie $n_{y,k} = n_{\eta,k}$, so beschreibt das diskrete, zeitvariante Zustandsraummodell

$$\mathbf{x}_{k+1} = \mathbf{I}\mathbf{x}_k$$

$$\tilde{y}_k = \mathbf{C}_k \mathbf{x}_k + n_{y,k}$$

den „LS-Prozess“ $\tilde{\eta}_k = \boldsymbol{\psi}_k^T \boldsymbol{\theta} + n_{\eta,k}$.

Der Korrekturschritt des Kalmanfilters für dieses System lautet

$$\mathbf{K}_k = \frac{\mathbf{P}_{k,-} \boldsymbol{\psi}_k}{\boldsymbol{\psi}_k^T \mathbf{P}_{k,-} \boldsymbol{\psi}_k + \sigma_{\eta,k}^2}$$

$$\hat{\boldsymbol{\theta}}_{k,+} = \hat{\boldsymbol{\theta}}_{k,-} + \mathbf{K}_k \cdot (\tilde{\eta}_k - \boldsymbol{\psi}_k^T \hat{\boldsymbol{\theta}}_{k,-})$$

$$\mathbf{P}_{k,+} = (\mathbf{I} - \mathbf{K}_k \boldsymbol{\psi}_k^T) \cdot \mathbf{P}_{k,-}$$

und der Prädiktionsschritt lautet

$$\hat{\boldsymbol{\theta}}_{k+1,-} = \hat{\boldsymbol{\theta}}_{k,+}$$

$$\mathbf{P}_{k+1,-} = \mathbf{P}_{k,+}.$$

Es ergibt sich genau das gewichtete rekursive LS-Verfahren mit $w_k = \frac{1}{\sigma_{\eta,k}^2}$.

Für den Vergessensfaktor λ des LS-Verfahrens gibt es keine Entsprechung beim Kalmanfilter. Dafür verfügt dieser über die Möglichkeit, ein Systemrauschen für jeden Zustand einzeln vorzugeben. Dieses kann auch dazu benutzt werden, ein dem Vergessen ähnliches Verhalten zu erzeugen.

⁶ Auch wenn dies in der Literatur häufig so bezeichnet wird.

- In die Berechnung der Fehlerkovarianzmatrix und der Korrekturmatrixt gehen lediglich die (möglicherweise zeitvariante) Ausgangsmatrix, die Kovarianzmatrix des Messrauschen sowie die Kovarianzmatrix des Systemrauschen ein. Die Eingangsgrößen sowie die gemessene Ausgangsgröße spielen hierbei keine Rolle. Daher kann man den Verlauf der Fehlerkovarianzmatrix schon anhand ihres Anfangswertes sowie R_k und Q_k bestimmen.

Anfangswerte

Es müssen Anfangswerte \hat{x}_0 für die Schätzwerte sowie P_0 für die Fehlerkovarianz vorgegeben werden.

Damit das Kalmanfilter biasfreie Schätzwerte liefert, muss

$$E\{\hat{x}_0\}_r = E\{x_0\}_r$$

gewählt werden, was praktisch nur möglich ist, wenn man $E\{x_0\}_r$ kennt.

Ist gewissen Vorwissen vorhanden, in welchem Bereich die einzelnen Parameter liegen, so sollte dieses für die Wahl der Startwerte verwendet werden. So kann man den Startwert \hat{x}_0 in die Mitte eines sinnvollen Bereichs legen, und kann sich für die Werte von P_0 an der Breite dieses Bereichs orientieren.

Existiert kein Vorwissen, so kann $\hat{x}_0 = \mathbf{0}$ gesetzt und für den Startwert der Fehlerkovarianz $P_0 = \alpha \cdot I$ mit einem sehr großen α gewählt werden. Damit „informiert“ man das Kalmanfilter, dass die Startwerte sehr unsicher sind und stark korrigiert werden müssen.

Zudem sollte man sich auch bewusst machen, dass die Biasfreiheit sowie die Eigenschaft der minimalen Varianz sich auf die Scharen, d. h. die Gesamtheit der Realisierungen, des Zufallsexperiments „Kalmanfilter“ beziehen, und nicht auf eine einzelne Realisierung. D. h. führt man sehr häufig die selbe Schätzung (mit anderen Startwerten und anderem Rauschen) durch, so stimmt bei einer biasfreien Schätzung für jeden Zeitpunkt der Mittelwert der einzelnen Schätzwerte mit dem realen Zustand überein.⁷ Und die Schätzfehler der einzelnen Realisierungen streuen untereinander minimal (im Sinne der Varianz).

Bei einer Realisierung können aber sehr große Schätzfehler auftreten, die sich auch nur langsam abbauen. Und wenn man immer $\hat{x}_0 = \mathbf{0}$ wählt, obwohl bekannt ist, dass die Zustände zu Beginn z. B. immer um $[10 \ 100]^T$ liegen, dann macht man damit einen systematischen Fehler. Das Kalmanfilter ist dann (über alle Realisierungen gesehen) biasbehaftet. Und jede praktische Durchführung beginnt mit einem großen Startfehler und ist unnötig schlecht.

Kovarianzmatrix des Systemrauschen

Liegt bei dem betrachteten System tatsächlich ein (wesentliches) Systemrauschen vor, so sollte Q_k möglichst aus dessen Eigenschaften bestimmt werden. Dabei kann die Annahme von Systemrauschen auch dazu dienen, Fehler bei der Modellierung und ggf. der Diskretisierung sowie Rauschen der Eingangsgrößen (siehe unten) zu erfassen.

Aber auch, wenn weder „echtes“ Systemrauschen noch eines der genannten Probleme vorliegt, kann es sinnvoll sein, Q_k ungleich Null zu wählen. Dies hängt mit dem zusammen, was bei der Wahl der Anfangswerte angesprochen wurde, nämlich dass die Optimalität des Kalmanfilters sich auf die Scharen bezieht. Es kann durchaus sein, dass man bei einer Durchführung der Filterung einen sehr schlechten Startwert erwischt, und dann das Rauschen der Ausgangsgröße zu Beginn vielleicht noch sehr ungünstige Werte hat, und so der Schätzwert sehr schlecht ist. Wenn kein Systemrauschen angenommen wird, d. h. $Q_k = \mathbf{0}$ ist, nehmen die Varianzen des Schätzfehlers aber monoton ab. Dies geschieht einfach aufgrund des angegebenen Startwerts P_0 sowie dem angegebenen Messrauschen R_k . Das Kalmanfilter vergleicht dazu nicht die Messwerte mit dem geschätzten Systemausgang. Aufgrund der abnehmenden Varianzen

⁷ Wenn auch Systemrauschen vorliegt, dann mit dem Mittelwert der realen Zustände über alle Realisierungen.

der Schätzfehler wird aber immer weniger korrigiert, d. h. die anfängliche schlechte Schätzung verbessert sich nur sehr langsam.

D. h. auch wenn alle Annahmen des Kalmanfilters erfüllt sind, kann es sein, dass man sich in einer schlechten Realisierung befindet, und es mag dann nicht viel helfen, wenn andere Realisierungen besser wären. Nimmt man ein gewisses Systemrauschen \mathbf{Q}_k an, so stellt dieses eine untere Schranke für die Varianzen der Schätzfehler nach der Prädiktion dar, da dabei \mathbf{Q}_k immer zu \mathbf{P}_k addiert wird. Damit kann erreicht werden, dass die Messwerte immer zu einem gewissen Maß zur Korrektur der Schätzwerte verwendet werden. Allerdings zu dem Preis, dass die Streuungen der Schätzwerte zunehmen.

Nimmt man ein Systemrauschen aus den genannten Gründen an, obwohl tatsächlich keines wirkt, so wird die Varianz des Kalmanfilters höher. Man erkauft sich ein besseres Verhalten für einzelne Realisierungen mit über alle Realisierungen gesehen schlechtere Eigenschaften.

Parallele Messungen zusammen oder einzeln verarbeiten

Besitzt das betrachtete System mehrere Ausgänge, so können diese allgemein aufgeteilt werden,

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{C}_1 \\ \mathbf{C}_2 \end{bmatrix} \cdot \mathbf{x}.$$

Sind die Messungen $\tilde{\mathbf{y}}_1$ und $\tilde{\mathbf{y}}_2$ der beiden Teilausgänge unkorreliert, d. h. gilt

$$\mathbf{R} = \begin{bmatrix} \mathbf{R}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_2 \end{bmatrix},$$

dann macht es keinen Unterschied ob man die Messungen komplett in einem Schritt berücksichtigt, oder ob man beide Messungen einzeln verwendet. Schematisch bedeutet dies, dass

$$\hat{\mathbf{x}}_- \xrightarrow{\begin{bmatrix} \tilde{\mathbf{y}}_1 \\ \tilde{\mathbf{y}}_2 \end{bmatrix}} \hat{\mathbf{x}}_+$$

und

$$\hat{\mathbf{x}}_- \xrightarrow{\tilde{\mathbf{y}}_1} \hat{\mathbf{x}}'_+ \xrightarrow{\tilde{\mathbf{y}}_2} \hat{\mathbf{x}}_+$$

äquivalent sind. Dies ist in Anhang F.2 gezeigt.

Diese Aussage lässt sich rekursiv auf beliebig viele unkorrelierte Messgrößen erweitern.

Die Bedeutung dieser Eigenschaft liegt darin, dass man so bei unkorrelierten Messungen die Dimension der in

$$\mathbf{K}_k = \mathbf{P}_{k,-} \mathbf{C}_k^T \cdot (\mathbf{C}_k \mathbf{P}_{k,-} \mathbf{C}_k^T + \mathbf{R}_k)^{-1}$$

zu invertierenden Matrix $(\mathbf{C}_k \mathbf{P}_{k,-} \mathbf{C}_k^T + \mathbf{R}_k)$ verringern kann. Sind alle Messgrößen unkorreliert, so müssen nur Kehrwerte von skalaren Werten bestimmt werden.

Rauschen am Eingang

Ist die gemessene Eingangsgröße verrauscht,

$$\tilde{\mathbf{u}}_k = \mathbf{u}_k + \mathbf{n}_{u,k},$$

so kann dies auch als Systemrauschen aufgefasst werden.

So ergibt sich für die Systemgleichung

$$\begin{aligned}\mathbf{x}_{k+1} &= \mathbf{A}_k \mathbf{x}_k + \mathbf{B}_k \cdot (\tilde{\mathbf{u}}_k - \mathbf{n}_{u,k}) + \mathbf{n}_{x,k} \\ &= \mathbf{A}_k \mathbf{x}_k + \mathbf{B}_k \tilde{\mathbf{u}}_k - \mathbf{B}_k \mathbf{n}_{u,k} + \mathbf{n}_{x,k}.\end{aligned}$$

Es tritt also das „neue“ Systemrauschen

$$\mathbf{n}'_{x,k} = \mathbf{n}_{x,k} - \mathbf{B}_k \mathbf{n}_{u,k}$$

auf. Wenn $\{\mathbf{n}_{x,k}\}_r$ und $\{\mathbf{n}_{u,k}\}_r$ unkorreliert sind, ergibt sich

$$\mathbf{Q}_k = \mathbf{P}_{x',k} = \mathbf{P}_{x,k} + \mathbf{B}_k \mathbf{P}_{u,k} \mathbf{B}_k^T$$

für die vom Kalmanfilter zu berücksichtigende Kovarianz des Systemrauschens.

Bei der Prädiktion zeitkontinuierlicher Systeme ist darauf zu achten, dass die Schrittweite der Simulation entsprechend gering gewählt wird, damit das Rauschen realistisch berücksichtigt wird. (Dies ist auch der Fall, wenn das tatsächliche – und nicht nur gemessene – Eingangssignal verrauscht ist.)

Beim erweiterten Kalmanfilter (siehe unten) kann die Eingangsgröße auch einen Einfluss auf die linearisierte System- und Ausgangsmatrix besitzen. Insbesondere wenn die Kovarianzmatrix des Schätzfehlers in größeren Schritten diskretisiert wird als die Systemgleichung ist dann auf das Rauschen zu achten.

10.3.4 Beispiele

Es wird als Beispiel das System

$$\begin{aligned}\mathbf{x}_{k+1} &= \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \mathbf{x}_k + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \mathbf{u}_k \\ \mathbf{y}_k &= \begin{bmatrix} 1 & 0 \end{bmatrix} \mathbf{x}_k\end{aligned}$$

betrachtet. Dieses ist in Abbildung 10.2 als Blockschaltbild dargestellt. Es handelt sich um zwei hintereinandergeschaltete (diskrete) Integratoren.

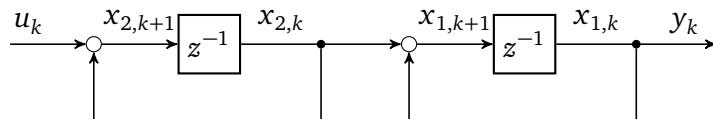


Abbildung 10.2: Beispielsystem

Der reale Anfangszustand ist

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

während der geschätzte Anfangszustand

$$\hat{\mathbf{x}}_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

ist.

Der Anfangswert der Fehlerkovarianzmatrix sei

$$\mathbf{P}_0 = \begin{bmatrix} p_{1,0} & 0 \\ 0 & p_{2,0} \end{bmatrix}$$

und die Varianz des Messrauschen beträgt $\mathbf{R} = r$.

Erste Messung

Es soll mit einer Messung begonnen werden. Es wird also $\hat{\mathbf{x}}_{1,-} = \hat{\mathbf{x}}_0$ und $\mathbf{P}_{1,-} = \mathbf{P}_0$ gesetzt. Für die Korrekturmatrixt ergibt sich

$$\mathbf{K}_1 = \mathbf{P}_{1,-} \mathbf{C}^T \cdot (\mathbf{C} \mathbf{P}_{1,-} \mathbf{C}^T + \mathbf{R})^{-1} = \begin{bmatrix} \frac{p_{1,0}}{p_{1,0}+r} \\ 0 \end{bmatrix}.$$

Für den korrigierten Zustandsvektor ergibt sich

$$\hat{\mathbf{x}}_{1,+} = \hat{\mathbf{x}}_{1,-} + \begin{bmatrix} \frac{p_{1,0}}{p_{1,0}+r} \\ 0 \end{bmatrix} \cdot (\tilde{y}_1 - \mathbf{C} \hat{\mathbf{x}}_{1,-}),$$

und die neue Kovarianzmatrix lautet

$$\mathbf{P}_{1,+} = \begin{bmatrix} \frac{p_{1,0}r}{p_{1,0}+r} & 0 \\ 0 & p_{2,0} \end{bmatrix}.$$

Ist das Vorwissen über den Zustand gegenüber den Messungen sehr unsicher, $p_{1,0} \gg r$, so ist

$$\mathbf{K}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

und

$$\mathbf{P}_{1,+} = \begin{bmatrix} r & 0 \\ 0 & p_{2,0} \end{bmatrix}.$$

Mit den oben angegebenen Zahlenwerten ergibt sich bei einer sicheren Messung $\tilde{y}_1 = x_{1,1} = 1$ ($r = 0$) dann

$$\hat{\mathbf{x}}_{1,+} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} \cdot (1 - 0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Damit wäre der erste Zustand exakt bekannt, was sich auch durch $(\mathbf{P}_{1,+})_{11} = r = 0$ zeigt.

Zweite Messung

Führt man direkt eine zweite (unabhängige) Messung durch ($\hat{\mathbf{x}}_{2,-} = \hat{\mathbf{x}}_{1,+}$ und $\mathbf{P}_{2,-} = \mathbf{P}_{1,+}$), dann ergibt sich

$$\mathbf{K}_2 = \begin{bmatrix} \frac{p_{1,0}}{2p_{1,0}+r} \\ 0 \end{bmatrix}.$$

Der zweite Zustand wird also weiterhin nicht korrigiert, und die Korrektur des ersten Zustandes fällt (relativ gesehen) geringer aus als zuvor. Es ergibt sich

$$\mathbf{P}_{2,+} = \begin{bmatrix} \frac{p_{1,0}r}{2p_{1,0}+r} & 0 \\ 0 & p_{2,0} \end{bmatrix}.$$

Mit $r \ll p_{1,0}$ erhält man weiter

$$\mathbf{P}_{2,+} = \begin{bmatrix} \frac{r}{2} & 0 \\ 0 & p_{2,0} \end{bmatrix},$$

und für die gegebenen Zahlenwerte ergibt sich

$$\hat{\mathbf{x}}_{2,+} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} \frac{1}{2} \\ 0 \end{bmatrix} \cdot (1 - 1) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Es ist an dieser Stelle abzusehen – was sich auch mit der Erwartung deckt – dass weitere Messungen zwar die Varianz des Schätzwertes des ersten Zustands verringert, aber sich an der Varianz des Schätzwertes des zweiten Zustands nichts ändert.

Zeitschritt (Prädiktion)

Zwischen den Messungen 2 und 3 findet ein Zeitschritt statt, wobei $\mathbf{u}_k = 0$ ist. Für den realen Systemzustand ergibt sich damit

$$\mathbf{x}_3 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}.$$

Es soll jetzt der Prädiktionsschritt von $k = 2$ auf $k = 3$ durchgeführt werden. Man erhält

$$\mathbf{P}_{3,-} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{r}{2} & 0 \\ 0 & p_{2,0} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} + \mathbf{Q}.$$

Für die Betrachtungen hier wird davon ausgegangen, dass kein Prozessrauschen vorliegt, d. h. $\mathbf{Q} = \mathbf{0}$ und es ergibt sich

$$\mathbf{P}_{3,-} = \begin{bmatrix} \frac{r}{2} + p_{2,0} & p_{2,0} \\ p_{2,0} & p_{2,0} \end{bmatrix}.$$

Der prädizierte Systemzustand ist

$$\hat{\mathbf{x}}_{3,-} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \hat{\mathbf{x}}_{2,+} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Dritte Messung

Liegt nach dem Prädiktionsschritt eine weitere Messung vor, kann wieder ein Korrekturschritt ausgeführt werden. Man erhält

$$\mathbf{K}_3 = \begin{bmatrix} \frac{p_{2,0} + \frac{r}{2}}{p_{2,0} + \frac{3}{2}r} \\ \frac{p_{2,0}}{p_{2,0} + \frac{3}{2}r} \end{bmatrix}$$

und damit

$$\mathbf{P}_{3,+} = \frac{1}{p_{2,0} + \frac{3}{2}r} \cdot \begin{bmatrix} \frac{1}{2}r^2 + p_{2,0}r & p_{2,0}r \\ p_{2,0}r & \frac{3}{2}p_{2,0}r \end{bmatrix}.$$

Ist die Messung auch deutlich sicherer als das Vorwissen über den zweiten Systemzustand, $r \ll p_{2,0}$, gilt

$$\mathbf{K}_3 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

und

$$\mathbf{P}_{3,+} = \begin{bmatrix} r & r \\ r & \frac{3}{2}r \end{bmatrix}.$$

Wären die Messungen exakt, $r = 0$, dann wären nach der zweiten Messung die Zustände exakt bekannt ($\mathbf{P}_{3,+} = \mathbf{0}$).

Mit den Zahlenwerten ergibt sich

$$\hat{\mathbf{x}}_{3,+} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} \cdot (2 - 1) = \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

Um alle Zustände schätzen zu können, muss der Schätzer das dynamische Verhalten, also die Kopplung der Zustände untereinander, ausnutzen. Diese Kopplungen und Abhängigkeiten spiegeln sich in den Kovarianzmatrizen wider und werden darüber auch in der Schätzung berücksichtigt.

Fehler in Prädiktion der Kovarianzmatrix

Geht man im Prädiktionsschritt der Fehlerkovarianz fälschlicherweise von

$$\mathbf{A} = \begin{bmatrix} 1 & 0,9 \\ 0 & 1 \end{bmatrix}$$

aus, so ergibt sich

$$\mathbf{P}_{3,-} = \begin{bmatrix} \frac{r}{2} + 0,81 \cdot p_{2,0} & 0,9 \cdot p_{2,0} \\ 0,9 \cdot p_{2,0} & p_{2,0} \end{bmatrix}$$

und daraus im nächsten Korrekturschritt ($r \ll p_{2,0}$)

$$\mathbf{K}_3 = \begin{bmatrix} 1 \\ 1,1 \end{bmatrix}.$$

Der zweite Zustand wird also falsch korrigiert,

$$\hat{\mathbf{x}}_{3,+} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \\ 1,1 \end{bmatrix} \cdot (2 - 1) = \begin{bmatrix} 2 \\ 1,1 \end{bmatrix}.$$

Dennoch ist der Schätzwert $\hat{\mathbf{x}}_{3,+}$ erwartungstreu. Dies ist er nämlich für jedes \mathbf{K} . Dies ist nochmal eine gute Gelegenheit, sich klar zu machen, was der Scharmittelwert, auf den sich die Erwartungstreue bezieht, ist.

Würde der gesamte Schätzvorgang wiederholt werden, und dabei von

$$\hat{\mathbf{x}}'_0 = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$$

ausgegangen werden (so dass der Mittelwert der beiden zweiten Einträge $\hat{x}_{2,0}$ und $\hat{x}'_{2,0}$ dem wahren Wert $x_{2,0}$ entspricht), so ergäbe sich

$$\hat{\mathbf{x}}'_{3,+} = \begin{bmatrix} 2 \\ 0,9 \end{bmatrix}.$$

Der Mittelwert der beiden Schätzwerte $\hat{\mathbf{x}}_{3,+}$ und $\hat{\mathbf{x}}'_{3,+}$ entspricht genau dem wahren Zustandswert \mathbf{x}_3 zu diesem Zeitpunkt.

Rechnet man die Fehlerkovarianzen aus, die sich nach diesem Korrekturschritt für $r = 0$ ergeben, so erhält man $\mathbf{P}'_{3,+} = \mathbf{0}$. Sollte dies korrekt sein, wäre dies ein Widerspruch zu den gerade gemachten Beobachtungen. Allerdings ist die Gleichung zur Prädiktion der Fehlerkovarianzmatrizen hier nicht gültig, da bei der Herleitung davon ausgegangen ist, dass die Prädiktion die exakte Systemmatrix verwendet.

Andere Ausgangsgröße

Es wird die in Abbildung 10.3 dargestellte Variante

$$\begin{aligned} \mathbf{x}_{k+1} &= \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \mathbf{x}_k + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \mathbf{u}_k \\ \mathbf{y}_k &= \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \end{bmatrix} \mathbf{x}_k \end{aligned}$$

des Beispielsystems betrachtet, die sich lediglich in der Ausgangsmatrix unterscheidet.

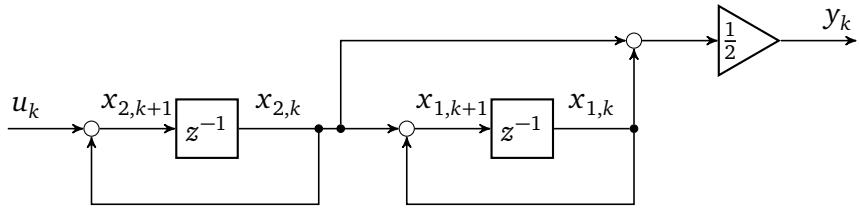


Abbildung 10.3: Beispielsystem (Variante)

Für die erste Messung kommt man für dieses System auf

$$\mathbf{K}_1 = \begin{bmatrix} \frac{2p_{1,0}}{p_{1,0} + p_{2,0} + 4r} \\ \frac{2p_{2,0}}{p_{1,0} + p_{2,0} + 4r} \end{bmatrix},$$

und damit erhält man

$$\mathbf{P}_{1,+} = \frac{1}{p_{1,0} + p_{2,0} + 4r} \cdot \begin{bmatrix} p_{1,0}p_{2,0} + 4p_{1,0}r & -p_{1,0}p_{2,0} \\ -p_{1,0}p_{2,0} & p_{1,0}p_{2,0} + 4p_{2,0}r \end{bmatrix}.$$

Auch bei einer perfekten Messung, $r = 0$, besitzen beide Schätzwerte eine Varianz,

$$\mathbf{P}_{1,+} = \frac{1}{p_{1,0} + p_{2,0}} \cdot \begin{bmatrix} p_{1,0}p_{2,0} & -p_{1,0}p_{2,0} \\ -p_{1,0}p_{2,0} & p_{1,0}p_{2,0} \end{bmatrix}.$$

Bestimmt man allerdings die Varianz der Ausgangsgröße, der sich mit dem geschätzten $\hat{x}_{1,+}$ ergibt,

$$\text{Var}\{\hat{y}\}_r = \mathbf{C}\mathbf{P}\mathbf{C}^T = \frac{(p_{1,0} + p_{2,0}) \cdot r}{p_{1,0} + p_{2,0} + 4r} = 0,$$

so sieht man, dass diese exakt wiedergegeben wird.

Auch hier führen weitere (perfekte) Messungen nicht dazu, dass die Parameter besser geschätzt werden, solange keine Zeitschritte dazwischen abgewartet werden.

Fazit

Prädiktion und Korrektur sind unabhängige Schritte. Aber wenn die Ausgangsmatrix nicht regulär ist, dann benötigt man zwingend die Prädiktion, um über das dynamische Verhalten des Systems auf alle Zustände schließen zu können.

10.4 Kontinuierlich-diskretes Kalmanfilter

Liegt ein System mit zeitkontinuierlicher Dynamik

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t) + \mathbf{n}_x(t)$$

mit weißem Rauschen $E\{\mathbf{n}_x(t)\mathbf{n}_x^T(t')\}_r = \mathbf{Q}(t) \cdot \delta(t-t')$ vor, bei dem die Messungen aber nur zu diskreten Zeitpunkten verfügbar sind,

$$\tilde{\mathbf{y}}_k = \mathbf{C}_k \mathbf{x}_k + \mathbf{D}_k \mathbf{u}_k + \mathbf{n}_{y,k},$$

dann bleiben die Korrekturgleichungen unverändert und für die Prädiktion ergibt sich

$$\begin{aligned}\dot{\hat{\mathbf{x}}}(t) &= \mathbf{A}(t)\hat{\mathbf{x}}(t) + \mathbf{B}(t)\mathbf{u}(t) \\ \dot{\mathbf{P}}(t) &= \mathbf{A}(t)\mathbf{P}(t) + \mathbf{P}(t)\mathbf{A}^T(t) + \mathbf{Q}(t)\end{aligned}$$

wobei diese Differentialgleichungen für die Zeit von $t = t_k$ bis $t = t_{k+1}$ gelöst werden müssen. Dabei ist ein geeignetes Simulationsverfahren zu wählen. Simulationsverfahren werden z.B. in der Vorlesung „Modellbildung und Simulation“ vorgestellt. Dort wird auch auf die Stabilität der Simulationsverfahren und die Simulation steifer Systeme (Systeme mit sich betragsmäßig sehr stark unterscheidenden Eigenwerten) eingegangen.

- Systemgleichungen

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t) + \mathbf{n}_x(t) \\ \tilde{\mathbf{y}}_k &= \mathbf{C}_k\mathbf{x}_k + \mathbf{D}_k\mathbf{u}_k + \mathbf{n}_{y,k}\end{aligned}$$

- Prädiktion

$$\begin{aligned}\dot{\hat{\mathbf{x}}}(t) &= \mathbf{A}(t)\hat{\mathbf{x}}(t) + \mathbf{B}(t)\mathbf{u}(t) & \hat{\mathbf{x}}(t_k) &= \hat{\mathbf{x}}_{k,+} \\ \dot{\mathbf{P}}(t) &= \mathbf{A}(t)\mathbf{P}(t) + \mathbf{P}(t)\mathbf{A}^T(t) + \mathbf{Q}(t) & \mathbf{P}(t_k) &= \mathbf{P}_{k,+} \\ \Rightarrow \hat{\mathbf{x}}_{k+1,-} &= \hat{\mathbf{x}}(t_{k+1}), & \mathbf{P}_{k+1,-} &= \mathbf{P}(t_{k+1})\end{aligned}$$

- Korrektur

$$\begin{aligned}\mathbf{K}_k &= \mathbf{P}_{k,-}\mathbf{C}_k^T \cdot (\mathbf{C}_k\mathbf{P}_{k,-}\mathbf{C}_k^T + \mathbf{R}_k)^{-1} \\ \hat{\mathbf{x}}_{k,+} &= \hat{\mathbf{x}}_{k,-} + \mathbf{K}_k \cdot (\tilde{\mathbf{y}}_k - \mathbf{h}(\hat{\mathbf{x}}_k, \hat{\mathbf{u}}_k)) \\ \mathbf{P}_{k,+} &= (\mathbf{I} - \mathbf{K}_k\mathbf{C}_k) \cdot \mathbf{P}_{k,-}\end{aligned}$$

Daneben wird in der Praxis bisweilen in diesem Fall auch der Systemzustand kontinuierlich prädiziert, während die Kovarianzmatrix zeitdiskret prädiziert wird. Bzw. werden bei der Simulation der beiden Differentialgleichungen unterschiedliche Rechenschrittweiten verwendet. Hierfür gibt es weniger eine gute mathematische als vielmehr eine pragmatische Rechtfertigung.

So ist der Berechnungsaufwand für die Prädiktion der Kovarianzmatrix um ein Vielfaches höher als der der Prädiktion des Zustandes. Wird das einfache Euler-Verfahren mit der Schrittweite Δt zur Simulation verwendet, so muss bei der Prädiktion des Schätzwertes und ohne Berücksichtigung der Eingangsgröße für jeden Zeitschritt der Simulation (nicht Zeitschritt zwischen zwei Messungen!) eine $n \times n$ -Matrix mit einem n -dimensionalen Vektor (n ist die Systemordnung) multipliziert und das Ergebnis zu einem n -dimensionalen Vektor addiert werden. Für die Matrix-Vektor-Multiplikation werden ungefähr n^2 Multiplikationen und Additionen benötigt.

Bei der Prädiktion der Kovarianzmatrix muss in jedem Zeitschritt eine Multiplikation zweier $n \times n$ -Matrizen ($\mathbf{A}(t)\mathbf{P}(t)$) ist das Transponierte von $\mathbf{P}(t)\mathbf{A}^T$) berechnet werden. Dazu sind ungefähr n^3 Multiplikationen und Additionen nötig. Insbesondere bei Systemen höherer Ordnung übersteigt damit der Aufwand der Prädiktion der Kovarianzmatrix deutlich den Aufwand der Prädiktion des Schätzwertes.

Da der Schätzwert in gewisser Weise noch als der wichtigere Wert angesehen werden kann, da ein Fehler bei der Berechnung sich „direkter“ auf die Güte der Schätzung auswirkt, während eine falsch bestimmte Kovarianzmatrix sich „nur“ in einer falschen Gewichtung äußert, wird teilweise auf diese Weise der Rechenaufwand reduziert.

Kontinuierliches Kalman-Bucy-Filer

Im Fall, dass auch die Messungen ständig vorliegen, geht man zum Kalman-Bucy-Filter über. Dies wird hier nicht mehr betrachtet.

10.5 Erweitertes Kalmanfilter

Liegt ein nichtlineares Systemmodell vor, so beruht die Ableitung des sogenannten erweiterten Kalmanfilters (Extended Kalmanfilter) auf einer linearisierten Darstellung. Damit ergeben sich Gleichungen, die sehr nahe an den bisherigen Kalmanfiltergleichungen liegen, jedoch ist durch die Linearisierung weder die Optimalität noch überhaupt die Stabilität sichergestellt. (Für die Herleitungen wird auf entsprechende Literatur verwiesen, z. B. [Kamen und Su 1999].)

Bei nichtlinearen Systemen wird auch die Wahl der Anfangswerte der Schätzwerte kritischer, und es erweist sich teilweise als schwierig, geeignete Startkovarianzen \mathbf{P}_0 sowie geeignete Werte für \mathbf{Q} zu finden.

10.5.1 Zeitdiskretes erweitertes Kalmanfilter

- Systemgleichungen

$$\begin{aligned}\mathbf{x}_{k+1} &= \mathbf{f}(\mathbf{x}_k, \mathbf{u}_k) + \mathbf{n}_{x,k} \\ \tilde{\mathbf{y}}_k &= \mathbf{h}(\mathbf{x}_k, \mathbf{u}_k) + \mathbf{n}_{y,k}\end{aligned}$$

- Linearisierte Matrizen

$$\mathbf{A}_k = \left. \frac{\partial \mathbf{f}(\mathbf{x}, \mathbf{u})}{\partial \mathbf{x}} \right|_{\hat{\mathbf{x}}_{k,-}, \mathbf{u}_k} \quad \mathbf{C}_k = \left. \frac{\partial \mathbf{h}(\mathbf{x}, \mathbf{u})}{\partial \mathbf{x}} \right|_{\hat{\mathbf{x}}_{k,-}, \mathbf{u}_k}$$

- Prädiktion

$$\begin{aligned}\hat{\mathbf{x}}_{k+1,-} &= \mathbf{f}(\hat{\mathbf{x}}_{k,+}, \mathbf{u}_k) \\ \mathbf{P}_{k+1,-} &= \mathbf{A}_k \mathbf{P}_{k,+} \mathbf{A}_k^T + \mathbf{Q}_k\end{aligned}$$

- Korrektur

$$\begin{aligned}\mathbf{K}_k &= \mathbf{P}_{k,-} \mathbf{C}_k^T \cdot (\mathbf{C}_k \mathbf{P}_{k,-} \mathbf{C}_k^T + \mathbf{R}_k)^{-1} \\ \hat{\mathbf{x}}_{k,+} &= \hat{\mathbf{x}}_{k,-} + \mathbf{K}_k \cdot (\tilde{\mathbf{y}}_k - \mathbf{h}(\hat{\mathbf{x}}_{k,-}, \mathbf{u}_k)) \\ \mathbf{P}_{k,+} &= (\mathbf{I} - \mathbf{K}_k \mathbf{C}_k) \cdot \mathbf{P}_{k,-}\end{aligned}$$

Anmerkungen

- Es wird nur die linearisierte System- und Ausgangsmatrix benötigt. In der Prädiktion des Zustandes wird direkt die nichtlineare Zustandsgleichung und in der Korrektur des Zustandes die nichtlineare Ausgangsgleichung verwendet.
- Die linearisierten Matrizen hängen vom geschätzten Systemzustand sowie der Eingangsgröße ab, und können daher nicht vorher berechnet werden.
- Die Simulation (Prädiktion) der nichtlinearen Zustandsgleichung muss wieder durch ein geeignetes Verfahren erfolgen.

10.5.2 Kontinuierlich-diskretes erweitertes Kalmanfilter

- Systemgleichungen

$$\dot{\hat{\mathbf{x}}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t)) + \mathbf{n}_x(t)$$

$$\tilde{\mathbf{y}}_k = \mathbf{h}(\mathbf{x}_k, \mathbf{u}_k) + \mathbf{n}_{y,k}$$

- Linearisierte Matrizen

$$\mathbf{A}(t) = \left. \frac{\partial \mathbf{f}(\mathbf{x}, \mathbf{u})}{\partial \mathbf{x}} \right|_{\hat{\mathbf{x}}(t), \mathbf{u}(t)} \quad \mathbf{C}_k = \left. \frac{\partial \mathbf{h}(\mathbf{x}, \mathbf{u})}{\partial \mathbf{x}} \right|_{\hat{\mathbf{x}}_{k,-}, \mathbf{u}_k}$$

- Prädiktion

$$\dot{\hat{\mathbf{x}}}(t) = \mathbf{f}(\hat{\mathbf{x}}(t), \mathbf{u}(t))$$

$$\hat{\mathbf{x}}(t_k) = \hat{\mathbf{x}}_{k,+}$$

$$\dot{\mathbf{P}}(t) = \mathbf{A}(t)\mathbf{P}(t) + \mathbf{P}(t)\mathbf{A}^T(t) + \mathbf{Q}(t)$$

$$\mathbf{P}(t_k) = \mathbf{P}_{k,+}$$

$$\Rightarrow \hat{\mathbf{x}}_{k+1,-} = \hat{\mathbf{x}}(t_{k+1}), \quad \mathbf{P}_{k+1,-} = \mathbf{P}(t_{k+1})$$

- Korrektur

$$\mathbf{K}_k = \mathbf{P}_{k,-} \mathbf{C}_k^T \cdot (\mathbf{C}_k \mathbf{P}_{k,-} \mathbf{C}_k^T + \mathbf{R}_k)^{-1}$$

$$\hat{\mathbf{x}}_{k,+} = \hat{\mathbf{x}}_{k,-} + \mathbf{K}_k \cdot (\tilde{\mathbf{y}}_k - \mathbf{h}(\hat{\mathbf{x}}_{k,-}, \mathbf{u}_k))$$

$$\mathbf{P}_{k,+} = (\mathbf{I} - \mathbf{K}_k \mathbf{C}_k) \cdot \mathbf{P}_{k,-}$$

Anmerkungen

- $\mathbf{A}(t)$ hängt im Allgemeinen von der aktuellen geschätzten Zustandsgröße ab. Soll dies für die Prädiktion der Fehlerkovarianz möglichst genau berücksichtigt werden, dann muss der Zeitverlauf von $\hat{\mathbf{x}}(t)$ innerhalb des Prädiktionsintervalls durch die Prädiktion der Zustände schon bekannt sein, oder die beiden Differentialgleichungen müssen parallel simuliert werden.
- Auch hier wird teilweise zur Reduktion des Rechenaufwands die Fehlerkovarianz diskret prädiziert oder zumindest mit größeren Rechenschrittweiten simuliert. Dies kann durch die Nichtlinearitäten aber schnell dazu führen, dass das Kalmanfilter gar nicht mehr konvergiert.

10.6 Beispiel

Es wird das Beispiel aus Abschnitt 10.1, die rotatorische Trägheit, betrachtet, wobei die Dämpfung durch ein Kalmanfilter identifiziert werden soll. Damit ergibt sich das erweiterte, nichtlineare Zustandsraummodell

$$\begin{bmatrix} \dot{\varphi}(t) \\ \dot{\omega}(t) \\ \dot{x}_d(t) \end{bmatrix} = \begin{bmatrix} \omega(t) \\ -\frac{1}{J} \cdot x_d(t) \cdot \omega(t) + \frac{1}{J} \cdot M(t) \\ 0 \end{bmatrix}$$

$$y(t) = \varphi(t)$$

für das ein Kalmanfilter zu entwerfen ist.

Die Ausgangsgleichung ist hier linear, und es lässt sich einfach

$$\mathbf{C}_k = [1 \ 0 \ 0]$$

angeben. Für die Systemmatrix erhält man durch Linearisierung

$$\mathbf{A}(t) = \left. \frac{\partial \mathbf{f}(\mathbf{x}, \mathbf{u})}{\partial \mathbf{x}} \right|_{\hat{\mathbf{x}}(t), \mathbf{u}(t)} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & -\frac{1}{J} \cdot \hat{x}_d(t) & -\frac{1}{J} \cdot \hat{\omega}(t) \\ 0 & 0 & 0 \end{bmatrix}.$$

Für dieses System ist die Identifikation mit dem Kalmanfilter simuliert. Zur Identifikation wird ein konstantes Eingangsmoment von 1 Nm verwendet, also die Sprungantwort betrachtet.

Der wahre Zustandswert zu Beginn ist

$$\mathbf{x}(t_0) = \begin{bmatrix} 0,1 \text{ rad} \\ 0,5 \text{ rad/s} \\ 0,05 \text{ Nm/(rad/s)} \end{bmatrix}.$$

Das System befindet sich also nicht in Ruhe. Der Schätzer wird jedoch mit

$$\hat{\mathbf{x}}_0 = [0 \ 0 \ 0]^T$$

initialisiert, und die Fehlervarianz wird zu Beginn auf

$$\mathbf{P}_0 = \begin{bmatrix} 1000 \text{ rad}^2 & & \\ & 1000 (\text{rad/s})^2 & \\ & & 1 (\text{Nm}/(\text{rad/s}))^2 \end{bmatrix}$$

gesetzt. Wenn bekannt ist, dass sich die Zustandswerte um Größenordnungen unterscheiden, dann sollte sich dies auch hier widerspiegeln.

Das Systemrauschen wird als Vielfaches des Anfangswertes der Fehlervarianzmatrix parametriert,

$$\mathbf{Q} = \gamma \cdot \mathbf{P}_0.$$

Es sind Simulationen für $\gamma = 1 \cdot 10^{-6}$ und $\gamma = 0$ durchgeführt.

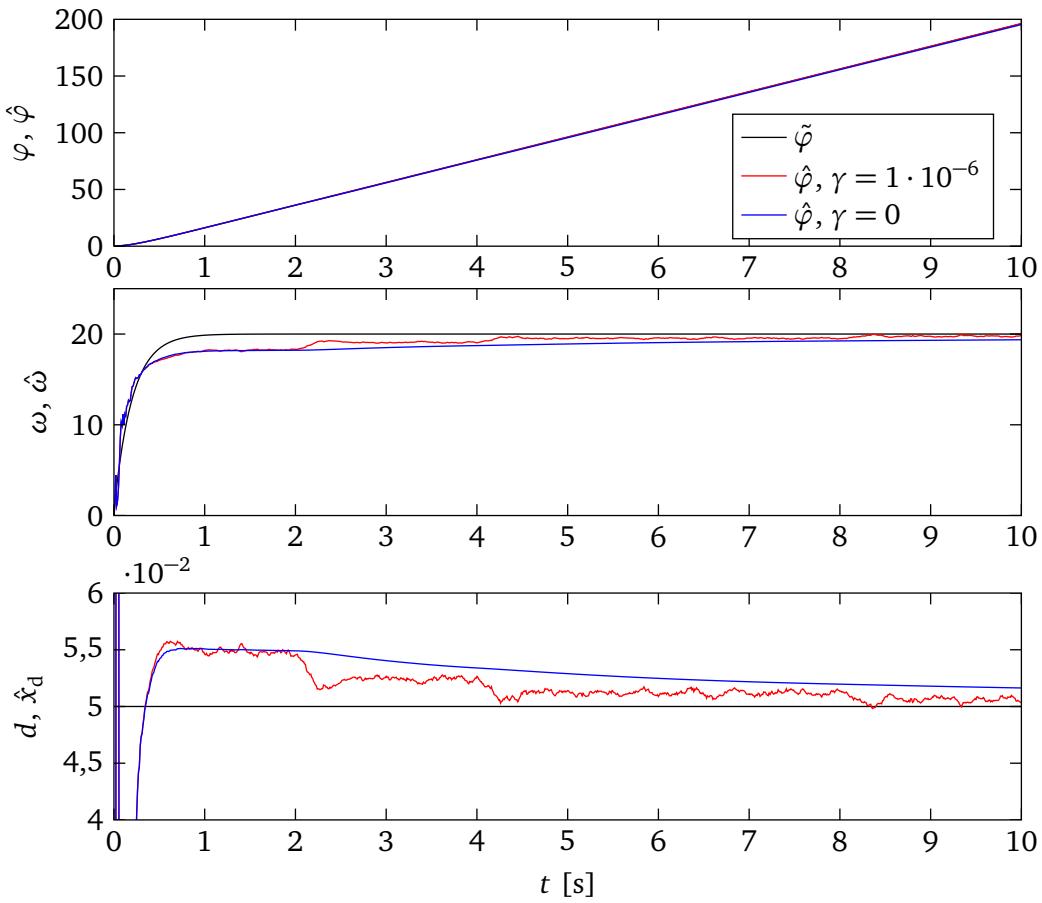


Abbildung 10.4: Identifikation der viskosen Dämpfung

Die Messungen, welche mit einer Abtastzeit von 10 ms vorliegen, sind mit einem diskretes weißes Rauschen mit einer Standardabweichung von $\sigma_y = 0,02 \text{ rad}$ (ca. 1°) simuliert. Entsprechend wird

$$\mathbf{R} = (0,02 \text{ rad})^2$$

verwendet.

Das Kalmanfilter ist als erweitertes, zeitkontinuierlich-diskretes Filter implementiert. Die Lösung der Differentialgleichungen für die Prädiktion erfolgt mit dem Runge-Kutta-4-Verfahren und einer Rechenschrittweite von $\Delta t = 1 \text{ ms}$.

In Abbildung 10.4 sind die Ergebnisse dargestellt.

Der obere Plot zeigt den gemessenen sowie den geschätzten Ausgang. Letzteren für die beiden Fälle $\gamma = 1 \cdot 10^{-6}$ und $\gamma = 0$. Der mittlere Plot zeigt die wahre und die geschätzte Geschwindigkeit, der untere den konstanten wahren Wert für die Dämpfung und die entsprechenden Schätzwerte.

Man erkennt, dass während der ersten halben Sekunde insbesondere für die Dämpfung extrem falsche Werte geschätzt werden, bis ein „stabiler“ Schätzwert erreicht wird, der jedoch nicht dem wahren Wert entspricht. Im weiteren Verlauf nähert sich der Schätzwert sehr langsam dem wahren Wert, wobei die Annäherung im Fall $\gamma = 0$, also ohne Annahme von Systemrauschen, deutlich langsamer ist als für den Fall $\gamma = 1 \cdot 10^{-6}$. Dafür sind die Schwankungen des Schätzwertes bei $\gamma = 1 \cdot 10^{-6}$ größer.

Das angenommene Systemrauschen bei $\gamma = 1 \cdot 10^{-6}$ begrenzt die Varianz des Schätzfehlers nach unten und ermöglicht so, dass immer eine gewisse Korrektur mit den Messwerten vorgenommen wird.

Es ist auch möglich, das Systemrauschen zeitvariabel vorzugeben. D. h. zu Beginn ein relativ hohes Systemrauschen anzunehmen, um sich möglichst schnell den wahren Parametern zu nähern, und dann auf

ein schwaches Systemrauschen umzuschalten, um die Streuungen (über die Zeit gesehen) im Schätzvektor zu minimieren.

Anmerkung

Zu dem Beispiel muss man auch erwähnen, dass die Einstellung von \hat{x}_0 , P_0 , Q und R sehr schwierig sein kann. Das erweiterte Kalmanfilter reagiert teilweise sehr empfindlich auf eine ungünstige Wahl dieser Werte, und auch die Prädiktion, d. h. die Simulation kann je nach System und gemessener Größen sehr kritisch sein. Dies bedeutet, dass bei einer ungünstigen Wahl der Schätzwert zu „Phantasiewerten“ konvergiert oder auch divergieren kann.

Dies hängt auch von den Messgrößen ab, die zur Verfügung stehen. Im vorliegenden Beispiel wäre es zum Beispiel deutlich besser, wenn die Geschwindigkeit und nicht die Position zurückgeführt werden würde.

Das Kalmanfilter ist damit eine Möglichkeit zur Identifikation, wenn online spezielle physikalische Systemparameter, die sich ggf. auch mit der Zeit ändern können, geschätzt werden sollen.

Im vorliegenden Fall hätte man auch auf das normale LS-Verfahren zurückgreifen können. Wenn man die Parameter c und J als bekannt voraussetzt, kann der Parameter d direkt geschätzt werden. Auch könnte man (für eine Offline-Identifikation) den Parameter über eine Ausgangsfehleroptimierung ermitteln. Bei nur einem Parameter geht dies schnell und ist auch „sicherer“ durchzuführen als die Kalmanfilterung.

A Stochastik

A.1 Varianz Schätzung des quadratischen Mittels

Auschreiben der Definition der Varianz für den Schätzer (4.9) führt über Ausmultiplizieren und Zusammenfassen

$$\begin{aligned}\text{Var}\{\hat{r}_{\text{xx},k}\}_r &= E\{(\hat{r}_{\text{xx},k} - E\{\hat{r}_{\text{xx},k}\}_r)^2\}_r \\ &= E\{(\hat{r}_{\text{xx},k} - \bar{r}_{\text{xx},k} - \sigma_x^2)^2\}_r \\ &= E\{\hat{r}_{\text{xx},k}^2\}_r + \bar{r}_{\text{xx},k}^2 + (\sigma_x^2)^2 - 2 \cdot E\{\hat{r}_{\text{xx},k}\}_r \cdot \bar{r}_{\text{xx},k} - 2 \cdot E\{\hat{r}_{\text{xx},k}\}_r \cdot \sigma_x^2 + 2 \cdot \bar{r}_{\text{xx},k} \cdot \sigma_x^2 \\ &= E\{\hat{r}_{\text{xx},k}^2\}_r - \bar{r}_{\text{xx},k}^2 - 2 \cdot \bar{r}_{\text{xx},k} \sigma_x^2 - (\sigma_x^2)^2,\end{aligned}\quad (\text{A.1})$$

wobei $E\{\hat{r}_{\text{xx},k}\}_r = \bar{r}_{\text{xx},k} + \sigma_x^2$ verwendet wurde. Der verbleibende Erwartungswert kann analog der Rechnung zur Varianz des einfachen Mittelwertschätzers als

$$E\{\hat{r}_{\text{xx},k}^2\}_r = \frac{1}{k^2} \cdot \left(\sum_{k'} E\{\tilde{x}_{k'}^4\}_r + \sum_{k'} \sum_{\substack{k'' \\ k'' \neq k'}} E\{\tilde{x}_{k'}^2\}_r \cdot E\{\tilde{x}_{k''}^2\}_r \right) \quad (\text{A.2})$$

geschrieben werden.

Für den ersten Erwartungswert ergibt sich

$$\begin{aligned}E\{\tilde{x}_{k'}^4\}_r &= E\{(x_{k'} + n_{k'})^4\}_r \\ &= x_{k'}^4 + 4x_{k'}^3 \cdot E\{n_{k'}\}_r + 6x_{k'}^2 \cdot E\{n_{k'}^2\}_r + 4x_{k'} \cdot E\{n_{k'}^3\}_r + E\{n_{k'}^4\}_r \\ &= x_{k'}^4 + 6x_{k'}^2 \cdot \sigma_x^2 + 4x_{k'} \cdot \mu_{\text{xxx}} + \mu_{\text{xxxx}},\end{aligned}$$

wobei μ_{xxx} und μ_{xxxx} das dritte bzw. vierte statistische Moment des Messrauschen bezeichnet. Diese Werte sind konstant und als endlich anzunehmen.

Setzt man den ersten Erwartungswert wieder in die Summe ein, so erhält man

$$\begin{aligned}\sum_{k'} E\{\tilde{x}_{k'}^4\}_r &= \sum_{k'} x_{k'}^4 + \sum_{k'} 6x_{k'}^2 \cdot \sigma_x^2 + \sum_{k'} 4x_{k'} \cdot \mu_{\text{xxx}} + \sum_{k'} \mu_{\text{xxxx}} \\ &= \sum_{k'} x_{k'}^4 + 6\sigma_x^2 \cdot \sum_{k'} x_{k'}^2 + 4 \cdot \mu_{\text{xxx}} \cdot \sum_{k'} x_{k'} + k \cdot \mu_{\text{xxxx}}.\end{aligned}$$

Die zweite Summe in (A.2) kann nach

$$\begin{aligned}\sum_{k'} \sum_{\substack{k'' \\ k'' \neq k'}} E\{\tilde{x}_{k'}^2\}_r \cdot E\{\tilde{x}_{k''}^2\}_r &= \sum_{k'} \sum_{\substack{k'' \\ k'' \neq k'}} (x_{k'}^2 + \sigma_x^2) \cdot (x_{k''}^2 + \sigma_x^2) \\ &= \sum_{k'} \sum_{\substack{k'' \\ k'' \neq k'}} (x_{k'}^2 x_{k''}^2 + \sigma_x^2 x_{k'}^2 + \sigma_x^2 x_{k''}^2 + (\sigma_x^2)^2) \\ &= \sum_{k'} \sum_{\substack{k'' \\ k'' \neq k'}} x_{k'}^2 x_{k''}^2 + (k-1) \cdot \sigma_x^2 \cdot \sum_{k'} x_{k'}^2 + \sigma_x^2 \cdot \sum_{k'} \sum_{\substack{k'' \\ k'' \neq k'}} x_{k''}^2 + k \cdot (k-1) \cdot (\sigma_x^2)^2\end{aligned}$$

umgeformt werden.

Für den Term in der Klammer von Gl. (A.2) ergibt sich damit

$$\begin{aligned}
& \sum_{k'} E\{\tilde{x}_{k'}^4\}_r + \sum_{k'} \sum_{\substack{k'' \\ k'' \neq k'}} E\{\tilde{x}_{k'}^2\}_r \cdot E\{\tilde{x}_{k''}^2\}_r \\
&= \underbrace{\sum_{k'} x_{k'}^4 + \sum_{k'} \sum_{\substack{k'' \\ k'' \neq k'}} x_{k'}^2 x_{k''}^2}_{\sum_{k'} \sum_{k''} x_{k'}^2 x_{k''}^2} + \underbrace{6\sigma_x^2 \cdot \sum_{k'} x_{k'}^2 + \sigma_x^2 \cdot \sum_{k'} \sum_{\substack{k'' \\ k'' \neq k'}} x_{k''}^2}_{5\sigma_x^2 \cdot \sum_{k'} x_{k'}^2 + \sigma_x^2 \cdot \sum_{k'} \sum_{k''} x_{k''}^2} \\
&\quad + (k-1) \cdot \sigma_x^2 \sum_{k'} x_{k'}^2 + k \cdot (k-1) \cdot (\sigma_x^2)^2 + 4 \cdot \mu_{xxx} \cdot \sum_{k'} x_{k'} + k \cdot \mu_{xxxx} \\
&= \sum_{k'} x_{k'}^2 \sum_{k''} x_{k''}^2 + 5\sigma_x^2 \cdot \sum_{k'} x_{k'}^2 + \sigma_x^2 \cdot k \cdot \sum_{k''} x_{k''}^2 + (k-1) \cdot \sigma_x^2 \sum_{k'} x_{k'}^2 \\
&\quad + k \cdot (k-1) \cdot (\sigma_x^2)^2 + 4 \cdot \mu_{xxx} \cdot \sum_{k'} x_{k'} + k \cdot \mu_{xxxx}.
\end{aligned}$$

Damit kann

$$\begin{aligned}
E\{\hat{r}_{xx,k}^2\}_r &= \frac{1}{k^2} \cdot \left(\sum_{k'} E\{\tilde{x}_{k'}^4\}_r + \sum_{k'} \sum_{\substack{k'' \\ k'' \neq k'}} E\{\tilde{x}_{k'}^2\}_r \cdot E\{\tilde{x}_{k''}^2\}_r \right) \\
&= \frac{1}{k} \cdot \sum_{k'} x_{k'}^2 \cdot \frac{1}{k} \cdot \sum_{k''} x_{k''}^2 + 5\sigma_x^2 \cdot \frac{1}{k} \cdot \frac{1}{k} \cdot \sum_{k'} x_{k'}^2 + \sigma_x^2 \cdot \frac{1}{k} \cdot \sum_{k''} x_{k''}^2 + \frac{(k-1)}{k} \cdot \sigma_x^2 \cdot \frac{1}{k} \cdot \sum_{k'} x_{k'}^2 \\
&\quad + \frac{k-1}{k} \cdot (\sigma_x^2)^2 + 4 \cdot \mu_{xxx} \cdot \frac{1}{k} \cdot \frac{1}{k} \cdot \sum_{k'} x_{k'} + \frac{1}{k} \cdot \mu_{xxxx} \\
&= \bar{r}_{xx,k}^2 + 5 \cdot \frac{1}{k} \cdot \sigma_x^2 \cdot \bar{r}_{xx,k} + \sigma_x^2 \cdot \bar{r}_{xx,k} + \frac{(k-1)}{k} \cdot \sigma_x^2 \cdot \bar{r}_{xx,k} \\
&\quad + \frac{k-1}{k} \cdot (\sigma_x^2)^2 + 4 \cdot \mu_{xxx} \cdot \frac{1}{k} \cdot \bar{x}_k + \frac{1}{k} \cdot \mu_{xxxx} \\
&= \bar{r}_{xx,k}^2 + 2\sigma_x^2 \cdot \bar{r}_{xx,k} + (\sigma_x^2)^2 + \frac{1}{k} \cdot (4 \cdot \sigma_x^2 \cdot \bar{r}_{xx,k} - (\sigma_x^2)^2 + 4 \cdot \mu_{xxx} \cdot \bar{x}_k + \mu_{xxxx})
\end{aligned}$$

angegeben werden. Dies in Gl. (A.1) führt auf den gesuchten Ausdruck

$$\begin{aligned}
\text{Var}\{\hat{r}_{xx,k}\}_r &= E\{\hat{r}_{xx,k}^2\}_r - \bar{r}_{xx,k}^2 - 2 \cdot \bar{r}_{xx,k} \sigma_x^2 - (\sigma_x^2)^2 \\
&= \frac{1}{k} \cdot (4 \cdot \sigma_x^2 \cdot \bar{r}_{xx,k} + 4 \cdot \mu_{xxx} \cdot \bar{x}_k + \mu_{xxxx} - (\sigma_x^2)^2)
\end{aligned}$$

für die Varianz.

B Signalverarbeitung

B.1 atan2

In Abbildung B.1 sind vier Punkte in einem zweidimensionalen kartesischen Koordinatensystem mit gleichem Betrag $|y/x|$ markiert. Gesucht sind die Winkel zur positiven reellen Achse. Liegt der – häufige – Fall vor, dass das Argument einer komplexen Zahl gesucht ist, so wäre x der Real- und y der Imaginärteil dieser Zahl.

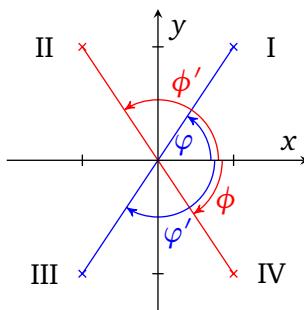


Abbildung B.1

Zur Umrechnung von (x, y) in einen Winkel kann nicht einfach der Arkustangens verwendet werden. Im gegebenen Beispiel haben der Tangens von φ und φ' den selben Wert. Gleiches gilt für den Tangens von ϕ und ϕ' . Anders ausgedrückt können die Quadranten I und III sowie II und IV nicht anhand des Quotienten y/x unterschieden werden, $\frac{y}{x} = \frac{-y}{-x}$ bzw. $\frac{-y}{x} = \frac{y}{-x}$. Entsprechend gibt die Arkustangensfunktion nur Werte von $(-\pi/2, \pi/2)$ zurück.

Bei der Umrechnung von kartesischen Koordinaten muss anhand der Vorzeichen von x und y noch eine Fallunterscheidung getroffen werden. Da dies eine häufige Aufgabe ist, ist dafür eine Arkustangensfunktion mit zwei Argumenten eingeführt, die auch jede höhere Programmiersprache zur Verfügung stellt. Dieser werden die Werte für x und y (in der Reihenfolge y - x !) einzeln übergeben. Darüberhinaus kann damit auch einfach der Fall $x = 0$ abgefangen werden, für den der Quotient y/x nicht definiert ist.

Die atan2-Funktion ist dann als

$$\text{atan2}(y, x) = \begin{cases} \text{atan}(y/x) & \text{für } x > 0 \text{ (Quadranten I + IV)} \\ \text{atan}(y/x) + \pi & \text{für } x < 0 \wedge y \geq 0 \text{ (Quadrant II)} \\ \text{atan}(y/x) - \pi & \text{für } x < 0 \wedge y < 0 \text{ (Quadrant III)} \\ \pi/2 & \text{für } x = 0 \wedge y > 0 \text{ (positive Ordinate)} \\ -\pi/2 & \text{für } x = 0 \wedge y < 0 \text{ (negative Ordinate)} \\ 0 & \text{für } x = 0 \wedge y = 0 \text{ (Ursprung)} \end{cases}$$

definiert. (Der Fall $x = 0 \wedge y = 0$ kann auch unterschiedlich behandelt werden.) Diese gibt also Werte zwischen $(-\pi, \pi]$ zurück.

B.2 Verbindung zwischen der DFT und der Fourier-Reihe bei periodischen Signalen

Falls die DFT von einem periodischen Signal bestimmt wird, kein Aliasing auftritt und NT_s einem ganzen Vielfachen n der Periodendauer T entspricht, $NT_s = nT$, dann kann ein Zusammenhang zwischen den Werten der DFT und den Fourier-Koeffizienten aufgezeigt werden.

Hierbei sei t_0 die Zeit des ersten Samples der N Daten über die die DFT berechnet wird. Dann gilt $k = \frac{t-t_0}{T_s}$, wobei T_s die Abtastzeit ist. Wenn nun der Wert $y(t)$ aus den DFT-Werten berechnet werden soll, dann muss zunächst k bestimmt werden, und für dieses dann die inverse DFT ausgewertet werden, d. h.

$$y(t) = y_{\frac{t-t_0}{T_s}} = \frac{1}{N} \cdot \sum_{l=0}^{N-1} e^{j2\pi \cdot \frac{t-t_0}{T_s \cdot N} \cdot l} \cdot Y_l, \quad \frac{t}{T_s} \in \mathbb{N}. \quad (\text{B.1})$$

Mit $\frac{2\pi}{T_s \cdot N} = \frac{2\pi}{nT} = \omega_{Tn}$ ergeben sich die folgenden Umformungen:

$$\begin{aligned} y(t) &= \frac{1}{N} \cdot \sum_{l=0}^{N-1} e^{j\omega_{Tn} \cdot l \cdot (t-t_0)} \cdot Y_l \\ &= \frac{1}{N} \cdot \sum_{l=0}^{N-1} \frac{1}{N} \cdot Y_l \cdot e^{j\omega_{Tn} \cdot l \cdot (t-t_0)} \\ &= \frac{1}{N} \cdot \sum_{l=0}^{N-1} Y_l \cdot e^{-j\omega_{Tn} \cdot l \cdot t_0} \cdot e^{j\omega_{Tn} \cdot l \cdot t} \\ &= \frac{1}{N} \cdot \sum_{l=0}^{N-1} \tilde{Y}_l \cdot e^{j\omega_{Tn} \cdot l \cdot t}. \end{aligned}$$

Die Summe wird jetzt in zwei Teile zerlegt (N gerade, o. B. d. A., da bei ungeraden N der Wert $Y_{N-1} = 0$ sein muss, da Aliasing ausgeschlossen wurde),

$$y(t) = \frac{1}{N} \cdot \sum_{l=0}^{N/2-1} \tilde{Y}_l \cdot e^{j\omega_{Tn} \cdot l \cdot t} + \frac{1}{N} \cdot \sum_{l=N/2}^{N-1} \tilde{Y}_l \cdot e^{j\omega_{Tn} \cdot l \cdot t}.$$

Mit der Substitution $l' = l - N$ der Laufvariablen der zweiten Summe ergibt sich

$$y(t) = \frac{1}{N} \cdot \sum_{l=0}^{N/2-1} \tilde{Y}_l \cdot e^{j\omega_{Tn} \cdot n \cdot t} + \frac{1}{N} \cdot \sum_{l'=-N/2}^{-1} \tilde{Y}_{N+l'} \cdot e^{j\omega_{Tn} \cdot (N+l') \cdot t}.$$

Da Y_l und $e^{j\omega_{Tn} \cdot l \cdot t}$ für $t/T_s \in \mathbb{N}$ N -periodisch sind, gilt

$$\begin{aligned} y(t) &= \frac{1}{N} \cdot \sum_{l=0}^{N/2-1} \tilde{Y}_l \cdot e^{j\omega_{Tn} \cdot l \cdot t} + \frac{1}{N} \cdot \sum_{l'=-N/2}^{-1} \tilde{Y}_{l'} \cdot e^{j\omega_{Tn} \cdot l' \cdot t} \\ &= \sum_{l=-N/2}^{N/2-1} \frac{\tilde{Y}_l}{N} \cdot e^{j\omega_{Tn} \cdot l \cdot t}. \end{aligned}$$

Wird die letzte Zeile mit der Gleichung (5.1) verglichen sieht man, dass diese für mit $\check{y}_l = \frac{\tilde{Y}_l}{N}$ identisch sind, wenn $\check{y}_l = 0 \forall \{l \mid |l| \geq N/2\}$ gilt.

Damit sind die Fourierkoeffizienten bezüglich der Periodendauer nT bzw. der Grundfrequenz ω_{Tn} gegeben. Sollen die Fourierkoeffizienten bezüglich der (minimalen) Periodendauer T bzw. der dazugehörigen Grundfrequenz ω_T des betrachteten Signals $y(t)$ angegeben werden, muss nur jeder n -te Wert der DFT betrachtet werden,

$$\check{y}_l = \frac{1}{N} \cdot \tilde{Y}_{nl}$$

Alle anderen Werte der DFT sind gleich Null.

Damit können für periodische Signale die Fourier-Koeffizienten \check{y}_l bestimmt werden, indem die DFT-Werte Y_{nl} mit $\frac{1}{N}$ skaliert werden, und ggf. die Phase mit dem Startpunkt t_0 korrigiert wird:

$$\check{y}_l = \frac{1}{N} \cdot Y_{nl} \cdot e^{-j\omega_T l t_0}. \quad (\text{B.2})$$

Daraus folgt auch, dass Gl. (B.1) auch für nichtganzzahlige $\frac{t-t_0}{T_s}$ ausgewertet werden kann, bzw. das die Gleichung der inversen DFT in diesem Fall auch für nicht ganzzahlige k sinnvolle Ergebnisse liefert.

B.3 Goertzel-Algorithmus

Der Goertzel-Algorithmus ist eine Methode zur Berechnung des Wertes X_l der DFT eines Signals x_k zu einem gegebenen Frequenzindex l . Die Definition von X_l ist

$$X(l) = \sum_{k=0}^{N-1} x_k W_N^{l \cdot k} \quad (\text{B.3})$$

mit

$$W_N = e^{-j \cdot \frac{2\pi}{N}}. \quad (\text{B.4})$$

Selbst wenn x_k ein N -periodisches Signal wäre, wird für den Zweck dieses Algorithmus davon ausgegangen, dass x_k Null außerhalb von $0 \leq k < N$ ist.

Eine Herleitung des ALgorithmus ist in [Oppenheim und Schafer 1995] zu finden. Dazu wird zunächst die Reihe

$$y_{l,r} = \sum_{k=0}^r x_k W_N^{-l \cdot (r-k)} \quad (\text{B.5})$$

definiert. Der erste Index l bezeichnet dabei den Frequenzindex und der zweite, hier r , das Reihenglied. Da $x_N = 0$ und Gl. (B.3) berücksichtigend ergibt sich

$$y_{l,N} = \sum_{k=0}^N x_k W_N^{-l \cdot (N-k)} = \sum_{k=0}^{N-1} x_k W_N^{-l \cdot (N-k)} = X_l \quad (\text{B.6})$$

wenn l eine Ganzzahl ist, da $W_N^{lN} = 1$ für alle Ganzzahlen l ist. D.h. $y_{l,N}$ entspricht dem Wert der DFT von x_k für den Frequenzindex l .

Im Gegensatz zu [Oppenheim und Schafer 1995] beruht die weitere Herleitung nicht auf systemtheoretischen Argumentationen sondern nur auf algebraischen Umformungen.

Das letzte Element der Summe in Gl. (B.5) wird abgespalten und so

$$y_{l,r} = x_r W_N^{-l \cdot 0} + \sum_{k=0}^{r-1} x_k W_N^{-l \cdot (r-k)}$$

erhalten. W_N^0 ist 1 und W_N^{-l} wird aus der Summe ausgeklammert was auf

$$y_{l,r} = x_r + W_N^{-l} \cdot \sum_{k=0}^{r-1} x_k W_N^{-l \cdot (r-1-k)}$$

führt. Die verbleibende Summe entspricht $y_{l,r-1}$ und es kann damit

$$y_{l,r} = x_r + W_N^{-l} y_{l,r-1} \quad (\text{B.7})$$

geschrieben werden. Für $r = 0$ erfordert Gl. (B.5) dass $y_{l,0} =! x_0$ gilt, was den Startwert

$$y_{l,-1} = 0$$

festlegt.

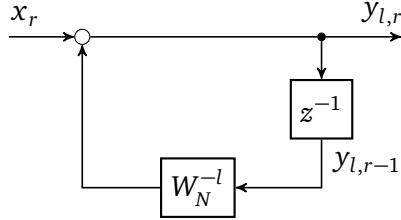


Abbildung B.2: Goertzel-Algorithmus als IIR-Filter erster Ordnung

Mit Gl. (B.7) ist die DFT für einen Frequenzindependency über ein IIR-Filter erster Ordnung gegeben, welches in Abbildung B.2 dargestellt ist. Es muss dabei lediglich ein Faktor W_N^l gespeichert werden. Da aber W_N^l im Allgemeinen eine komplexe Zahl ist, müssen je Iteration zwei Multiplikationen durchgeführt werden (wenn x_k ein reelles Signal ist).

Mit Gl. (B.7) kann der Ausdruck $y_{l,r} - W_N^l \cdot y_{l,r-1}$ auch als

$$y_{l,r} - W_N^l \cdot y_{l,r-1} = x_r + W_N^{-l} y_{l,r-1} - W_N^l x_{r-1} - W_N^l W_N^{-l} y_{l,r-2}$$

geschrieben werden. Bringt man $-W_N^l \cdot y_{l,r-1}$ auf die andere Seite und unter Berücksichtigung, dass $W_N^l + W_N^{-l} = 2 \cdot \operatorname{Re}(W_N^l) = 2 \cdot \cos \frac{2\pi l}{N} = a_1$ gilt (wobei a_1 als Abkürzung verwendet wird), dann kann $y_{l,r}$ als IIR-Filter zweiter Ordnung

$$y_{l,r} = x_r - W_N^l x_{r-1} + a_1 \cdot y_{l,r-1} - y_{l,r-2}$$

geschrieben werden, wobei sich die Anfangswerte

$$y_{l,-1} = y_{l,-2} = 0$$

aus den Bedingungen $y_{l,0} =! x_0$ und $y_{l,1} =! x_0 W_N^{-l} + x_1$, die aus Gl. (B.5) folgen, ergeben.

Um die komplexe Multiplikation $W_N^l x_{r-1}$ in jedem Schritt zu vermeiden wird der Filter in einen IIR-Teil

$$\tilde{y}_{l,r} = x_r + a_1 \tilde{y}_{l,r-1} - \tilde{y}_{l,r-2} \quad (\text{B.8})$$

mit

$$\tilde{y}_{l,-1} = \tilde{y}_{l,-2} = 0 \quad (\text{B.9})$$

und einen FIR-Teil

$$y_{l,r} = \tilde{y}_{l,r} - W_N^l \tilde{y}_{l,r-1} \quad (\text{B.10})$$

zerlegt. Die Gültigkeit dieser Zerlegung kann überprüft werden, indem $\tilde{y}_{l,r}$ und $\tilde{y}_{l,r-1}$ in Gl. (B.10) durch Gl. (B.8) ausgedrückt werden.

Der durch Gl. (B.8) und (B.10) definierte Algorithmus ist in Abbildung B.3 dargestellt. Gl. (B.10) muss dabei nur beim letzten Schritt verwendet werden. In allen anderen Schritten wird nur Gl. (B.8), die in Abbildung B.3 in grau hinterlegt ist, ausgewertet. Da a_1 sowie x_r reellwertig sind müssen in diesen Schritten jeweils nur eine Multiplikation und zwei Additionen durchgeführt werden.

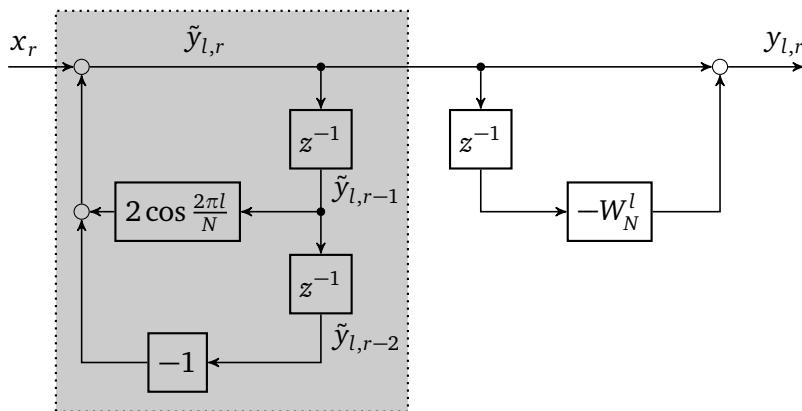


Abbildung B.3: Goertzel-Algorithmus

Stabilität

Die Pole $z_{\{1,2\}}$ des Filters (B.8) liegen bei

$$z_{\{1,2\}} = \cos\left(\frac{2\pi l}{N}\right) \pm j \cdot \sin\left(\frac{2\pi l}{N}\right)$$

und damit gilt $|z_{\{1,2\}}| = 1$. Das bedeutet, dass numerische Rundungsfehler nicht gedämpft werden. Daher muss geprüft werden, ob die anhand des Goertzel-Algorithmus berechneten Werte für die gegebene Länge N der DFT und die verwendete Zahlendarstellung verlässlich sind.

B.3.1 Goertzel-Algorithmus für nicht-ganzzahlige l

Es kann sinnvoll sein, Gl. (B.3) auch für nicht-ganzzahlige l auszuwerten. Dies ist der Fall, wenn nicht die DFT sondern die DTFT eines zeitbegrenzten Signals, welches für $k < 0$ und $k \geq N$ Null ist, bestimmt werden soll. Wenn die Frequenz ω dann als reell-wertiges l ausgedrückt wird, dann kann die DTFT in der Form (B.3) geschrieben werden.

Es kann nützlich sein den Goertzel-Algorithmus auch in diesem Fall anzuwenden. Dazu muss die Gl. (B.6) nochmals betrachtet werden. Da $W_N^{-l \cdot N}$ nicht 1 entspricht verbleibt dieser Faktor auf der rechten Seite,

$$y_{l,N} = \sum_{k=0}^N x_k W_N^{-l \cdot (N-k)} = W_N^{-l \cdot N} \cdot \sum_{k=0}^{N-1} x_k W_N^{lk} = W_N^{-l \cdot N} \cdot X_l .$$

Daher ergibt sich

$$X_l = W_N^{l,N} \cdot y_{l,N} = e^{-j2\pi l} \cdot y_{l,N}. \quad (\text{B.11})$$

D.h. $y_{l,N}$ wird wie für ganzzahlige l bestimmt. Dann muss lediglich noch Gl. (B.11) verwendet werden um die Phase zu korrigieren.

B.4 Nebenrechnungen Fourierreihe

Gerade und ungerade Signale

Es werden die Integrale $\int_0^T f(t) \sin(\omega t) dt$ und $\int_0^T f(t) \cos(\omega t) dt$ für gerade und ungerade $f(t)$ betrachtet.

Hier wird von der Freiheit der Wahl des Anfangspunktes der Integration Gebrauch gemacht, und so kann für das erste Integral auch

$$I_s = \int_{-\frac{T}{2}}^{\frac{T}{2}} f(t) \sin(\omega t) dt = \int_{-\frac{T}{2}}^0 f(t) \sin(\omega t) dt + \int_0^{\frac{T}{2}} f(t) \sin(\omega t) dt$$

geschrieben werden. Substituiert man im Integral des ersten Summanden $t' = -t$, so ergibt sich

$$I_s = - \int_{\frac{T}{2}}^0 f(-t') \cdot \sin(-\omega t') dt' + \int_0^{\frac{T}{2}} f(t) \sin(\omega t) dt$$

und nach dem Tauschen der Integrationsgrenzen des ersten Integrals erhält man durch Zusammenfassen der beiden Integrale

$$I_s = \int_0^{\frac{T}{2}} (f(t) - f(-t)) \cdot \sin(\omega t) dt.$$

Daraus folgt

$$\int_0^T f(t) \sin(\omega t) dt = \begin{cases} 2 \cdot \int_0^{\frac{T}{2}} f(t) \sin(\omega t) dt & \text{für ungerade Funktionen, } f(t) = -f(-t) \\ 0 & \text{für gerade Funktionen, } f(t) = f(-t). \end{cases}$$

Analog ergibt sich

$$\int_0^T f(t) \cos(\omega t) dt = \begin{cases} 0 & \text{für ungerade Funktionen, } f(t) = -f(-t) \\ 2 \cdot \int_0^{\frac{T}{2}} f(t) \cos(\omega t) dt & \text{für gerade Funktionen, } f(t) = f(-t). \end{cases}$$

Monofrequentes Signal

Es sollen die Fourierkoeffizienten des Signals

$$y(t) = \cos(\omega t) + 2 \cdot \sin(\omega t)$$

bestimmt werden. Aufgrund der Orthogonalität der Winkelfunktionen sind alle Koeffizienten \check{y}_l für $l \neq 1$ gleich null. Für $l = 1$ ergibt sich

$$\begin{aligned}\check{y}_1 &= \frac{1}{T} \cdot \int_0^T (\cos(\omega t) + 2 \cdot \sin(\omega t)) \cdot (\cos(l\omega t) - j \cdot \sin(l\omega t)) dt \\ &= \frac{1}{T} \cdot \int_0^T \cos^2(\omega t) dt + \frac{1}{T} \cdot \int_0^T \cos(\omega t) \sin(\omega t) dt \\ &\quad - j \cdot \frac{1}{T} \cdot \int_0^T \cos(\omega t) \sin(\omega t) dt - j \cdot \frac{2}{T} \cdot \int_0^T \sin^2(\omega t) dt \\ &= \frac{1}{T} \cdot \frac{T}{2} - j \cdot \frac{2}{T} \cdot \frac{T}{2} \\ &= \frac{1}{2} - j \cdot 1.\end{aligned}$$

Rechteckschwingung

Es sollen die Fourierkoeffizienten des Signals

$$y(t) = \begin{cases} h & 0 \leq t < T/2 \\ -h & T/2 \leq t < T \end{cases}, \text{ periodisch fortgesetzt, } y(k \cdot T + t) = y(t),$$

bestimmt werden. Es ergibt sich

$$\begin{aligned}\check{y}_l &= \frac{1}{T} \cdot \int_0^T y(t) \cdot e^{-jl\omega t} dt \\ &= \frac{1}{T} \cdot \int_0^T y(t) \cdot (\cos(l\omega t) - j \cdot \sin(l\omega t)) dt\end{aligned}$$

und weiter, da $y(t)$ ungerade ist,

$$\begin{aligned}&= -j \cdot \frac{1}{T} \cdot \int_0^T y(t) \cdot \sin(l\omega t) dt \\ &= -j \cdot \frac{2h}{T} \cdot \int_0^{\frac{T}{2}} \sin(l\omega t) dt\end{aligned}$$

was durch Lösen auf

$$\begin{aligned}&= -j \cdot \frac{2h}{T} \cdot \left[-\frac{1}{l\omega} \cdot \cos(l\omega t) \right]_0^{\frac{T}{2}} \\ &= -j \cdot \frac{2h}{T} \cdot \frac{Tl}{2\pi} \cdot \left[-\cos\left(\frac{2\pi}{T} \cdot lt\right) \right]_0^{\frac{T}{2}} \\ &= j \cdot \frac{-h}{l\pi} \cdot (1 - \cos(l\pi)) \\ &= \begin{cases} j \cdot \frac{-2h}{l\pi} & l \text{ gerade} \\ 0 & l \text{ ungerade} \end{cases}\end{aligned}$$

führt.

Orthogonalität der Winkelfunktionen

Es werden die Integrale $\int_0^T \sin(\omega l t) \sin(\omega k t) dt$, $\int_0^T \cos(\omega l t) \cos(\omega k t) dt$ und $\int_0^T \sin(\omega l t) \cos(\omega k t) dt$ betrachtet, wobei $\omega = \frac{2\pi}{T}$ ist und l und k Ganzzahlen sind.

Wird auf das erste Integral

$$I_{ss} = \int_0^T \sin(\omega l t) \sin(\omega k t) dt$$

die Regeln der partiellen Integration mit $u = \sin(\omega l t)$ und $v' = \sin(\omega k t)$ angewandt, so ergibt sich

$$I_{ss} = \left[-\frac{1}{\omega k} \sin(\omega l t) \cos(\omega k t) \right]_0^T - \int_0^T (-1) \cdot \frac{l}{k} \cdot \cos(\omega l t) \cos(\omega k t) dt$$

und weiter, da $\sin(\omega l T) = 0$ für alle l ist,

$$I_{ss} = \frac{l}{k} \cdot \int_0^T \cos(\omega l t) \cos(\omega k t) dt . \quad (\text{B.12})$$

Abermalige partielle Integration mit $u = \cos(\omega l t)$ und $v' = \cos(\omega k t)$ führt auf

$$I_{ss} = \frac{l}{k} \cdot \left[\frac{1}{\omega k} \sin(\omega k t) \cos(\omega l t) \right]_0^T - \frac{l}{k} \cdot \int_0^T (-1) \cdot \frac{l}{k} \cdot \sin(\omega l t) \sin(\omega k t) dt .$$

Auch hier ergibt die Auswertung des ersten Summanden immer Null, und nachdem alle konstanten Faktoren aus dem Integral gezogen sind bleibt genau das eigentlich gesuchte Integral über, und die Gleichung lautet

$$I_{ss} = \left(\frac{l}{k} \right)^2 \cdot I_{ss} .$$

Für $l = k$ kann dieser Gleichung keine Information entnommen werden, aber für $l \neq k$ (und $l > 0$ und $k > 0$) folgt daraus, dass $I_{ss} = 0$ sein muss,

$$\int_0^T \sin(\omega l t) \sin(\omega k t) dt = 0 \quad \forall l \neq k .$$

Für $l = k$ kann wieder bei Gl. (B.12) angesetzt werden, die dann

$$\int_0^T \sin^2(\omega l t) dt = \int_0^T \cos^2(\omega l t) dt$$

lautet. Mit $\cos^2(\omega l t) = 1 - \sin^2(\omega l t)$ folgt daraus

$$\int_0^T \sin^2(\omega l t) dt = \frac{1}{2} \cdot \int_0^T dt = \frac{T}{2} = \int_0^T \cos^2(\omega l t) dt .$$

Für das Integral $I_{cc} = \int_0^T \cos(\omega l t) \cos(\omega k t) dt$ führt ein entsprechender Weg zum gleichen Ergebnis, bzw. kann auch hier direkt Gl. (B.12) verwendet werden.

Zum Auswerten des Integrals

$$I_{sc} = \int_0^T \sin(\omega l t) \cos(\omega k t) dt$$

wird wieder mit partieller Integration, $u = \sin(\omega l t)$ und $v' = \cos(\omega k t)$, begonnen,

$$I_{sc} = \left[\frac{1}{\omega k} \sin(\omega l t) \sin(\omega k t) \right]_0^T - \int_0^T \frac{l}{k} \cdot \cos(\omega l t) \sin(\omega k t) dt = -\frac{l}{k} \cdot \int_0^T \cos(\omega l t) \sin(\omega k t) dt .$$

Im Fall $l = k$ steht hier $I_{sc} = -I_{sc}$ was nur für $I_{sc} = 0$ erfüllt sein kann. Für den Fall $l \neq k$ führt eine letzte partielle Integration mit $u = \cos(\omega l t)$ und $v' = \sin(\omega k t)$ letztlich wieder auf

$$I_{sc} = \left(\frac{l}{k} \right)^2 \cdot I_{sc}$$

und damit ebenfalls auf $I_{sc} = 0$. Im Fall des „gemischten“ Integrals ist das Ergebnis also immer null, selbst wenn $l = k$ ist.

Zusammenfassend ergibt sich (die noch fehlenden Fälle $l = 0$ und/oder $k = 0$ sollten leicht nachvollziehbar sein)

$$\int_0^T \sin(\omega l t) \sin(\omega k t) dt = \begin{cases} 0 & l = k = 0 \\ \frac{T}{2} & l = k \neq 0 \\ 0 & l \neq k , \end{cases}$$

$$\int_0^T \cos(\omega l t) \cos(\omega k t) dt = \begin{cases} T & l = k = 0 \\ \frac{T}{2} & l = k \neq 0 \\ 0 & l \neq k , \end{cases}$$

und

$$\int_0^T \sin(\omega l t) \cos(\omega k t) dt = 0 \quad \forall l, k .$$

C Signale

C.1 Zur Berechnung der Varianz bei Rauschen

[Eykhoff 1974, 532f]

Zunächst wird dazu das Integral I einer Funktion von zwei Variablen, t und t' über dem quadratischen Gebiet $[0, T] \times [0, T]$ betrachtet,

$$I = \int_0^T \int_0^T f(t, t') dt' dt .$$

Das Integrationsgebiet kann wie in Abbildung C.1a in zwei Teilgebiete zerlegt werden,

$$\int_0^T \int_0^T f(t, t') dt' dt = \underbrace{\int_0^T \int_0^t f(t, t') dt' dt}_{I_1} + \underbrace{\int_0^T \int_t^T f(t, t') dt' dt}_{I_2} .$$

Dabei bezeichnet I_1 und I_2 nicht die Teilgebiete (oder deren Flächen), sondern den Wert des Integrals über den jeweiligen Teilgebieten.

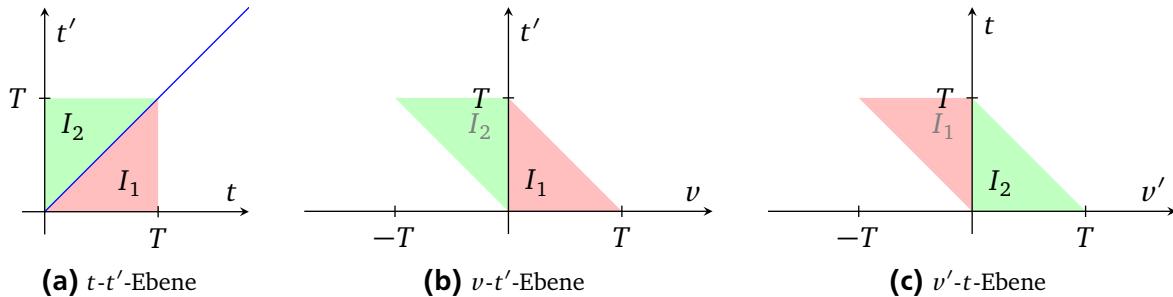


Abbildung C.1: Zur Integration über das Gebiet $[0, T] \times [0, T]$ in der t - t' -Ebene

Zur Auswertung des ersten Integrals I_1 wird die Substitution $t = t' + v$ verwendet. Die Gebiete in der v - t' -Ebene sind in Abbildung C.1a dargestellt. Damit kann

$$\begin{aligned} I_1 &= \int_0^T \int_0^t f(t, t') dt' dt = \int_0^T \int_{t'}^T f(t, t') dt' dt \\ &= \int_0^T \int_0^{T-t'} f(t' + v, t') dv dt' = \int_0^T \int_0^{T-v} f(t' + v, t') dv dt' \end{aligned}$$

umgeformt werden. Dabei sind im ersten Schritt die Integrationsvariablen t' und t vertauscht (womit auch die Parametrierungen der Integralgrenzen anzupassen sind, was am einfachsten anhand der Skizze in Abbildung C.1a nachvollzogen werden kann) und im zweiten Schritt ist die schon angekündigte Substitution $t = t' + v$ durchgeführt. Zuletzt ist wieder die Reihenfolge der Integrationsvariablen, nun t' und v , vertauscht.

Für das zweite Teilintegral I_2 ergibt sich mit der Substitution $v' = t' - t$

$$I_2 = \int_0^T \int_t^T f(t, t') dt' dt = \int_0^T \int_0^{T-t} f(t, t + v') dv' dt = \int_0^T \int_0^{T-v'} f(t, t + v') dt dv'.$$

Damit kann das gesamte Integral allgemein auch als

$$I = \int_0^T \int_t^T f(t, t') dt' dt = \int_0^T \int_0^{T-v} f(t + v, t) dt dv + \int_0^T \int_0^{T-v} f(t, t + v) dt dv$$

geschrieben werden.

Angewendet auf das vorliegende Problem ergibt sich

$$\text{Var}\{\text{Re}(N(\omega))\}_r = 2 \cdot \int_0^T \int_0^{T-v} R_{nn}(v) \cdot \cos(\omega \cdot (v + t)) \cdot \cos(\omega t) dt dv,$$

wobei von $R_{nn}(v) = R_{nn}(-v)$ Gebrauch gemacht wurde.

Mit dem allgemeinen Zusammenhang $\cos(\alpha) \cdot \cos(\beta) = \frac{1}{2} \cdot (\cos(\alpha - \beta) + \cos(\alpha + \beta))$ folgt weiter

$$\begin{aligned} \text{Var}\{\text{Re}(N(\omega))\}_r &= \int_0^{NT} R_{nn}(v) \cdot \int_0^{T-v} (\cos(\omega v) + \cos(\omega \cdot (2t + v))) dt dv \\ &= \int_0^{NT} R_{nn}(v) \cdot \left[\cos(\omega v) \cdot t + \frac{1}{\omega} \cdot \sin(\omega \cdot (2t + v)) \right]_{t=0}^{T-v} dv \\ &= \int_0^{NT} R_{nn}(v) \cdot \left(\cos(\omega v) \cdot (T - v) - \frac{\sin(\omega \cdot (2T - v)) - \sin(\omega v)}{\omega} \right) dv. \end{aligned}$$

Zuletzt wird noch T ausgeklammert, so dass

$$\begin{aligned} \text{Var}\{\text{Re}(N(\omega))\}_r &= T \cdot \int_0^T R_{nn}(v) \cdot \left(\left(1 - \frac{v}{T}\right) \cdot \cos(\omega v) + \frac{1}{\omega T} \cdot (\sin(\omega \cdot (2T - v)) - \sin(\omega v)) \right) dv \end{aligned}$$

entsteht.

Für den Imaginärteil ergibt sich aus

$$\text{Var}\{\text{Im}(N(\omega))\}_r = \int_0^T \int_0^T R_{nn}(t - t') \cdot \sin(\omega t) \cdot \sin(\omega t') \cdot dt' dt$$

über den gleichen Rechenweg

$$\text{Var}\{\text{Im}(N(\omega))\}_r = T \cdot \int_0^T R_{nn}(v) \cdot \left(\left(1 - \frac{v}{T}\right) \cdot \cos(\omega v) - \frac{1}{\omega T} \cdot (\sin(\omega \cdot (2T - v)) - \sin(\omega v)) \right) dv,$$

wobei der entsprechende Zusammenhang der Winkelfunktionen hier $\sin(\alpha) \sin(\beta) = \frac{1}{2} \cdot (\cos(\alpha - \beta) - \cos(\alpha + \beta))$ lautet. (Es ändert sich lediglich das Vorzeichen des zweiten Summanden des Integranden.)

Für die Kovarianz ergibt sich aus

$$\begin{aligned}\text{Cov}(\{\text{Re}(N(\omega))\}_r, \{\text{Im}(N(\omega))\}_r) &= E\{\text{Re}(N(\omega)) \cdot \text{Im}(N(\omega))\}_r \\ &= \int_0^T \int_0^T R_{nn}(t-t') \cdot \sin(\omega t) \cdot \cos(\omega t') \cdot dt' dt\end{aligned}$$

über den gleichen Rechenweg

$$E\{\text{Re}(N(\omega)) \cdot \text{Im}(N(\omega))\}_r = T \cdot \int_0^T R_{nn}(\nu) \cdot \frac{1}{2\omega} \cdot (\cos(\omega \cdot (2T - \nu)) - \cos(\omega \nu)) \cdot d\nu,$$

wobei der entsprechende Zusammenhang der Winkelfunktionen hier $\sin(\alpha) \cdot \cos(\beta) = \frac{1}{2} \cdot (\sin(\alpha - \beta) + \sin(\alpha + \beta))$ lautet.

Für den Fall, dass die Messdauer T ein ganzes Vielfaches von $\frac{2\pi}{\omega}$ ist, ergeben sich die etwas einfacheren Zusammenhänge

$$\begin{aligned}\text{Var}\{\text{Re}(N(\omega))\}_r &= T \cdot \int_0^T R_{nn}(\nu) \cdot \left(\left(1 - \frac{\nu}{T}\right) \cdot \cos(\omega \nu) - \frac{1}{2\omega T} \cdot \sin(\omega \nu) \right) d\nu \\ \text{Var}\{\text{Im}(N(\omega))\}_r &= T \cdot \int_0^T R_{nn}(\nu) \cdot \left(\left(1 - \frac{\nu}{T}\right) \cdot \cos(\omega \nu) + \frac{1}{2\omega T} \cdot \sin(\omega \nu) \right) d\nu\end{aligned}$$

$$\text{Cov}(\{\text{Re}(N(\omega))\}_r, \{\text{Im}(N(\omega))\}_r) = 0.$$

D Methode der kleinsten Quadrate

D.1 Rekursives LS-Verfahren

D.1.1 Gewichtetes LS-Verfahren

Mit der Anwendung der Sherman-Morrison-Formel erhält man

$$\begin{aligned}
 \hat{\theta}_{k+1} &= (\mathbf{L}_k + \boldsymbol{\Psi}_{k+1} w_{k+1} \boldsymbol{\Psi}_{k+1}^T)^{-1} \cdot (\mathbf{r}_k + \boldsymbol{\Psi}_{k+1} w_{k+1} \tilde{\eta}_{k+1}) \\
 &= \left(\mathbf{L}_k^{-1} - \frac{\mathbf{L}_k^{-1} \boldsymbol{\Psi}_{k+1} \boldsymbol{\Psi}_{k+1}^T \mathbf{L}_k^{-1}}{\frac{1}{w_{k+1}} + \boldsymbol{\Psi}_{k+1}^T \mathbf{L}_k^{-1} \boldsymbol{\Psi}_{k+1}} \right) \cdot (\mathbf{r}_k + \boldsymbol{\Psi}_{k+1} w_{k+1} \tilde{\eta}_{k+1}) \\
 &= \left(\mathbf{L}_k^{-1} - \frac{\mathbf{L}_k^{-1} \boldsymbol{\Psi}_{k+1} \boldsymbol{\Psi}_{k+1}^T \mathbf{L}_k^{-1}}{\frac{1}{w_{k+1}} + \boldsymbol{\Psi}_{k+1}^T \mathbf{L}_k^{-1} \boldsymbol{\Psi}_{k+1}} \right) \cdot \mathbf{r}_k + \left(\mathbf{L}_k^{-1} - \frac{\mathbf{L}_k^{-1} \boldsymbol{\Psi}_{k+1} \boldsymbol{\Psi}_{k+1}^T \mathbf{L}_k^{-1}}{\frac{1}{w_{k+1}} + \boldsymbol{\Psi}_{k+1}^T \mathbf{L}_k^{-1} \boldsymbol{\Psi}_{k+1}} \right) \cdot \boldsymbol{\Psi}_{k+1} w_{k+1} \tilde{\eta}_{k+1} \\
 &= \mathbf{L}_k^{-1} \mathbf{r}_k - \frac{\mathbf{L}_k^{-1} \boldsymbol{\Psi}_{k+1}}{\frac{1}{w_{k+1}} + \boldsymbol{\Psi}_{k+1}^T \mathbf{L}_k^{-1} \boldsymbol{\Psi}_{k+1}} \boldsymbol{\Psi}_{k+1}^T \mathbf{L}_k^{-1} \mathbf{r}_k \\
 &\quad + \mathbf{L}_k^{-1} \boldsymbol{\Psi}_{k+1} \cdot \left(1 - \frac{\boldsymbol{\Psi}_{k+1}^T \mathbf{L}_k^{-1} \boldsymbol{\Psi}_{k+1}}{\frac{1}{w_{k+1}} + \boldsymbol{\Psi}_{k+1}^T \mathbf{L}_k^{-1} \boldsymbol{\Psi}_{k+1}} \right) \cdot w_{k+1} \tilde{\eta}_{k+1}.
 \end{aligned}$$

Mit der Abkürzung

$$\gamma_k = \frac{\mathbf{L}_k^{-1} \boldsymbol{\Psi}_{k+1}}{\boldsymbol{\Psi}_{k+1}^T \mathbf{L}_k^{-1} \boldsymbol{\Psi}_{k+1} + \frac{1}{w_{k+1}}} = \frac{\mathbf{P}'_k \boldsymbol{\Psi}_{k+1}}{\boldsymbol{\Psi}_{k+1}^T \mathbf{P}'_k \boldsymbol{\Psi}_{k+1} + \frac{1}{w_{k+1}}}$$

erhält man

$$\begin{aligned}
 \hat{\theta}_{k+1} &= \hat{\theta}_k - \gamma_k \boldsymbol{\Psi}_{k+1}^T \hat{\theta}_k + \mathbf{L}_k^{-1} \boldsymbol{\Psi}_{k+1} \cdot \left(\frac{\frac{1}{w_{k+1}}}{\frac{1}{w_{k+1}} + \boldsymbol{\Psi}_{k+1}^T \mathbf{L}_k^{-1} \boldsymbol{\Psi}_{k+1}} \right) \cdot w_{k+1} \tilde{\eta}_{k+1} \\
 &= \hat{\theta}_k - \gamma_k \boldsymbol{\Psi}_{k+1}^T \hat{\theta}_k + \gamma_k \tilde{\eta}_{k+1} \\
 &= \hat{\theta}_k + \gamma_k \cdot (\tilde{\eta}_{k+1} - \boldsymbol{\Psi}_{k+1}^T \hat{\theta}_k).
 \end{aligned}$$

\mathbf{P}'_{k+1} ergibt sich analog zum ungewichteten Fall. Damit erhält man

$$\begin{aligned}
 \gamma_k &= \frac{\mathbf{P}'_k \boldsymbol{\Psi}_{k+1}}{\boldsymbol{\Psi}_{k+1}^T \mathbf{P}'_k \boldsymbol{\Psi}_{k+1} + \frac{1}{w_{k+1}}} \\
 \hat{\theta}_{k+1} &= \hat{\theta}_k + \gamma_k \cdot (\tilde{\eta}_{k+1} - \boldsymbol{\Psi}_{k+1}^T \hat{\theta}_k) \\
 \mathbf{P}'_{k+1} &= (\mathbf{I} - \gamma_k \boldsymbol{\Psi}_{k+1}^T) \cdot \mathbf{P}'_k.
 \end{aligned}$$

D.1.2 LS-Verfahren mit exponentiellem Vergessen

Mit der Anwendung der Sherman-Morrison-Formel erhält man

$$\begin{aligned}\hat{\theta}_{k+1} &= (\lambda_{k+1} \mathbf{L}_k + \boldsymbol{\Psi}_{k+1} \boldsymbol{\Psi}_{k+1}^T)^{-1} \cdot (\lambda_{k+1} \mathbf{r}_k + \boldsymbol{\Psi}_{k+1} \tilde{\eta}_{k+1}) \\ &= \left(\mathbf{L}_k + \frac{1}{\lambda_{k+1}} \cdot \boldsymbol{\Psi}_{k+1} \boldsymbol{\Psi}_{k+1}^T \right)^{-1} \cdot \left(\mathbf{r}_k + \frac{1}{\lambda_{k+1}} \cdot \boldsymbol{\Psi}_{k+1} \tilde{\eta}_{k+1} \right).\end{aligned}$$

Für γ_k und $\hat{\theta}_{k+1}$ weiter wie bei LS mit Gewichtungsfaktor $w_{k+1} = \frac{1}{\lambda_{k+1}}$,

$$\begin{aligned}\gamma_k &= \frac{\mathbf{P}'_k \boldsymbol{\Psi}_{k+1}}{\boldsymbol{\Psi}_{k+1}^T \mathbf{P}'_k \boldsymbol{\Psi}_{k+1} + \lambda_{k+1}} \\ \hat{\theta}_{k+1} &= \hat{\theta}_k + \gamma_k \cdot (\tilde{\eta}_{k+1} - \boldsymbol{\Psi}_{k+1}^T \hat{\theta}_k).\end{aligned}$$

Was sich ändert ist die Bestimmung von $\mathbf{P}'_{k+1} = \mathbf{L}_{k+1}^{-1}$. Aus

$$\mathbf{L}_{k+1} = \lambda_{k+1} \mathbf{L}_k + \boldsymbol{\Psi}_{k+1} \boldsymbol{\Psi}_{k+1}^T$$

folgt

$$\begin{aligned}\mathbf{L}_{k+1}^{-1} &= (\lambda_{k+1} \mathbf{L}_k + \boldsymbol{\Psi}_{k+1} \boldsymbol{\Psi}_{k+1}^T)^{-1} \\ &= \frac{1}{\lambda_{k+1}} \cdot \mathbf{L}_k^{-1} - \frac{\frac{1}{\lambda_{k+1}} \cdot \mathbf{L}_k^{-1} \cdot \boldsymbol{\Psi}_{k+1} \boldsymbol{\Psi}_{k+1}^T \cdot \frac{1}{\lambda_{k+1}} \cdot \mathbf{L}_k^{-1}}{1 + \boldsymbol{\Psi}_{k+1}^T \cdot \frac{1}{\lambda_{k+1}} \cdot \mathbf{L}_k^{-1} \cdot \boldsymbol{\Psi}_{k+1}} \\ &= \frac{1}{\lambda_{k+1}} \cdot \left(\mathbf{L}_k^{-1} - \frac{\mathbf{L}_k^{-1} \cdot \boldsymbol{\Psi}_{k+1} \boldsymbol{\Psi}_{k+1}^T \cdot \mathbf{L}_k^{-1}}{1 + \boldsymbol{\Psi}_{k+1}^T \cdot \frac{1}{\lambda_{k+1}} \cdot \mathbf{L}_k^{-1} \cdot \boldsymbol{\Psi}_{k+1}} \right) \\ &= \frac{1}{\lambda_{k+1}} \cdot (\mathbf{L}_k^{-1} - \gamma_k \boldsymbol{\Psi}_{k+1}^T \mathbf{L}_k^{-1}) \\ &= (\mathbf{I} - \gamma_k \boldsymbol{\Psi}_{k+1}^T) \cdot \mathbf{L}_k^{-1} \cdot \frac{1}{\lambda_{k+1}} \\ &= (\mathbf{I} - \gamma_k \boldsymbol{\Psi}_{k+1}^T) \cdot \mathbf{P}'_k \cdot \frac{1}{\lambda_{k+1}}.\end{aligned}$$

Zusammengefasst lautet der Algorithmus damit

$$\begin{aligned}\gamma_k &= \frac{\mathbf{P}'_k \boldsymbol{\Psi}_{k+1}}{\boldsymbol{\Psi}_{k+1}^T \mathbf{P}'_k \boldsymbol{\Psi}_{k+1} + \lambda_{k+1}} \\ \hat{\theta}_{k+1} &= \hat{\theta}_k + \gamma_k \cdot (\tilde{\eta}_{k+1} - \boldsymbol{\Psi}_{k+1}^T \hat{\theta}_k) \\ \mathbf{P}'_{k+1} &= (\mathbf{I} - \gamma_k \boldsymbol{\Psi}_{k+1}^T) \cdot \mathbf{P}'_k \cdot \frac{1}{\lambda_{k+1}}.\end{aligned}$$

D.2 QR-Zerlegung durch Householder-Transformationen

(Aus „Ergänzungen zur Vorlesung Identifikation dynamischer Systeme – Numerische Verfahren zur Identifikation linearer Systeme“ von Michael Vogt.)

Bei einer Householder-Transformation wird ein Vektor \mathbf{a} von links mit einer orthogonalen Matrix \mathbf{Q}_1 multipliziert, die so konstruiert ist, dass alle Elemente des transformierten Vektors bis auf das erste Null werden. Die Matrix \mathbf{Q}_1 wird aus den beiden Forderungen

$$(a) \quad \mathbf{Q}_1^T \mathbf{Q}_1 = \mathbf{I} \quad (\text{Orthogonalität})$$

$$(b) \quad \mathbf{Q}_1 \cdot \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{bmatrix} = \begin{bmatrix} \|\mathbf{a}\|_2 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

bestimmt. Die Orthogonalität von \mathbf{Q}_1 sorgt dafür, dass die (euklidische) Norm von \mathbf{a} erhalten bleibt. Dies ist der Grund für die guten numerischen Eigenschaften der Householder-Transformation. \mathbf{Q}_1 berechnet sich als

$$\mathbf{Q}_1 = \mathbf{I} - \beta_1 \mathbf{u}_1 \mathbf{u}_1^T \quad (\text{D.1})$$

wobei

$$\mathbf{u}_1 = [\sigma(a_1) \cdot (|a_1| + \|\mathbf{a}\|_2) \quad a_2 \quad \dots \quad a_N]^T \quad \text{mit } \sigma(x) = \begin{cases} 1 & \text{für } x \geq 0 \\ -1 & \text{für } x < 0 \end{cases} \quad (\text{D.2})$$

und

$$\beta_1 = \frac{1}{\|\mathbf{a}\|_2 \cdot (|a_1| + \|\mathbf{a}\|_2)} \quad (\text{D.3})$$

ist.

Zum Erzeugen der QR-Zerlegung werden nun p Transformationen auf Ψ angewendet – und zwar die i -te Transformation auf die Zeilen i, \dots, k und die Spalten i, \dots, p . Schematisch ergibt sich folgende Kette von Transformationen, wobei „*“ einen nicht näher spezifizierten Wert und „•“ einen Wert darstellt, der sich im Weiteren nicht mehr ändert:

$$\begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \end{bmatrix} \xrightarrow{\mathbf{Q}_1} \begin{bmatrix} \bullet & \bullet & \bullet \\ 0 & * & * \\ 0 & * & * \\ 0 & * & * \\ 0 & * & * \end{bmatrix} \xrightarrow{\mathbf{Q}_2} \begin{bmatrix} \bullet & \bullet & \bullet \\ 0 & \bullet & \bullet \\ 0 & 0 & * \\ 0 & 0 & * \\ 0 & 0 & * \end{bmatrix} \xrightarrow{\mathbf{Q}_3} \begin{bmatrix} \bullet & \bullet & \bullet \\ 0 & \bullet & \bullet \\ 0 & 0 & \bullet \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} .$$

Die p Transformationen, die in gleicher Weise wie \mathbf{Q}_1 konstruiert werden, lassen sich formal als $k \times k$ -Matrizen $\mathbf{Q}_1, \dots, \mathbf{Q}_p$ schreiben¹ und zur Matrix \mathbf{Q} zusammenfassen:

$$\underbrace{\prod_{i=1}^p \mathbf{Q}_i}_{\mathbf{Q}} \cdot \Psi = \begin{bmatrix} \mathbf{R} \\ \mathbf{0} \end{bmatrix} .$$

Mit jeder Transformation \mathbf{Q}_i wird also jeweils die i -te Spalte und die i -te Zeile von $\begin{bmatrix} \mathbf{R} \\ \mathbf{0} \end{bmatrix}$ fertig gestellt.

¹ Mit der oben angegebenen Berechnungsvorschrift (D.1) - (D.3) erhält man in jedem Schritt eine $(k-i+1 \times k-i+1)$ -Matrix. Diese ist um eine Einheitsmatrix passender Dimension zu erweitern, also z. B. $\mathbf{Q}_3 = \begin{bmatrix} \mathbf{I}_2 & \\ & \mathbf{Q}'_3 \end{bmatrix}$, wobei \mathbf{Q}'_3 die nach der oben angegebenen Vorschrift berechnete (3×3) -Matrix für den dritten Transformationsschritt ist.

D.3 DSFI-Verfahren

(Aus „Ergänzungen zur Vorlesung Identifikation dynamischer Systeme – Numerische Verfahren zur Identifikation linearer Systeme“ von Michael Vogt, teilweise ergänzt.)

Auch auf der Basis der QR-Zerlegung lässt sich ein rekursiver Algorithmus konstruieren, der jedoch die sehr guten numerischen Eigenschaften der Orthogonalisierungsverfahren erhält. Dieser wird in der Regelungstechnik als *DSFI-Verfahren* (Discrete-Square-Root Filtering in Information Form) bezeichnet und berechnet die Größen \mathbf{R}_k , \mathbf{c}_k und ggf. \mathbf{Q}_k aus den „alten“ Größen \mathbf{R}_{k-1} , \mathbf{c}_{k-1} bzw. \mathbf{Q}_{k-1} .

Das DSFI-Verfahren wird – wie auch das in Abschnitt 8.6.2 vorgestellte rekursive Least-Squares-Verfahren – in der Regel mit exponentieller Gewichtung benutzt, um zeitvariante Parameter schätzen zu können. Wird das Verfahren *ohne* Gewichtung auf den kompletten Datensatz angewendet, dann kann es an Stelle der Householder-Transformation zur Berechnung der QR-Zerlegung benutzt werden. Weil der Datensatz in diesem Fall zeilenweise abgearbeitet wird, müssen (wegen $k \gg p$) wesentlich weniger Daten im Speicher gehalten werden, was auch die Bearbeitung extrem großer Datensätze erlaubt.

Der Einfachheit halber wird zunächst der Algorithmus ohne Gewichtung hergeleitet. Man nimmt dazu wieder an, die QR-Zerlegung von Ψ_{k-1} sei aus dem vorangegangenen Schritt bekannt und sucht dann die QR-Zerlegung der Matrix

$$\Psi_k = \begin{bmatrix} \Psi_{k-1} \\ \psi_k^T \end{bmatrix},$$

die sich durch Hinzufügen des neuen Eingangsvektors ψ_k^T ergibt. Es ist also ein Weg gesucht, aus dem Zusammenhang

$$\begin{bmatrix} \mathbf{Q}_{k-1} & \mathbf{0} \\ \mathbf{0}^T & 1 \end{bmatrix} \begin{bmatrix} \Psi_{k-1} \\ \psi_k^T \end{bmatrix} = \begin{bmatrix} \mathbf{R}_{k-1} \\ \mathbf{0} \\ \psi_k^T \end{bmatrix} \quad \begin{array}{l} \} \quad p \text{ Zeilen} \\ \} \quad k-p-1 \text{ Zeilen} \\ \} \quad 1 \text{ Zeile} \end{array} \quad (\text{D.4})$$

wieder eine QR-Zerlegung für Ψ_k zu erhalten.

Benötigt wird jetzt eine Folge von orthogonalen Transformationen, die ψ_k^T auf der rechten Seite der Gleichung in eine Nullzeile überführt, so dass die gesuchte QR-Zerlegung von Ψ_k entsteht. Diese Aufgabe lässt sich mit *Givens-Transformationen* lösen, die im zweidimensionalen Fall einer ebenen Rotation entsprechen. Die Givens-Rotation ist dazu geeignet, Nullen in einer Matrix zu erzeugen. Im Gegensatz zur Householder-Transformation wird aber mit jeder Givens-Rotation nur ein einziges Element zu Null gemacht.

Givens-Rotation

Es sei zunächst \mathbf{M} eine beliebige $2 \times \mu$ -Matrix. Diese soll mit einer orthogonalen 2×2 -Givens-Matrix

$$\mathbf{G} = \begin{bmatrix} \gamma & \sigma \\ -\sigma & \gamma \end{bmatrix}$$

multipliziert werden, so dass das Element m'_{21} der transformierten Matrix $\mathbf{M}' = \mathbf{GM}$ verschwindet, also

$$\begin{bmatrix} \gamma & \sigma \\ -\sigma & \gamma \end{bmatrix} \begin{bmatrix} m_{11} & m_{12} & \cdots \\ m_{21} & m_{22} & \cdots \end{bmatrix} = \begin{bmatrix} \gamma m_{11} + \sigma m_{21} & \gamma m_{12} + \sigma m_{22} & \cdots \\ -\sigma m_{11} + \gamma m_{21} & -\sigma m_{12} + \gamma m_{22} & \cdots \end{bmatrix} = \begin{bmatrix} m'_{11} & m'_{12} & \cdots \\ 0 & m'_{22} & \cdots \end{bmatrix}.$$

Abbildung D.1 zeigt dies für einen zweidimensionalen Vektor $\mathbf{m} = [m_{11} \quad m_{21}]^T$.

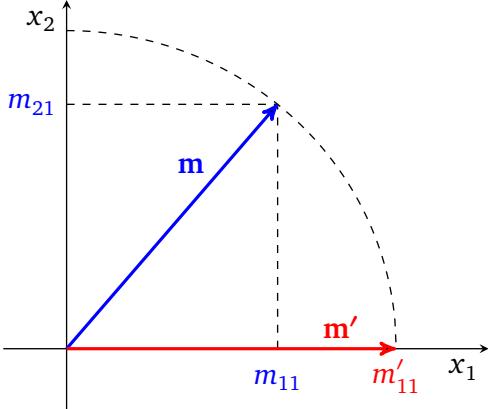


Abbildung D.1: Givens-Rotation eines Vektors

Aus den beiden Forderungen

$$\det(\mathbf{G}) = \gamma^2 + \sigma^2 = 1 \quad (\text{Normierung})$$

$$m'_{21} = -\sigma m_{11} + \gamma m_{21} = 0 \quad (\text{Eliminierung von } m'_{21})$$

erhält man die Rotationsparameter

$$\gamma = \frac{m_{11}}{\sqrt{m_{11}^2 + m_{21}^2}} \quad \text{und} \quad \sigma = \frac{m_{21}}{\sqrt{m_{11}^2 + m_{21}^2}},$$

und die Elemente der transformierten Matrix ergeben sich zu

$$m'_{11} = \sqrt{m_{11}^2 + m_{21}^2}, \quad m'_{21} = 0$$

sowie

$$\left. \begin{array}{l} m'_{1j} = \gamma m_{1j} + \sigma m_{2j} \\ m'_{2j} = -\sigma m_{1j} + \gamma m_{2j} \end{array} \right\} j = 2, \dots, \mu.$$

D. h. während m'_{21} zu Null wird, ändern alle übrigen Elemente ihren Wert.

Anwenden der Givens-Rotation zur QR-Zerlegung

Um die QR-Zerlegung von Ψ_k zu bestimmen, werden nun nacheinander Givens-Rotationen auf die Gleichung (D.4) angewendet. Zuerst wird der neue Eingangsvektor Ψ_k^T mit der ersten Zeile von \mathbf{R}_{k-1} kombiniert und so das erste Element von Ψ_k^T eliminiert, dann werden die umgerechneten Eingangswerte mit der zweiten Zeile von \mathbf{R}_{k-1} kombiniert, um das zweite Element zu entfernen usw.:

$$\begin{bmatrix} * & * & * \\ 0 & * & * \\ 0 & 0 & * \\ 0 & 0 & 0 \\ * & * & * \end{bmatrix} \xrightarrow{\Omega_1} \begin{bmatrix} * & * & * \\ 0 & * & * \\ 0 & 0 & * \\ 0 & 0 & 0 \\ 0 & * & * \end{bmatrix} \xrightarrow{\Omega_2} \begin{bmatrix} * & * & * \\ 0 & * & * \\ 0 & 0 & * \\ 0 & 0 & 0 \\ 0 & 0 & * \end{bmatrix} \xrightarrow{\Omega_3} \begin{bmatrix} * & * & * \\ 0 & * & * \\ 0 & 0 & * \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Alle Givens-Rotationen können formal als orthogonale $k \times k$ -Matrizen $\Omega_1, \dots, \Omega_p$ aufgefasst werden. Die jeweils nicht involvierten Zeilen bleiben unverändert. So lauten die Matrizen hier

$$\Omega_1 = \begin{bmatrix} \gamma & & & \sigma \\ & 1 & & \\ & & 1 & \\ -\sigma & & & \gamma \end{bmatrix}, \quad \Omega_2 = \begin{bmatrix} 1 & & & \sigma \\ & \gamma & & \\ & & 1 & \\ & -\sigma & & \gamma \end{bmatrix} \quad \text{und} \quad \Omega_3 = \begin{bmatrix} 1 & & & \sigma \\ & 1 & & \\ & & \gamma & \\ -\sigma & & & \gamma \end{bmatrix},$$

wobei die Nullelemente weggelassen sind und γ sowie σ für jede Matrix andere Werte besitzen.

Diese Matrizen können mit \mathbf{Q}_{k-1} zu der neuen Matrix

$$\mathbf{Q}_k = \prod_{i=1}^p \Omega_{i,k} \cdot \begin{bmatrix} \mathbf{Q}_{k-1} & \mathbf{0} \\ \mathbf{0}^T & 1 \end{bmatrix}$$

zusammengefasst werden, so dass schließlich

$$\mathbf{Q}_k \Psi_k = \begin{bmatrix} \mathbf{R}_k \\ \mathbf{0} \end{bmatrix}$$

gilt.

Gewichtete Methode der kleinsten Quadrate

In gleicher Weise kann das mit $\mathbf{W} = \mathbf{D}^T \mathbf{D}$ gewichtete Problem mit den Größen $\Psi_w = \mathbf{D}\Psi$ und $\tilde{\eta}_w = \mathbf{D}\tilde{\eta}$ gelöst werden, indem die Gütfunktion

$$\begin{aligned} V_w &= \|\mathbf{D}\hat{\mathbf{e}}\|_2^2 = \|\mathbf{Q}_w \mathbf{D}\hat{\mathbf{e}}\|_2^2 \\ &= \|\mathbf{Q}_w \Psi_w \hat{\theta} - \mathbf{Q}_w \tilde{\eta}_w\|_2^2 = \dots = \|\mathbf{R}_w \hat{\theta} - \mathbf{c}_w\|_2^2 + \|\tilde{\mathbf{c}}_w\|_2^2 \end{aligned}$$

minimiert wird. Bei exponentiellem Vergessen mit konstantem λ lautet \mathbf{D} in rekursiver Form (vgl. Gl. (8.10) auf Seite 240)

$$\mathbf{D}_k = \begin{bmatrix} \sqrt{\lambda} \mathbf{D}_{k-1} & \\ & 1 \end{bmatrix}.$$

Die rekursive Berechnung der Größen \mathbf{R}_w und \mathbf{c}_w der QR-Zerlegung

$$\mathbf{Q}_w \Psi_w = \begin{bmatrix} \mathbf{R}_w \\ \mathbf{0} \end{bmatrix}$$

erfolgt ähnlich wie beim ungewichteten Algorithmus. Man geht dazu von der QR-Zerlegung

$$\mathbf{Q}_{w,k-1} \mathbf{D}_{k-1} \Psi_{k-1} = \begin{bmatrix} \mathbf{R}_{w,k-1} \\ \mathbf{0} \end{bmatrix}$$

zum vorherigen Zeitschritt $k-1$ aus und multipliziert diese mit $\sqrt{\lambda}$ und ergänzt eine Zeile mit dem neuen Eingangsvektor Ψ_k^T , so dass man die Gleichung

$$\left[\begin{array}{cc} \mathbf{Q}_{w,k-1} & \mathbf{0} \\ \mathbf{0}^T & 1 \end{array} \right] \underbrace{\left[\begin{array}{cc} \sqrt{\lambda} \mathbf{D}_{k-1} & \mathbf{0} \\ \mathbf{0}^T & 1 \end{array} \right]}_{\mathbf{D}_k} \underbrace{\left[\begin{array}{c} \Psi_{k-1} \\ \Psi_k^T \end{array} \right]}_{\Psi_k} = \left[\begin{array}{c} \sqrt{\lambda} \mathbf{R}_{w,k-1} \\ \mathbf{0} \\ \Psi_k^T \end{array} \right]$$

erhält. Wie beim ungewichteten Verfahren wird nun wieder durch Anwendung von p Givens-Rotationen die QR-Zerlegung

$$\mathbf{Q}_{w,k} \underbrace{\mathbf{D}_k \Psi_k}_{\Psi_{w,k}} = \begin{bmatrix} \mathbf{R}_{w,k} \\ \mathbf{0} \end{bmatrix}$$

zum Zeitpunkt k erzeugt. Beim Algorithmus mit exponentieller Gewichtung muss also \mathbf{R}_w lediglich vor jedem Schritt mit $\sqrt{\lambda}$ multipliziert werden. Für $\lambda = 1$ geht der Algorithmus in das ungewichtete Verfahren über.

Besonders bei der rekursiven Berechnung ist es von Vorteil, den DSFI-Algorithmus nicht auf $\mathbf{R}_{w,k}$ sondern auf $\mathbf{S}_k = [\mathbf{R}_{w,k} \quad \mathbf{c}_{w,k}]$ anzuwenden, was schon für den nichtrekursiven Fall erläutert wurde:

$$\mathbf{Q}_{w,k} \cdot [\Psi_{w,k} \quad \tilde{\eta}_{w,k}] = \begin{bmatrix} \mathbf{R}_{w,k} & \mathbf{c}_{w,k} \\ \mathbf{0} & \tilde{\mathbf{c}}_{w,k} \end{bmatrix}.$$

Da man in der Praxis nur $\mathbf{R}_{w,k}$ und $\mathbf{c}_{w,k}$ zur Bestimmung der Lösung aus

$$\mathbf{R}_{w,k} \hat{\theta}_k = \mathbf{c}_{w,k}$$

benötigt, wird auf die Berechnung von $\mathbf{Q}_{w,k}$ und $\tilde{\mathbf{c}}_{w,k}$ verzichtet. Der Eingangsvektor Ψ_k^T und der Ausgang $\tilde{\eta}_k$ werden dazu entsprechend $\mathbf{S}_k = [\mathbf{R}_{w,k} \quad \mathbf{c}_{w,k}]$ in einem neuen Vektor

$$\Phi_k^T = [\Psi_k^T \quad \tilde{\eta}_k]$$

kombiniert.

Dieser Vektor Φ_k^T wird unten an $\sqrt{\lambda}\mathbf{S}_{k-1} = [\sqrt{\lambda}\mathbf{R}_{w,k-1} \quad \sqrt{\lambda}\mathbf{c}_{w,k-1}]$ angehängt, wobei der Faktor $\sqrt{\lambda}$ vor $\mathbf{c}_{w,k-1}$ daher kommt, dass $\begin{bmatrix} \mathbf{c}_{w,k-1} \\ \tilde{\mathbf{c}}_{w,k-1} \end{bmatrix} = \mathbf{Q}_{w,k-1} \tilde{\eta}_{w,k-1}$ ist, und sich $\mathbf{c}_{w,k-1}$ im k -ten Schritt damit wie $\tilde{\eta}_{w,k-1} = \mathbf{D}_{w,k-1} \tilde{\eta}_{k-1}$ verhält,

$$\tilde{\eta}_{w,k} = \begin{bmatrix} \lambda \tilde{\eta}_{w,k-1} \\ \tilde{\eta}_{w,k} \end{bmatrix}.$$

Schematisch ergeben sich die Transformationsschritte

$$\begin{bmatrix} \sqrt{\lambda}\mathbf{R}_{w,k-1} & \sqrt{\lambda}\mathbf{c}_{w,k-1} \\ \mathbf{0} & \sqrt{\lambda}\tilde{\mathbf{c}}_{w,k-1} \\ \Psi_k^T & \tilde{\eta}_k \end{bmatrix} = \begin{bmatrix} * & * & * & * \\ 0 & * & * & * \\ 0 & 0 & * & * \\ 0 & 0 & 0 & * \\ * & * & * & * \end{bmatrix} \xrightarrow{\Omega_1} \begin{bmatrix} \bullet & \bullet & \bullet & \bullet \\ 0 & * & * & * \\ 0 & 0 & * & * \\ 0 & 0 & 0 & * \\ 0 & * & * & * \end{bmatrix} \xrightarrow{\Omega_2} \begin{bmatrix} \bullet & \bullet & \bullet & \bullet \\ 0 & \bullet & \bullet & \bullet \\ 0 & 0 & * & * \\ 0 & 0 & 0 & * \\ 0 & 0 & * & * \end{bmatrix} \xrightarrow{\Omega_3} \begin{bmatrix} \bullet & \bullet & \bullet & \bullet \\ 0 & \bullet & \bullet & \bullet \\ 0 & 0 & \bullet & \bullet \\ 0 & 0 & 0 & * \\ 0 & 0 & 0 & * \end{bmatrix} = \begin{bmatrix} \mathbf{R}_{w,k} & \mathbf{c}_{w,k} \\ \mathbf{0} & \tilde{\mathbf{c}}_{w,k} \end{bmatrix},$$

wobei die Teil-Matrix $[\mathbf{0} \quad \sqrt{\lambda}\tilde{\mathbf{c}}_{w,k-1}]$ (bzw. $[\mathbf{0} \quad \tilde{\mathbf{c}}_{w,k}]$ nach dem letzten Schritt) in grau dargestellt und nur zur Übersicht angegeben ist. Praktisch werden die Werte nicht berechnet, und sie sind auch nicht Bestandteil der Matrix $\sqrt{\lambda}\mathbf{S}_{k-1}$ bzw. \mathbf{S}_k .

Einen Schritt des DSFI-Verfahrens kann man also wie folgt zusammenfassen:

- Setze:

$$\begin{aligned} \mathbf{S}_{k-1} &= [\mathbf{R}_{w,k-1} \quad \mathbf{c}_{w,k-1}] \\ \Phi_k^{(0)} &= [\Psi_k^T \quad \tilde{\eta}_k]^T \end{aligned}$$

- Berechne für $i = 1, \dots, p$:

- Wenn $(\Phi_k^{(i-1)})_i = 0$:

$$\begin{aligned} (\mathbf{S}_k)_{i,:} &= (\mathbf{S}_{k-1})_{i,:} \\ \Phi_k^{(i)} &= \Phi_k^{(i-1)} \end{aligned}$$

- Wenn $(\Phi_k^{(i-1)})_i \neq 0$:

$$\begin{aligned} (\mathbf{s}_k)_{ii} &= \sqrt{\lambda \cdot (\mathbf{s}_{k-1})_{ii}^2 + (\Phi_k^{(i-1)})_i^2} \\ \gamma &= \frac{\sqrt{\lambda} \cdot (\mathbf{s}_{k-1})_{ii}}{(\mathbf{s}_k)_{ii}}, \quad \sigma = \frac{(\Phi_k^{(i-1)})_i}{(\mathbf{s}_k)_{ii}}, \quad ((\Phi_k^{(i)})_i = 0) \end{aligned}$$

* Berechne für $j = i+1, \dots, p+1$:

$$\begin{aligned} (\mathbf{s}_k)_{ij} &= \gamma \cdot \sqrt{\lambda} \cdot (\mathbf{s}_{k-1})_{ij} + \sigma \cdot (\Phi_k^{(i-1)})_j \\ (\Phi_k^{(i)})_j &= -\sigma \cdot \sqrt{\lambda} \cdot (\mathbf{s}_{k-1})_{ij} + \gamma \cdot (\Phi_k^{(i-1)})_j \end{aligned}$$

- Berechne $\hat{\theta}_k$ durch Rückeinsetzen (bei Bedarf):
 - Wenn $\det((\mathbf{s}_k)_{1:p,1:p}) \neq 0$, d. h. $(\mathbf{s}_k)_{ii} \neq 0 \forall i = 1, \dots, p$:

$$\underbrace{(\mathbf{s}_k)_{1:p,1:p}}_{\mathbf{R}_{w,k}} \cdot \hat{\theta}_k = \underbrace{(\mathbf{s}_k)_{p+1,:}}_{\mathbf{c}_{w,k}}$$

Der Zweig zur Bedingung $(\Phi_k^{(i-1)})_i = 0$ sagt aus, dass keine Givens-Rotation nötig ist, wenn der entsprechende Eintrag in der (transformierten) angefügten Zeile schon Null ist. Damit ist auch der Fall abgefangen, dass sowohl $(\mathbf{s}_{k-1})_{ii} = 0$ als auch $(\Phi_k^{(i-1)})_i = 0$ ist. Dadurch kann der DSFI-Algorithmus mit $\mathbf{R}_0 = \mathbf{0}$ und $\mathbf{c}_0 = \mathbf{0}$, also

$$\mathbf{s}_0 = \mathbf{0}$$

gestartet werden. Der Schätzwert $\hat{\theta}_k$ kann erst dann bestimmt werden, wenn die Matrix $\mathbf{R}_k = (\mathbf{s}_k)_{1:p,1:p}$ vollen Rang erreicht hat, d. h. alle Diagonalelemente ungleich Null sind. Dieses Startverfahren ist exakt, d. h. wenn $\hat{\theta}_k$ bestimmt werden kann, entspricht es dem Wert, der sich auch bei der Anwendung eines geeigneten Verfahrens auf den gesamten Datensatz ergeben würde.

Schematisch kann man das Verfahren wie in Abbildung D.2 darstellen.

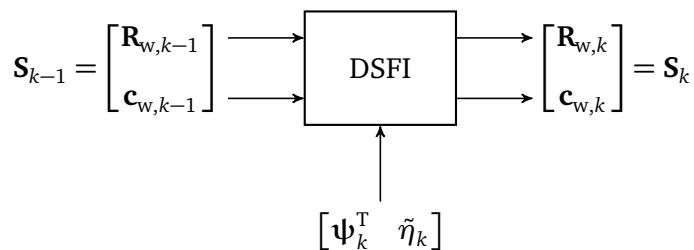


Abbildung D.2: DSFI-Algorithmus

Der DSFI-Algorithmus lässt sich relativ leicht in eine Matlab-Funktion umsetzen, wenn man die erweiterte Matrix \mathbf{S} verwendet und die benötigten Givens-Matrizen mit der `givens`-Funktion berechnet. Der Parametervektor kann in jedem Schritt als `theta=S(:,1:p)\S(:,p+1)` berechnet werden. Eine Funktion, die den Algorithmus auf diese Weise implementiert, ist in Listing D.1 gezeigt.

Listing D.1: Matlab-Funktion zur DSFI-Methode

```
function S = dsfi(S_old, psi, eta, lambda)

3    p = length(psi);
    phi = [psi.', eta];
    S_old = sqrt(lambda) * S_old;
    S = zeros(size(S_old));

8    for i = 1:p
        % GIVENS gibt auch fuer (S_old(i,i) == 0) && (phi(i) == 0)
        % eine sinnvolle Matrix G zurueck
        G = givens(S_old(i,i), phi(i));
        M = G * [S_old(i,i:p+1); phi(i:p+1)];
        S(i,i:p+1) = M(1,:);
        phi(i:p+1) = M(2,:);

13   end

18   % if all(diag(S(:,1:p)) ~= 0)
       %     theta = S(:,1:p) \ S(:,p+1);
       % end

end
```

E Parametrische Identifikation

E.1 Zeitkontinuierliches Butterworth-Filter

Die Übertragungsfunktion eines Butterworth-Filters der Ordnung n_F und der Grenzkreisfrequenz ω_F lautet

$$G_F(s) = \frac{1}{\prod_k \left(1 + \alpha_k \frac{s}{\omega_F} + \beta_k \frac{s^2}{\omega_F^2}\right)},$$

wobei sich die Koeffizienten α_k und β_k sowie die Anzahl der Faktoren sich in Abhängigkeit von n_F wie folgt ergeben.

Für gerade Ordnungen n_F gilt

$$\alpha_k = 2 \cdot \cos \frac{(2k-1) \cdot \pi}{2n_F}$$

$$\beta_k = 1$$

für $k = 1, \dots, \frac{n_F}{2}$.

Für ungerade Ordnungen n_F gilt

$$\alpha_1 = 1$$

$$\beta_1 = 0$$

und

$$\alpha_k = 2 \cdot \cos \frac{(k-1) \cdot \pi}{n_F}$$

$$\beta_k = 1$$

für $k = 2, \dots, \frac{n_F+1}{2}$.

E.2 Alternatives Aufstellen der Matrizen bei Schätzung aus Frequenzbereichsdaten

Der Frequenzgang kann auch in Real- und Imaginärteil aufgeteilt geschrieben werden,

$$H(\omega) = H_R(\omega) + jH_I(\omega).$$

Diese Schreibweise wird später nur für den gemessenen Frequenzgang $\tilde{H}(\omega)$ verwendet, so dass sich an dieser Stelle keine Gedanken darüber gemacht werden müssen, wie $H_R(\omega)$ und $H_I(\omega)$ von den Koeffizienten a_j und b_j abhängen.

Führt man in

$$H(\omega) = \frac{b_0 + jb_1\omega + \cdots + j^m b_m}{1 + ja_1\omega + \cdots + j^n a_n\omega^n}$$

die Potenzierungen der imaginären Einheit j aus, so erhält man

$$\begin{aligned} H(\omega) &= \frac{b_0 + jb_1\omega + j^2 b_2\omega^2 + j^3 b_3\omega^3 + j^4 b_4\omega^4 + j^5 b_5\omega^5 + j^6 b_6\omega^6 + \cdots}{1 + ja_1\omega + j^2 a_2\omega^2 + j^3 a_3\omega^3 + j^4 a_4\omega^4 + j^5 a_5\omega^5 + j^6 a_6\omega^6 + \cdots} \\ &= \frac{b_0 + jb_1\omega - b_2\omega^2 - jb_3\omega^3 + b_4\omega^4 + jb_5\omega^5 - b_6\omega^6 + \cdots}{1 + ja_1\omega - a_2\omega^2 - ja_3\omega^3 + a_4\omega^4 + ja_5\omega^5 - a_6\omega^6 + \cdots}. \end{aligned}$$

In der schon oben angegebenen Form

$$H(\omega) = \frac{B(j\omega)}{A(j\omega)}$$

lassen sich $B(j\omega)$ und $A(j\omega)$ jeweils in Real- und Imaginärteil aufteilen, so dass man

$$H(\omega) = \frac{B_R(\omega) + j \cdot B_I(\omega)}{A_R(\omega) + j \cdot A_I(\omega)}$$

erhält, wobei alle Polynome ($A_R(\omega)$, $A_I(\omega)$, $B_R(\omega)$ und $B_I(\omega)$) reell sind. Ausgeschrieben lauten die Polynome

$$A_R(\omega) = 1 - a_2\omega^2 + a_4\omega^4 - a_6\omega^6 + \cdots, \quad (\text{E.1})$$

$$A_I(\omega) = a_1\omega - a_3\omega^3 + a_5\omega^5 - \cdots, \quad (\text{E.2})$$

$$B_R(\omega) = b_0 - b_2\omega^2 + b_4\omega^4 - b_6\omega^6 + \cdots \quad \text{und} \quad (\text{E.3})$$

$$B_I(\omega) = b_1\omega - b_3\omega^3 + b_5\omega^5 - \cdots. \quad (\text{E.4})$$

Um den Fehler $\hat{e}_{k'}$ in Real- und Imaginärteil aufzuteilen, werden alle in Gl. (9.14) auftretenden komplexen Größen in Real- und Imaginärteil geschrieben,

$$\begin{aligned} \hat{e}_{k'} &= \tilde{H}_R(\omega_{k'}) + j\tilde{H}_I(\omega_{k'}) \\ &\quad - ((1 - \hat{A}_R(\omega_{k'})) \cdot (\tilde{H}_R(\omega_{k'}) + j\tilde{H}_I(\omega_{k'})) + (\hat{B}_R(\omega_{k'}) + j\hat{B}_I(\omega_{k'}))) \\ &= \tilde{H}_R(\omega_{k'}) + j\tilde{H}_I(\omega_{k'}) \\ &\quad - ((1 - \hat{A}_R(\omega_{k'})) \cdot \tilde{H}_R(\omega_{k'}) + \hat{A}_I(\omega_{k'}) \cdot \tilde{H}_I(\omega_{k'}) + \hat{B}_R(\omega_{k'})) \\ &\quad - j \cdot ((1 - \hat{A}_R(\omega_{k'})) \cdot \tilde{H}_I(\omega_{k'}) - \hat{A}_I(\omega_{k'}) \cdot \tilde{H}_R(\omega_{k'}) + \hat{B}_I(\omega_{k'})). \end{aligned}$$

Hieraus lässt sich jetzt einfach

$$\hat{e}_{R,k'} = \tilde{H}_R(\omega_{k'}) - \underbrace{((1 - \hat{A}_R(\omega_{k'})) \cdot \tilde{H}_R(\omega_{k'}) + \hat{A}_I(\omega_{k'}) \cdot \tilde{H}_I(\omega_{k'}) + \hat{B}_R(\omega_{k'}))}_{\hat{H}_R(\omega)}$$

und

$$\hat{e}_{I,k'} = \tilde{H}_I(\omega_{k'}) - \underbrace{((1 - \hat{A}_R(\omega_{k'})) \cdot \tilde{H}_I(\omega_{k'}) - \hat{A}_I(\omega_{k'}) \cdot \tilde{H}_R(\omega_{k'}) + \hat{B}_I(\omega_{k'}))}_{\hat{H}_I(\omega)}$$

ablesen.

Mit Gl. (E.1) bis (E.4) erhält man

$$\begin{aligned}\hat{H}_R(\omega) &= \hat{a}_1 \cdot \omega \tilde{H}_I(\omega) + \hat{a}_2 \cdot \omega^2 \tilde{H}_R(\omega) - \hat{a}_3 \cdot \omega^3 \tilde{H}_I(\omega) - \hat{a}_4 \cdot \omega^4 \tilde{H}_R(\omega) + \dots \\ &\quad + \hat{b}_0 - \hat{b}_2 \omega^2 + \hat{b}_4 \omega^4 - \hat{b}_6 \omega^6 + \dots\end{aligned}$$

und

$$\begin{aligned}\hat{H}_I(\omega) &= -\hat{a}_1 \cdot \omega \tilde{H}_R(\omega) + \hat{a}_2 \cdot \omega^2 \tilde{H}_I(\omega) + \hat{a}_3 \cdot \omega^3 \tilde{H}_R(\omega) - \hat{a}_4 \cdot \omega^4 \tilde{H}_I(\omega) - \dots \\ &\quad + \hat{b}_1 \omega - \hat{b}_3 \omega^3 + \hat{b}_5 \omega^5 - \dots.\end{aligned}$$

Damit können jetzt die Ausgangs- und Eingangsgrößen des LS-Problems für den Real- und Imaginärteil des Fehlers angegeben werden. Der Parametervektor ist in beiden Fällen identisch und lautet¹

$$\boldsymbol{\theta} = [a_1 \ a_2 \ a_3 \ \dots \ a_n \ b_0 \ b_1 \ b_2 \ \dots \ b_m]^T.$$

Für den Realteil gilt

$$\tilde{\eta}_{R,k'} = \tilde{H}_R(\omega_{k'})$$

und

$$\tilde{\Psi}_{R,k'}^T = [\omega_{k'} \tilde{H}_I(\omega_{k'}) \ \omega_{k'}^2 \tilde{H}_R(\omega_{k'}) \ -\omega_{k'}^3 \tilde{H}_I(\omega_{k'}) \ -\omega_{k'}^4 \tilde{H}_R(\omega_{k'}) \ \dots \ 1 \ 0 \ -\omega_{k'}^2 \ 0 \ \omega_{k'}^4 \ \dots].$$

Analog ergibt sich für den Imaginärteil

$$\tilde{\eta}_{I,k'} = \tilde{H}_I(\omega_{k'})$$

und

$$\tilde{\Psi}_{I,k'}^T = [-\omega_{k'} \tilde{H}_R(\omega_{k'}) \ \omega_{k'}^2 \tilde{H}_I(\omega_{k'}) \ \omega_{k'}^3 \tilde{H}_R(\omega_{k'}) \ -\omega_{k'}^4 \tilde{H}_I(\omega_{k'}) \ \dots \ 0 \ \omega_{k'}^2 \ 0 \ -\omega_{k'}^3 \ 0 \ \dots].$$

Fasst man die Daten für k Frequenzpunkte zusammen, so erhält man insgesamt $2k$ Gleichungen, die über

$$\tilde{\eta}_k = \begin{bmatrix} \tilde{\eta}_{R,1} \\ \tilde{\eta}_{I,1} \\ \tilde{\eta}_{R,2} \\ \tilde{\eta}_{I,2} \\ \vdots \\ \tilde{\eta}_{R,k} \\ \tilde{\eta}_{I,k} \end{bmatrix} = \begin{bmatrix} \tilde{H}_R(\omega_1) \\ \tilde{H}_I(\omega_1) \\ \tilde{H}_R(\omega_2) \\ \tilde{H}_I(\omega_2) \\ \vdots \\ \tilde{H}_R(\omega_k) \\ \tilde{H}_I(\omega_k) \end{bmatrix} \quad \text{und} \quad \tilde{\Psi}_k = \begin{bmatrix} \tilde{\Psi}_{R,1}^T \\ \tilde{\Psi}_{I,1}^T \\ \tilde{\Psi}_{R,2}^T \\ \tilde{\Psi}_{I,2}^T \\ \vdots \\ \tilde{\Psi}_{R,k}^T \\ \tilde{\Psi}_{I,k}^T \end{bmatrix}$$

sowie die Gewichtsmatrix

$$\mathbf{W}_k = \begin{bmatrix} w_1 & & & & \\ & w_1 & & & \\ & & w_2 & & \\ & & & w_2 & \\ & & & & \ddots \\ & & & & \\ & & & & w_k & \\ & & & & & w_k \end{bmatrix}$$

zusammengefasst werden.

¹ Im Vergleich zum Aufbau im Haupttext ist die Reihenfolge der Parameter hier anders!

E.3 Überlagerung von harmonischen Schwingungen

Die Summe von zwei harmonischen Schwingungen gleicher Frequenz ist wieder eine harmonische Schwingung mit der selben Frequenz, d. h. mit $x_1(t) = A_1 \cdot \cos(\omega t + \varphi_1)$ und $x_2(t) = A_2 \cdot \cos(\omega t + \varphi_2)$ gilt

$$\begin{aligned} x(t) &= \alpha_1 x_1(t) + \alpha_2 x_2(t) \\ &= A \cdot \cos(\omega t + \varphi), \end{aligned}$$

wobei α_1 und α_2 hier frei wählbare Gewichtungsfaktoren sind. Mit diesen beiden Faktoren lässt sich unter geringen Voraussetzungen erreichen, dass die sich ergebende Schwingung eine beliebige Amplitude und Phase besitzt.

Um dies zu sehen, werden die harmonischen Schwingungen in einer komplexen Zeigerdarstellung geschrieben. Es gilt

$$\begin{aligned} x(t) &= \alpha_1 A_1 \cdot \cos(\omega t + \varphi_1) + \alpha_2 A_2 \cdot \cos(\omega t + \varphi_2) \\ &= \alpha_1 \cdot \frac{A_1}{2} \cdot (e^{j(\omega t + \varphi_1)} + e^{-j(\omega t + \varphi_1)}) + \alpha_2 \cdot \frac{A_2}{2} \cdot (e^{j(\omega t + \varphi_2)} + e^{-j(\omega t + \varphi_2)}) \end{aligned}$$

und damit weiter

$$x(t) = e^{j\omega t} \cdot \underbrace{\left(\alpha_1 \cdot \frac{A_1}{2} e^{j\varphi_1} + \alpha_2 \cdot \frac{A_2}{2} e^{j\varphi_2} \right)}_{X} + e^{-j\omega t} \cdot \underbrace{\left(\alpha_1 \cdot \frac{A_1}{2} e^{-j\varphi_1} + \alpha_2 \cdot \frac{A_2}{2} e^{-j\varphi_2} \right)}_{X^*}.$$

In der komplexen Größe

$$X = \alpha_1 \cdot \frac{A_1}{2} e^{j\varphi_1} + \alpha_2 \cdot \frac{A_2}{2} e^{j\varphi_2} = \alpha_1 X_1 + \alpha_2 X_2$$

steckt die Information über die Amplitude und Phasenlage der sich ergebenden Schwingung $x(t)$.

Die Addition der beiden Teile zu X ist in Abbildung E.1 grafisch veranschaulicht. Solange A_1 und A_2 (und damit $|X_1|$ und $|X_2|$) ungleich Null sind und X_1 und X_2 in unterschiedliche Richtungen zeigen, d. h. die Differenz $\varphi_2 - \varphi_1$ kein ganzes Vielfaches von π ist, kann durch eine geeignete Wahl von α_1 und α_2 jede beliebige Schwingung gleicher Frequenz aus $x_1(t)$ und $x_2(t)$ erzeugt werden.

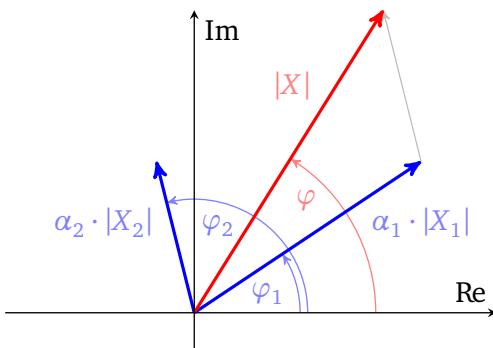


Abbildung E.1: Zur Addition $X = \alpha_1 X_1 + \alpha_2 X_2$

F Kalmanfilter

F.1 Betrachtung des Korrektur- und Prädiktionsschritt bezüglich des MSE

Korrekturschritt

Für den Schätzfehler über einen Korrekturschritt gilt

$$\Delta \hat{x}_{i,+} = (\mathbf{e}_i^T - \mathbf{k}_i^T \mathbf{C}) \cdot \Delta \hat{\mathbf{x}}_- - \mathbf{k}_i^T \mathbf{n}_y .$$

Es soll der mittlere quadrierte Fehler aller Komponenten $\hat{x}_{i,+}$ minimiert werden. Aufgrund dessen, dass über die \mathbf{k}_i^T die Mittelung unabhängig stattfindet, kann

$$\text{MSE}\{\Delta \hat{x}_{i,+}\}_r = E\{\Delta \hat{x}_{i,+}^2\}_r \rightarrow \min$$

für jedes i gefordert werden.

Das Quadrat des Schätzfehlers nach der Korrektur ist

$$\begin{aligned} \Delta \hat{x}_{i,+}^2 &= ((\mathbf{e}_i^T - \mathbf{k}_i^T \mathbf{C}) \cdot \Delta \hat{\mathbf{x}}_- - \mathbf{k}_i^T \mathbf{n}_y)^2 \\ &= (\mathbf{e}_i^T - \mathbf{k}_i^T \mathbf{C}) \cdot \Delta \hat{\mathbf{x}}_- \Delta \hat{\mathbf{x}}_-^T \cdot (\mathbf{e}_i - \mathbf{C}^T \mathbf{k}_i) + 2 \cdot (\mathbf{e}_i^T - \mathbf{k}_i^T \mathbf{C}) \cdot \Delta \hat{\mathbf{x}}_- \mathbf{n}_y^T \mathbf{k}_i + \mathbf{k}_i^T \mathbf{n}_y \mathbf{n}_y^T \mathbf{k}_i \end{aligned}$$

(($\mathbf{e}_i^T - \mathbf{k}_i^T \mathbf{C}$) \cdot $\Delta \hat{\mathbf{x}}_-$ und $\mathbf{k}_i^T \mathbf{n}_y$ sind Skalare und können daher immer transponiert werden) und damit ergibt sich für dessen Erwartungswert

$$\begin{aligned} E\{\Delta \hat{x}_{i,+}^2\}_r &= (\mathbf{e}_i^T - \mathbf{k}_i^T \mathbf{C}) \cdot E\{\Delta \hat{\mathbf{x}}_- \Delta \hat{\mathbf{x}}_-^T\}_r \cdot (\mathbf{e}_i - \mathbf{C}^T \mathbf{k}_i) \\ &\quad + 2 \cdot (\mathbf{e}_i^T - \mathbf{k}_i^T \mathbf{C}) \cdot E\{\Delta \hat{\mathbf{x}}_-\}_r \cdot E\{\mathbf{n}_y^T\}_r \cdot \mathbf{k}_i \\ &\quad + \mathbf{k}_i^T \cdot E\{\mathbf{n}_y \mathbf{n}_y^T\}_r \cdot \mathbf{k}_i . \end{aligned}$$

Der Erwartungswert von $E\{\Delta \hat{\mathbf{x}}_- \Delta \hat{\mathbf{x}}_-^T\}_r$ ist die „Korrelationsmatrix“ und wird hier mit $\mathbf{P}'_{\Delta,-}$ bezeichnet,

$$\mathbf{P}'_{\Delta,-} = E\{\Delta \hat{\mathbf{x}}_- \Delta \hat{\mathbf{x}}_-^T\}_r .$$

Die Diagonalelemente $(\mathbf{P}'_{\Delta,-})_{ii} = E\{\Delta \hat{x}_{ii,-}^2\}_r$ entsprechen den mittleren quadrierten Fehler der einzelnen Komponenten des Schätzfehlers.

Da das Rauschen weiterhin als mittelwertfrei angenommen wird, gilt $E\{\Delta \hat{\mathbf{x}}_-\}_r \cdot E\{\mathbf{n}_y^T\}_r = 0$ und damit

$$E\{\Delta \hat{x}_{i,+}^2\}_r = (\mathbf{e}_i^T - \mathbf{k}_i^T \mathbf{C}) \cdot \mathbf{P}'_{\Delta,-} \cdot (\mathbf{e}_i - \mathbf{C}^T \mathbf{k}_i) + \mathbf{k}_i^T \cdot \mathbf{P}_y \cdot \mathbf{k}_i .$$

Diese Gleichung hat dieselbe Struktur wie Gl. (10.7), und damit ergibt sich zur Minimierung der linken Seite auch der gleiche Ausdruck für \mathbf{k}_i^T , nur mit $\mathbf{P}'_{\Delta,-}$ anstelle von $\mathbf{P}_{\Delta,-}$. Es folgt also

$$\mathbf{K} = \mathbf{P}'_{\Delta,-} \mathbf{C}^T \cdot (\mathbf{C} \mathbf{P}'_{\Delta,-} \mathbf{C}^T + \mathbf{P}_y)^{-1} .$$

Zur Bestimmung der Kalmanverstärkung wird neben der Kovarianzmatrix des Messrauschen hier auch die „Korrelationsmatrix“ des Schätzfehlers benötigt. Daher muss noch bestimmt werden, wie sich dieser bei der Korrektur und Prädiktion verhält.

Die Korrelationsmatrix des Schätzfehlers nach der Korrektur ist

$$\mathbf{P}'_{\Delta,+} = E\{\Delta\hat{\mathbf{x}}_+ \cdot \Delta\hat{\mathbf{x}}_+^T\}_r .$$

Für den Schätzfehler nach der Korrektur gilt

$$\begin{aligned}\Delta\hat{\mathbf{x}}_+ &= \mathbf{x} - \hat{\mathbf{x}}_+ = \mathbf{x} - (\mathbf{I} - \mathbf{KC}) \cdot \hat{\mathbf{x}}_- - \mathbf{K}\tilde{\mathbf{y}} \\ &= \mathbf{x} - (\mathbf{I} - \mathbf{KC}) \cdot \hat{\mathbf{x}}_- - \mathbf{KCx} - \mathbf{Kn}_y \\ &= (\mathbf{I} - \mathbf{KC}) \cdot (\mathbf{x} - \hat{\mathbf{x}}_-) - \mathbf{Kn}_y \\ &= (\mathbf{I} - \mathbf{KC}) \cdot \Delta\hat{\mathbf{x}}_- - \mathbf{Kn}_y .\end{aligned}$$

Damit folgt

$$\begin{aligned}\Delta\hat{\mathbf{x}}_+ \cdot \Delta\hat{\mathbf{x}}_+^T &= (\mathbf{I} - \mathbf{KC}) \cdot \Delta\hat{\mathbf{x}}_- \cdot \Delta\hat{\mathbf{x}}_-^T \cdot (\mathbf{I} - \mathbf{C}^T \mathbf{K}^T) \\ &\quad - (\mathbf{I} - \mathbf{KC}) \cdot \Delta\hat{\mathbf{x}}_- \cdot \mathbf{n}_y^T \cdot \mathbf{K}^T \\ &\quad - \mathbf{K} \cdot \mathbf{n}_y \cdot \Delta\hat{\mathbf{x}}_-^T \cdot (\mathbf{I} - \mathbf{C}^T \mathbf{K}^T) \\ &\quad + \mathbf{Kn}_y \mathbf{n}_y^T \mathbf{K}^T\end{aligned}$$

und weiter ergibt sich für das gesuchte $\mathbf{P}'_{\Delta,+}$

$$\mathbf{P}'_{\Delta,+} = E\{\Delta\hat{\mathbf{x}}_+ \cdot \Delta\hat{\mathbf{x}}_+^T\}_r = (\mathbf{I} - \mathbf{KC}) \cdot \mathbf{P}'_{\Delta,-} \cdot (\mathbf{I} - \mathbf{C}^T \mathbf{K}^T) + \mathbf{K} \cdot \mathbf{P}_y \cdot \mathbf{K}^T$$

was ebenfalls, für das oben bestimmte \mathbf{K} zu

$$\mathbf{P}'_{\Delta,+} = (\mathbf{I} - \mathbf{KC}) \cdot \mathbf{P}'_{\Delta,-}$$

vereinfacht werden kann.

Prädiktionsschritt

Die Prädiktion erfolgt mit der gleichen Gleichung wie im Fall der minimalen Varianz, so dass sich auch für den Schätzfehler derselbe Ausdruck

$$\Delta\hat{\mathbf{x}}_{k+1} = \mathbf{A}_k \cdot \Delta\hat{\mathbf{x}}_k + \mathbf{n}_{x,k}$$

ergibt.

Die Korrelationsmatrix des Schätzfehlers nach dem Prädiktionsschritt ist

$$\mathbf{P}'_{\Delta,k+1} = E\{\Delta\hat{\mathbf{x}}_{k+1} \cdot \Delta\hat{\mathbf{x}}_{k+1}^T\}_r .$$

Für das Produkt im Erwartungswert gilt

$$\Delta\hat{\mathbf{x}}_{k+1} \cdot \Delta\hat{\mathbf{x}}_{k+1}^T = \mathbf{A}_k \cdot \Delta\hat{\mathbf{x}}_k \cdot \Delta\hat{\mathbf{x}}_k^T \cdot \mathbf{A}_k^T + \mathbf{A}_k \cdot \Delta\hat{\mathbf{x}}_k \cdot \mathbf{n}_{x,k}^T + \mathbf{n}_{x,k} \cdot \Delta\hat{\mathbf{x}}_k^T \cdot \mathbf{A}_k^T + \mathbf{n}_{x,k} \mathbf{n}_{x,k}^T$$

und damit erhält man

$$\mathbf{P}'_{\Delta,k+1} = \mathbf{A}_k \cdot \mathbf{P}'_{\Delta,k+1} \cdot \mathbf{A}_k^T + \mathbf{P}_{x,k} .$$

F.2 Parallele Messungen zusammen oder einzeln verarbeiten

Im Folgenden wird gezeigt, dass es mathematisch äquivalent ist, die Messungen \tilde{y}_1 und \tilde{y}_2 zweier Ausgangsgrößen(vektoren)

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} \cdot x$$

einzeln oder zusammen im Kalmanfilter zu berücksichtigen, solange diese unkorreliert sind, d. h.

$$R = \begin{bmatrix} R_1 & 0 \\ 0 & R_2 \end{bmatrix}$$

gilt.

Dazu wird zunächst der Schätzwert \hat{x}_2^* sowie die Fehlerkovarianzmatrix P_2^* bestimmt, der sich aus der gemeinsamen Berücksichtigung der beiden Messungen ergibt.

Im Anschluss werden die beiden Messungen einzeln berücksichtigt. Dazu wird zunächst der Schätzwert \hat{x}_1 bestimmt, der sich nach der Korrektur mit \tilde{y}_1 ergibt. Dieser wird dann im einen weiteren Schritt mit \tilde{y}_2 zu \hat{x}_2 korrigiert.

Es wird sich zeigen, dass $\hat{x}_2^* = \hat{x}_2$ und $P_2^* = P_2$ ist, womit die aufgestellte Behauptung bewiesen ist.

Gemeinsame Berücksichtigung

Fasst man beide Teilmessungen \tilde{y}_1 und \tilde{y}_2 zu einem Vektor zusammen und berücksichtigt diesen gemeinsam, dann ergibt sich die Korrekturmatrixt

$$\begin{aligned} K^* &= P_0 \cdot [C_1^T \quad C_2^T] \cdot \left(\begin{bmatrix} C_1 \\ C_2 \end{bmatrix} P_0 \begin{bmatrix} C_1^T & C_2^T \end{bmatrix} + \begin{bmatrix} R_1 & 0 \\ 0 & R_2 \end{bmatrix} \right)^{-1} \\ &= P_0 \cdot [C_1^T \quad C_2^T] \cdot \begin{bmatrix} C_1 P_0 C_1^T + R_1 & C_1 P_0 C_2^T \\ C_2 P_0 C_1^T & C_2 P_0 C_2^T + R_2 \end{bmatrix}^{-1}. \end{aligned}$$

Mit den Abkürzungen

$$X_{11} = C_1 P_0 C_1^T + R_1, \quad X_{12} = C_1 P_0 C_2^T, \quad X_{21} = C_2 P_0 C_1^T \quad \text{und} \quad X_{22} = C_2 P_0 C_2^T + R_2$$

kann dies auch

$$K^* = P_0 \cdot [C_1^T \quad C_2^T] \cdot \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix}^{-1}$$

geschrieben werden.

Die Inverse kann unter der Bedingung, dass $X_{11} = C_1 P_0 C_1^T + R_1$ invertierbar ist (was ohnehin nötig wäre, um auch die einzelne Auswertung durchführen zu können) als

$$\begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix}^{-1} = \begin{bmatrix} X_{11}^{-1} + X_{11}^{-1} X_{12} \tilde{X}^{-1} X_{21} X_{11}^{-1} & -X_{11}^{-1} X_{12} \tilde{X}^{-1} \\ \tilde{X}^{-1} X_{21} X_{11}^{-1} & \tilde{X}^{-1} \end{bmatrix}$$

mit

$$\tilde{X} = X_{22} - X_{21} X_{11}^{-1} X_{12}$$

geschrieben werden. (Siehe [Zurmühl und Falk 1984, S. 303].)

Damit erhält man den korrigierten Schätzwert

$$\begin{aligned}\hat{x}_2^* &= \hat{x}_0 + K^* \cdot \left(\begin{bmatrix} \tilde{y}_1 \\ \tilde{y}_2 \end{bmatrix} - \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} \cdot \hat{x}_0 \right) \\ &= \hat{x}_0 + P_0 \cdot \begin{bmatrix} C_1^T & C_2^T \end{bmatrix} \cdot \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix}^{-1} \cdot \left(\begin{bmatrix} \tilde{y}_1 \\ \tilde{y}_2 \end{bmatrix} - \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} \cdot \hat{x}_0 \right)\end{aligned}$$

und die neue Fehlerkovarianzmatrix

$$\begin{aligned}P_2^* &= P_0 - K^* \cdot \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} \cdot P_0 \\ &= P_0 - P_0 \cdot \begin{bmatrix} C_1^T & C_2^T \end{bmatrix} \cdot \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix}^{-1} \cdot \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} \cdot P_0.\end{aligned}$$

Einzelne Berücksichtigung

Für die Korrektur mit der ersten Messung lautet die Kalmanverstärkung

$$\begin{aligned}K_1 &= P_0 C_1^T \cdot (C_1 P_0 C_1^T + R_1)^{-1} \\ &= P_0 C_1^T \cdot X_{11}^{-1},\end{aligned}$$

wobei X_{11} dieselbe Abkürzung wie oben ist. Es folgt daraus die Fehlerkovarianzmatrix

$$P_1 = P_0 - K_1 \cdot C_1 P_0 \tag{F.1}$$

$$= P_0 - P_0 C_1^T \cdot X_{11}^{-1} \cdot C_1 P_0 \tag{F.2}$$

nach dem ersten Schritt.

Mit dieser und der Kovarianzmatrix R_2 der Messung von y_2 beträgt die Kalmanverstärkung für den zweiten Schritt

$$K_2 = P_1 C_2^T \cdot (C_2 P_1 C_2^T + R_2)^{-1}. \tag{F.3}$$

Für die zu invertierende Matrix $(C_2 P_1 C_2^T + R_2)$ erhält man mit (F.2)

$$\begin{aligned}C_2 P_1 C_2^T + R_2 &= C_2 \cdot (P_0 - P_0 C_1^T \cdot X_{11}^{-1} \cdot C_1 P_0) \cdot C_2^T + R_2 \\ &= X_{22} - X_{21} X_{11}^{-1} X_{12}\end{aligned}$$

und damit

$$C_2 P_1 C_2^T + R_2 = \tilde{X}.$$

Dieses wieder in Gl. (F.3) eingesetzt ergibt

$$\begin{aligned}K_2 &= P_1 C_2^T \cdot \tilde{X}^{-1} \\ &= P_0 C_2^T \tilde{X}^{-1} - P_0 X_{11}^{-1} X_{12} \tilde{X}^{-1}.\end{aligned}$$

Mit den Kalmanverstärkungen K_1 und K_2 erhält man für die beiden Korrekturschritte der Schätzwerte

$$\begin{aligned}\hat{x}_1 &= \hat{x}_0 + K_1 \cdot (\tilde{y}_1 - C_1 \hat{x}_0) \quad \text{und} \\ \hat{x}_2 &= \hat{x}_1 + K_2 \cdot (\tilde{y}_2 - C_2 \hat{x}_1).\end{aligned}$$

Substituiert man die beiden $\hat{\mathbf{x}}_1$ in der zweiten Gleichung durch die erste erhält man

$$\begin{aligned}
 \hat{\mathbf{x}}_2 &= \hat{\mathbf{x}}_0 + \mathbf{K}_1 \cdot (\tilde{\mathbf{y}}_1 - \mathbf{C}_1 \hat{\mathbf{x}}_0) + \mathbf{K}_2 \cdot (\tilde{\mathbf{y}}_2 - \mathbf{C}_2 \cdot (\hat{\mathbf{x}}_0 + \mathbf{K}_1 \cdot (\tilde{\mathbf{y}}_1 - \mathbf{C}_1 \hat{\mathbf{x}}_0))) \\
 &= \hat{\mathbf{x}}_0 + \mathbf{K}_1 \cdot (\tilde{\mathbf{y}}_1 - \mathbf{C}_1 \hat{\mathbf{x}}_0) + \mathbf{K}_2 \cdot (\tilde{\mathbf{y}}_2 - \mathbf{C}_2 \hat{\mathbf{x}}_0) - \mathbf{K}_2 \mathbf{C}_2 \mathbf{K}_1 \cdot (\tilde{\mathbf{y}}_1 - \mathbf{C}_1 \hat{\mathbf{x}}_0) \\
 &= \hat{\mathbf{x}}_0 + [\mathbf{K}_1 - \mathbf{K}_2 \mathbf{C}_2 \mathbf{K}_1 \quad \mathbf{K}_2] \cdot \begin{bmatrix} \tilde{\mathbf{y}}_1 - \mathbf{C}_1 \hat{\mathbf{x}}_0 \\ \tilde{\mathbf{y}}_2 - \mathbf{C}_2 \hat{\mathbf{x}}_0 \end{bmatrix} \\
 &= \hat{\mathbf{x}}_0 + \underbrace{[\mathbf{K}_1 - \mathbf{K}_2 \mathbf{C}_2 \mathbf{K}_1 \quad \mathbf{K}_2]}_{\mathbf{K}} \cdot \left(\begin{bmatrix} \tilde{\mathbf{y}}_1 \\ \tilde{\mathbf{y}}_2 \end{bmatrix} - \begin{bmatrix} \mathbf{C}_1 \\ \mathbf{C}_2 \end{bmatrix} \hat{\mathbf{x}}_0 \right).
 \end{aligned}$$

Die einzelne Berücksichtigung der Messungen entspricht demnach einer gemeinsamen Berücksichtigung mit der Korrekturmatrixt

$$\begin{aligned}
 \mathbf{K} &= [\mathbf{K}_1 - \mathbf{K}_2 \mathbf{C}_2 \mathbf{K}_1 \quad \mathbf{K}_2] \\
 &= [\mathbf{P}_0 \mathbf{C}_1^T \cdot \mathbf{X}_{11}^{-1} - (\mathbf{P}_0 \mathbf{C}_2^T \tilde{\mathbf{X}}^{-1} - \mathbf{P}_0 \mathbf{X}_{11}^{-1} \mathbf{X}_{12} \tilde{\mathbf{X}}^{-1}) \cdot \mathbf{C}_2 \mathbf{P}_0 \mathbf{C}_1^T \cdot \mathbf{X}_{11}^{-1} \quad \mathbf{P}_0 \mathbf{C}_2^T \tilde{\mathbf{X}}^{-1} - \mathbf{P}_0 \mathbf{X}_{11}^{-1} \mathbf{X}_{12} \tilde{\mathbf{X}}^{-1}] \\
 &= \mathbf{P}_0 \cdot [\mathbf{C}_1^T \quad \mathbf{C}_2^T] \cdot \begin{bmatrix} \mathbf{X}_{11}^{-1} + \mathbf{X}_{11}^{-1} \mathbf{X}_{12} \tilde{\mathbf{X}}^{-1} \mathbf{X}_{21} \mathbf{X}_{11}^{-1} & -\mathbf{X}_{11}^{-1} \mathbf{X}_{12} \tilde{\mathbf{X}}^{-1} \\ -\tilde{\mathbf{X}}^{-1} \mathbf{X}_{21} \mathbf{X}_{11}^{-1} & \tilde{\mathbf{X}}^{-1} \end{bmatrix}.
 \end{aligned}$$

Dieses entspricht dem \mathbf{K}^* , welches sich bei der gemeinsamen Berücksichtigung der Messungen ergab.

Da effektiv das gleiche \mathbf{K} verwendet wird, ist auch die Kovarianzmatrix

$$\mathbf{P}_2 = \mathbf{P}_1 - \mathbf{K}_2 \cdot \mathbf{C}_2 \mathbf{P}_1$$

identisch mit \mathbf{P}_2^* , wovon man sich auch durch analoges Einsetzen und Partitionieren überzeugen kann.

Literatur

- Brammer, Karl und Gerhard Siffling (1975). *Stochastische Grundlagen des Kalman-Bucy-Filters Wahrscheinlichkeitsrechnung und Zufallsprozesse*. Methoden der Regelungstechnik. München, Wien: R. Oldenbourg Verlag.
- (1989). *Kalman-Bucy-Filter Deterministische Beobachtung und Filterung*. 3. Methoden der Regelungstechnik. München, Wien: R. Oldenbourg Verlag.
- Eykhoff, Pieter (1974). *System Identification Parameter and State Estimation*. London, etc.: Wiley-Interscience.
- Föllinger, Otto (1977). *Laplace- und Fourier-Transformation*. Berlin: Elitera.
- Freund, Rudolf J., William J. Wilson und Ping Sa (2006). *Regression Analysis*. 2. Burlington, MA, etc.: Elsevier.
- Golub, Gene H. und Charles F. Van Loan (2013). *Matrix Computations*. 4. The Johns Hopkins University Press.
- Hänsler, Eberhard (2001). *Statistische Signale*. 3. Berlin, etc.: Springer.
- Hoffmann, Armin, Bernd Marx und Werner Vogt (2006). *Mathematik für Ingenieure 2*. München: Pearson Studium.
- Isermann, R. und M. Münchhof (2011). *Identification of Dynamic Systems*. Berlin, Heidelberg: Springer.
- Kamen, E. W. und J. K. Su (1999). *Introduction to Optimal Estimation*. London, etc.: Springer.
- Kammeyer, Karl-Dirk und Kristian Kroschel (2006). *Digital Signalverarbeitung*. 6. Wiesbaden: Teubner.
- Ljung, Lennart (1999). *System Identification Theory for the User*. 2. Upper Saddle River, NJ: Prentice Hall.
- Natke, Hans Günther (1992). *Einführung in Theorie und Praxis der Zeitreihen- und Modalanalyse*. 3. Braunschweig, Wiesbaden: vieweg.
- Oppenheim, Alan V. und Ronald W. Schafer (1995). *Zeitdiskrete Signalverarbeitung*. 2. München, Wien: Oldenbourg.
- Pintelon, Rik und Johan Schoukens (2001). *System Identification A Frequency Domain Approach*. New York: IEEE Press.
- Rinne, Horst (2008). *Taschenbuch der Statistik*. 3. Frankfurt am Main: Harri Deutsch Verlag.
- Wooldridge, Jeffrey M. (2009). *Introductory Econometrics*. 5. Internationale Ausgabe. South-Western, Cengage Learning.
- Zurmühl, Rudolf und Sigurd Falk (1984). *Matrizen und ihre Anwendungen für Mathematiker, Physiker und Ingenieure Teil 1: Grundlagen*. 5. Berlin, etc.: Springer.