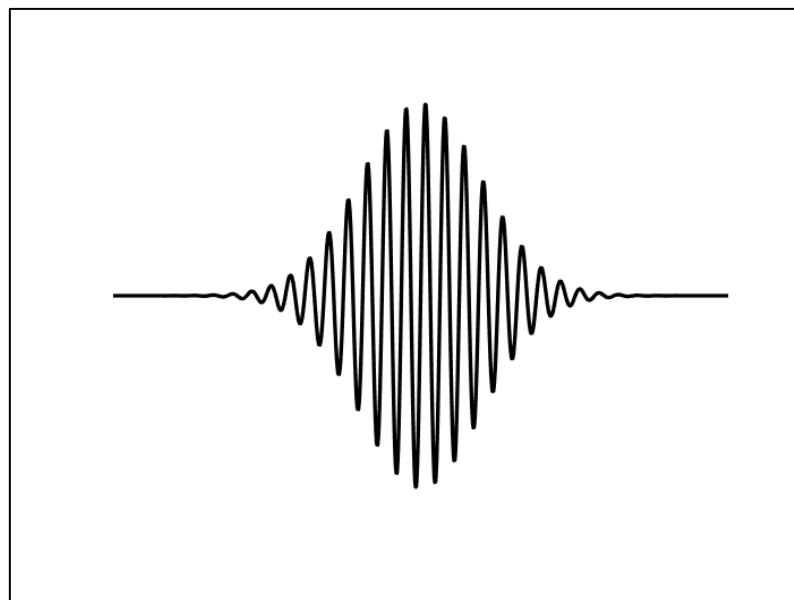


# Mathematische Methoden der Signal- und Bilddatenverarbeitung

## — Signalverarbeitung —



Prof. Dr.-Ing. habil. Dr. h. c. Uwe Hampel

Institut für Energietechnik

Fakultät Maschinenwesen

Technische Universität Dresden

2020

## Inhalt

<b>1 Einführung.....</b>	<b>3</b>
<b>2 Ziele der Vorlesung.....</b>	<b>4</b>
<b>3 Grundlagen.....</b>	<b>5</b>
3.1 Signale.....	5
3.2 Wichtige Grundsignalformen.....	8
3.3 Signal und Information .....	10
3.4 Signalverarbeitendes System .....	11
3.5 Technische Signaldigitalisierung.....	12
<b>4 Digitale Signalverarbeitung im Zeitbereich .....</b>	<b>15</b>
4.1 Interpolation: Polynome .....	15
4.2 Interpolation: Splines .....	17
4.3 Ausgleichsfunktionen .....	20
4.4 Numerisches Differenzieren.....	22
4.5 Numerisches Integrieren.....	23
4.6 Mittelungsverfahren: Arithmetische Mittelung .....	25
4.7 Mittelungsverfahren: Zentralwertbildung .....	27
<b>5 Statistische Signalanalyse .....</b>	<b>28</b>
5.1 Verteilungen und Kenngrößen .....	28
5.2 Wichtige Wahrscheinlichkeitsverteilungen .....	32
5.3 Auto- und Kreuzkorrelation .....	35
5.4 Lineare Regression.....	36
<b>6 Grundlagen der Signalverarbeitung im Frequenzbereich .....</b>	<b>39</b>
6.1 Fourier-Reihen.....	39
6.2 Fourier-Transformation .....	40
6.3 Signalleistung .....	44
6.4 Differenziale und Integrale im Frequenzbereich .....	46

6.5	Zeitverschobene Signale im Frequenzbereich.....	47
6.6	Faltung.....	48
6.7	Faltung im Frequenzraum.....	49
6.8	Korrelation und Faltung.....	50
6.9	Abtastung analoger Signale: Abtasttheorem und Aliasing .....	51
6.10	Abtastung von unendlichen Signalen.....	55
6.11	Fenster für die Begrenzung unendlicher Signale .....	60
<b>7</b>	<b>Analogsignalverarbeitung mit linearen zeitinvarianten Systemen .....</b>	<b>62</b>
7.1	Das Antwortverhalten linearer zeitinvarianter Systeme.....	62
7.2	Systemübertragungsfunktion und Systemantwortfunktion .....	64
7.3	Die Laplace-Transformation.....	65
7.4	Analoge Filter.....	68
7.5	Rauschen, Rauschleistung und die Wirkung von Filtern.....	72
7.6	Passive analoge Filter.....	75
7.7	Aktive analoge Filter .....	82
<b>8</b>	<b>Digitale Signalfilter.....</b>	<b>86</b>
8.1	Nichtrekursive digitale Filter.....	86
8.2	Die z-Transformation .....	88
8.3	Rekursive digitale Filter.....	90
8.4	Blockschaltbilder nichtrekursiver und rekursiver Filter .....	92

# 1 Einführung

Für die Automatisierung und Überwachung von Prozessen der Energie- und Verfahrenstechnik sind vielfältige Messinstrumente im Einsatz. Die betriebliche Instrumentierung besteht etwa aus Messstellen für Temperatur, Druck, Füllstand und Durchfluss. Des Weiteren kommen Sensoren zur Zustandsüberwachung oder für die stoffbezogene Analytik zum Einsatz.

Allen Messinstrumenten ist der unten gezeigte generelle Aufbau gemein (Abbildung 1.1). Ein Sensor wandelt eine zumeist nichtelektrische Prozessgröße (zum Beispiel die Temperatur oder den Druck) in eine elektrische Größe (Strom, Spannung, Ladung) um. Im analogen Teil der nachfolgenden Signalverarbeitung wird daraus ein Spannungssignal generiert. Dieses Spannungssignal wird durch Verstärker und Filterstufen in Amplitude und Frequenz an die Anforderungen der nachfolgenden digitalen Signalverarbeitung angepasst. Die digitale Signalverarbeitung beginnt mit der Digitalisierung des analogen Signals durch einen Analog-Digital-Wandler. Nachfolgend führt die digitale Signalverarbeitungsstufe komplexere Analysen des Signals durch. Sie ist zumeist durch Mikrocontroller, Mikroprozessoren oder programmierbare Logikschaltkreise (Field Programmable Gate Array, FPGA) realisiert. Schlussendlich werden die digitalen Daten auf Monitoren für die Prozessvisualisierung dargestellt, im Leitsystem als Betriebsdaten gespeichert bzw. der Prozesssteuerung und Regelung zugeführt.

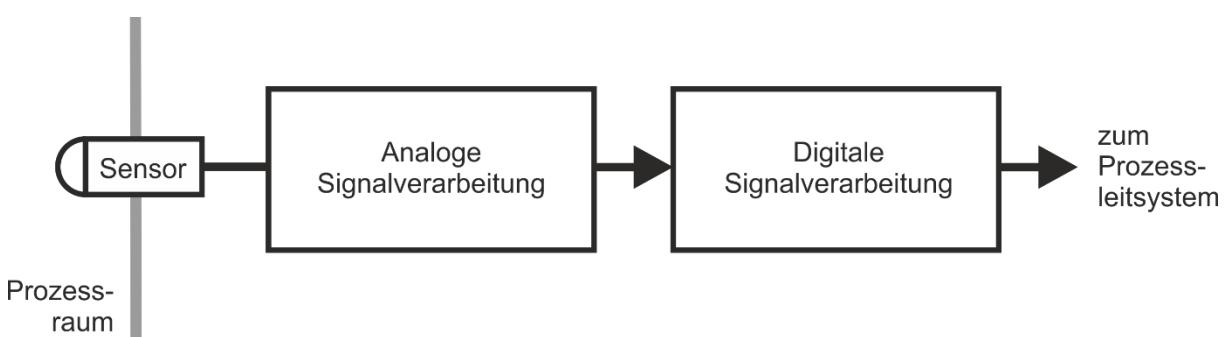


Abbildung 1.1: Sensor und signalverarbeitendes System

## 2 Ziele der Vorlesung

Die Vorlesung hat das Ziel, grundlegende und fortgeschrittene Methoden der analogen und digitalen Signalverarbeitung in der Energie- und Verfahrenstechnik für Maschinenbauingenieur:innen zu vermitteln. Die Vorlesung baut dabei auf den Grundlagen der Mess- und Automatisierungstechnik sowie der Mathematik des Grundstudiums auf.

## 3 Grundlagen

### 3.1 Signale

Als Signal bezeichnen wir allgemein den Verlauf einer informationstragenden Größe über der Zeit. Beispiele für Signale physikalischer Größen sind:

- die lokale Temperatur als Funktion der Zeit  $\vartheta(t)$
- der Absolutdruck an einem Druckwandler als Funktion der Zeit  $p(t)$
- die Schalldruckamplitude an einem Mikrofon  $\Delta p(t)$
- der Füllstand in einem Behälter als Funktion der Zeit  $h(t)$
- der von einem Coriolis-Durchflussmesser aufgezeichnete Massestrom  $\dot{m}(t)$
- ...

Da die Signalverarbeitung solche Signale mit einem einheitlichen mathematischen Konzept behandelt und zudem im analogelektronischen Teil des signalverarbeitenden Systems Signale immer als elektrische Spannung vorliegen, abstrahieren wir das analoge Signal im mathematischen Sinne zu einer eindimensionalen kontinuierlichen einheitenlosen Funktion  $f(t)$ .

**Beachte:** Wenngleich die Mathematik der Signalverarbeitung mit einheitenlosen Zeitfunktionen arbeitet, sollte nicht vergessen werden, dass die Interpretation der Signale in Bezug auf den Prozess immer der Berücksichtigung der physikalischen Zusammenhänge bedarf!

Hinsichtlich ihres Definitionsbereichs und des Wertebereichs unterscheiden wir folgende Signalarten (Abbildung 3.1):

**Analoges Signal:** Das Signal hat einen kontinuierlichen Wertebereich. Der Wertebereich kann dabei begrenzt sein und ist dies in der Praxis natürlich immer. Innerhalb der Grenzen ist aber jeder Wert zulässig.

**Digitales Signal:** Das Signal hat einen diskreten Wertebereich, es kann also nur eine endliche Anzahl an Werten annehmen. Ein Binärsignal hat den kleinsten Wertevorrat. Es kodiert die Information von 1 Bit. Ein Beispiel dafür ist das Ausgangssignal eines Füllstandbegrenzgebers. Es kann nur die Werte 0 (Füllstand nicht erreicht) oder 1 (Füllstand erreicht) annehmen. Allgemein finden wir digitale Signale in der Signalverarbeitungskette nach der Analog-Digital-Umsetzung.

**Zeitkontinuierliches Signal:** Zu jedem beliebigen Zeitpunkt liegt ein Signalwert vor. Das Signal ist über den gesamten Definitionsbereich, gegebenenfalls innerhalb bestimmter Grenzen, definiert.

**Zeitdiskretes Signal:** Der Signalwert liegt nur zu bestimmten abzählbaren Zeitpunkten vor. Man spricht auch von einem abgetasteten Signal. Zwischen den Abtastzeitpunkten ist der Signalwert nicht definiert. In der Praxis erfolgt die zeitliche Abtastung fast immer in äquidistanten Zeitschritten.

**Endliches und unendliches Signal:** Ist der Definitionsbereich (Zeitbereich) mindestens positiv unendlich, sprechen wir von einem unendlichen Signal, sonst von einem endlichen Signal. Diese Unterscheidung ist für eine formal-mathematische Behandlung wichtig. Praktisch sind Signale natürlich immer zeitbegrenzt, also endlich.

Aus den obigen Definitionen ergeben sich entsprechende Kombinationen.

**Zeitkontinuierliche analoge Signale:** Sie sind für die Analogsignalverarbeitung von Bedeutung. Wenn wir für diese Signalart von mathematischen Methoden sprechen, dann sind das in der Regel mathematische Modelle für die Analogsignalverarbeitung.

**Zeitdiskrete analoge Signale:** Sie werden im Rahmen der mathematischen Behandlung abgetasteter Analogsignale betrachtet und sind deshalb vorrangig für die theoretische Behandlung zeitdiskreter Signale von Bedeutung. Praktisch sind sie allerdings kaum von Bedeutung, da für die digitale Informationsverarbeitung immer auch die Signalamplitude diskretisiert werden muss.

**Zeitdiskrete digitale Signale:** Sie sind der digitalen Signalverarbeitung zugänglich und die mathematischen Methoden finden sich direkt in Softwarealgorithmen der Signalverarbeitung wieder.

**Zeitkontinuierliche digitale Signale:** Sie können mit digitaler Elektronik verarbeitet werden. Hier spielen vor allem Logikschaltungen eine Rolle. Sie sind nicht Gegenstand der Vorlesung.

Die Tatsache, dass insbesondere die digitale Signalverarbeitung nur endliche Signale verarbeiten kann, führt zu wichtigen Aussagen bei der Wahl von Fensterfunktionen für die Begrenzung des Zeitbereichs analoger Signale. Ebenso führt das Abtasten analoger Signale zu Phänomenen, die einer genaueren Betrachtung bedürfen. Beides wird in dieser Vorlesung diskutiert.

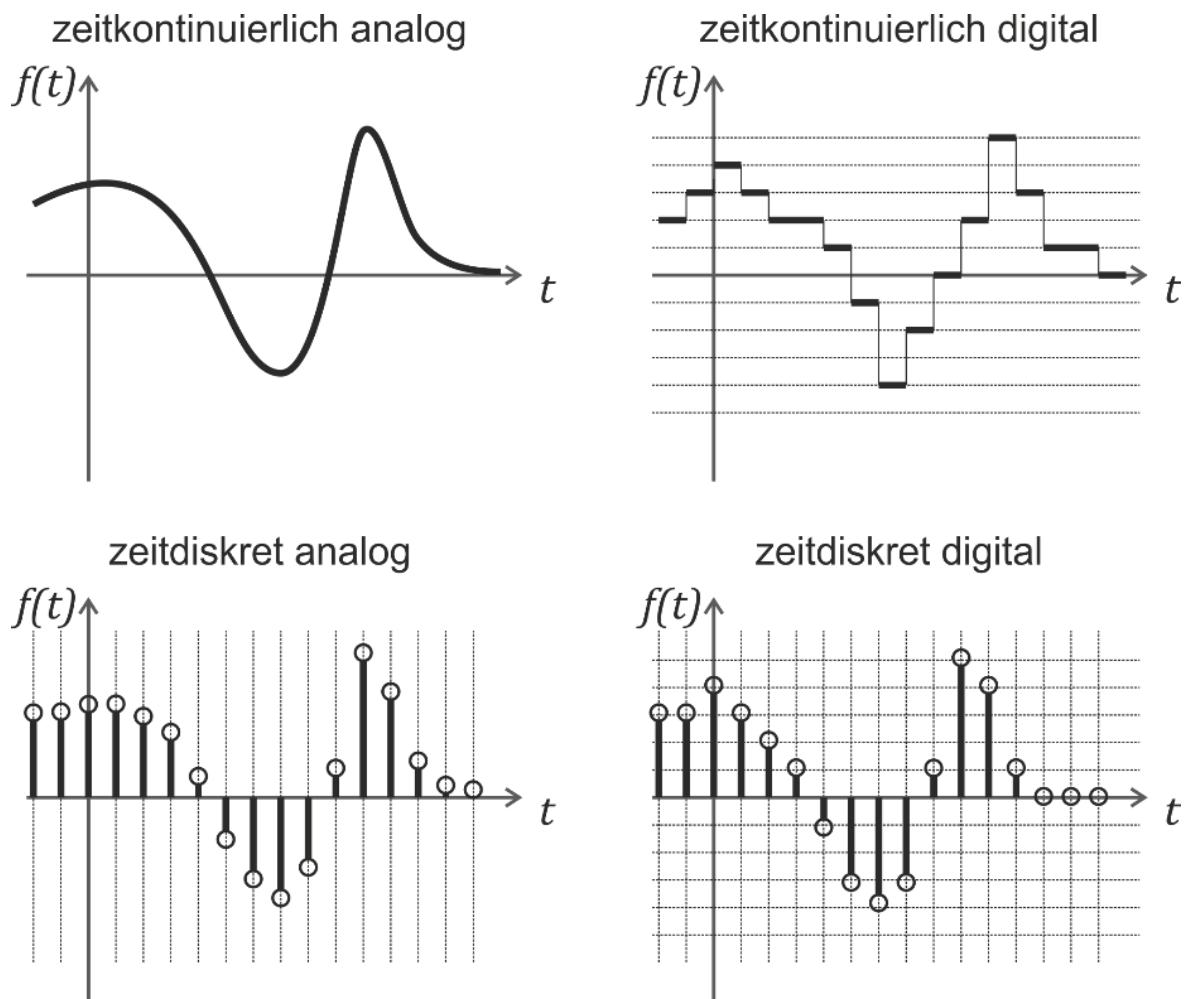


Abbildung 3.1: Signaltypen

### 3.2 Wichtige Grundsignalformen

Für die mathematische Behandlung der Signalverarbeitung gibt es einige grundlegende Signalformen, die im Folgenden kurz erläutert werden. Ihre prinzipiellen Kurvenformen sind in Abbildung 3.2 zusammenfassend dargestellt.

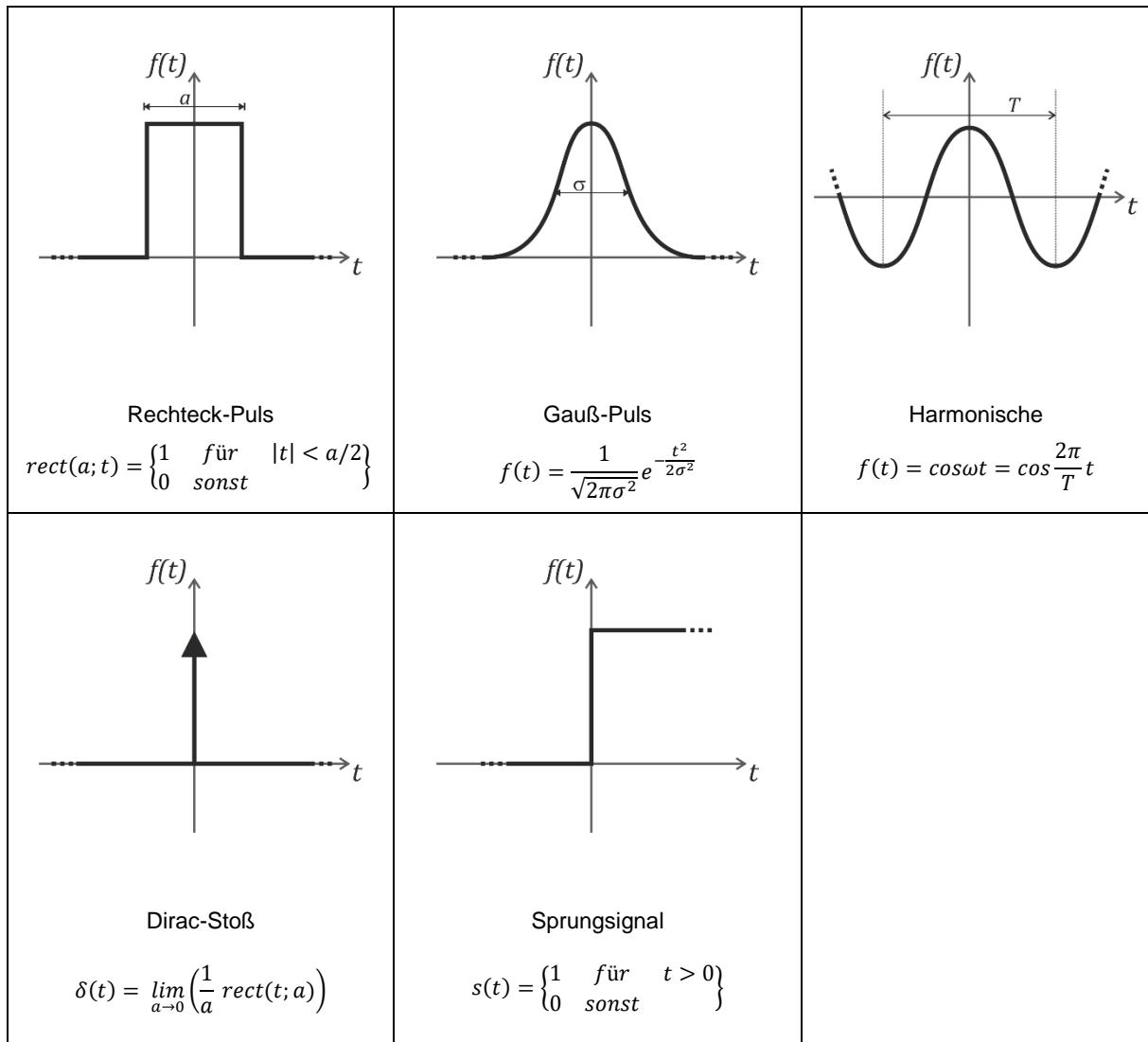


Abbildung 3.2: Wichtige Grundsignalformen.

**Rechteck-Puls** und **Gauß-Puls** sind pulsförmige Signale einer gegebenen zeitlichen Ausdehnung, für die  $\lim_{t \rightarrow \pm\infty} f(t) = 0$  gilt. **Harmonische** sind Kosinus- bzw. Sinusfunktionen. Die Grundsignalform ist das bezüglich  $t = 0$  symmetrische **Kosinus-Signal**. Dieses kann durch Phasenverschiebung in das Sinussignal überführt werden. Der **Dirac-Stoß** oder die sogenannte Dirac-Distribution  $\delta(t)$  stellt den Grenzfall eines Signals dar, welches einen nur zu einem einzigen Zeitpunkt wirksamen physikalischen Effekt mit normierter Wirkungsstärke repräsentiert. Es handelt sich um eine mathematische Abstraktion. Die Bezeichnung *Distribution* geht darauf zurück, dass es

sich hier um den Grenzfall (Limes) einer Funktionenserie handelt. Die Dirac-Distribution lässt sich als Grenzfall verschiedener Funktionenserien darstellen. Am bekanntesten ist der Grenzfall einer Rechteck-Pulses, dessen Breite unter Beibehaltung des Flächenintegrals von Eins gegen Null strebt (siehe Abbildung 3.2):

$$\delta(t) = \lim_{a \rightarrow 0} \left( \frac{1}{a} \operatorname{rect}(t; a) \right). \quad (3.1)$$

Alternativ lässt sich  $\delta(t)$  aber zum Beispiel auch aus der Normalverteilung ableiten:

$$\delta(t) = \lim_{\sigma \rightarrow 0} \left( \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{t^2}{2\sigma^2}} \right). \quad (3.2)$$

Durch bestimmte Integration der Dirac-Distribution erhält man das **Sprungsignal**:

$$s(t) = \int_{-\infty}^t \delta(t') dt' \quad (3.3)$$

**Beachte:** Im Folgenden werden wir die gestrichene Zeitvariable  $t'$  in Integralen verwenden, wenn zur Bildung eines Ausgangssignals (hier  $s(t)$ ) über ein Eingangssignal (hier  $\delta(t)$ ) integriert wird.

### 3.3 Signal und Information

Hat ein digitales Signal einen endlichen Wertevorrat, das heißt, es kann nur einen von  $N$  Werten aus  $\{x_i; i \in \mathbb{N}, 0 \leq i < N\}$  annehmen und ist die Auftrittswahrscheinlichkeit jedes Wertes  $x_i$  gleich groß, dann ist der Informationsgehalt eines einzelnen Signalwertes

$$I(x_i) = \log_2 N \quad (3.4)$$

Der Informationsgehalt wird in Bit angegeben. Anschaulich kann man sagen

- (1) Information ist beseitigte Ungewissheit
- (2) Der Informationsgehalt eines Signalwertes in Bit drückt aus, wie viele Fragen man stellen muss, um die Information herauszubekommen.

Beispiel zu (2): Ein Signal kann 8 diskrete Spannungswerte zwischen 0V und 5V annehmen. Es braucht genau 3 Fragen ( $\log_2 8 = 3$ ), um den Spannungswert sicher zu ermitteln:

Frage 1: Ist die Spannung größer oder kleiner 2,5 V? Antwort: kleiner

Frage 2: Ist die Spannung größer oder kleiner 1,25 V? Antwort: größer

Frage 3: Ist die Spannung größer oder kleiner 1,875 V?

Das oben gesagte gilt für einen Datenpunkt. Für ein zeitdiskretes Signal mit  $M$  unabhängigen Datenpunkten ist der Informationsgehalt dann entsprechend  $M \log_2 N$ .

**Bemerkung:** In der Tat sind die Dinge formal noch etwas komplizierter. Nach der von Claude Shannon begründeten Informationstheorie, nach welcher Information als beseitigte Ungewissheit definiert ist, geht auch die statistische Häufigkeit des Auftretens eines Wertes in die Betrachtung ein. Hier haben wir festgelegt, dass Werte eines Signals mit gleicher Wahrscheinlichkeit auftreten, oder anders gesagt, dass nichts a-priori über das Signal bekannt ist. Auftrittswahrscheinlichkeiten für Signalwerte sagen etwas über bereits vorhandenes Wissen aus, welches den Informationsgehalt (als beseitigte Ungewissheit) reduziert. Machen wir ein anschauliches Beispiel: Ein elektronisches Thermometer misst die Außentemperatur in einem Wertebereich von  $-50^{\circ}\text{C}$  bis  $+50^{\circ}\text{C}$  in Schritten von  $1\text{ K}$ . Intuitiv wissen wir, dass extreme Temperaturen eher unwahrscheinlich sind. Das heißt, Werte im Bereich von etwa  $> 40^{\circ}\text{C}$  tragen formal weniger Information als Werte um  $20^{\circ}\text{C}$ . Der interessierte Studierende sei hier auf weiterführende Literatur verwiesen.

### 3.4 Signalverarbeitendes System

Ein signalverarbeitendes System transformiert ein Eingangssignal in ein Ausgangssignal. Es sei mit  $f(t)$  das Eingangssignal und mit  $g(t)$  das Ausgangssignal bezeichnet. Die Transformation kann dann mathematisch allgemein als

$$g(t) = \mathcal{H}\{f(t)\} \quad (3.5)$$

notiert werden. Aus dem Signal  $f(t)$  produziert das signalverarbeitende System eine **Systemantwort**  $g(t)$ .

Von besonderer Bedeutung sind **lineare** Transformationen. Eine Transformation ist linear, wenn

$$\mathcal{H}\{af(t) + bg(t)\} = a \mathcal{H}\{f(t)\} + b \mathcal{H}\{g(t)\} \quad (3.6)$$

mit beliebigen Konstanten  $a$  und  $b$  gilt.

Weiterhin gilt für viele signalverarbeitende Systeme die Eigenschaft der **Zeitinvarianz**. Diese ist gegeben, wenn die Transformation des Eingangssignals in das Ausgangssignal zu jedem Zeitpunkt gleich ist, also die Systemantwort immer dieselbe ist und nicht von der Zeit abhängt:

$$\mathcal{H}\{f(t - t')\} = g(t - t'). \quad (3.7)$$

Ein signalverarbeitendes System, das (3.6) und (3.7) erfüllt, wird als Lineares Zeitinvariantes System (engl.: Linear Time-invariant System, LTI) bezeichnet.

Für die theoretische Behandlung linearer zeitinvarianter Systeme ist es erforderlich, deren Eigenschaften im Frequenzraum zu verstehen und analysieren zu können. Dies ist Gegenstand des zweiten Teils der Vorlesung (Kapitel 6).

### 3.5 Technische Signaldigitalisierung

Nachdem im Analogsignalverarbeitungsteil einer Messschaltung das Analogsignal im Wertebereich und Zeitbereich bzw. Frequenzbereich durch Verstärkung und Filterung geformt wurde, übernimmt ein Analog-Digital-Wandler (engl.: Analog-to-Digital Converter, ADC) die Digitalisierung des Signals. Dessen grundlegender Aufbau ist in Abbildung 3.3 gezeigt.

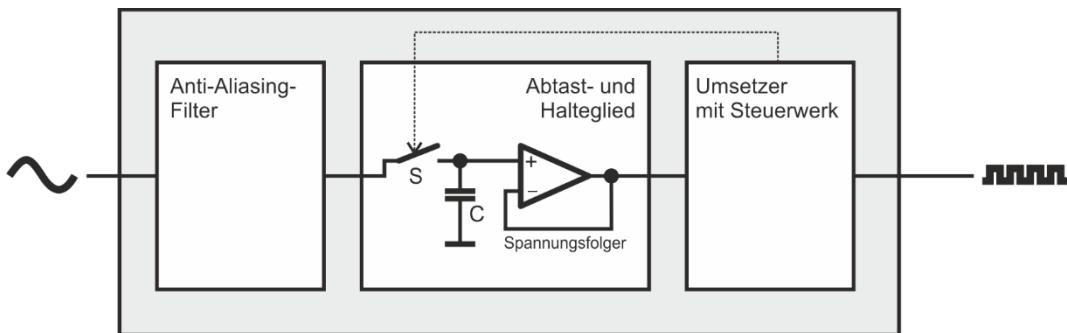


Abbildung 3.3: Grundlegender Aufbau eines Analog-Digital-Konverters.

Das analoge Eingangssignal wird zunächst mittels eines Anti-Aliasing-Filters bandbegrenzt. Prinzip und Notwendigkeit dieses Filters werden in Abschnitt 6.9 erklärt. Dem folgt ein sogenanntes Abtast- und Halteglied (engl.: Sample-and-Hold, S&H). Es sorgt dafür, dass während einer Wertedigitalisierung die Eingangsspannung am Umsetzer konstant bleibt. Durch Schließen eines Transistorschalters  $S$  wird der Haltekondensator  $C$  aufgeladen. Nach Öffnen des Transistorschalters bleibt die Ladung und damit auch die Spannung über dem Kondensator für den Zeitraum der Wandlung konstant. Der hochohmige Spannungsfolger trennt den Kondensator vom Umsetzer, so dass der Kondensator nicht durch Stromfluss in den Umsetzer entladen wird. Bei den Umsetzern kommen heute im Wesentlichen drei Funktionsprinzipien zur Anwendung. Diese werden nachfolgend kurz schematisch beschrieben.

**Integrierender Umsetzer:** Ein Sägezahngenerator produziert einen linear ansteigenden Spannungsverlauf  $U_Z(t)$  ② mit konstantem Anstieg  $dU_Z/dt$  zwischen der unteren und der oberen Spannungsgrenze, zum Beispiel  $0\text{ V}$  und  $U_{ref}$ . Diese Sägezahnspannung wird über einen Komparator gegen die gehaltene Eingangsspannung  $U_E$  ① verglichen. Der Komparator-Ausgang ist solange logisch 1, wie  $U_Z \leq U_E$  gilt. Die bis zum Umschalten des Komparators verstrichende Zeit ist damit ein Maß für die Größe von  $U_E$ . Diese Zeitspanne wird in diskreter Form durch

Zählen von Pulsen (3) eines synchron laufenden Oszillators erfasst, dessen Pulsausgang durch UND-Verschaltung mit dem Komparator-Ausgang maskiert ist (4).

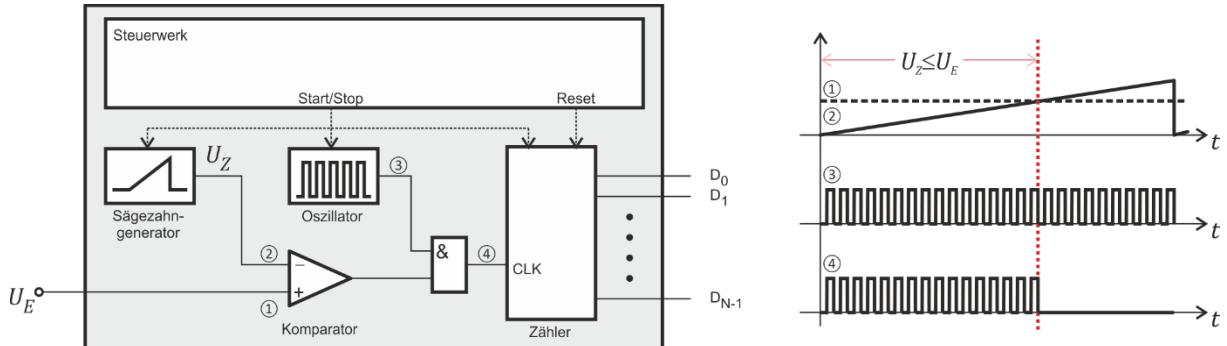


Abbildung 3.4: Integrierender Analog-Digital-Umsetzer.

**Umsetzer nach dem Prinzip der sukzessiven Approximation:** In diesem Umsetzer wird durch das Steuerwerk und einen Digital-Analog-Umsetzer eine Vergleichsspannung  $U_V$  (2) erzeugt und mit der Eingangsspannung  $U_E$  (1) verglichen. Die Vergleichsspannung wird direkt aus den digitalen Ausgängen des Steuerwerks abgeleitet. Diese werden initial gelöscht, also alle auf 0 gesetzt. Im ersten Schritt wird das Digit  $D_0 = 1$  gesetzt. Dadurch wird  $U_V = U_{ref}/2$ . Zeigt der Komparator für diese Vergleichsspannung  $U_V < U_E$  an, so wird das Digit wieder gelöscht ( $D_0 = 0$ ), anderenfalls beibehalten. Im zweiten Schritt wird das zweite Digit gesetzt ( $D_1 = 1$ ). Die Vergleichsspannung ist jetzt  $D_0 U_{ref}/2 + U_{ref}/4$ . Zeigt der Komparator für diese Vergleichsspannung  $U_V < U_E$  an, so wird das Digit wieder gelöscht ( $D_1 = 0$ ), anderenfalls beibehalten. Dieser Prozess wird bis zum letzten Digit fortgesetzt. Die Vergleichsspannung  $U_V$  nähert sich so  $U_E$  immer mehr an. Für einen  $n$ -Bit-Umsetzer wird das Ergebnis in  $n$  Schritten erreicht. Damit ist das Verfahren sehr schnell.

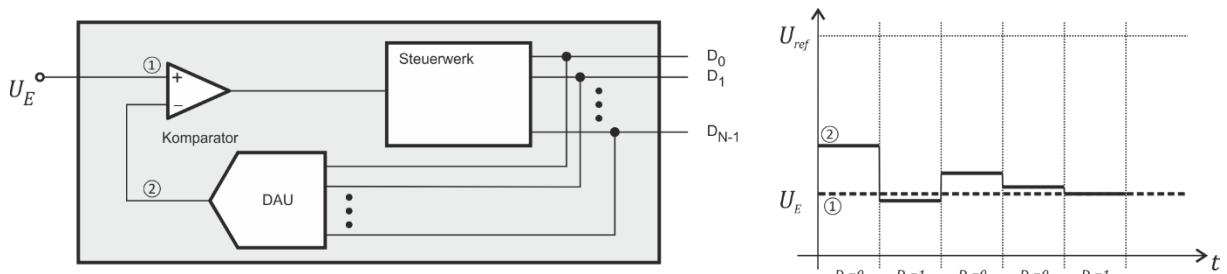


Abbildung 3.5: Analog-Digital-Umsetzer mit sukzessiver Approximation.

**Sigma-Delta-Wandler:** Herzstück dieses Wandlers ist ein Bitstrommodulator. Abbildung 3.6 zeigt das Wandlerschema als Prinzip, hier für einen einstufigen Sigma-Delta-Wandler. Die Eingangsspannung wird mit der Ausgangsspannung eines

einstufigen DA-Wandlers, im Prinzip ein einfacher Umschalter für positive und negative Referenzspannung, summiert, das Ergebnis integriert und im Komparator bezgl. des Vorzeichens ausgewertet. Unterstellen wir, das am negativen Summierer-Eingang und am Integratorausgang die negative Referenzspannung  $-U_{ref}$  anliegt. Damit erzeugt der Integrator einen Spannungsanstieg mit der Rate  $U_E + U_{ref}$ . Wird die Spannung am Integratorausgang positiv, so schaltet der Binär-DAC um und der Integrator integriert in Gegenrichtung mit der Rate  $U_E - U_{ref}$ . Wie Abbildung 3.6 rechts zeigt, weicht die Integratorausgangsspannung beim Umschalten des Komparators nur wenig von der Vergleichsspannung (0V) ab, während sie für die steilere Flanke etwas stärker von 0V abweicht. Daraus ergeben sich am Digitalfilter unterschiedlich lange Binärfolgen von Nullen oder Einsen, aus welchen die Eingangsspannung diskretisiert werden kann. Zudem mittelt das digitale Filter den Wert über viele Pulsperioden (Überabtastung). Ein Sigma-Delta-Wandler arbeitet mit mindestens 20-facher Überabtastung, um eine genügend genaue Mittelung zu erreichen. Der Integrator sorgt für eine sogenannte Rauschverschiebung (Noise Shaping) hin zu hochfrequenten Anteilen. Eine mehrstufige Ausführung der Summier- und Integrationsstufe kann das Rauschverhalten des Wandlers zusätzlich verbessern. Sigma-Delta-Wandler werden eingesetzt, wenn hohe Anforderungen an die Signalqualität und Bit-Auflösung bestehen. Dies ist zum Beispiel in digitalen Audiosystemen der Fall. Sie sind aufgrund der Überabtastung aber im Allgemeinen langsamer, als andere Analog-Digital-Wandler.

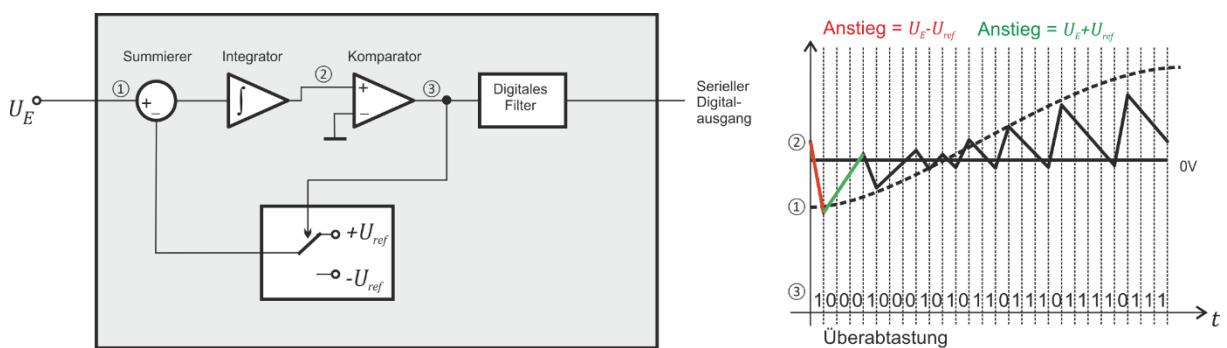


Abbildung 3.6: Sigma-Delta-Wandler.

## 4 Digitale Signalverarbeitung im Zeitbereich

Wir werden uns in diesem Kapitel mit den wichtigsten elementaren Bearbeitungsverfahren für digitale Signaldaten befassen. Es handelt sich dabei um Verfahren und Algorithmen zur Bestimmung charakteristischer Werte aus den Amplitudenverläufen in der Umgebung eines Datenpunktes.

### 4.1 Interpolation: Polynome

Die Interpolation von abgetasteten Signalen ist oft erforderlich, wenn die Abtastdichte für eine bestimmte Analyseaufgabe zu gering ist. Ein Beispiel ist die Integration von Signalen. Bei starken Amplitudenschwankungen ist die Annäherung der Integration durch eine einfache Summe möglicherweise zu ungenau. Deshalb ist es erforderlich, Datenpunkte durch nichtlineare Funktionen zu verbinden, die bestimmten Anforderungen, wie Stetigkeit und Glattheit genügen. Ein anderes Beispiel ist die Überbrückung von ungewollten Abtastlücken oder zeitlichen Bereichen mit offenkundigen Signaldegradationsen.

Eine naheliegende Vorgehensweise ist es,  $N$  Datenpunkte  $\{t_i, y_i(t_i)\}; 0 \leq i < N$  durch ein Polynom  $(N - 1)$ -ter Ordnung zu verbinden. Der Polynomansatz dazu lautet:

$$y(t) = a_0 + a_1 t + a_2 t^2 + \dots \quad (4.1)$$

Mit  $N$  Datenpunkten ergibt sich daraus das nach den Koeffizienten  $a_i$  zu lösende Gleichungssystem

$$\begin{bmatrix} 1 & t_0 & t_0^2 & \dots & t_0^{N-1} \\ 1 & t_1 & t_1^2 & \dots & t_1^{N-1} \\ 1 & t_2 & t_2^2 & \dots & t_2^{N-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & t_{N-1} & t_{N-1}^2 & \dots & t_{N-1}^{N-1} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_{N-1} \end{bmatrix} = \begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ \vdots \\ y_{N-1} \end{bmatrix}. \quad (4.2)$$

Wendet man diese Methode an, zeigt sich bei Polynomen höheren Grades ein oft wenig befriedigendes Ergebnis. Wie Abbildung 4.1 zeigt, tendieren Polynome zu starken Oszillationen zwischen den Stützstellen. Auch wenn alle Stützstellen getroffen werden, sind die Funktionsverläufe zwischen den Stützstellen eher „unphysikalisch“, da der hohe Grad des Polynoms starke Krümmungen begünstigt. Besser ist es, Polynome niedrigeren Grades stückweise und über wenige Stützstellen mit entsprechenden Glattheitsbedingungen an den Stoßstellen zu legen. Dies wird elegant

durch sogenannte Splines gelöst. Wir diskutieren dies am Beispiel kubischer Splines, welche sehr häufig für Interpolationszwecke genutzt werden.

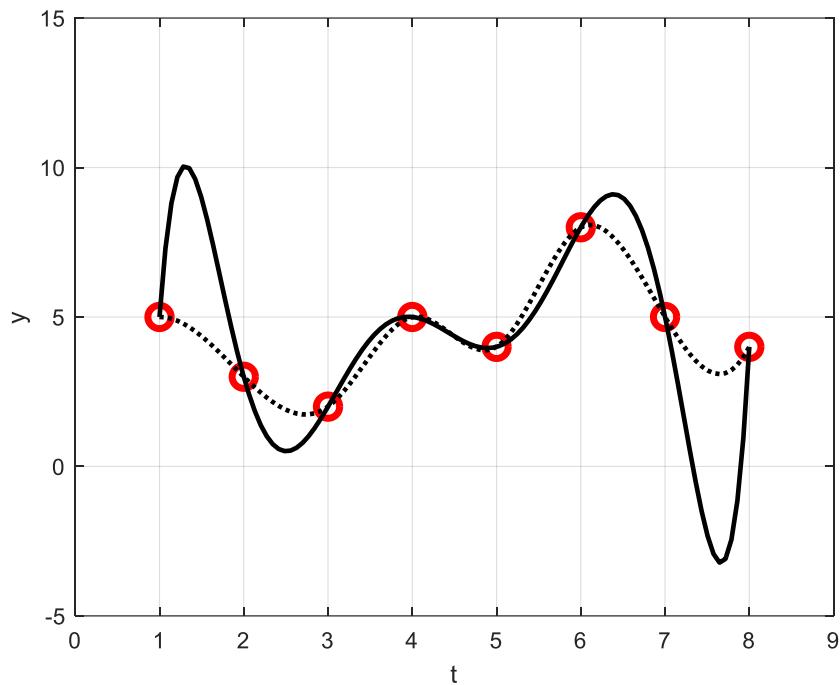


Abbildung 4.1: Dateninterpolation: Die acht roten Datenpunkte wurden hier einmal mit einem Polynom 7ter Ordnung (durchgezogene Linie) und zum anderen mit einem kubischen Spline (gestrichelte Linie) interpoliert. Klar erkennbar ist die Tendenz der Polynome, stark um die Datenpunkte zu oszillieren.

## 4.2 Interpolation: Splines

Ein kubischer Spline verbindet  $N$  Datenpunkte mit folgenden Bedingungen:

1. Jeweils zwei Datenpunkte werden durch ein Polynom dritten Grades verbunden.
2. An den Stoßstellen zu den Nachbarpunkten sind die 1. und die 2. Ableitung identisch.

Da die folgende Herleitung viele Rechenschritte umfasst, wird zur Verbesserung der Verständlichkeit vorab die Lösungsstrategie zusammenfassend erläutert.

Jeweils zwischen zwei Stützstellen soll ein Polynom der Form

$$s(t) = a + bt + ct^2 + dt^3 \quad (4.3)$$

gelegt werden. Im ersten Schritt schreiben wir die sechs Gleichungen für die nullte, erste und zweite Ableitung jeweils an der linken und rechten Stützstelle. Im zweiten Schritt nehmen wir davon die vier Gleichung, die keine ersten Ableitungen enthalten und stellen diese nach den gesuchten Polynomkoeffizienten um. Es entstehen vier Gleichungen mit unbekannten zweiten Ableitungen. Im dritten Schritt nutzen wir die Festlegung gleicher erster Ableitungen an den Stützstellen, um ein Gleichungssystem zu bilden, aus welchem die zweiten Ableitungen ermittelt werden können. Aus diesen können dann im finalen vierten Schritt die gesuchten Polynomkoeffizienten bestimmt werden. Als Konvention bezgl. der Indizierung legen wir fest:  $s_i(t)$  ist der  $i$ -te Spline zwischen den Stützstellen  $(t_i, y_i)$  und  $(t_{i+1}, y_{i+1})$ . Fernerhin sei  $h_i$  der Stützstellenabstand des  $i$ -ten Splines. Die Indizierung beginnt mit Null. Es gibt  $N$  Datenpunkte und  $N - 1$  Splines.

**Schritt 1:** Aufstellen der Gleichungen für die nullte (Datenwert  $y$ ), erste und zweite Ableitung an den beiden Stützstellen der  $i$ -ten Spline:

$$\text{Datenwert linke Stützstelle: } s_i(t_i) = y_i = a_i \quad (4.4)$$

$$\text{Datenwert rechte Stützstelle: } s_i(t_{i+1}) = y_{i+1} = a_i + b_i h_i + c_i h_i^2 + d_i h_i^3 \quad (4.5)$$

$$1. \text{ Ableitung linke Stützstelle: } s_i'(t_i) = y'_i = b_i \quad (4.6)$$

$$1. \text{ Ableitung rechte Stützstelle: } s_i'(t_{i+1}) = y'_{i+1} = b_i + 2c_i h_i + 3d_i h_i^2 \quad (4.7)$$

$$2. \text{ Ableitung linke Stützstelle: } s_i''(t_i) = y''_i = 2c_i \quad (4.8)$$

$$2. \text{ Ableitung rechte Stützstelle: } s_i''(t_{i+1}) = y''_{i+1} = 2c_i + 6d_i h_i \quad (4.9)$$

**Schritt 2:** Ableitung der Gleichungen für die Polynomkoeffizienten:

$$\text{aus (4.4):} \quad a_i = y_i \quad (4.10)$$

$$\text{aus (4.8):} \quad c_i = \frac{1}{2}y_i'' \quad (4.11)$$

$$\text{aus (4.8) und (4.9):} \quad d_i = \frac{1}{6h_i}(y_{i+1}'' - y_i'') \quad (4.12)$$

$$\text{aus (4.4), (4.5), (4.8) und (4.9):} \quad b_i = \frac{1}{h_i}(y_{i+1} - y_i) - \frac{h_i}{6}(y_{i+1}'' + 2y_i'') \quad (4.13)$$

**Schritt 3:** Aufstellen einer Beziehung zwischen den Stützstellenkoordinaten und den zweiten Ableitungen durch Gleichsetzen der ersten Ableitung an den Stützstellen. Hierzu nehmen wir zunächst (4.7) für die erste Ableitung an der rechten Stützstelle des Splines  $s_{i-1}$ . Das Einsetzen der Koeffizientengleichungen (4.11) - (4.13) in (4.7) ergibt:

$$s'_{i-1}(t_i) = \frac{1}{h_i}(y_i - y_{i-1}) + \frac{h_{i-1}}{6}(2y_i'' + y_{i-1}''). \quad (4.14)$$

Für die linke Stützstelle des nachfolgenden Splines gilt gemäß (4.6):

$$s'_i(t_i) = b_i \quad (4.15)$$

Gleichsetzen von (4.14) und (4.15) sowie etwas Umarrangieren ergibt:

$$\begin{aligned} h_{i-1}y_{i-1}'' + 2(h_{i-1} + h_i)y_i'' + h_iy_{i+1}'' = \\ -\frac{6}{h_{i-1}}(y_i - y_{i-1}) + \frac{6}{h_i}(y_{i+1} - y_i). \end{aligned} \quad (4.16)$$

Aus dieser Gleichung lässt sich ein Gleichungssystem formulieren, das die gesuchten zweiten Ableitungen zu den Stützstellenwerten in Beziehung setzt. Da es aber zur nullten Stützstelle keinen linken Nachbarn und Nten Stützstelle keinen rechten Nachbarn gibt, müssen dort die zweiten Ableitungen festgelegt werden. Dazu kommen folgende Ansätze in Frage:

1. Natürlicher Spline:  $y_0'' = 0$  und  $y_N'' = 0$ .
2. Zyklischer Spline  $y_0'' = y_N''$ .
3. Not-a-knot-Spline: Anstelle eines Splines wird durch die ersten und die letzten drei Stützstellen ein Polynom dritter Ordnung gelegt und die zweiten Ableitungen der Polynome an der dritten und der drittletzten Stützstelle für die sich anschließenden Splines verwendet.

Mit diesen Randbedingungen wird das Gleichungssystem für die Indizes  $1 \leq i \leq N-1$  aufgestellt. Es lautet:

$$\begin{bmatrix} 2(h_0 + h_1) & h_2 & 0 & \cdots & y_1'' \\ h_0 & 2(h_0 + h_1) & h_2 & \cdots & y_2'' \\ 0 & h_1 & 2(h_1 + h_2) & \cdots & y_3'' \\ 0 & 0 & h_2 & \cdots & y_4'' \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & h_{N-2} & 2(h_{N-2} + h_{N-1}) & y_{N-1}'' \end{bmatrix} \quad (4.17)$$

$$= \begin{bmatrix} -\frac{6}{h_0}(y_1 - y_0) + \frac{6}{h_1}(y_2 - y_1) \\ -\frac{6}{h_1}(y_2 - y_1) + \frac{6}{h_2}(y_3 - y_2) \\ -\frac{6}{h_2}(y_3 - y_2) + \frac{6}{h_3}(y_4 - y_3) \\ -\frac{6}{h_3}(y_4 - y_3) + \frac{6}{h_4}(y_5 - y_4) \\ \vdots \\ -\frac{6}{h_{N-2}}(y_{N-1} - y_{N-2}) + \frac{6}{h_{N-1}}(y_N - y_{N-1}) \end{bmatrix}$$

**Schritt 4:** Durch Lösen von (4.17) erhält man die zweiten Ableitungen an allen Stützstellen mit Index  $i = 1 \dots N-1$ . Einsetzen der  $y_i$  und  $y_i''$  in (4.10)-(4.13) ergibt die gesuchten Polynomkoeffizienten.

**Bemerkung:** Das Gleichungssystem (4.17) besitzt eine diagonaldominante Matrix und verhält sich damit im Hinblick auf die inverse Lösung „gutmütig“. Es lässt sich effizient durch ein Triangulationsverfahren oder iteratives Verfahren, wie Gauß-Seidel-Zerlegung oder das „Successive Over-Relaxation“-Verfahren, lösen.

### 4.3 Ausgleichsfunktionen

Ausgleichsfunktionen dienen der Datenreduktion. Häufig werden Signale mit sehr hoher Rate abgetastet, man ist aber innerhalb eines bestimmten Zeitraums eher an einem Trend des Datenverlaufs interessiert. So will man vielleicht beispielsweise aus Daten eines Wegsensors erfahren, wie die mittlere Geschwindigkeit oder Beschleunigung des überwachten Objektes innerhalb eines Zeitfensters ist. Ein anderes Beispiel sind modellbasierte Datenanalysen. Prozessmodelle sind häufig von geringer Ordnung. Das heißt, sie lassen sich durch Funktionen oder Differenzialgleichungen mit wenigen Parametern beschreiben. Hier ist es von Interesse, welche Modellparameter am besten zu den gemessenen Daten passen.

Die Approximationen von Signalen durch einfachere Funktionen verbinden wir mit Begriffen wie Ausgleichsrechnung, Regressionsanalyse, Fitting, u.a. In der Tat gibt es vielfältige Möglichkeiten des „Anfittens“ von niedrigparametrischen Kurven an Datenserien. Eine weitere Möglichkeit sind Fourier-Reihen (siehe Kapitel 6.1). Mit dieser lässt sich ein Signal durch eine gewichtete Summe orthonormaler harmonischer Basisfunktionen approximieren. Sie ist vor allem dann sinnvoll, wenn Signale auf periodischen Mustern oder Schwingungen aufbauen. Dies ist zum Beispiel in der Akustik der Fall. Ein anderes Beispiel ist das Ermitteln von Zeitkonstanten für exponentielle Abkling- und Wachstumssignale. Ein Beispiel ist die Dämpfung einer Schwingungsamplitude oder das Abklingen der Temperatur eines Körpers. Hier ist es sinnvoll, Exponentialfunktionen an die Daten zu fitten. Für Trendanalysen oder die oben genannte Frage nach Geschwindigkeit und Beschleunigung eines Wegsensors sind Polynomausgleiche interessant. Wir wollen diesen Fall daher als Beispiel für Ausgleichsanalysen heranziehen.

Gegeben sei ein abgetastetes Signal  $f_A(t)$ . Dieses besteht aus einer Menge von Signalwerten  $y_i$  zu den Abtastpunkten  $t_i$ . Gesucht wird ein Polynom  $M$ -ter Ordnung der Form  $p(t) = \sum_{m=0}^{M-1} a_m t^m$ , welches das Signal innerhalb der Zeitpunkte  $n = 1 \dots N-1$  bestmöglich approximiert. In der Regel wird die Forderung „bestmöglich“ durch Minimierung der Quadrate der Abweichungen spezifiziert. Dann ist

$$\sum_{i=0}^{N-1} [y_i - p(t_i)]^2 \xrightarrow{!} \text{Min.} \quad (4.18)$$

Zur Lösung setzt man die Polynomfunktion ein:

$$\sum_{i=0}^{N-1} [y_i - a_0 - a_1 t_i - \dots - a_{M-1} t_i^{M-1}]^2 \xrightarrow{!} \text{Min.} \quad (4.19)$$

Gesucht werden Koeffizienten  $a_m$  für die der Ausdruck in Gl. (4.19) minimal wird. D.h., die ersten Ableitungen nach den Variablen  $a_m$  müssen zu Null werden:

$$\begin{aligned} \frac{\partial}{\partial a_0} \sum_{i=0}^{N-1} [y_i - a_0 - a_1 t_i - \dots - a_{M-1} t_i^{M-1}]^2 &= 0 \\ &\vdots \\ \frac{\partial}{\partial a_{M-1}} \sum_{i=0}^{N-1} [y_i - a_0 - a_1 t_i - \dots - a_{M-1} t_i^{M-1}]^2 &= 0 \end{aligned} \quad (4.20)$$

Nach Bildung der partiellen Ableitungen wird

$$\begin{aligned} \sum_{i=0}^{N-1} 2[y_i - a_0 - a_1 t_i - \dots - a_{M-1} t_i^{M-1}] &= 0 \\ &\vdots \\ \sum_{i=0}^{N-1} 2t_i^{M-1}[y_i - a_0 - a_1 t_i - \dots - a_{M-1} t_i^{M-1}] &= 0 \end{aligned} \quad (4.21)$$

Der konstante Faktor 2 in den Gleichungen kann gestrichen werden. Daraus ergibt sich folgendes zu lösendes Gleichungssystem in Matrixschreibweise:

$$\left[ \begin{array}{cccc} N & \sum_{i=0}^{N-1} t_i & \dots & \sum_{i=0}^{N-1} t_i^{M-1} \\ \sum_{i=0}^{N-1} t_i & \sum_{i=0}^{N-1} t_i^2 & \dots & \sum_{i=0}^{N-1} t_i^M \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=0}^{N-1} t_i^{M-1} & \sum_{i=0}^{N-1} t_i^M & \dots & \sum_{i=0}^{N-1} t_i^{2(M-1)} \end{array} \right] \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_{M-1} \end{bmatrix} = \begin{bmatrix} \sum_{i=0}^{N-1} y_i \\ \sum_{i=0}^{N-1} t_i y_i \\ \vdots \\ \sum_{i=0}^{N-1} t_i^{M-1} y_i \end{bmatrix} \quad (4.22)$$

Um die gesuchten Koeffizienten  $a_0$  bis  $a_{n-1}$  zu erhalten, ist dieses Gleichungssystem zu lösen.

#### 4.4 Numerisches Differenzieren

Die einfachste Form des numerischen Differenzierens ist besteht in der Ersetzung des Differentialquotienten  $\frac{df}{dt}\Big|_{t_i}$  an der Abtaststelle  $t_i$  durch einen Differenzenquotienten:

$$\text{Linksseitiger Differenzenquotient: } \frac{df}{dt}\Big|_{t_i} \approx \frac{f(t_i) - f(t_{i-1})}{h} \quad (4.23)$$

$$\text{Rechtsseitiger Differenzenquotient: } \frac{df}{dt}\Big|_{t_i} \approx \frac{f(t_{i+1}) - f(t_i)}{h} \quad (4.24)$$

$$\text{Beidseitiger Differenzenquotient: } \frac{df}{dt}\Big|_{t_i} \approx \frac{f(t_{i+1}) - f(t_{i-1})}{2h} \quad (4.25)$$

Diese recht einfache Vorgehensweise führt natürlich zu größeren Fehlern bei verrauschten Daten. In der Regel erfolgt das Differenzieren durch Anlegen eines Ausgleichspolynoms bzw. eines Splines. Die Vorgehensweise für beide Fälle ist oben beschrieben. Die ersten Ableitungen ergeben sich für ein Ausgleichspolynom zu

$$\frac{df}{dt}\Big|_{t_i} \approx \frac{dp}{dt}\Big|_{t_i} = \sum_{m=0}^{M-2} m a_m t_i^m. \quad (4.26)$$

und für eine Spline-Interpolation zu

$$\frac{df}{dt}\Big|_{t_i} \approx \frac{ds_i}{dt}\Big|_0 = b_i. \quad (4.27)$$

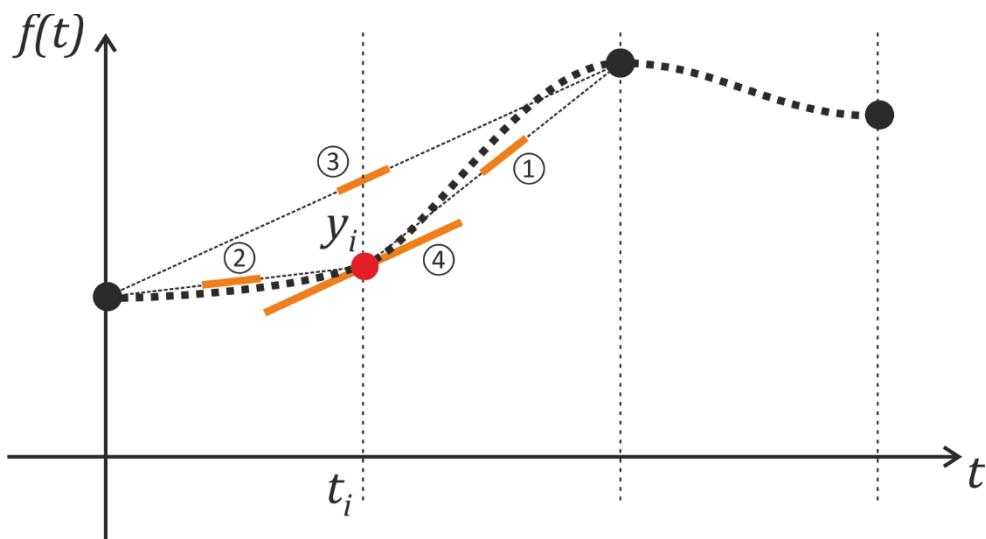


Abbildung 4.2: Numerisches Differenzieren mit Anstieg des rechtsseitigen ①, linksseitigen ② und beidseitigen ③ Differenzenquotienten bzw. Anstieg des angefitteten Splines ④.

## 4.5 Numerisches Integrieren

Beim Integrieren von Signalen ist der Flächeninhalt unter einem Signal zu bestimmen. Sind die Anforderungen an die Genauigkeit gering, können die Rechteckregel oder die Trapezregel angewendet werden (Abbildung 4.3). Für die Rechteckregel ergibt sich für das Integral zwischen zwei Datenpunkten

$$\int_{t_i}^{t_{i+1}} f(t)dt \approx y_i h \quad (4.28)$$

Daraus folgt für das Integral von  $t_0$  bis  $t_N$ :

$$\int_{t_0}^{t_N} f(t)dt \approx \sum_{i=0}^{N-1} y_i h \quad (4.29)$$

Für die Trapezregel gilt

$$\int_{t_i}^{t_{i+1}} f(t)dt \approx \frac{y_i + y_{i+1}}{2} h. \quad (4.30)$$

$$\int_{t_0}^{t_N} f(t)dt \approx h \left( \frac{1}{2}(y_0 + y_N) + \sum_{i=1}^{N-1} y_i \right) \quad (4.31)$$

Die Keplersche Fassregel ergibt sich, wenn durch drei aufeinanderfolgende Stützstellen ein Polynom zweiter Ordnung gelegt wird. Der Polynomansatz lautet

$$f(t_i) = y_i = a + bt_i + ct_i^2. \quad (4.32)$$

Mit

$$y_0 = a \quad (4.33)$$

$$y_1 = a + bh + ch^2 \quad (4.34)$$

$$y_2 = a + 2bh + 4ch^2 \quad (4.35)$$

ergeben sich die gesuchten Koeffizienten zu

$$a = y_0 \quad (4.36)$$

$$b = -\frac{3}{2h} \left( y_0 - \frac{4}{3}y_1 + \frac{1}{3}y_2 \right) \quad (4.37)$$

$$c = -\frac{1}{2h^2} (-y_0 + 2y_1 - y_2) \quad (4.38)$$

Das Integral über zwei Abtastpunkte ist

$$\int_{t_i}^{t_{i+2}} f(t) dt = \int_{t_i}^{t_{i+2}} (a + bt + ct^2) dt = 2ah + 2bh^2 + \frac{8}{3}ch^3 \quad (4.39)$$

Einsetzen der Koeffizienten ergibt die Keplersche Fassregel

$$\int_{t_i}^{t_{i+2}} f(t) dt = \frac{1}{3}h(y_i + 4y_{i+1} + y_{i+2}) \quad (4.40)$$

Für eine ungerade Zahl an Abtastpunkten (mindestens 3) ergibt sich mit  $M = (N-1)/2$  durch Summierung die sogenannte Simpson-Regel:

$$\int_{t_0}^{t_N} f(t) dt = \frac{1}{3}h(y_0 + 4y_1 + 2y_2 + 4y_3 + 2y_4 + \dots + y_N). \quad (4.41)$$

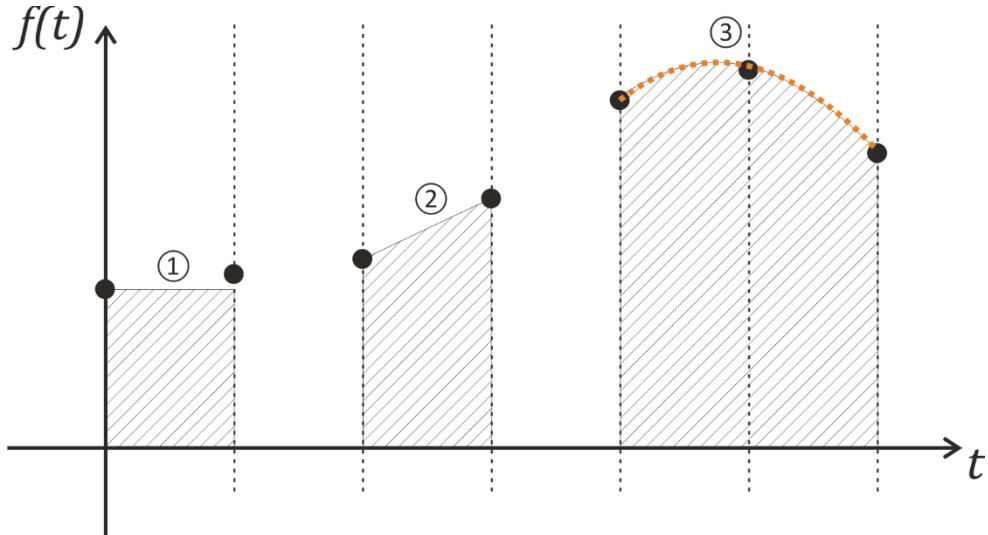


Abbildung 4.3: Numerisches Integrieren nach Rechteckregel ①, Trapezregel ② und Keplerscher Fassregel ③.

## 4.6 Mittelungsverfahren: Arithmetische Mittelung

Eine häufige Aufgabe der numerischen Datenverarbeitung besteht in der Mittelung von Messwerten um einen zentralen Messwert. Dies kann durch ungewichtete oder gewichtete Integration erfolgen. Weiterhin ist es möglich, die Zentralwertbildung (Median-Mittelung) anzuwenden. Bei der ungewichteten Integration wird ein analoges Signal formal über einen Bereich  $t - \Delta t \dots t + \Delta t$  integriert:

$$m = \frac{1}{2\Delta t} \int_{t-\Delta t}^{t+\Delta t} f(t) dt \quad (4.42)$$

Bei abgetasteten Signalen erfolgt die Mittelwertbildung dann durch eines der obigen Integrationsverfahren

$$m = \frac{1}{(2N+1)h} I_{t_{i-N}}^{t_{i+N}} \quad (4.43)$$

wobei  $I$  das mit einer der obigen Methoden bestimmte numerische Integral bezeichnet.

Bei gewichteter Mittelwertbildung wird den weiter vom Zentralwert entfernten Messwerten sukzessive weniger Gewicht in der Mittelwertbildung zugewiesen. Damit haben Datenpunkte nahe des lokalen Wertes mehr Einfluss auf den Mittelwert, als entfernte Datenpunkte. Numerisch führt dies zu einer gewichteten Summation

$$m_j = \sum_{i=-N}^N a_{|i|} y_{j+i} \quad (4.44)$$

Die Gewichte sind in der Regel symmetrisch und fernerhin so zu wählen, dass  $\sum a_i = 1$  gilt. Für analoge Signale entspricht die gewichtete Mittelwertbildung einer Faltung mit einer Mittelungsfunktion  $h(t)$  entsprechend

$$m(t) = \int_{-\infty}^{+\infty} h(t-t') f(t') dt \quad (4.45)$$

wobei

$$\int_{-\infty}^{+\infty} m(t) dt = 1. \quad (4.46)$$

Eine zu beachtende Besonderheit der numerischen Faltung und damit auch der gewichteten Mittelwertbildung ist die, dass für eine Filterfunktion (Gewichtevektor) der Länge  $2L + 1$  die Operation formal nur zwischen den Abtastpunkten  $L$  und  $N - L$  möglich ist, da für alle anderen Werte jeweils links- oder rechtsseitige Messwerte für

die Mittelwertbildung fehlen. Entweder nimmt man in Kauf, dass die Ergebnissequenz sich um die Länge  $2L + 1$  reduziert, oder man muss die „fehlenden“ Messwerte durch Annahmen ersetzen. Möglichkeiten sind:

1. Annahme von Messwerten  $y_i = 0$ . Diese Maßnahme wird als Zero-Padding (Auffüllen mit Nullen) bezeichnet. Je nach aktuellen Endwerten  $y_0$  bzw.  $y_N$  ergibt sich daraus allerdings ein oft unphysikalischer ansteigender bzw. abfallender Endwerteverlauf in der neuen Sequenz.
2. Extrapolation an den Enden der Ausgangssequenz. Die fehlenden Messwerte können durch Extrapolation der ersten bzw. letzten Messwerte der Ausgangssequenz berechnet werden. Am einfachsten ist eine konstante Interpolation bei der die Endwerte  $y_0$  bzw.  $y_N$  auch über das Ende der Sequenz hinaus weiterverwendet werden. Alternativ kann linear oder mit Polynomen höherer Ordnung extrapoliert werden.

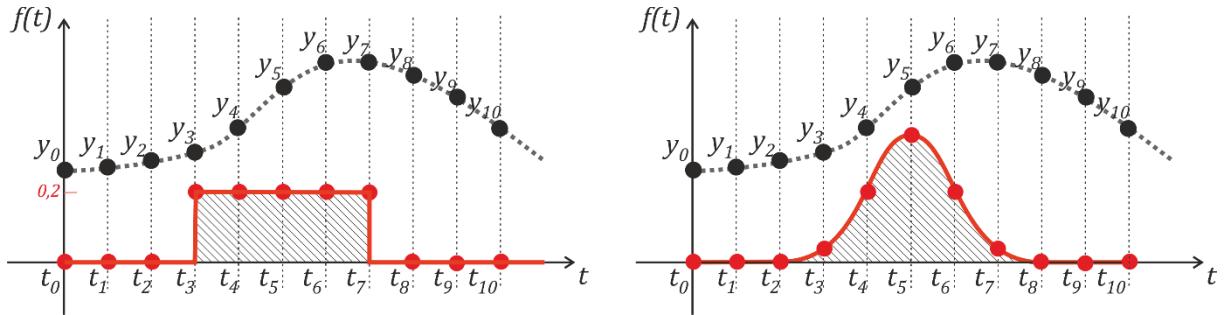


Abbildung 4.4: Ungewichtete (links) und gewichtete Mittelung. Rot ist die Gewichtsverteilung dargestellt. Die Mittelung entspricht formal einer Faltung des Signals mit der roten Funktion.

## 4.7 Mittelungsverfahren: Zentralwertbildung

Bei der Zentralwertbildung oder Median-Mittelung wird ein Rangordnungsverfahren angewendet. Soll der Medianwert für die Datenpunkte  $(t_0, y_0) \dots (t_{N-1}, y_{N-1})$  bestimmt werden, so werden die Werte  $y_0 \dots y_{N-1}$  zunächst der Größe nach sortiert. Der Medianwert ist der zentrale Wert der so sortierten Folge, also  $y_{(N-1)/2}$  für eine ungerade Zahl von Datenpunkten bzw.  $0,5 \cdot (y_{N/2} + y_{N/2+1})$  für eine gerade Zahl von Datenpunkten. Der Zentralwert ist unempfindlich gegen Extremwerte. Dies kann vor allem dann hilfreich sein, wenn die Signalfolge physikalisch unplausible Ausreißer enthält. Der Medianwert vermeidet zu starke Glättung der Signale und eliminiert effektiv den Effekt extremer Datenpunkte auf die Mittelung.

## 5 Statistische Signalanalyse

Durch die statistische Signalanalyse werden Verteilungen von Werten oder Merkmalen in einem Signal untersucht oder statistische Beziehungen zwischen verschiedenen Signalen ermittelt. Verteilungen können dabei gute Einblicke in Zusammenhänge geben, die aus dem Signalverlauf selbst nicht ohne Weiteres ablesbar sind. Die Zusammenhänge können sich dabei im Zeitbereich als auch im Amplitudenzbereich auftun. Oft ist eine statistische Analyse erst möglich oder sinnvoll, wenn bestimmte Merkmale aus Signaldaten extrahiert wurden, da die Analyse für diese Merkmale durchgeführt wird. Statistische Signalanalyse ist eng mit statistischen Modellen und Wahrscheinlichkeitsverteilungen verbunden.

### 5.1 Verteilungen und Kenngrößen

Eine Häufigkeitsverteilung  $H(x_i)$  gibt an, wie oft ein Merkmal  $x_i$  innerhalb einer Stichprobe mit Umfang  $N$  vorkommt. Dazu muss der Merkmalsraum eine abzählbar endliche oder unendliche Menge sein. Zum Beispiel könnten wir uns fragen, wie oft ein Signalwert bestimmter Amplitude innerhalb eines endlichen zeitdiskreten Signals auftritt. Das ist nur für ein amplitudendiskretes Signal sinnvoll, also ist eine Amplitudenklassifizierung vorzunehmen. Dazu wird die Signalamplitude in verschiedene Klassen oder Intervalle eingeteilt. Dann gibt  $H(x_i)$  die absolute Häufigkeit dafür an, dass ein Amplitudenwert  $x_i = x(t_i)$  mit  $i \in \{1, \dots, N\}$  im Intervall  $X_j$  mit  $j \in \{1, \dots, M\}$  ( $M$  ist die Klassenzahl) liegt. Gebräuchlich ist eine lineare Klasseneinteilung in äquidistante Intervalle. Für digitalisierte Signale liegt eine Klasseneinteilung im Prinzip bereits bereit vor.

Dividiert man die Häufigkeiten pro Klassenintervall durch den Umfang der Stichprobe, so erhält man die relative Häufigkeitsverteilung  $h(x_i) = H(x_i)/N$ . Summiert man die Häufigkeit über  $H(x_i)$  vom kleinsten Wert bis zu einem Wert  $k$  im Merkmalsraum, so erhält man die kumulative Häufigkeit

$$H_c(x_k) = \sum_{n=1}^k H(x_i) \quad (5.1)$$

bzw. die relative kumulative Häufigkeit

$$h_c(x_k) = \sum_{n=1}^k h(x_i). \quad (5.2)$$

Im Sinne der Wahrscheinlichkeitsrechnung ist die erfasste Merkmalsgesamtheit eine Stichprobe einer Grundgesamtheit. Der Grundgesamtheit liegt eine Wahrscheinlichkeitsverteilung zugrunde. Es sei  $x_i$  ein durch Messung gewonnener Merkmalswert. Die Wahrscheinlichkeitsdichtefunktion  $w(x)$  beschreibt die Wahrscheinlichkeit, dass dieser Wert im Intervall  $x + dx$  liegt. Es gilt

$$\int_{-\infty}^{+\infty} w(x)dx = 1. \quad (5.3)$$

Die Verteilungsfunktion  $W(a)$  gibt die Wahrscheinlichkeit an, dass dieser Wert kleiner als  $a$  ist, also

$$W(x_i < a) = \int_{-\infty}^a w(x)dx. \quad (5.4)$$

**Gesetz der großen Zahlen:** Die relative Häufigkeitsverteilung geht im Grenzfall einer unendlich großen Stichprobe ( $N \rightarrow \infty$ ) in die Wahrscheinlichkeitsdichtefunktion über. Die relative kumulative Häufigkeitsverteilung in die Verteilungsfunktion:

$$\lim_{N \rightarrow \infty} h(x_i) = w(x)$$

$$\lim_{N \rightarrow \infty} H(x_i) = W(x)$$

Die übliche Darstellungsform der Häufigkeitsverteilung ist das Histogramm (Abbildung 5.1). In ihm sind auf der Abszisse die Merkmalsklassen durch geschlossene Intervalle festgelegt. Auf der Ordinate sind die absoluten oder relativen Häufigkeiten des Vorkommens des Merkmals abgetragen. Im Beispiel Abbildung 5.1 sehen wir ein zeitdiskretes Signal. Es handelt sich um eine konstante Spannung  $U = 2 V$  die von Rauschen überlagert ist. Die Histogramme rechts zeigen die relative Amplitudenhäufigkeit sowie die relative kumulative Amplitudenhäufigkeit, eingeteilt in Klassen mit Klassenbreite 50 mV.

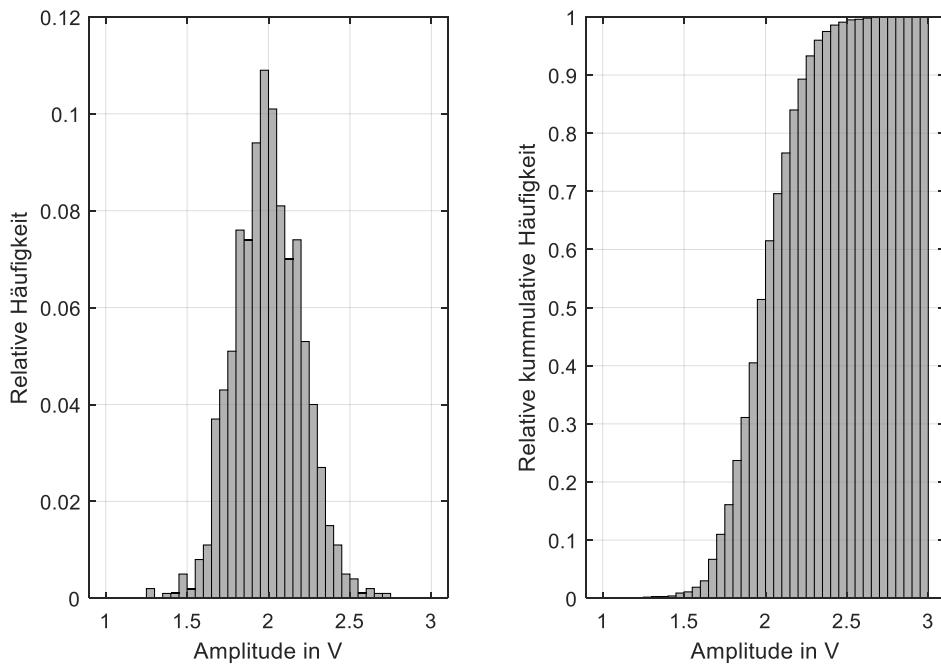
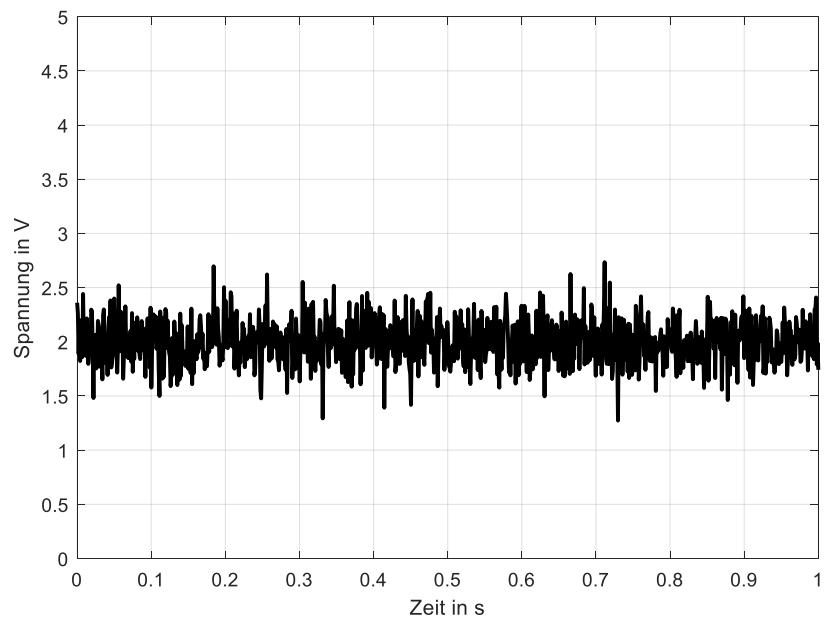


Abbildung 5.1: Zeitreihe und Histogramm eines zeitdiskreten Signals.

Verteilungen werden durch die in Tabelle 5.1 aufgezählten charakteristischen Kenngrößen beschrieben. Wir beschränken uns dabei auf die charakteristischen Größen von Häufigkeitsverteilungen. In Spalte 1 sind die gebräuchlichsten Namen der Größen angegeben. Der Ausdruck in Klammern bezieht sich auf den korrespondierenden Parameter der Wahrscheinlichkeitsverteilung.

Tabelle 5.1: Wichtige Kenngrößen von Häufigkeitsverteilungen.

Stichprobenmittelwert, Arithmetischer Mittelwert, Empirischer Mittelwert (Erwartungswert)	$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i$
Stichprobenvarianz, Empirische Varianz (Varianz)	$s^2 = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2$
Stichprobenstandardabweichung, Empirische Standardabweichung (Standardabweichung)	$s = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2}$
Variationskoeffizient	$v = \frac{s}{\bar{x}}$
Median	$x_{med} = \begin{cases} \hat{x}_{(N+1)/2} & \text{für } N \text{ ungerade} \\ \frac{1}{2}(\hat{x}_{N/2} + \hat{x}_{N/2+1}) & \text{für } N \text{ gerade} \end{cases}$ ist der zentrale Wert der sortierten Folge $\hat{x}_i$
Modalwert	$x_{mod} = argmax(H(x_i))$ ist der am häufigsten vorkommende Wert
Spanne	$R = x_{max} - x_{min}$ ist der Abstand zwischen kleinstem und größtem beobachteten Wert
Quantil	$x_p = \begin{cases} \frac{1}{2}(\hat{x}_{Np} + \hat{x}_{Np+1}) & Np \text{ ganzzahlig} \\ \hat{x}_{[Np]+1} & Np \text{ nicht ganzzahlig} \end{cases}$ ist der Wert $x_p$ des $p$ -Quantils ( $0 < p \leq 1$ ), für welchen $(100 * p)$ Prozent aller Werte kleiner als $x_p$ sind.

## 5.2 Wichtige Wahrscheinlichkeitsverteilungen

Das Messen von Prozessgrößen führt zunächst auf Signale, welche zu verarbeiten sind. Aus den Signalen können Merkmale durch Signalverarbeitung extrahiert werden. Die Signalamplitude ist ein Merkmal, für welches keine Signalverarbeitung erforderlich ist. Komplexere Merkmale, welche über Signalverarbeitung extrahiert werden müssen, sind etwa der momentane arithmetische Mittelwert, momentane Ableitungen (Anstieg, Krümmung), Abstände zwischen Ereignissen, etwa an einem Detektor erfasste Strahlungsimpulse, oder die Stärke spektraler Muster. Die Häufigkeit, mit der solche Merkmale gefunden werden, unterliegt einer durch die Spezifika des beobachteten Prozesses zugrundliegenden Wahrscheinlichkeitsverteilung. Die Aufgaben der statistischen Signalverarbeitung kann es einerseits sein, auf Basis eines bekannten Prozessmodells charakteristische Parameter anhand statistischer Merkmale oder aber allgemeine statistische Größen zur Identifikation des Prozessmodells zu ermitteln. Im Folgenden werden daher wichtige Wahrscheinlichkeitsverteilungen diskutiert. Sie sind des Weiteren in Tabelle 5.2 übersichtlich zusammengefasst.

**Gleichverteilung:** Die Wahrscheinlichkeit des Auftretens eines beliebigen Wertes im Signal ist gleichgroß. Es gilt:  $w(x) = \text{const}$  innerhalb eines Intervalls  $[x_{\min}, x_{\max}]$  und  $\bar{x} = \frac{1}{2}(x_{\min} + x_{\max})$ .

**Binomialverteilung:** Sie beschreibt die Wahrscheinlichkeit, dass ein binärer Wert bei  $N$ -fach Messung genau  $M$ -fach auftritt, wenn dessen Auftrittswahrscheinlichkeit durch  $p$  gegeben ist.

**Beispiel:** In einem Rohr strömen zwei nicht-mischbare Flüssigkeiten. Flüssigkeit A hat einen Volumenanteil von 10%, Flüssigkeit B einen Volumenanteil von 90%. Unterstellen wir eine ausreichende Homogenität der Verteilung der Phasen im Rohrquerschnitt, so ist die Wahrscheinlichkeit, zu einem bestimmten Zeitpunkt Flüssigkeit A an einer lokalen Messsonde im Rohr nachzuweisen  $p_A = 0.1$ , die Wahrscheinlichkeit, zu einem bestimmten Zeitpunkt Flüssigkeit B an einer lokalen Messsonde im Rohr nachzuweisen  $p_B = (1 - p_A) = 0.9$ . Für  $N$  Abtastpunkte ergibt sich die Wahrscheinlichkeit, Flüssigkeit A  $M$ -fach anzutreffen aus der Binomialverteilung zu

$$w(M) = \binom{N}{M} p_A^M (1 - p_A)^{N-M}$$

$$\binom{N}{M} = \frac{N!}{M!(N-M)!}$$

**Exponentialverteilung:** Sie beschreibt insbesondere die Wahrscheinlichkeit des Eintretens eines Übergangsereignisses von einem Zustand A in eine Zustand B, wenn die Eintrittswahrscheinlichkeit des Übergangs proportional zur Häufigkeit des Zustands A ist.

**Beispiel:** In einem Gefäß wird eine Substanz aufgeschäumt. Nach dem Aufschäumen zerfällt der Schaum allmählich. Dabei kommt es zum Platzen von Blasen im Schaum. Ein akustischer Sensor zeichnet das Geräusch des Platzens auf. Der Zustand einer Einzelblase ist entweder A: „Blase stabil“, oder B: „Blase geplatzt“. Die Anzahl  $N$  der pro Zeiteinheit zerfallenden Blasen ist proportional zur Gesamtzahl vorhandener Blasen. Es gilt

$$\frac{dN}{dt} = -\lambda N.$$

Mit der Anfangsbedingung  $N(t = 0) = N_0$  ( $N_0$  ist die initiale Blasenzahl) ergibt sich als Lösung der Differenzialgleichung

$$N(t) = N_0 e^{-\lambda t}.$$

Die (kumulative) Wahrscheinlichkeit, dass eine einzelne Blase bis zum Zeitpunkt  $t$  nicht geplatzt ist, ergibt sich aus dem Verhältnis existierender Blasen (Zustand A) zum Zeitpunkt  $t$  zur initialen Gesamtblasenzahl, also

$$W_A(t) = e^{-\lambda t}.$$

Die (kumulative) Wahrscheinlichkeit, dass eine einzelne Blase bis zum Zeitpunkt  $t$  platzt, ist

$$W_B(t) = 1 - e^{-\lambda t}.$$

Die Wahrscheinlichkeitsdichtefunktion

$$w_B(t) = \frac{dW_B(t)}{dt} = \lambda e^{-\lambda t}$$

beschreibt die Wahrscheinlichkeit, mit der eine Blase im Intervall  $t \dots t + dt$  platzt.

**Poisson-Verteilung:** Sie beschreibt die Häufigkeit von zufälligen Ereignissen innerhalb eines Zeitfensters, in dessen Dauer im Mittel  $\alpha$  Ereignisse eintreten.

**Beispiel:** Wir betrachten wieder den Blasenzerfall oben. Zu einem bestimmten Zeitpunkt ergibt die Verteilung den durchschnittlichen Zerfall von  $\alpha = 5,7$  Blasen pro Sekunde. Die Wahrscheinlichkeit, dass in einer Sekunde  $n$  Blasenzerfälle detektiert werden, ist durch die Poisson-Verteilung

$$w(n) = \frac{\alpha^n}{n!} e^{-\alpha}$$

gegeben. Für 10 Blasenzerfälle pro Sekunde ist die Wahrscheinlichkeit  $w(10) = 0.0334$ . Die Wahrscheinlichkeit, dass in einer Sekunde kein Blasenzerfall detektiert wird, ist  $w(0) = 0.0033$ , also etwa 10-mal geringer.

**Gaußsche Normalverteilung:** Normalverteilt sind Parameter immer dann, wenn sie durch vielen unabhängige Größen additiv beeinflusst werden. Dies trifft für viele stochastische Prozesse zu.

**Beispiel:** Thermisches Rauschen in einem Widerstand erzeugt ein normalverteiltes Amplitudenspektrum der Form

$$p(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{u}{u_{eff}}\right)^2},$$

mit dem Effektivwert der Rauschspannung  $u_{eff}$ .

**Logarithmische Normalverteilung:** Streuen Messwerte über Größenordnungen und können nur positive Werte vorkommen, dann ist es oft sinnvoll, den Parameter zu logarithmieren. Eine Zufallsgröße heißt dann logarithmisch normalverteilt, wenn die  $\ln x_i$  normalverteilt sind. Normalverteilt sind Parameter immer dann, wenn sie durch vielen unabhängige Größen multiplikativ beeinflusst werden.

Tabelle 5.2: Wichtige Wahrscheinlichkeitsverteilungen und deren Kenngrößen.

	Wahrscheinlichkeitsdichte	Erwartungswert	Standardabweichung
Binomialverteilung	$w(M) = \binom{N}{M} p^M (1-p)^{N-M}$	$\bar{x} = Np$	$s = \sqrt{Np(1-p)}$
Exponentialverteilung	$w(x) = \lambda e^{-\lambda x}$	$\bar{x} = \frac{1}{\lambda}$	$s = \frac{1}{\lambda}$
Poisson-Verteilung	$w(x) = \frac{\alpha^x}{x!} e^{-\alpha}$	$\bar{x} = \alpha$	$s = \sqrt{\alpha}$
Normalverteilung	$w(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\bar{x})^2}{2\sigma^2}}$	$\bar{x}$	$s$
Log-Normalverteilung	$w(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\ln x - \bar{x})^2}{2\sigma^2}}$	$\bar{x}$	$s$

### 5.3 Auto- und Kreuzkorrelation

Korrelationstechniken werden angewendet, wenn lokale Ähnlichkeiten der Signalform entweder zwischen zwei Signalen (Kreuzkorrelation) oder innerhalb eines Signals von Interesse sind. Für analoge Signale gilt:

$$\text{Kreuzkorrelationsfunktion: } r_{fg}(\tau) = \int_{-\Delta T/2}^{+\Delta T/2} f(t)g(t + \tau)dt \quad (5.5)$$

$$\text{Autokorrelationsfunktion: } r_{ff}(\tau) = \int_{-\Delta T/2}^{+\Delta T/2} f(t)f(t + \tau)dt \quad (5.6)$$

Die Kreuzkorrelationsfunktion gibt an, wie ähnlich sich die Signale  $f(t)$  und  $g(t)$  bei einer Verschiebung von  $g(t)$  um den Betrag  $\tau$  innerhalb des Korrelationsbereichs  $\Delta T$  sind. Große Werte stehen für hohe Ähnlichkeit. Analog gibt die Autokorrelationsfunktion an, wie ähnlich das Signal  $f(t)$  sich selbst bei einer Verschiebung um  $\tau$  innerhalb des Korrelationsbereichs  $\Delta T$  ist.

Die diskreten Pendants der beiden Funktionen sind:

$$\text{Diskrete Kreuzkorrelationsfunktion: } r_{fg}(k) = \sum_{i=-N}^N f_i g_{i+k} \quad (5.7)$$

$$\text{Diskrete Autokorrelationsfunktion: } r_{ff}(k) = \sum_{i=-N}^N f_i f_{i+k} \quad (5.8)$$

Abbildung 5.2 zeigt beispielhaft die Ermittlung von Laufzeiten von Ultraschallpulsen. Hier wird das stark verrauschte Empfängersignal mit der bekannten Signalform des Pulses kreuzkorreliert. Im Ergebnis werden die Pulsformen im Empfängersignal sicher gefunden.

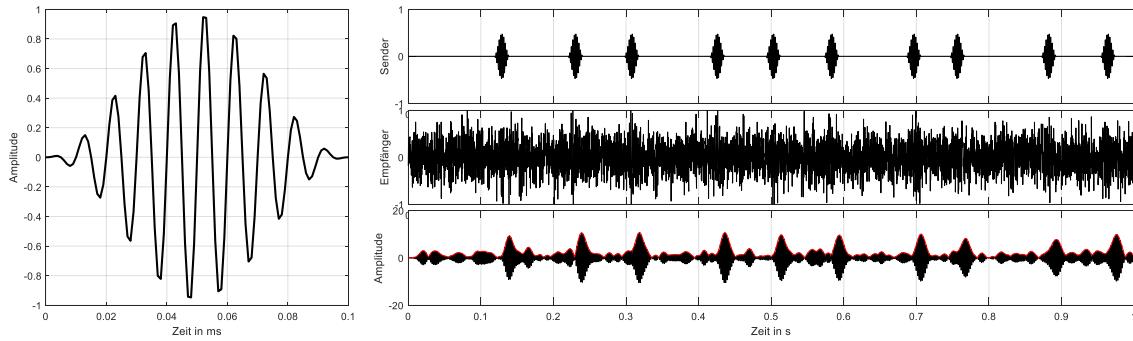


Abbildung 5.2: Kreuzkorrelation von Signalen. Im linken Bild ist die Signalform eines kurzen Ultraschallpulses zu sehen. Der Plot rechts oben zeigt die ungestörte Signalform unregelmäßig abgegebener Ultraschall-Pulse. Darunter ist die von einem Empfänger aufgezeichnete Sequenz abgebildet. Dem ursprünglichen Signal ist Rauschen in der Größenordnung der Signalleistung überlagert. Die ursprünglichen Pulse sind so kaum zu erkennen. Kreuzkorreliert man das Ausgangssignal links mit dem verrauschten Eingangssignal, so erhält man deutliche Ausschläge an den Stellen des Pulses (schwarze Kurve). Durch Bestimmung der Einhüllenden des Signals (roter Kurve) lassen sich die Pulsankunftszeiten über eine Maximalwertbestimmung ermitteln.

## 5.4 Lineare Regression

Ziel der Regressionsanalyse ist es, innerhalb einer „Datenwolke“ wahrscheinliche funktionale Zusammenhänge zu ermitteln. Dabei wird ein funktionaler Zusammenhang  $f(x)$  zwischen einer unabhängigen Variable  $x$  und einer abhängigen Variable  $y$  angenommen. In der Regel wird der Modellzusammenhang in Form einer parametrisierten Funktion vorgegeben, wobei deren freie Parameter zu schätzen sind. Das Schätzverfahren basiert auf der Minimierung einer Abweichung, den sogenannten Residuen, zwischen Modelfunktion und Datenwerten. Wie diese Abweichungen bestimmt werden, hängt vom zugrunde gelegten Modell ab. Wir betrachten das Beispiel der linearen Regression, welche sehr häufig zur Anwendung kommt.

Das der linearen Regression zugrundeliegende Modell ist das eines linearen Zusammenhangs zwischen unabhängigen und abhängigen Variablen. Lassen Sie uns annehmen, wir erhalten zwei synchrone zeitdiskrete Signale  $x_i = x(t_i)$  und  $y_i = y(t_i)$  aus einem Prozess und wir wollen herausfinden, ob das Signal  $y$  linear zum Signal  $x$  in Beziehung steht. Wir unterstellen also einen linearen Zusammenhang der Form

$$y_i = ax_i + b \quad (5.9)$$

Als Maß der Abweichung wählen das der Fehlerquadrate und fordern, dass die Summe der Quadrate der Abweichungen  $e_i = y_i - f(x_i)$  minimiert wird:

$$\sum_i e^2 = \sum_i (y_i - ax_i - b)^2 \stackrel{!}{\rightarrow} \text{Min} \quad (5.10)$$

Wir erhalten die Lösung, wenn wir Anleihe an das Lösungsschema der Ausgleichspolynome nehmen (Abschnitt 4.1). Dort setzen wir ein Polynom erster Ordnung an und ersetzen  $t$  durch  $x$  sowie  $a_1$  durch  $a$  und  $a_0$  durch  $b$  und erhalten:

$$\begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} N & \sum_{i=0}^{N-1} x_i \\ \sum_{i=0}^{N-1} x_i & \sum_{i=0}^{N-1} x_i^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{i=0}^{N-1} y_i \\ \sum_{i=0}^{N-1} x_i y_i \end{bmatrix} \quad (5.11)$$

mit der Lösung

$$\begin{aligned} a &= \frac{N \sum_{i=0}^{N-1} x_i y_i}{N \sum_{i=0}^{N-1} x_i^2 - (\sum_{i=0}^{N-1} x_i)^2} - \frac{(\sum_{i=0}^{N-1} x_i)(\sum_{i=0}^{N-1} y_i)}{N \sum_{i=0}^{N-1} x_i^2 - (\sum_{i=0}^{N-1} x_i)^2} \\ b &= \frac{(\sum_{i=0}^{N-1} x_i^2)(\sum_{i=0}^{N-1} y_i)}{N \sum_{i=0}^{N-1} x_i^2 - (\sum_{i=0}^{N-1} x_i)^2} - \frac{(\sum_{i=0}^{N-1} x_i)(\sum_{i=0}^{N-1} x_i y_i)}{N \sum_{i=0}^{N-1} x_i^2 - (\sum_{i=0}^{N-1} x_i)^2} \end{aligned} \quad (5.12)$$

Wie Abbildung 5.3 zeigt, erhalten wir mit den Geradenparametern  $a$  und  $b$  eine sogenannte Ausgleichsgerade, welche den im Sinne der kleinsten Fehlerquadrate bestmöglichen linearen Funktionszusammenhang zwischen den  $x_i$  und den  $y_i$  repräsentiert. Als Gütemaß für die Regression gilt das Bestimmtheitsmaß:

$$R^2 = \frac{\sum_i (y_i - \bar{y})^2}{\sum_i (f(x_i) - \bar{y})^2}. \quad (5.13)$$

Es ist das Verhältnis der Streuung der Messwerte  $y_i$  um ihren Mittelwert zur Streuung der Modellwerte um ihren Mittelwert.

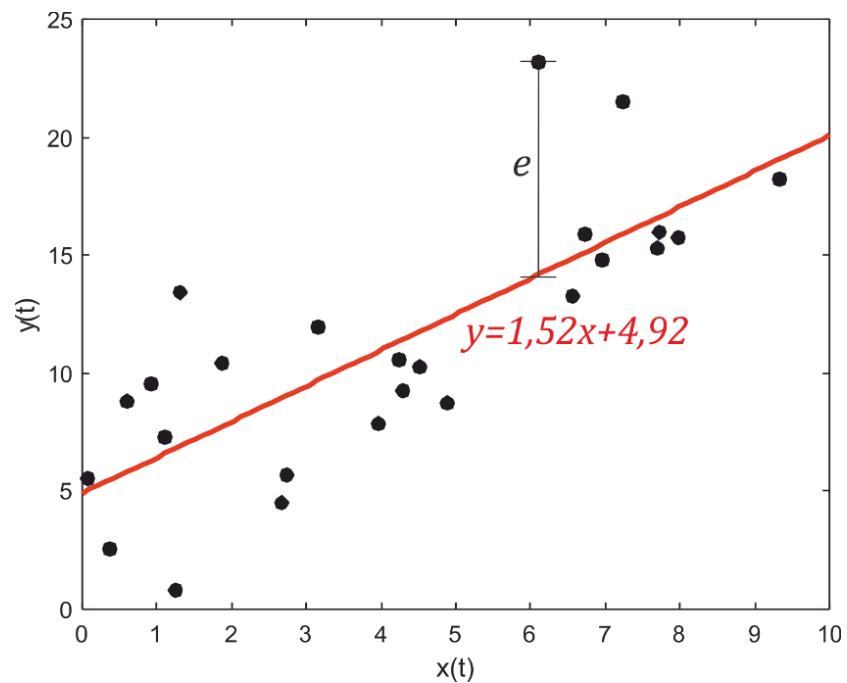


Abbildung 5.3: Lineare Regression.

## 6 Grundlagen der Signalverarbeitung im Frequenzbereich

### 6.1 Fourier-Reihen

Jede periodische Funktion lässt sich durch eine Fourier-Reihe approximieren. Es gilt

$$f(t) = \sum_{n=-\infty}^{+\infty} F_n e^{j\omega_n t} \quad (6.1)$$

mit den Grundfrequenzen  $\omega_n = 2\pi n/T$ , der Periodendauer  $T$  und den Fourier-Koeffizienten

$$F_n = \frac{1}{T} \int_{-T/2}^{+T/2} f(t) e^{-j\omega_n t} dt \quad (6.2)$$

Über die Eulersche Beziehung

$$e^{jx} = \cos x + j \sin x \quad (6.3)$$

lässt sich die Fourier-Reihe auch über die zugehörigen Cosinus- und Sinus-Reihen darstellen:

$$f(t) = \frac{A_0}{2} + \sum_{n=1}^{+\infty} A_n \cos(\omega_n t) + \sum_{n=1}^{+\infty} B_n \sin(\omega_n t). \quad (6.4)$$

Die Fourier-Koeffizienten  $F_n$  der komplexen Reihe sind dabei allgemein komplexwertig, die Koeffizienten  $A_n$  und  $B_n$  reellwertig.

## 6.2 Fourier-Transformation

In den Frequenzraum eines analogen Signals gelangt man durch Fourier-Transformation des Signals. Diese ist formal durch

$$F(\omega) = \mathcal{F}\{f(t)\} = \int_{-\infty}^{+\infty} f(t) e^{-j\omega t} dt \quad (6.5)$$

gegeben. Mit  $\omega = 2\pi f$  ist die Kreisfrequenz gegeben ( $f$  bezeichnet hier die Frequenz, nicht die Signalfunktion). Die Funktion  $F(\omega)$  bezeichnet das Signalspektrum. Die Rücktransformation aus dem Frequenzbereich in den Zeitbereich ist durch

$$f(t) = \mathcal{F}^{-1}\{F(\omega)\} = \frac{1}{2\pi} \int_{-\infty}^{+\infty} F(\omega) e^{j\omega t} d\omega \quad (6.6)$$

gegeben. Existiert das Fourier-Integral eines Signals  $f(t)$ , so hat  $f(t)$  die korrespondierende Spektralfunktion  $F(\omega)$  und umgekehrt. Wir kennzeichnen diese Korrespondenz mit dem Symbol  $\leftrightarrow$  und schreiben kurz:

$$f(t) \leftrightarrow F(\omega) \quad (6.7)$$

Das Fourier-Spektrum ist im allgemeinen komplexwertig. Jeder Frequenz  $\omega$  ist ein komplexer Spektralwert  $F(\omega) = a + jb$  zugeordnet. Dieser gibt die Amplitude

$$|F(\omega)| = A = \sqrt{a^2 + b^2} \quad (6.8)$$

und die Phasenlage

$$\angle F(\omega) = \varphi = \arccos\left(\frac{a}{\sqrt{a^2 + b^2}}\right) \quad (6.9)$$

des Signals bei der jeweiligen Frequenz  $\omega$  an. Die aus der Fourier-Transformation resultierende komplexe Amplitude ist abstrakt und daher in der Regel schwer zu verstehen. Hilfreich ist ein Blick auf das Zeigerdiagramm in der komplexen Zahlenebene (Abbildung 6.1). Die Harmonische

$$F(\omega)e^{j\omega t} = [Ae^{j\varphi}]e^{j\omega t} \quad (6.10)$$

wird durch einen mit der Kreisfrequenz rotierenden (durchgezogenen) Zeiger mit Länge  $A$  und Phasenoffset  $\varphi$  repräsentiert. Der gestrichelte Zeiger ist der zum durchgezogenen Zeiger konjugiert komplexe Zeiger und wird gleich erläutert.

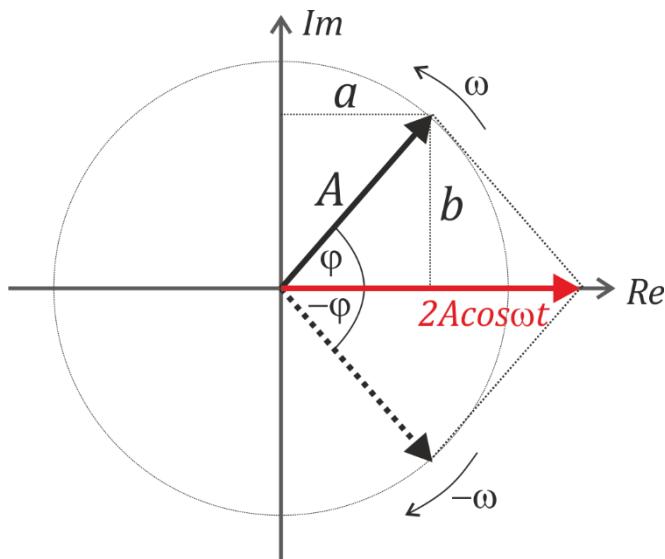


Abbildung 6.1: Zeigerdiagramm

Auch die Fourier-Transformation kann vermöge der Eulerschen Beziehungen wieder als Summe einer Sinustransformation und einer Kosinustransformation dargestellt werden. Es gilt

$$e^{j\omega t} = \cos \omega t + j \sin \omega t \quad (6.11)$$

$$F(\omega) = \int_{-\infty}^{+\infty} f(t) \cos \omega t dt + j \int_{-\infty}^{+\infty} f(t) \sin \omega t dt. \quad (6.12)$$

Daraus folgen die Eigenschaften (3) und (4) in

Tabelle 6.2: Um den Ursprung symmetrische Signale haben ein rein reelles Fourier-Spektrum, um den Ursprung antisymmetrische Signale ein rein imaginäres Frequenzspektrum.

Weiterhin gilt für ein reellwertiges Signal:

$$F(\omega) = F^*(-\omega). \quad (6.13)$$

Das heißt, der Spektralwert bei  $-\omega$  ist zum Spektralwert bei  $\omega$  konjugiert komplex.

**Beweis:**

$$\begin{aligned} F(-\omega) &= \int_{-\infty}^{+\infty} f(t) \cos(-\omega t) dt + j \int_{-\infty}^{+\infty} f(t) \sin(-\omega t) dt \\ &= \int_{-\infty}^{+\infty} f(t) \cos(\omega t) dt - j \int_{-\infty}^{+\infty} f(t) \sin(\omega t) dt \\ &= F^*(\omega) \end{aligned}$$

Dies bedeutet, dass es genügt, den positiven (oder negativen) Teil des Spektrums zu kennen. Der übrige Teil ergibt sich aus Gl. (6.13).

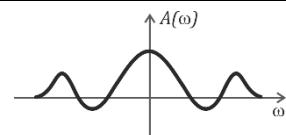
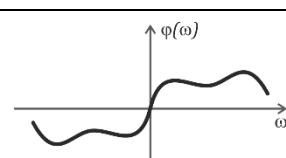
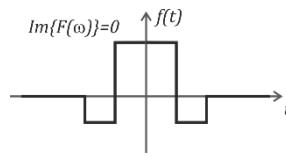
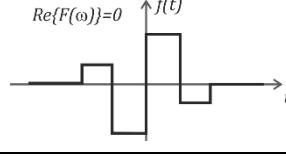
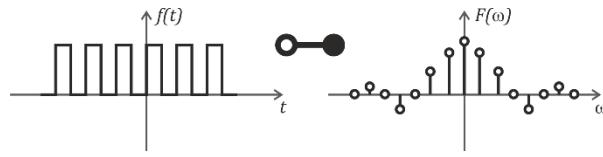
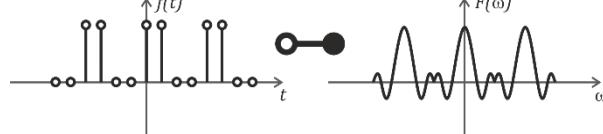
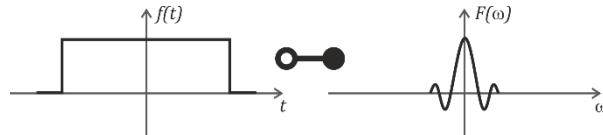
**Bemerkung:** Diese Symmetrieeigenschaft mag zunächst etwas befremdlich erscheinen, da die Hälfte des Spektrums bezgl. des Informationsgehaltes offensichtlich redundant ist. Dies ist auch in der Tat so. Dazu muss man sich vergegenwärtigen, dass das Signalspektrum, anders als das Signal im Zeitbereich, komplexe Werte annimmt. Daher verdoppelt sich der Informationsgehalt der Spektralwerte, sie bestehen aus je zwei reellen Zahlen  $a$  und  $b$ , während das Signal im Zeitbereich nur je eine reelle Zahl für jeden Zeitpunkt umfasst. Entsprechend der Verdopplung der Information im Wertebereich halbiert sich die Information im Definitionsbereich. Gäbe es komplexwertige Signale, würde deren Frequenzspektrum im allgemeinen keine Symmetrie aufweisen.

Die Tatsache, dass die Fourier-Transformierte eines reellen Signals konjugiert komplexe Werte im positiven und negativen Frequenzbereich haben muss, lässt sich wieder anschaulich am Zeigerdiagramm (Abbildung 6.1) erläutern. Zu jedem Zeiger  $F(\omega)e^{j\omega t} = [Ae^{j\varphi}]e^{j\omega t}$  gibt es einen gegenläufig rotierenden Zeiger  $F^*(-\omega)e^{j\omega t} = [Ae^{-j\varphi}]e^{-j\omega t}$  mit gleicher Amplitude  $A$  und entgegengesetztem Phasenoffset  $\varphi$ . Die Summe beider Zeiger ergibt stets einen reellen Zeiger der Länge  $2A \cos\varphi$ . Mit anderen Worten: da die Frequenz  $\omega$  in reellen Signalen nur mit einer reellwertigen Amplitude vertreten sei kann, müssen die zu  $\omega$  gehörigen beiden komplexen Zeiger zueinander konjugiert komplex sein!

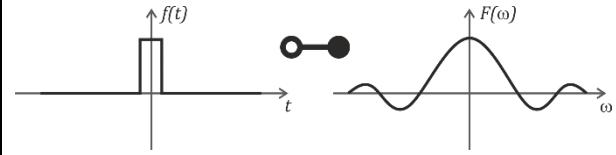
Tabelle 6.1: Einige wichtige Korrespondenzen der Fourier-Transformation

Stoßfunktion	$\delta(t) \leftrightarrow 1$	
Harmonische	$\sin(\omega_0 t) \leftrightarrow e^{j\omega} \delta(\omega \pm \omega_0)$ $\cos(\omega_0 t) \leftrightarrow \delta(\omega \pm \omega_0)$	
Rechteck-Puls	$\text{rect}(t; a) \leftrightarrow \frac{\sin \omega}{\omega} = \text{asinc}(\omega a)$	
Gauß-Puls	$\frac{1}{\sqrt{2\pi}\sigma^2} e^{-\frac{t^2}{2\sigma^2}} \leftrightarrow e^{-\frac{\sigma^2 t^2}{2}}$	

Tabelle 6.2: Wichtige Eigenschaften der Fourier-Transformation für reellwertige Signale

Die Fourier-Transformierte ist im Allgemeinen komplexwertig	$F(\omega) = a(\omega) + jb(\omega) = A(\omega)e^{j\varphi(\omega)}$
Die Werte der Fourier-Transformierten positiver Frequenzen sind zu den Werten der negativen Frequenzen konjugiert komplex	$F(-\omega) = F^*(\omega)$
Das Amplitudenspektrum ist immer symmetrisch (gerade)	
Das Phasenspektrum ist immer antisymmetrisch (ungerade)	
Bezgl. $t = 0$ symmetrische (gerade) Signale haben ein rein reelles Spektrum	
Bezgl. $t = 0$ antisymmetrische (ungerade) Signale haben ein rein imaginäres Spektrum	
Periodische Signale haben ein diskretes Spektrum	
Diskrete Signale haben ein periodisches Spektrum	
Breite Signale haben schmale Spektren	

Schmale Signale haben breite Spektren



### 6.3 Signalleistung

In einer elektrischen Schaltung würde ein Signal der Spannung  $u(t)$  an einem Widerstand eine Verlustleistung von  $p(t) = u(t)^2/R$  erzeugen. Deshalb wird das Leistungssignal mathematisch als das Signal der an einem Einheitswiderstand ( $R = 1 \Omega$ ) abfallenden Verlustleistung definiert:

$$p(t) = \int_{-\infty}^{+\infty} u(t)^2 dt \quad (6.14)$$

Entsprechend ergibt sich das Leistungsspektrum zu

$$P(\omega) = F(\omega)^2. \quad (6.15)$$

**Bemerkung:** Diese Definition hat ihren Ursprung im elektrotechnischen Grundverständnis, dass Signale üblicherweise Spannungsfolgen sind. Die elektrische Leistung ist durch  $P = U \cdot I$  gegeben. Die Stromstärke in einen Ohmschen Widerstand der Größe  $R$  leitet sich aus dem Ohmschen Gesetz ( $R = U/I$ ) zu  $I = U/R$  ab. Zusammengefasst ist also  $P = U^2/R$ .

Die Leistung eines elektrischen Signals ist eine wichtige Größe, wenn es darum geht zu bewerten, wie gut ein Signal von Rauschen unterschieden werden kann oder wie gut ein Signal übertragen werden kann. In beiden Fällen ist die energetische Wirkung des Signals wichtig. Stellen wir uns vor, die Spannungsfolge soll über eine Antenne übertragen werden. Dann ist die Spannungsfolge in eine elektrische Leistung umzusetzen. Diese ist proportional zu  $u(t)^2$ .

Anders verhält es sich, wenn wir beispielsweise optische Signale betrachten. Ist bei solchen die Signalamplitude durch die Lichtstärke gegeben, so ist diese bereits selbst eine Leistung, womit es sich von vornherein um ein Leistungssignal handelt. Dies ist im jeweiligen Kontext zu berücksichtigen.

Signalamplituden werden häufig durch logarithmische Maße miteinander verglichen. Die dafür gebräuchliche Einheit ist das Bel ( $B$ ) bzw. das Dezibel ( $dB$ ). 1 Bel steht für ein Leistungsverhältnis von  $V = p_1(t)/p_2(t) = 1$ , ein Dezibel für  $V = p_1(t)/p_2(t) = 0,1$ . Allgemein gilt also:

$$V = 10 \log_{10} \frac{p_1(t)}{p_2(t)} dB. \quad (6.16)$$

Für Effektivgrößen, oder sogenannte Leistungswurzelgrößen, wie die Spannung, gilt dann entsprechend

$$V = 20 \log_{10} \frac{u_1(t)}{u_2(t)} dB. \quad (6.17)$$



## 6.4 Differenziale und Integrale im Frequenzbereich

Eine interessante Eigenschaft der Fourier-Transformation, die sich bereits ihr Namensgeber, Jean-Baptiste Joseph Fourier, im Rahmen der Fourier-Analyse zur Lösung der Wärmeleitungsgleichung zunutze machte, besteht darin, dass die Fourier-Transformation die Transformationen der Differentiation und der Integration von Signalen im Frequenzbereich zu einfachen algebraischen Operationen macht. Es gilt:

$$\frac{d}{dt} f(t) \leftrightarrow j\omega F(\omega) \quad (6.18)$$

$$\int f(t) dt \leftrightarrow \frac{1}{j\omega} F(\omega) + C \quad (6.19)$$

Die Integrationskonstante konkretisiert sich dabei fallspezifisch entsprechend der Randbedingungen einer bestimmten Integration.

**Beweis zu GI. (6.20):**

$$\begin{aligned} \frac{d}{dt} f(t) &= \frac{d}{dt} \mathcal{F}^{-1}\{F(\omega)\} = \frac{d}{dt} \frac{1}{2\pi} \int_{-\infty}^{+\infty} F(\omega) e^{j\omega t} d\omega = \frac{1}{2\pi} \int_{-\infty}^{+\infty} F(\omega) \frac{d}{dt} e^{j\omega t} d\omega \\ &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} F(\omega) j\omega e^{j\omega t} d\omega = \mathcal{F}^{-1}\{j\omega F(\omega)\} \end{aligned}$$

**Beweis zu GI. (6.19):**

$$\begin{aligned} \int f(t) dt &= \int \mathcal{F}^{-1}\{F(\omega)\} dt = \int \left( \frac{1}{2\pi} \int_{-\infty}^{+\infty} F(\omega) e^{j\omega t} d\omega \right) dt = \frac{1}{2\pi} \int_{-\infty}^{+\infty} F(\omega) \left[ \int (e^{j\omega t}) dt \right] d\omega \\ &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} F(\omega) \left( \frac{1}{j\omega} e^{j\omega t} + C \right) d\omega = \mathcal{F}^{-1}\left\{\frac{1}{j\omega} F(\omega)\right\} + C \end{aligned}$$

## 6.5 Zeitverschobene Signale im Frequenzbereich

Wird das Signal im Zeitbereich um  $\Delta t$  verschoben, so ist

$$f(t - t_0) \circledast e^{-j\omega\Delta t} F(\omega). \quad (6.20)$$

*Beweis:*

$$\begin{aligned}\mathcal{F}\{f(t - \Delta t)\} &= \int_{-\infty}^{+\infty} f(t - \Delta t) e^{-j\omega t} dt = \int_{-\infty}^{+\infty} f(t) e^{-j\omega(t + \Delta t)} dt \\ &= e^{-j\omega\Delta t} \int_{-\infty}^{+\infty} f(t) e^{-j\omega t} dt = e^{-j\omega\Delta t} \mathcal{F}\{f(t - \Delta t)\} \\ &= e^{-j\omega\Delta t} F(\omega)\end{aligned}$$

Bei der Vertauschung der Verschiebung in den beiden Faktoren unter dem Integral (Zeile 1) wird von der Tatsache Gebrauch gemacht, dass es innerhalb des unendlichen Integrals (und zwar nur für den Fall eines unendlichen Integrals) egal ist, ob die Funktion  $f(t)$  um  $+\Delta t$  oder die Funktion  $e^{-j\omega t}$  um  $-\Delta t$  verschoben wird.

## 6.6 Faltung

Die Faltung ist eine spezielle Integraltransformation. Sie ist formal beschrieben durch

$$g(t) = \int_{-\infty}^{+\infty} h(t - t') f(t') dt' \equiv f(t) * g(t). \quad (6.21)$$

Um die mathematische Operation der Faltung intuitiv zu verstehen, sei auf Abbildung 6.2 verwiesen. Der Faltungskern  $h(t)$  wird linear über das Signal  $f(t)$  „geschoben“. Für jede Verschiebung  $t - t'$  wird über den gesamten Definitionsbereich das Produkt aus Faltungskern und Eingangssignal integriert und daraus jeweils ein Wert des Ausgangssignals gebildet.

Die Faltung ist eine der wichtigsten Operationen der Systemtheorie. Sie dient formal der Erzeugung von Antwortfunktionen linearer zeitinvarianter Systeme. Sie wird darüber hinaus zur Lösung gewöhnlicher und partieller linearer Differentialgleichungen genutzt (siehe auch: *retardierte Potenziale, Greensche Funktionen*). Um ihre Bedeutung im Rahmen der Signalverarbeitung zu zeigen, wird nachfolgend das Beispiel der Signalglättung erläutert.

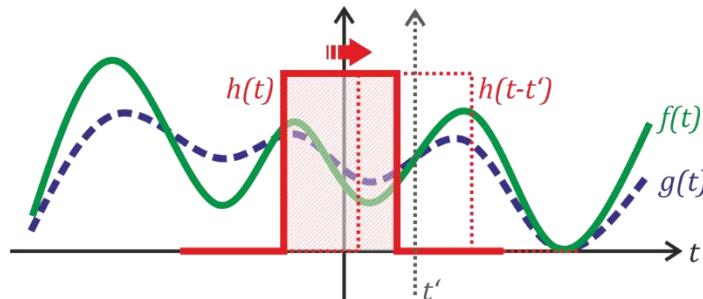


Abbildung 6.2: Funktionsweise der Faltungsoperation im Zeitbereich.

### Beispiel: Glättung eines Signals

Signale werden geglättet, um hochfrequentes Rauschen zu unterdrücken. Ein Signal  $f(t)$  kann geglättet werden, indem ein neues Signal  $g(t)$  so gebildet wird, dass für jeden Zeitpunkt  $t$  die Werte von  $f(t)$  innerhalb eines schmalen Zeitfensters mit Breite  $\Delta T$  gemittelt werden. Es folgt:

$$g(t) = \frac{1}{\Delta T} \int_{-\infty}^{+\infty} \text{rect}(t - t'; \Delta T) f(t') dt'.$$

Die Rechteckfunktion ist hier der Kern des Faltungsintegrals.

## 6.7 Faltung im Frequenzraum

Wir werden jetzt Gleichung (6.21) durch Fourier-Transformation in den Frequenzraum überführen und dort analysieren. Dafür bedienen wir uns der Korrespondenz für zeitverschobene Signale

$$f(t - t') \circledast e^{-j\omega t'} F(\omega) \quad (6.22)$$

und erhalten

$$f(t) * g(t) \circledast G(\omega)F(\omega) \quad (6.23)$$

### Beweis:

$$\begin{aligned} H(\omega) &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \left\{ \int_{-\infty}^{+\infty} g(t - t') f(t') dt' \right\} e^{-j\omega t} d\omega \\ &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} f(t') \left\{ \int_{-\infty}^{+\infty} g(t - t') e^{-j\omega t} d\omega \right\} dt' \\ &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} f(t') \mathcal{F}\{g(t - t')\} dt' \\ &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} f(t') e^{-j\omega t'} G(\omega) dt' \\ &= G(\omega) \frac{1}{2\pi} \int_{-\infty}^{+\infty} f(t') e^{-j\omega t'} dt' \\ &= G(\omega)F(\omega) \end{aligned}$$

Beim Übergang von Zeile 1 zu Zeile 2 wurde die Integrationsreihenfolge (Faltungsintegral und Fourier-Integral) vertauscht, beim Übergang von Zeile 3 zu Zeile 4 die Korrespondenzregel angewendet.

## 6.8 Korrelation und Faltung

Das Korrelationsintegral sieht dem Faltungsintegral sehr ähnlich. In der Tat gibt es zwischen beiden auch einen engen Zusammenhang. Wir führen die Diskussion anhand analoger Signale. Für diskrete Signale gelten die Herleitungen entsprechend.

Die vollständige Kreuzkorrelationsfunktion ist

$$r_{fg}(\tau) = \int_{-\infty}^{+\infty} f(t)g(t + \tau)dt. \quad (6.24)$$

Setzen wir  $\hat{t} = -t$  entsteht ein Faltungsintegral

$$r_{fg} = \int_{-\infty}^{+\infty} f(-\hat{t})g(\tau - \hat{t})d\hat{t} = f(-\hat{t}) * g(\hat{t}). \quad (6.25)$$

Beachten wir die Korrespondenzen  $f(-t) \leftrightarrow F(-\omega)$  sowie  $F(-\omega) = F^*(\omega)$  gilt für das Spektrum der Kreuzkorrelationsfunktion

$$R_{fg}(\omega) = F^*(\omega)G(\omega). \quad (6.26)$$

Sowie für das Spektrum der Autokorrelationsfunktion

$$R_{ff}(\omega) = F^*(\omega)F(\omega). \quad (6.27)$$

Diese Herleitungen haben interessante Konsequenzen:

1. Die Kreuz- und Autokorrelationsfunktionen lassen sich ebenso, wie die Faltung, durch Fourier-Transformation, Multiplikation und anschließende Rücktransformation berechnen. Das kann bei großen Datenreihen effizient sein, da die diskreten Fourier-Transformationen als schnelle Fourier-Transformationen (Fast Fourier Transform, FFT) mit Berechnungskomplexität  $N \log N$  berechnet werden, wohingegen Korrelationssummen eine Berechnungskomplexität  $N^2$  haben.
2. Mit  $F^*(\omega)F(\omega) = |F(\omega)|^2$  folgt, dass das quadratische Betragsspektrum oder Energiedichtetespektrum  $|F(\omega)|^2$  gleich der Fourier-Transformierten der Autokorrelationsfunktion ist. Diese als Wiener-Chintschin-Theorem bekannte Tatsache ermöglicht es, die Autokorrelationsfunktion stochastischer Signale über ihr Energiedichtetespektrum  $|F(\omega)|^2$  zu bestimmen.
3. Mit der totalen Energie  $E = \int_{-\infty}^{+\infty} |F(\omega)|^2 d\omega$  des Spektrums folgt das Parsevalsche Theorem:  $\int_{-\infty}^{+\infty} |F(\omega)|^2 d\omega = E = \int_{-\infty}^{+\infty} f^2(t)dt$ .

## 6.9 Abtastung analoger Signale: Abtasttheorem und Aliasing

Das zeitlich äquidistante Abtasten eines Signals lässt sich formal durch Multiplikation des Signals mit einem Dirac-Kamm darstellen:

$$f_A(t) = f(t) \sum_{n=-\infty}^{+\infty} \delta(t - nT). \quad (6.28)$$

Hierbei werden Signalwerte an den Zeitpunkten  $nT$  ausgeblendet, alle anderen Signalanteile werden Null. Unterzieht man das resultierende abgetastete Signal  $f_A(t)$  einer Fourier-Transformation, erhält man

$$\begin{aligned} F_A(t) &= \mathcal{F}\{f_A(t)\} \\ &= \mathcal{F}\left\{f(t) \cdot \left[ \sum_{n=-\infty}^{+\infty} \delta(t - nT) \right] \right\} \\ &= F(\omega) * \mathcal{F}\left\{ \sum_{n=-\infty}^{+\infty} \delta(t - nT) \right\} \\ &= F(\omega) * \left[ \sum_{n=-\infty}^{+\infty} \mathcal{F}\{\delta(t - nT)\} \right] \\ &= F(\omega) * \left[ \sum_{n=-\infty}^{+\infty} e^{-j\omega nT} \mathcal{F}\{\delta(t)\} \right] \\ &= F(\omega) * \left[ \sum_{n=-\infty}^{+\infty} e^{-j\omega nT} \right] \end{aligned} \quad (6.29)$$

Der Ausdruck  $\sum_{n=-\infty}^{+\infty} e^{-j\omega nT}$  stellt wiederum eine Dirac-Impulsfolge, diesmal im Frequenzraum, mit Frequenzabstand  $\omega_n = 2\pi n/T$  dar. Die Faltung des Spektrums  $F(\omega)$  des nicht abgetasteten Signals mit dem Impulskamm  $\sum_{n=-\infty}^{+\infty} e^{-j\omega nT}$  im Frequenzraum ergibt eine überlagernde Spiegelung des Spektrums an allen Frequenzpunkten  $\omega_n$ . Formal wird damit eine Eigenschaft der Fourier-Transformation aus

Tabelle 6.2 bewiesen: zeitdiskrete Signale haben periodische Spektren, periodische Signale frequenzdiskrete Spektren. Letzteres ist uns durch die Fourier-Reihe bekannt. Ein periodisches Signal hat nur Fourier-Koeffizienten für die Vielfachen der Grundfrequenz  $\omega_n = 2\pi n/T$ . Die Tatsache, dass abgetastete Signale periodische Spektren haben, hat eine wesentliche praktische Implikation, die unter dem Begriff Aliasing bekannt ist.

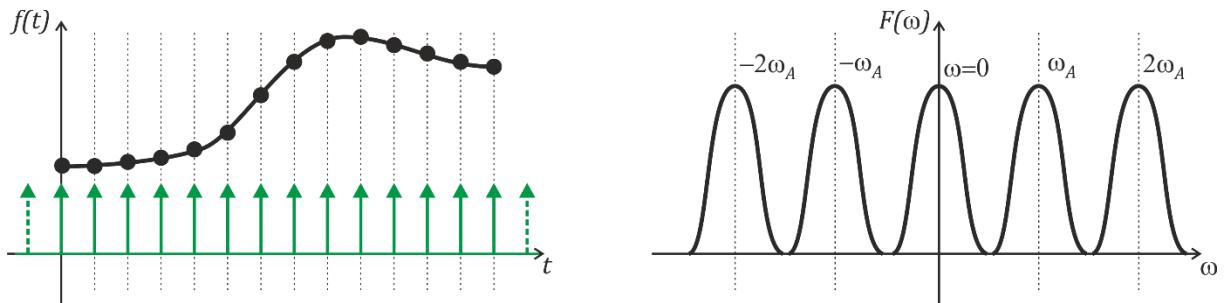


Abbildung 6.3: Abtastung eines Analogsignals mit Abtastintervallen  $T$ . Im Zeitbereich entspricht dies der Multiplikation des analogen Signals mit einem Dirac-Kamm (grün). Im Frequenzbereich ergibt sich daraus ein sich mit der Abtastfrequenz  $\omega_A = 2\pi f_A = 2\pi/T$  wiederholendes Spektrum.

Das Wort Aliasing kommt aus dem Englischen. Man beachte die phonetische Aussprache: 'eɪliəsɪŋ' (ey-lee-uh-sing). Alias im Latein bedeutet „Synonym“ oder „etwas Anderes“. Man kann die Begriffsbildung bezüglich der Signalverarbeitung so verstehen, dass Informationen aus anderen Frequenzbändern in das betrachtete Frequenzband eindringen und damit die Signale verfälschen bzw. stören.

Abbildung 6.4 zeigt das Problem im Frequenzbereich. Gegeben sei ein Signal  $f(t)$ , welches frequenzbegrenzt ist. Es hat also keine oder nur sehr kleine Signalanteile oberhalb einer Grenzfrequenz  $\omega_G$ . Wird dieses Signal mit einer Frequenz  $\omega_A > 2\omega_G$  abgetastet, so ergibt sich das Bild Abbildung 6.4 links. Die Periodizität des Frequenzspektrums des abgetasteten Signals ist dergestalt, dass sich die periodischen Frequenzbänder des Signals nicht überlappen. Ist hingegen  $\omega_A < 2\omega_G$  (Abbildung 6.4 rechts), so kommt es zur Überlappung der Frequenzbänder. Wie Gl. (6.29) zeigt, addieren sich die Signalanteile in den Überlappungsbereichen, so dass das abgetastete Signal spektral verfälscht wird. Dies äußert sich dann auch im Zeitbereich, wobei die Art der Störung dort nicht so leicht zu erkennen ist.

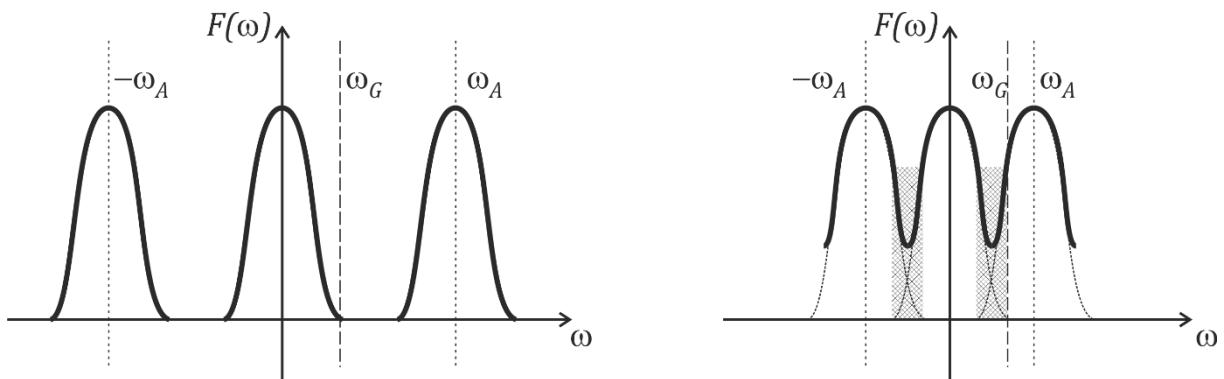


Abbildung 6.4: Aliasing im Frequenzbereich.

Die Kernaussage dieses Frequenzraumbildes ist die folgende:

**Nyquist-Shannon-Abtasttheorem:** Soll ein Signal unverfälscht abgetastet werden, so muss die Abtastfrequenz  $\omega_A$  mindestens 2x größer als die höchste vorkommende Frequenz  $\omega_G$  im Spektrum des Signals sein. Es gilt:

$$\omega_A > 2\omega_G.$$

Als Nyquist-Frequenz

$$\omega_{Ny} = \frac{1}{2} \omega_A.$$

(nach Harry Nyquist, schwedischer Elektrotechnik-Ingenieur) bezeichnet man die halbe Abtastfrequenz. Sie stellt die kritische Frequenz für unverzerrtes Abtasten dar.

Eine Illustration und Diskussion des Aliasing-Effekts im Zeitbereich kann anhand eines harmonischen Signals erfolgen (Abbildung 6.5). Wird ein Sinus-Signal der Frequenz  $\omega$  mit einer Abtastfrequenz  $\omega_A < 2\omega$  abgetastet, so wird die eigentlich negative Spektrallinie des Signal an der ersten positiven Abtastfrequenz so in das Spektrum hineingespiegelt, dass sie unterhalb der originalen positiven Spektrallinie auftaucht. Damit taucht im abgetasteten Signal eine Harmonische der Frequenz  $\omega_{Ny} - \omega$  auf, die im zeitkontinuierlichen Ausgangssignal nicht vorhanden ist. Im Zeitbereich erkennt man, dass das Unterabtasten der Sinuswelle zu einer (zeitdiskreten) Sinuswelle geringerer Frequenz führt.

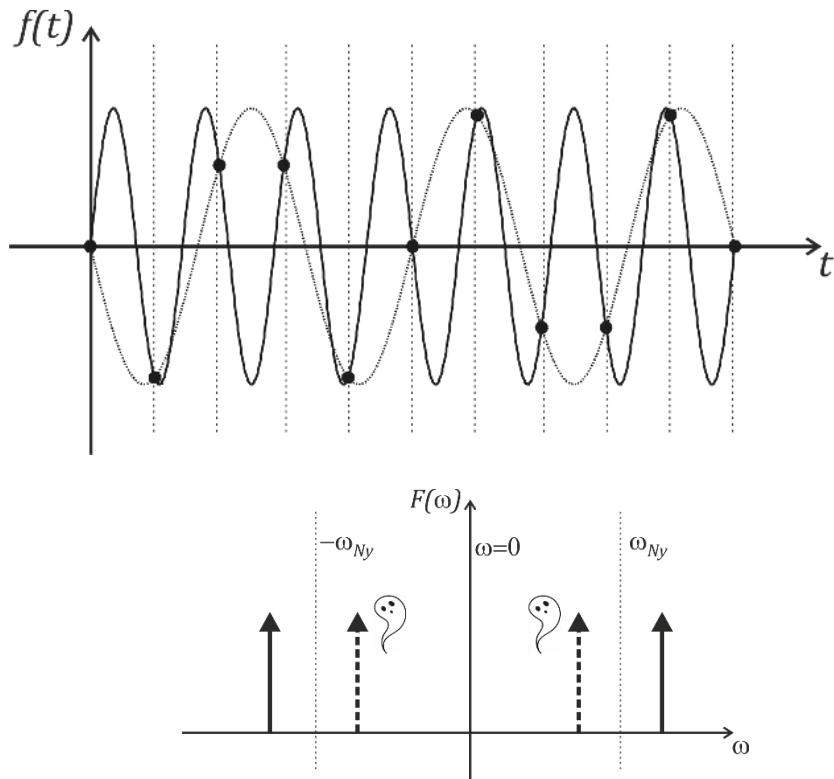


Abbildung 6.5: Das Phänomen des Aliasing im Zeit- und im Frequenzbereich am Beispiel einer einzigen Harmonischen.

**Beispiel für Aliasing im Kino:** Haben Sie schon mal in einem Film festgestellt, dass sich das Rad eines Fahrzeugs vermeintlich rückwärts dreht? Wenn ja, dann ist dies ein Fall von unvermeidbarem Aliasing. Kinofilme haben eine typische Bildfrequenz von 24 Hz. Rotiert das Rad eines Autos mit 24 Umdrehungen pro Sekunde, sieht es im Film aus, als würde sich das Rad gar nicht drehen. Rotiert es leicht langsamer, kommt es uns vor, als würde es sich rückwärts bewegen. Stellen wir uns die Winkelposition des Ventils im Uhrzeigersystem vor. Rotiert das Rad mit 22 Hz (entgegen dem Uhrzeigersinn), dann erscheint das Ventil in Bild 0 auf 12 Uhr, in Bild 1 auf 1 Uhr, in Bild 2 auf 2 Uhr, in Bild 3 auf 3 Uhr, usw. Bei schnellerer Bildfolge scheint das Rad rückwärts zu drehen. Leider lässt sich dieses Aliasing nicht vermeiden, da die Bildfrequenz des Films nicht erhöht werden kann.

In der Signalverarbeitung kann man Aliasing unterdrücken, indem man spektrale Anteile des Signals oberhalb der Grenzfrequenz  $\omega_G$  durch ein Tiefpassfilter unterdrückt (Abbildung 6.6). Solche sogenannten Anti-Aliasing-Filter finden sich in allen Signalverarbeitungssystemen vor dem Analog-Digital-Wandler, dessen Abtastrate festgelegt ist. Oft sind sie bereits in ADC-Schaltkreise integriert.

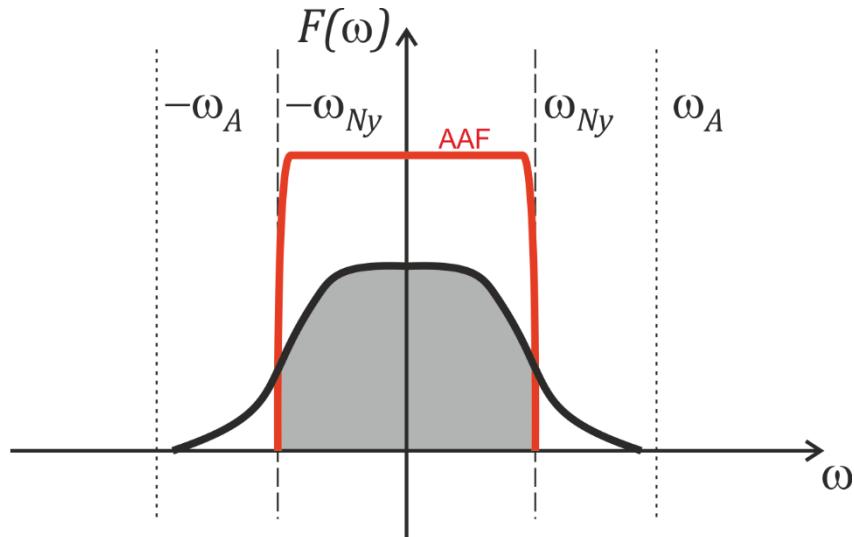


Abbildung 6.6: Ein Anti-Aliasing-Filter (AAF, rote Kurve) schneidet alle Anteile des Signals (schwarze Kurve) außerhalb des Bereichs  $-\omega_{Ny} \dots + \omega_{Ny}$  ab.

## 6.10 Abtastung von unendlichen Signalen

In der Theorie der Signalverarbeitung sind unendliche Signale eine Abstraktion. Sie hat ihre Begründung darin, dass man unterstellt, das eine erfasste physikalische Größe, etwa die Temperatur an einer Messstelle einer Prozessanlage, immer oder zumindest für sehr lange Zeit präsent ist. Die Aufgabe der Messtechnik ist es, diese Größe auch für einen langen Zeitraum zu erfassen. Deren Verarbeitung, etwa im Rahmen der Prozesssteuerung und -regelung, kann aber nur in einem begrenzten Zeitraum erfolgen. Dieser ist zunächst einmal fundamental durch das Speichervolumen und die Verarbeitungszeit im Prozessrechner begrenzt. Allerdings ist es auch praktisch in der Tat so, dass die Signalverarbeitung innerhalb relativ kurzer Zeitfenster erfolgt, die nur so lang sein müssen, bis die interessierende Größe aus dem Signal extrahiert ist. Wie lang dieses Zeitfenster ist, hängt von der Prozessgröße und dem Messsystem ab. Zwei Beispiele:

1. Ein Ofen wird beheizt und dessen Innentemperatur gemessen. Aus der Kenntnis der thermischen Trägheit des Ofens und seines Inventars lässt sich das Zeitfenster abschätzen, in dem ein Anstieg des Temperatursignals zu 99% abgeschlossen ist.
2. Ein Ultraschallpuls durchläuft ein Prozessvolumen. Bei Kenntnis der Schallgeschwindigkeit und maximalen Lauflänge lässt sich ein Zeitfenster finden, indem der reflektierte Puls sicher erfassbar ist.

Das „Herausgreifen“ von Messwerten innerhalb eines endlichen Zeitfensters aus einer eigentlich unendlichen Sequenz kann zu Problemen führen, wenn man anschließend signalverarbeitende Operationen im Frequenzraum durchführt. Dies wird zunächst formal anhand einer Frequenzraumbetrachtung diskutiert und anschließend durch eine Zeitbereichsbetrachtung verständlich gemacht.

Wir haben weiter oben kennengelernt, dass sich das Spektrum eines abgetasteten Signals bei jedem Vielfachen der Abtastfrequenz wiederholt. Jetzt wollen wir das unendliche diskrete Signal  $f_A(t)$  mit einem endlich langen Rechteckfenster multiplizieren (Abbildung 6.7). Dieses blendet jetzt nur die Signalanteile innerhalb des Messfensters  $[-\Delta T, +\Delta T]$  aus. Die Operation der Multiplikation im Zeitbereich entspricht laut Faltungstheorem einer Faltung mit dem Spektrum der Fensterfunktion im Zeitbereich. Für ein Rechteckfenster ist dies die Sinc-Funktion:

$$\mathcal{F}\{rect(t; \Delta T)\} = \Delta T sinc(\omega \Delta T) = \frac{\sin \omega \Delta T}{\omega}. \quad (6.30)$$

Abbildung 6.7 zeigt dies beispielhaft. In den linken Bildern ist das mit  $1 \text{ kHz}$  abgetastete Signal gezeigt, welches mit einem Rechteckfenster von  $20 \text{ ms}$  Dauer multipliziert wurde. In den rechten Bildern ist das zugehörige Spektrum dargestellt. Aufgrund der Abtastung wiederholt sich das Spektrum mit jedem ganzzahligen Vielfachen der Abtastfrequenz. Das Begrenzungsfenster führt zu einer Faltung des Signals mit der Sinc-Funktion. In der obersten Zeile ist dies beispielhaft für ein konstantes Signal dargestellt. Dieses ergibt die Funktion  $\delta(\omega) * sinc(\omega \Delta T) = sinc(\omega \Delta T)$ . Nimmt man jetzt nur die Abtastwerte innerhalb des Zeitfensters und führt diese einer diskreten Fourier-Transformation zu, so ergibt sich ein diskretes Spektrum mit Abtastpunkten an den Vielfachen von  $1/\Delta T$  (im Beispiel  $50 \text{ Hz}$ ), da  $\Delta T$  genau die größte Periode und  $1/\Delta T$  die dieser Periode zugeordnete kleinste diskrete Frequenz im Spektrum ist. Nicht ganz zufällig liegen die Abtastpunkte an den Nullstellen der Sinc-Funktion. Damit ist das diskrete Spektrum dahingehend „sauber“, dass Frequenzbeiträge an den diskreten Frequenzstellen außerhalb  $f = 0$  Null werden. Nun analysieren wir das Ganze für ein harmonisches Signal mit einer Frequenz  $n/\Delta T$ , wobei  $n$  ganzzahlig sein soll. Im Beispiel in Abbildung 6.7 sind dies  $200 \text{ Hz}$ , also  $n = 4$ . Diesmal ist  $\delta(\omega - 2\pi n/\Delta T) * sinc(\omega \Delta T) = sinc(\Delta T(\omega - 2\pi n))$ , das heißt, das diskrete Spektrum hat die gleiche Form, wie für das konstante Signal, nur um  $2\pi n/\Delta T$  verschoben. Auch hier liegen die durch die Fensterbreite festgelegten Frequenzabtastpunkte genau an den Nullstellen der Sinc-Funktion, so dass das diskrete Spektrum „sauber“ ist. Im dritten Fall wollen wir nun ein harmonisches Signal mit einem nichtganzzahligen  $n$  betrachten. In Abbildung 6.7, untere Zeile, ist dies ein harmonisches Signal mit  $230 \text{ Hz}$ . Die Sinc-Funktion ist damit um  $230 \text{ Hz}$  verschoben und die Frequenzabtastpunkte, die ja nicht von der Frequenz des Eingangssignals sondern nur von der Fensterbreite abhängen, liegen außerhalb der Nullstellen der verschobenen Sinc-Funktion. Dies hat zur Folge, dass im diskreten Spektrum plötzlich Frequenzanteile auftauchen, die in der originalen abgetasteten Eingangsfunktion  $f_A(t)$  gar nicht existieren. Sie werden durch das Fenster eingetragen. Diese Tatsache ist vor allem im Frequenzbild verblüffend, lässt sich aber durch eine Zeitbereichsbetrachtung sehr schnell erklären. Wir haben hier zunächst die Erläuterung im Frequenzraum bemüht, die zwar nicht so intuitiv ist, wie die Erklärung im Zeitbereich, aber das Problem quantitativ beschreibt. Das heißt, die durch die „Fensterung“ eingetragenen Fehler lassen sich, auch für ein

komplizierteres Eingangssignal, exakt beschreiben. Bevor wir das Problem im Zeitbereich beschreiben, hier der Merksatz:

**Merke:** Blendet man aus einem abgetasteten Zeitsignal  $N$  Werte aus und führt diese einer diskreten Fourier-Transformation zu, dann unterstellt man damit automatisch, dass die diskrete Zeitreihe periodisch mit der Periode  $N$  ist. Damit können auch nur Frequenzanteile „sauber“ diskretisiert werden, deren Frequenzen einem Vielfachen der inversen Fensterbreite entsprechen. Alle anderen Frequenzen werden entsprechend einer durch die Sinc-Funktion des Fensters definierten Gewichtung auf benachbarte Fenster-Frequenzen  $n/\Delta T$  aufgesplittet. Das kann zu erheblichen Signalverfälschungen führen.

Diese Erkenntnis hat weitreichende Konsequenzen. In Computern kann immer nur eine diskrete Fourier-Transformation endlicher Signale durchgeführt werden. Das dadurch notwendige Begrenzen eines eigentlich unendlichen Signals muss daher besonders dann, wenn das Begrenzungsfenster kurz ist, mit einiger Vorsicht und Kontrolle erfolgen. Wie das Beispiel in Abbildung 6.7 unten zeigt, wird eine Zwischenfrequenz auf zwei Nachbarfrequenzen aufgeteilt und kann damit im Spektrum erheblich kleinere Signalleistung „vortäuschen“, als eigentlich im Ausgangssignal vorhanden. Dies kann zum Beispiel zu Problemen in einer Spektralanalyse, zum Beispiel von Daten von Schwingungsaufnehmern führen. Bevor wir Maßnahmen zur Minderung des Problems diskutieren, illustriert Abbildung 6.8 das Problem im Zeitbereich. Wird das harmonische  $230\text{ Hz}$  Signal mittels eines  $20\text{ ms}$  Fensters begrenzt, und die Zeitreihe auf die innerhalb des Fensters liegenden Datenpunkte beschränkt, so ist das gleichbedeutend mit einer „Zwangsperiodisierung“ des analogen Ausgangssignals. Ist dessen Frequenz kein Vielfaches der inversen Fensterbreite  $1/\Delta T$ , dann entstehen Unstetigkeiten des nunmehr periodischen Signals an den Fenstergrenzen, die zu Geisterfrequenzen führen.

Es gibt drei Möglichkeiten, diesem Problem zu begegnen:

1. Das Fenster wird so breit wie möglich gewählt. Damit wird die Frequenzabtastung enger und die Abweichungen für nicht in das Fenster passende Frequenzen geringer. Nachteil: die Zahl der Abtastpunkte im Fenster wird größer und damit der Speicher- und Signalverarbeitungsaufwand.
2. Gibt es im Signal eine dominante Frequenz und ist diese bekannt, so wählt man die Fensterbreite so, dass sie ein Vielfaches der Periode der dominanten Frequenz ist. Ein Beispiel wäre eine Ultraschalldatenverarbeitung, bei der die Arbeitsfrequenz des Schwingers bekannt ist.

3. Es wird eine andere Form von Begrenzungsfenster gewählt. Im Allgemeinen handelt es sich um ein Fenster, dessen Grenzen gegen Null streben und damit das Signal an den Rändern des Ausblendungsbereichs einheitlich auf Null „abdrücken“. Damit werden die Signalwerte und auch erste und höhere Ableitungen an den Fenstergrenzen angeglichen, womit die Unstetigkeiten an den Stoßstellen verschwinden. Wir behandeln diese Methode im folgenden Kapitel.

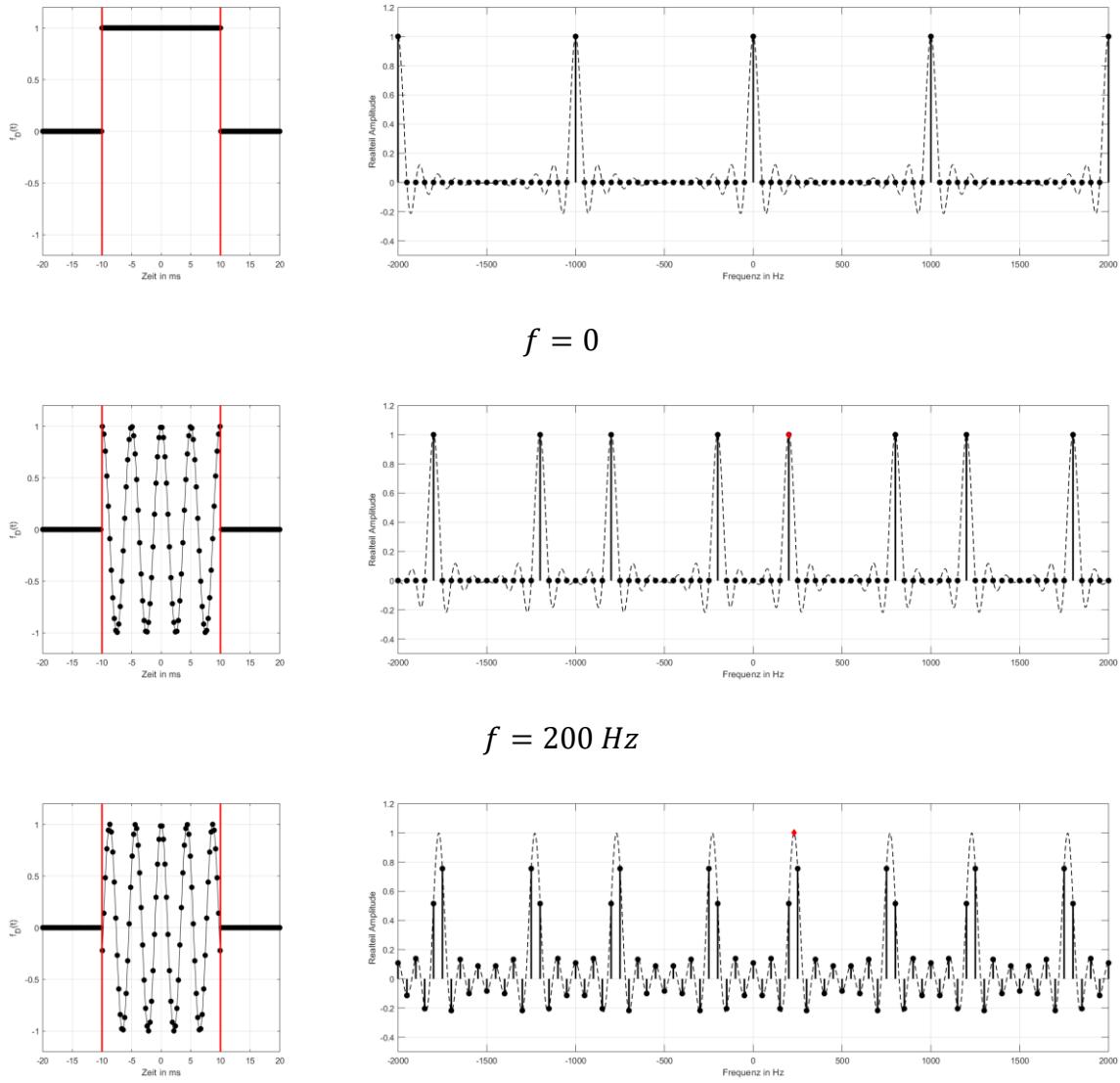


Abbildung 6.7: Abtastung unendlicher Signale innerhalb eines Zeitfensters (rote Berandung links). Ein konstantes Signal bzw. ein harmonisches Signal wurde mit 1 kHz abgetastet und durch ein 20 ms langes Zeitfenster begrenzt. Die linken Bilder zeigen das abgetastete Signal nach Multiplikation mit einem Rechteckfenster von 20 ms Breite. Die rechten Bilder zeigen das jeweils zugehörige Spektrum. Weitere Erklärungen im Text.

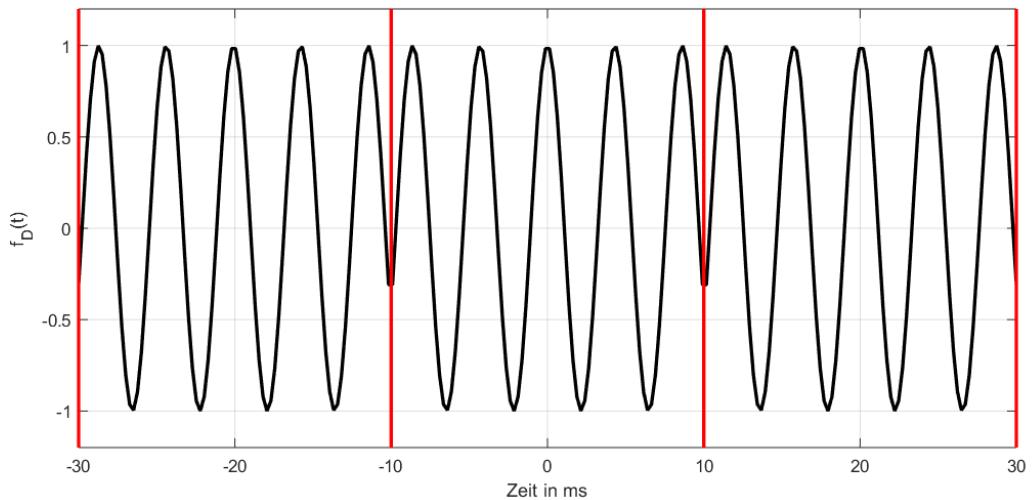
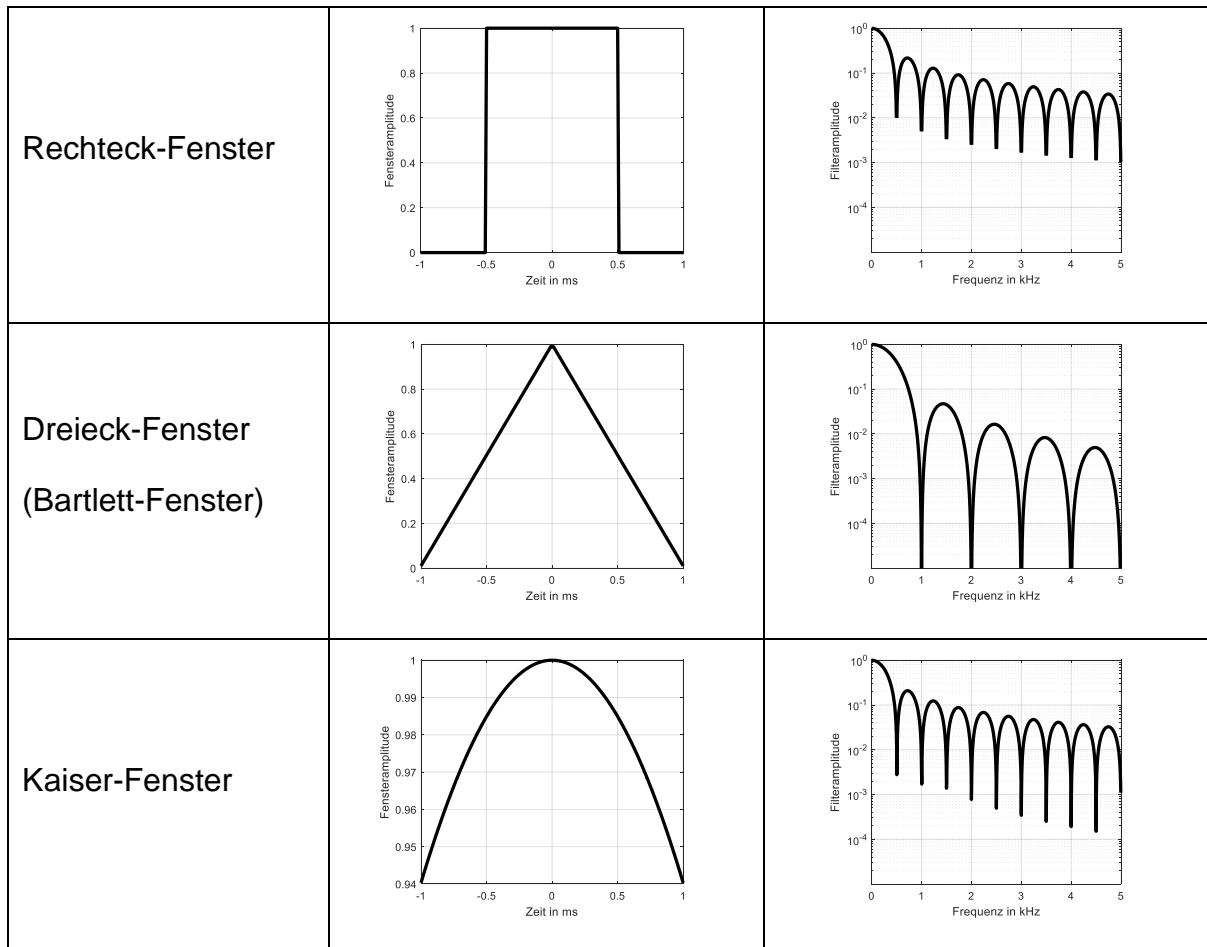


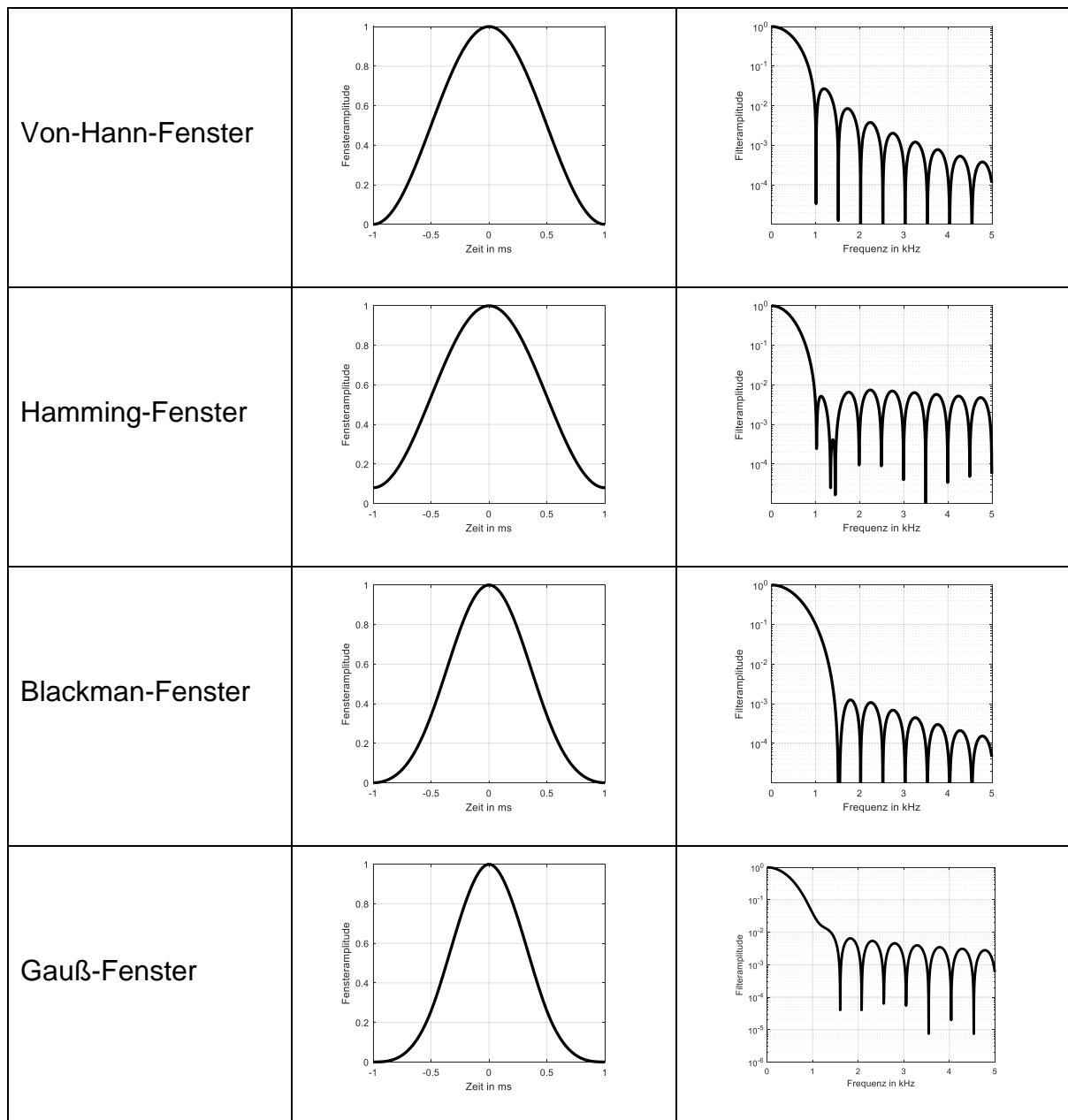
Abbildung 6.8: Durch die Fensterung mit  $\Delta T = 20 \text{ ms}$  wird eine  $20 \text{ ms}$  Sequenz aus dem ursprünglichen Signal (hier eine  $230 \text{ Hz}$  Harmonische) „zwangsperiodisiert“. Für Frequenzanteile, die keine Vielfachen der Grundfrequenz  $1/\Delta T$  des Fensters sind, führt dies zu diskontinuierlichen Funktionsverläufen an den Stoßstellen der periodischen Wiederholungen. Diese erzeugen Geisterfrequenzen im diskreten Fourier-Spektrum.

## 6.11 Fenster für die Begrenzung unendlicher Signale

Zur Vermeidung von Geisterfrequenzen werden verschiedene Fensterformen verwendet. Beispiele sind in Tabelle 6.3 aufgeführt. Allen Fensterformen ist gemein, dass sie an der linken und rechten Fenstergrenze jeweils gegen Null streben. Unterschiede gibt es vor allem im Verlauf der Ableitungen. Die in der Tabelle ebenso gezeigten Spektren der Fenster veranschaulichen deren Wirkung im Frequenzbereich. Dazu sind die spektralen Absolutwerte logarithmisch aufgetragen. Die Formung der Kanten des Fensters, also das allmähliche Zu-Null-Führen der Werte an den Rändern, bewirkt gegenüber dem Rechteck-Fenster, welches in der obersten Zeile der Tabelle als Vergleich angeführt ist, eine Absenkung der Stärke der Nebengipfel der Sinc-Funktion. Damit werden Geisterfrequenzen effektiv unterdrückt.

Tabelle 6.3: Fenster-Funktionen.





## 7 Analogsignalverarbeitung mit linearen zeitinvarianten Systemen

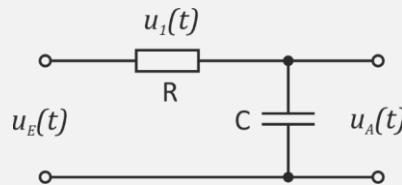
### 7.1 Das Antwortverhalten linearer zeitinvarianter Systeme

Lineare zeitinvariante Systeme transformieren ein Eingangssignal  $f(t)$  in ein Ausgangssignal  $g(t)$  durch Faltung. Da sehr viele analoge Signalverarbeitungskomponenten lineare zeitinvariante Systeme darstellen, ist die Faltung für die Signalverarbeitung und deren Systeme zentral. Wir werden die Begründung dafür zunächst im Zeitbereich anhand der Theorie gewöhnlicher inhomogener linearer Differenzialgleichungen erarbeiten. Ein lineares zeitinvariantes System kann grundsätzlich durch eine gewöhnliche inhomogene lineare Differenzialgleichung der allgemeinen Form

$$f(t) = a_0 g(t) + a_1 \frac{dg(t)}{dt} + a_2 \frac{d^2 g(t)}{dt^2} + \dots = \sum_i a_i \frac{d^i g(t)}{dt^i} \quad (7.1)$$

beschrieben werden. Wird dem System die Eingangsfunktion  $f(t)$  aufgeprägt, reagiert es mit der Ausgangsfunktion  $g(t)$ . Wir wollen dies anhand eines Beispiels aus der Elektrotechnik illustrieren.

#### Beispiel: RC-Glied



Wir legen an den Eingang der Schaltung die Spannung  $u_E(t)$  an. Durch den Widerstand fließt dann der Strom

$$i(t) = \frac{1}{R} u_1(t)$$

Wir unterstellen, dass kein Strom zum Ausgang abfließt (hochohmiger Ausgang). Damit muss  $i(t)$  als Verschiebestrom durch den Kondensator abfließen (1. Kirchhoffsches Gesetz). Der Verschiebestrom ist  $i(t) = C \frac{d}{dt} u_A(t)$ . Zusammengefasst ist also

$$\frac{1}{R} u_1(t) = C \frac{d}{dt} u_A(t).$$

Fernerhin ist  $u_E(t) = u_A(t) + u_1(t)$  (2. Kirchhoffsches Gesetz). Es folgt daraus die gewöhnliche inhomogene Differenzialgleichung

$$u_E(t) = RC \frac{d}{dt} u_A(t) + u_A(t).$$

Nun wollen wir Gleichung (7.1) in den Frequenzraum überführen. Mit Hilfe der Korrespondenzregel für Differenziale erhalten wir

$$F(\omega) = a_0 G(\omega) + a_1 j\omega G(\omega) + a_2 (j\omega)^2 G(\omega) + \dots = G(\omega) \sum_i a_i (j\omega)^i \quad (7.2)$$

Wir erkennen, dass aus den Differentialen im Frequenzraum ein komplexes Polynom entsteht, das sogenannte charakteristische Polynom. Es ist eine komplexe Funktion der Frequenz. Sein Kehrwert stellt die sogenannte **Übertragungsfunktion** des transformierenden Systems dar. Es gilt:

$$G(\omega) = \frac{1}{\sum_i a_i (j\omega)^i} F(\omega) = H(\omega) F(\omega). \quad (7.3)$$

Im Zeitbereich gilt entsprechend

$$h(t) = \mathcal{F}^{-1} \left\{ \frac{1}{\sum_i a_i (j\omega)^i} \right\}. \quad (7.4)$$

Im Zeitbereich bezeichnen wir  $h(t)$  als die **Systemantwortfunktion**.

**Beachte:** Für die Existenz der Fourier-Transformation muss das Fourier-Integral für die beteiligten Funktionen existieren. Ist dies nicht der Fall, so kann das Problem mittels der Laplace-Transformation behandelt werden. Die Ergebnisse sind aber sehr ähnlich.

## 7.2 Systemübertragungsfunktion und Systemantwortfunktion

Signale werden durch signalverarbeitende Systeme transformiert. Für analoge Signale sind dies oft elektronische Schaltungen mit linearem zeitinvariantem Antwortverhalten. Für diese gelten die oben abgeleiteten Beziehungen, die in Abbildung 7.1 noch einmal grafisch veranschaulicht sind.

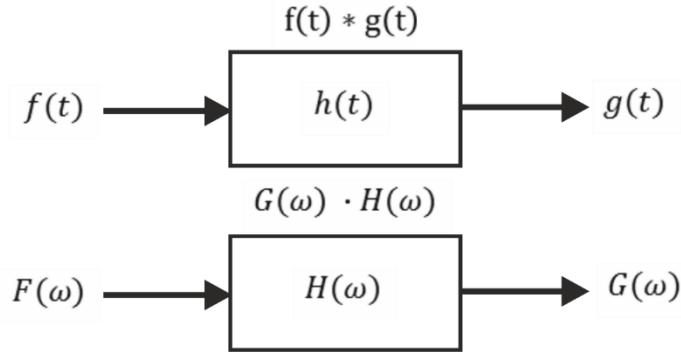


Abbildung 7.1: Die Systemübertragungsfunktion  $h(t)$  und Frequenzraumübertragungsfunktion  $H(\omega)$  eines signalverarbeitenden Systems.

Wie oben gezeigt wurde, lässt sich die Übertragungsfunktion über das charakteristische Polynom der systembestimmenden Differenzialgleichung ermitteln. Dafür muss diese Differenzialgleichung über Modellbildung ableitbar sein. Nicht immer ist dies möglich. Umgekehrt kann man den Weg der Systemidentifikation gehen, in dem man sich die linearen und zeitinvarianten Übertragungseigenschaften zunutze macht. Regt man das System mit einem breiten, ideal unendlich, gleichförmigen Spektralgemisch an, so liefert es die Übertragungsfunktion  $H(\omega)$ . Im Zeitbereich korrespondiert dies mit der Anregung mit einer Dirac-Distribution. Wie leicht zu sehen ist:

$$g(t) = \int_{-\infty}^{+\infty} h(t-t')\delta(t') dt' = h(t). \quad (7.5)$$

Das heißt, die Ausgangsfunktion entsteht durch Faltung der Eingangsfunktion mit der Systemantwortfunktion.

### 7.3 Die Laplace-Transformation

Die Laplace-Transformation kann man sich im Bereich der Signalverarbeitung als eine Erweiterung der Fourier-Transformation vorstellen. Sie hilft weiter, wenn das Fourier-Integral nicht konvergiert. Das ist für Signale selten der Fall, da diese in der Regel energetisch begrenzt sind. Es trifft aber oft auf allgemein zeitabhängige Funktionen zu. Daher brauchen wir die Laplace-Transformation eigentlich eher für Probleme der Analysis, etwa der Lösung von zeitabhängigen Differentialgleichungen. Wir wollen sie aber hier der Vollständigkeit halber mit einführen, insbesondere, da die am Schluss dieses Skriptes genutzte z-Transformation u. a. mit dieser im Zusammenhang steht.

Bei der Laplace-Transformation wird die Größe  $j\omega$  in der Fourier-Transformation durch die Größe  $s = \sigma + j\omega$  ersetzt. Aus der Harmonischen  $e^{j\omega}$  wird für  $\sigma < 0$  die gedämpfte Harmonische  $e^{\sigma+j\omega} = e^\sigma e^{j\omega}$ . Die Laplace-Transformation

$$F(s) = \int_0^{+\infty} f(t) e^{-st} dt. \quad (7.6)$$

ist also eine Fourier-Transformation, bei der die zu transformierende Funktion  $f$  beim Integrieren mit  $e^\sigma$  gedämpft wird. Sollte die Funktion  $f$  also nicht im Fourierschen Sinne integrabel sein, ist sie das sehr wahrscheinlich bei irgendeinem  $\sigma$ . Ein weiterer Unterschied besteht in der unteren Integrationsgrenze. Bei Laplace-Transformationen nehmen wir an, dass die zu transformierende Funktion als Zeitfunktion kausal ist und damit alle Werte für  $t < 0$  verschwinden. Abbildung 7.2 verdeutlicht die Konsequenzen in der komplexen Zahlenebene. Die Fourier-Transformation erzeugt eine Funktion für Argumente  $j\omega$ , also für alle Argumente auf der imaginären Achse der komplexen Zahlenebene. Die Laplace-Transformation erzeugt dagegen Funktionswerte für alle komplexen Zahlen. Für zwei Argumente mit gleichem Imaginärteil unterscheidet sich die Laplace-Transformierte nur um den Dämpfungsfaktor  $e^\sigma$ . Die Rücktransformation

$$f(t) = \frac{1}{2\pi j} \int_{a-j\infty}^{a+j\infty} F(s) e^{st} ds. \quad (7.7)$$

erfolgt dann auch entlang einer beliebigen Linie parallel zur imaginären Achse. Diese Linie muss so liegen, dass alle Pole der Funktion  $F(s)$  links von ihr liegen. Praktisch

wird die Rücktransformation über tabellierte Funktionen gemacht, da die komplexe Integration in Gl. (7.7) im Allgemeinen schwierig ist.

**Merke:** Wenn es gemeinsame Transformierte gibt, dann unterscheidet sich die Fourier-Transformation von der Laplace-Transformation tatsächlich nur durch die Ersetzung  $j\omega \Leftrightarrow s$ . Ebenfalls sind alle Korrespondenzregeln gleich.

Die Laplace-Transformierte gibt uns einen erweiterten Blick auf Übertragungsfunktionen. Hier hat sie ihre Stärke. Wir diskutieren dies an einem Beispiel.

Weiter unten in Tabelle 7.3 finden wir folgende Übertragungsfunktion für einen RC-Hochpassfilter:

$$H(\omega) = \frac{j\omega RC}{1 + j\omega RC}.$$

Nach dem oben gesagten, können wir hierfür äquivalent die Laplace-Transformierte mit

$$H(s) = \frac{sRC}{1 + sRC}$$

angeben. Das Bild unten zeigt links die Laplace-Transformierte für  $RC = 1$ . Die Fourier-Transformierte (rechts) ergibt sich aus den Werten entlang der imaginären Achse, also  $s = j\omega$ . Die ist im linken Plot rot hervorgehoben. Durch die Laplace-Transformierte sehen wir sehr viel besser, wie diese Form der Fourier-Transformierten eigentlich zustande kommt. Die Übertragungsfunktion hat genau eine Polstelle bei  $s = -1/RC$  sowie eine Nullstelle bei  $s = 0$ , also beide auf der reellen Achse. Durch die Nullstelle wird die Filterfunktion tatsächlich bei  $\omega = 0$  auf Null gedrückt, wie wir dies von einem Hochpassfilter erwarten. Die Polstelle hinter der Nullstelle sorgt dafür, dass die Funktionswerte mit  $\omega$  wieder „hochgedrückt“ werden. Funktionswerte zur Polstelle hin wachsen ins Unendliche.

Wir erkennen, dass die Lage der Null- und Polstellen das Verhalten unseres Filters bestimmt. Würden wir die Polstelle entlang der negativen Achse in Richtung  $-\infty$  verschieben, dann würde die Flanke des Filters zu größeren  $\omega$  verschoben, der Sperrbereich des Filters wäre größer! Physikalisch erreichen wir dies durch Verkleinern von  $RC$ . Wir könnten auch einen zweiten Pol hinzufügen, dann würde der Anstieg steiler werden und die Filtergüte wäre besser. Physikalisch erreichen wir dies durch Nachschalten einer weiteren RC-Filterstufe! Würden wir die Nullstelle aus dem Ursprung verschieben, würde das Filter bei  $\omega = 0$  Signal durchlassen. Dies alles verrät uns die wunderbare Laplace-Transformierte des Filters!

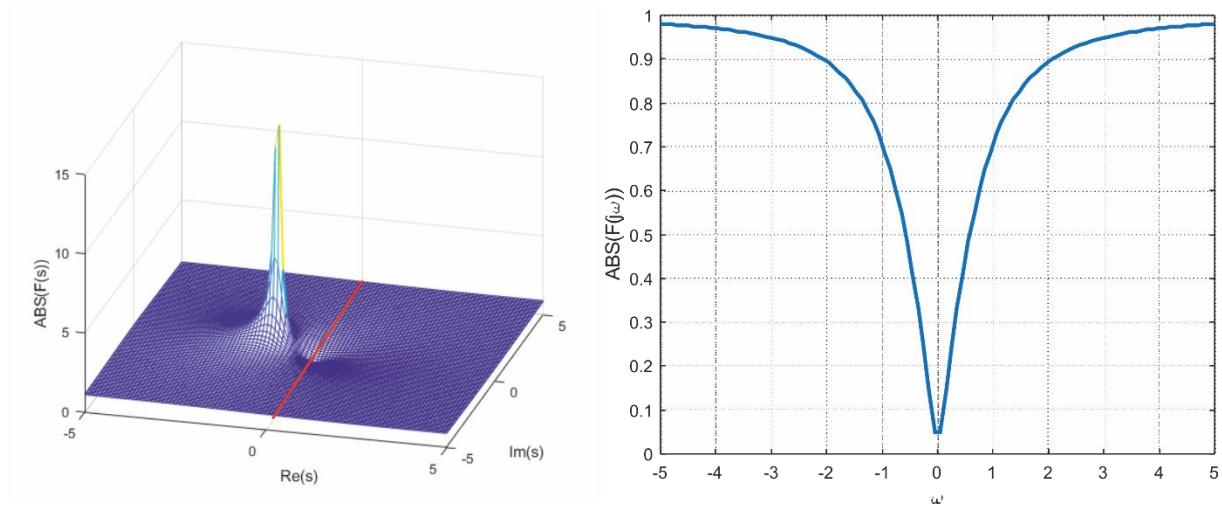


Abbildung 7.2: Laplace-Transformierte (links) und Fourier-Transformierte (rechts) der Hochpassfilterfunktion  $H(s) = 1/(1 + s)$ . Die rote Linie kennzeichnet die imaginäre Achse.

## 7.4 Analoge Filter

Ein Filter dient der Umformung eines Signals mit dem Zweck der Elimination unerwünschter Signalanteile im Frequenzraum. Bei linearen Filtern geschieht dies durch Multiplikation mit einer Filterfunktion im Frequenzraum. Entsprechend dem oben gesagten ist die Filterfunktion damit die Übertragungsfunktion des Filtersystems und die Operation selbst entspricht einer Faltung im Zeitbereich.

**Bemerkung:** Da wir im Folgenden oft physikalische Beispiele diskutieren, wird fortan die Frequenz  $f$  in Hz und nicht die Kreisfrequenz  $\omega$  in rad/s benutzt.

Bezgl. ihres Frequenzverhaltens werden im Allgemeinen vier Filtertypen unterschieden, deren praktische Bedeutung jeweils anhand eines Beispiels erläutert wird.

**Tiefpassfilter:** Es dient der Elimination hochfrequenter Signalanteile oberhalb einer Grenzfrequenz  $f_G$ .

**Beispiel:** Es wird die Temperatur eines Fluides mit Hilfe eines Widerstandsthermometers erfasst. Widerstandsthermometer sind thermisch träge, denn bei Änderung der Temperatur muss erst die gesamte Masse des Temperaturwiderstandes erwärmt oder abgekühlt werden. Hat das Widerstandsthermometer eine thermische Ansprechzeit  $\tau$ , so kann man davon ausgehen, dass Signalfrequenzen etwas oberhalb der Grenzfrequenz von  $f_G = 1/\tau$  kaum von Temperaturänderungen herrühren werden. Es ist also sinnvoll, alle Spektralanteile des Signals kurz oberhalb der Grenzfrequenz zu eliminieren. Ein weiteres Anwendungsbeispiel ist das Anti-Aliasing-Filter, das weiter unten beschrieben wird.

**Hochpassfilter:** Es dient der Elimination niederfrequenter Signalanteile unterhalb einer Grenzfrequenz  $f_G$  und insbesondere der Unterdrückung des Gleichanteils ( $f = 0$ ).

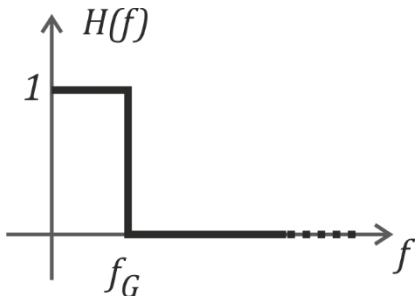
**Beispiel:** Ein Photodetektor soll kurze und schwache Impulse einer Lichtquelle empfangen und verarbeiten. Mittels eines Hochpassfilters kann der nur langsam veränderliche Umgebungslichtanteil (Tageslicht, Kunstlicht) im Signal unterdrückt werden.

**Bandpassfilter:** Es dient der Elimination von Signalanteilen außerhalb eines Frequenzbandes  $f_{G,u} < f < f_{G,o}$ .

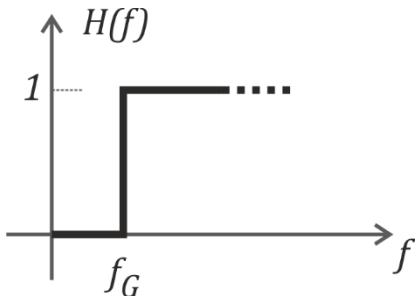
**Beispiel:** Ein Ultraschall-Doppler-System arbeitet mit Schallpulsen der zentralen Schallfrequenz  $f_z$ . Bei einem Dopplerstreuereignis kann die Frequenz des empfangenen Pulses nur innerhalb eines Bandes von  $f_z \pm f_{D,max}$  liegen. Die maximale Dopplerfrequenz ergibt sich zu  $f_{D,max} = f_z(1 - v_{max}/c)$  mit der maximal möglichen Objektgeschwindigkeit  $v_{max}$  und der Schallgeschwindigkeit  $c$ . Alle Frequenzen außerhalb des Bandes können eliminiert werden, sie tragen nichts zur Bestimmung der Objektgeschwindigkeit bei.

**Bandsperrfilter:** Es dient der Elimination von Signalanteilen innerhalb eines Frequenzbandes  $f_{G,u} < f < f_{G,o}$ .

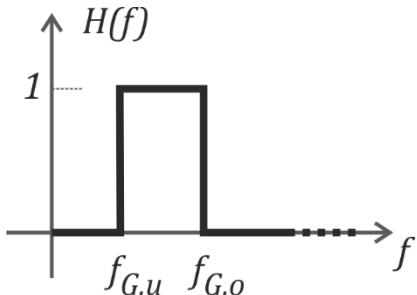
**Beispiel:** Koppelt in ein signalverarbeitendes System eine harmonische Störung bekannter Frequenz ein, so kann diese ggf. durch eine Bandsperre unterdrückt werden. Sehr prominent ist das Unterdrücken des 50 Hz Netzbrummens, welches oft über Spulen von Transformatoren induktiv in Signalleitungen eingekoppelt wird.



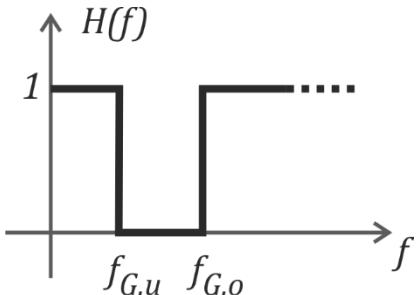
Ideales Tiefpassfilter



Ideales Hochpassfilter



Ideales Bandpassfilter



Ideales Bandsperrenfilter

Abbildung 7.3: Arten von Analogfiltern. Gezeigt sind ideale Filterformen.

Abbildung 7.3 zeigt ideale Formen der oben genannten Filtertypen. Die Bezeichnung ideal bezieht sich darauf, dass das Filter:

1. im Durchlassbereich alle Signalanteile unverändert lässt ( $H(f) = 1$ )
2. im Sperrbereich alle Signalanteile komplett löscht ( $H(f) = 0$ ) und
3. der Übergang zwischen Durchlassbereich und Sperrbereich sehr scharf erfolgt.

So stellt ein ideales Tiefpassfilter (Küpfmüller-Tiefpass) eine Rechteckfunktion im Frequenzraum dar. Die Bezeichnung „ideal“ sollte aber nicht missverstanden werden. Das Wort ideal bezieht sich auf die scharfe Frequenzausblendung. Praktisch ist so ein Tiefpassfilter kaum realisierbar und verfügt über eine keinesfalls ideale Systemübertragungsfunktion im Zeitbereich, da diese eine zeitlich sehr weit ausgedehnte Sinc-Funktion ist.

Fernerhin ist zu erkennen, dass sich Hochpassfilter und Tiefpassfilter leicht durch  $H_{HP} = 1 - H_{TP}$  ineinander überführen lassen. Selbiges gilt für Bandpass und Bandsperre. Weiterhin stellt ein Bandpassfilter ein frequenzverschobenes Tiefpassfilter und ein Bandsperrenfilter ein frequenzverschobenes Hochpassfilter dar. Wegen dieser Ähnlichkeiten werden die folgenden Diskussionen exemplarisch anhand des Tiefpassfilters geführt. Sie sind leicht auf die anderen Filtertypen übertragbar.

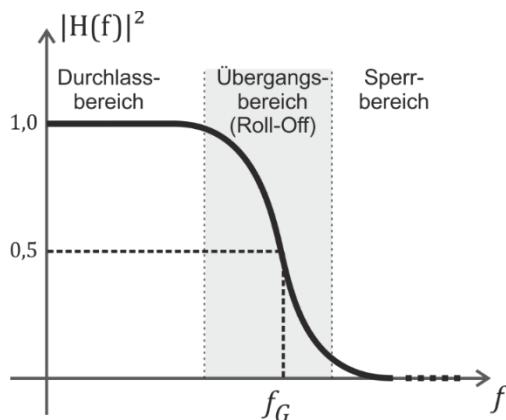
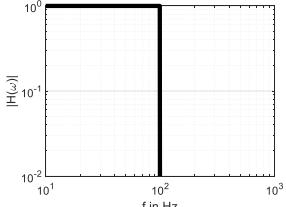
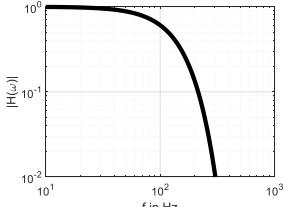
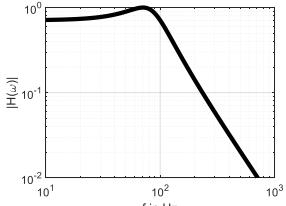
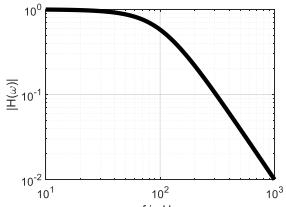
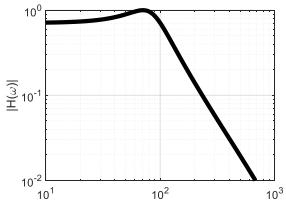
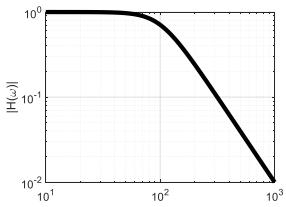


Abbildung 7.4: Reales Tiefpassfilter.

Reale Filter haben typischerweise einen Übergangsbereich zwischen Durchlassbereich und Sperrbereich (Abbildung 7.4). Die Weite des praktisch immer vorhandenen Übergangsbereichs charakterisiert die sogenannte Trennschärfe des Filters. Reale Filter werden in der analogelektronischen Signalverarbeitung durch passive und aktive Netzwerke mit Widerständen, Kondensatoren und Induktivitäten realisiert. In der digitalen Signalverarbeitung spielen hingegen mathematische Filterfunktionen eine Rolle. Wichtige mathematische Filterfunktionen sind in Tabelle 7.1 zusammengefasst.

Tabelle 7.1: Ausgewählte analoge Filterfunktionen

Name	Filterfunktion	Modulationsübertragungsfunktion
Küpfmüller-Filter (idealer Tiefpass)	$H(f) = \text{rect}(f; 2f_G)$	
Gauß-Filter	$ H(f)  = e^{-f^2/(2f_G)^2}$	
Tschebyscheff-Filter	$ H(f)  = \frac{1}{\sqrt{1 + \varepsilon^2 T_n^2(f/f_G)}}$	
Bessel-Filter	$ H(f)  = \frac{\Theta_n(0)}{\Theta_n(f/f_G)}$	
Cauer-Filter	$ H(f)  = \frac{1}{\sqrt{1 + \varepsilon^2 R_n^2(\xi, f/f_G)}}$	
Butterworth-Filter	$ H(f)  = \frac{1}{\sqrt{1 + (f/f_G)^{2n}}}$	
$\varepsilon$	-	Welligkeitsfaktor (hier 3 dB)
$\xi$	-	Empfindlichkeitsfaktor (hier 60 dB)
$n$	-	Filterordnung (hier 3)
$f_G$	-	Grenzfrequenz (hier 100 Hz)
$\Theta_n$	-	Bessel-Polynom n-ter Ordnung
$R_n$	-	Rational-elliptische Funktion n-ter Ordnung
$T_n$	-	Tschebyscheff-Polynom n-ter Ordnung

## 7.5 Rauschen, Rauschleistung und die Wirkung von Filtern

Technische Signale sind üblicherweise von Schwankungen begleitet, die man gemeinhin als Rauschen bezeichnet. Es gibt vielfältige Ursachen von Rauschen. Im Folgenden werden beispielhaft einige ausgewählte Rauscharten und Rauschprozesse genannt.

**Thermisches Rauschen** (auch Wärmerauschen, Widerstandsräuschen, Nyquist-Rauschen, Johnson-Rauschen oder Johnson-Nyquist-Rauschen): Es wird durch thermische Bewegung der Leitungselektronen in Leitern (Widerständen) erzeugt. Da jede Bewegung von Ladungen einen elektrischen Strom entspricht, entspricht auch die ungeordnete Bewegung von Elektronen einer chaotischen Verteilung kleiner Ströme und Spannungen, die sich überlagern.

**Schrotrauschen** (auch Poissonsches Schrotrauschen oder Schottky-Rauschen): ist das Rauschen was entsteht, wenn Ladungen eine Potenzialbarriere überwinden müssen. Da das Überwinden der Barriere ein stochastischer Prozess ist, äußert ich dieser in einer mikroskopisch schwankenden Stromstärke.

**Photonen-Rauschen:** Entsteht in Photodetektoren oder Strahlungsdetektoren bei geringen Licht- bzw. Strahlungsintensitäten. Der Lichtstrom schwankt auf Quantenebene durch die Statistik des Emissionsprozesses, welche immer eine Poisson-Statistik ist. Damit ist das Photonerauschen dem Schrotrauschen vergleichbar.

**Prozessrauschen:** Der durch einen Messwertaufnehmer beobachtete Prozess kann inhärenten Prozessschwankungen unterliegen. Sind diese stochastischer Natur, so kann die beobachtete Schwankung ebenfalls als Rauschen eingestuft werden. Ein Beispiel wäre die Messung der Fluidtemperatur in einem turbulent durchströmten Wärmeübertragerrohr. Diese ist von stochastischen Schwankungen begleitet, die aus der Natur des turbulenten Transports der Wärme von den Rohrwänden in das Fluid herrührt. Der Prozess ist stochastisch.

**Merke:** Rauschen ist kein Signal! Bei Abtastung des Rauschens kann jeder mögliche Zeitverlauf innerhalb vorgegebener energetischer Grenzen und Wahrscheinlichkeitsverteilungen vorkommen. Die einzige Kenntnis die man beim Rauschen bestenfalls hat, ist dessen spektrale Leistungsdichteverteilung.

Rauschen ist kein Signal! Es lässt sich einzig und allein über seine spektrale Leistungsdichte definieren. Als spektrale Leistungsdichte bezeichnen wir die Leistung pro Frequenzintervall. Die Rauschleistungsdichte wird in Watt/Hertz angegeben. Viele Rauschprozesse haben ein weißes Leistungsdichtespektrum. Das bedeutet, die Rauschleistung ist gleichmäßig über alle Frequenzen verteilt. Ist dem nicht so, sprechen wir von farbigem Rauschen (Abbildung 7.5).

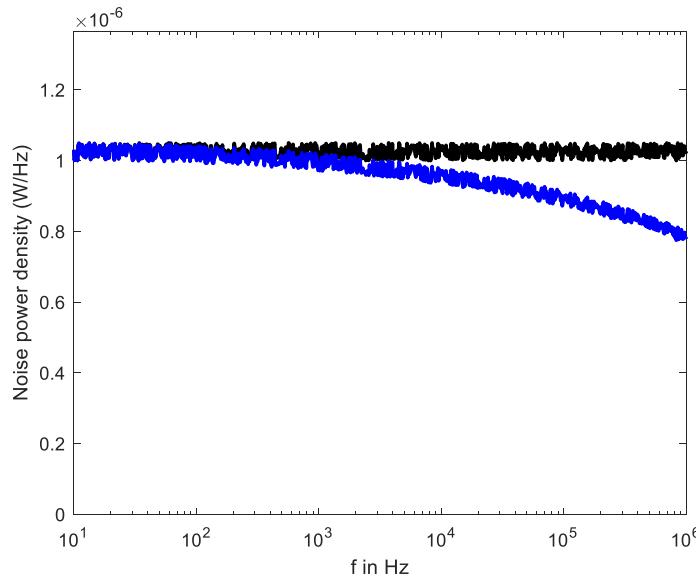


Abbildung 7.5: Leistungsdichtespektrum weißes Rauschen (schwarze Kurve) und farbiges Rauschen (blaue Kurve, schematisch).

**Beachte:** Das Rauschleistungsdichtespektrum charakterisiert das Rauschen im Frequenzbereich. Aus ihm lassen sich nur stochastische Aussagen im Zeitbereich ableiten, etwa, dass für bandbegrenztes farbiges Rauschen die Schwankungsamplitude im Zeitbereich bei höheren Abtastraten abnimmt. Es ist keinesfalls so, dass man durch Fourier-Rücktransformation des Rauschleistungsspektrums ein „Rauschsignal“ im Zeitbereich erhält!

Nach dem Wiener-Chintschin-Theorem ist die spektrale Leistungsdichte eines stationären Zufallsprozesses gleich der Fourier-Transformierten der korrespondierenden Autokorrelationsfunktion (Abschnitt 6.8). Es gilt für das Rauschleistungsdichtespektrum  $N(\omega)$ :

$$N(\omega) = \mathcal{F}\{r_{xx}(x(t))\} = R_{xx}(X(\omega)). \quad (7.8)$$

Zur Charakterisierung des Rauschens in Bezug auf ein Signal führen wir das sogenannte Signal-zu-Rausch-Verhältnis SNR als Verhältnis der Signalleistung zur Rauschleistung ein:

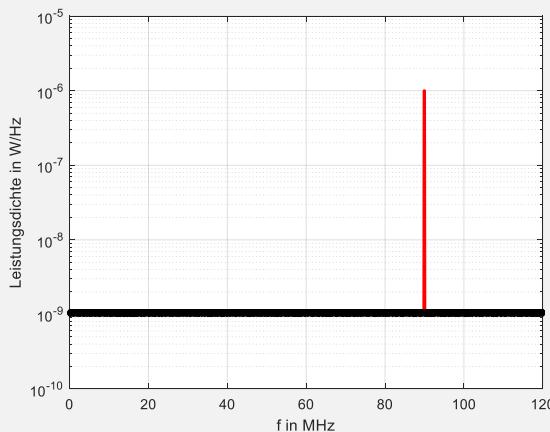
$$SNR = \frac{P_{sig}}{P_{rausch}}. \quad (7.9)$$

Filter helfen, dass SNR durch Eliminieren von Rauschanteilen außerhalb des Spektralbereichs des Signals zu verbessern. Wir diskutieren die Wirkung von Filtern anhand eines Beispiels aus der Rundfunktechnik.

**Beispiel:** Ein Rundfunksender sendet auf der UKW-Frequenz 92 MHz. Das Signal des Senders ist frequenzmoduliert und nimmt einen Bandbereich von  $\pm 20$  kHz um die Trägerfrequenz ein. Die Abbildung unten zeigt die Signal- und die Rauschleistungsdichte an einem Empfänger. Das Signal hat am Empfänger eine Gesamtleistung von  $10^{-6}W$ . Die Rauschleistungsdichte beträgt  $P(\omega) = 10^{-9} W/Hz$ . Im Empfänger sorgt ein Bandpassfilter für eine Elimination des Rauschens außerhalb des Empfängerbandes. Die im gefilterten Signal enthaltene Rauschleistung steht zu dem im gesamten Bereich integriertem Rauschen bis zur Nyquist-Frequenz (hier 100 MHz) im Verhältnis

$$\frac{P_{rausch,gefiltert}}{P_{rausch,ungefiltert}} = \frac{\int_{\omega_0-\omega_m}^{\omega_0+\omega_m} P(\omega)d\omega}{\int_0^{\omega_{Ny}} P(\omega)d\omega} = \frac{40kHz \cdot 10^{-9}W}{100MHz \cdot 10^{-9}W} = 2 \cdot 10^{-4}$$

$$SNR = 0.5 \cdot 10^4 \approx 37dB$$



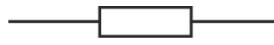
Rauschleistungsdichte (schwarze Kurve) und Signalfilterbereich (rot) eines Rundfunkempfängers.

## 7.6 Passive analoge Filter

Die Realisierung von Filtern in analogen elektronischen Signalverarbeitungsschaltungen erfolgt durch Verschaltung von Widerständen (Symbol  $R$ ), Kondensatoren (Symbol  $C$ ) und/oder Induktivitäten (Symbol  $L$ ). Schaltungen dieser Art werden oft entsprechend ihrer Symbolbezeichner benannt. So bedeutet etwa ein RC-Kreis eine Schaltung aus Widerständen und Kondensatoren. In aktiven Schaltungen kommen noch Operationsverstärker hinzu. In dieser Vorlesung werden dazu nur die wichtigsten Grundkenntnisse vermittelt. Für weiterführende Studien seien interessierte Studierende auf die elektrotechnische Literatur verwiesen.

Passive elektronische Schaltungen bestehen ausschließlich aus Widerständen, Kondensatoren und/oder Induktivitäten. Für sie gelten die in Tabelle 7.2 zusammengefassten Strom-Spannungs-Zusammenhänge, die nachfolgend noch einmal einzeln kurz erläutert werden.

Tabelle 7.2: Charakteristika passiver elektronischer Bauteile

	Ohmscher Widerstand	Kondensator	Induktivität
Symbol und Maßeinheit	$R$ in $\Omega$	$C$ in $F$	$L$ in $H$
Schalsymbol			
U/I-Zusammenhang im Zeitbereich	$u(t) = Ri(t)$	$i(t) = C \frac{du(t)}{dt}$	$u(t) = L \frac{di(t)}{dt}$
U/I-Zusammenhang im Frequenzbereich	$U(\omega) = R \cdot I(\omega)$	$I(\omega) = j\omega C \cdot U(\omega)$	$U(\omega) = j\omega L \cdot I(\omega)$
Impedanz	$Z_R(\omega) = \frac{U(\omega)}{I(\omega)} = R$	$Z_C(\omega) = \frac{1}{j\omega C}$	$Z_L(\omega) = j\omega L$

Für den Widerstand gilt das Ohmsche Gesetz. Strom und Spannung sind stets zueinander proportional.

In einem Kondensator sind zwei Elektroden durch ein nichtleitendes Dielektrikum getrennt. Liegt stationär eine Gleichspannung, so fließt kein Strom. Ändert sich die angelegte Spannung allerdings, so ändert sich das elektrische Feld im Kondensator

und es wird eine Kraft auf die Ladungen ausgeübt, die zu Ladungsverschiebungen führt. Es wird ein sogenannter Verschiebestrom  $i(t) = C du(t)/dt$  mit der Kapazität  $C$  induziert.

*Die **Kapazität** eines Kondensators gibt dessen Ladungsspeichervermögen an. Legt man eine Gleichspannung  $U$  an die Elektroden des Kondensator an, so speichert der Kondensator die Ladung*

$$Q = CU.$$

*Die Kapazität ergibt sich zu*

$$C = \varepsilon \frac{A}{d}$$

*mit der Permittivität  $\varepsilon$  des Dielektrikums, dem Plattenabstand  $d$  und der Plattenfläche  $A$ .*

Eine Spule besteht aus einem gewickelten elektrischen Leiter. Liegt eine Gleichspannung an, so fließt ein Strom. Die ideale Induktivität hat keinen Innenwiderstand, daher ist der Strom unendlich und der Spannungsabfall über der Spule Null. Ändert sich die Stromstärke in der Spule, ändert sich das Magnetfeld und es kommt zur Induktion einer Gegenspannung, die den Stromfluss begrenzt. Es gilt  $u(t) = L di(t)/dt$  mit der Selbstinduktivität  $L$ .

*Fließt durch eine **Spule** ein Gleichstrom  $I$  so erzeugt dieser einen magnetischen Fluss*

$$\Psi = LI.$$

*Die **Induktivität**  $L$  der Spule ist über geometrische Parameter und Materialparameter festgelegt. Für eine ideale lange Zylinderspule der Länge  $l$  und  $N$  gleichartigen Windungen mit Querschnittsfläche  $A$  gilt*

$$L = N^2 \mu A / l.$$

*Die Größe  $\mu$  ist die magnetische Permeabilität des Spulenkerne.*

Entsprechend der Korrespondenzregeln für die Fourier-Transformation lassen sich aus den differenziellen Beziehungen im Zeitbereich die in Tabelle 7.2 angegebenen Beziehungen im Frequenzbereich ableiten. Sie ähneln dem Ohmschen Gesetz. Die komplexen Widerstände  $Z_R$ ,  $Z_C$  und  $Z_L$  werden allgemein als Impedanzen bezeichnet. Im speziellen sprechen wir bei Kondensatoren und Induktivitäten von kapazitiver oder induktiver Reaktanz bzw. kapazitivem oder induktivem Blindwiderstand. Reaktanzen sind frequenzabhängig. Weiterhin verschieben sie die Signalphase zwischen Strom und Spannung um  $-90^\circ$  (Spule) bzw.  $+90^\circ$  (Kondensator). Die Bezeichnung Blindwiderstand resultiert aus der Tatsache, dass die Bauteile Signalspannungen nur

frequenzselektiv durchlassen und daher in bestimmten Frequenzbereichen wie Widerstände wirken. In der Tat wandeln sie aber keine elektrische Signalleistung in Wärme um.

Für den Aufbau von passiven und aktiven Analogfiltern gibt es eine große Anzahl möglicher Filtertopologien. Die wichtigsten sind in Abbildung 7.6 gezeigt. In der Vorlesung wollen wir uns auf die gebräuchlichsten Topologien beschränken. Für die passiven Filter ist dies die L-Topologie. Für diese schauen wir uns im Folgenden die Berechnung von Übertragungsfunktionen an. Für andere Passivfiltertopologien ist die Vorgehensweise dabei analog.

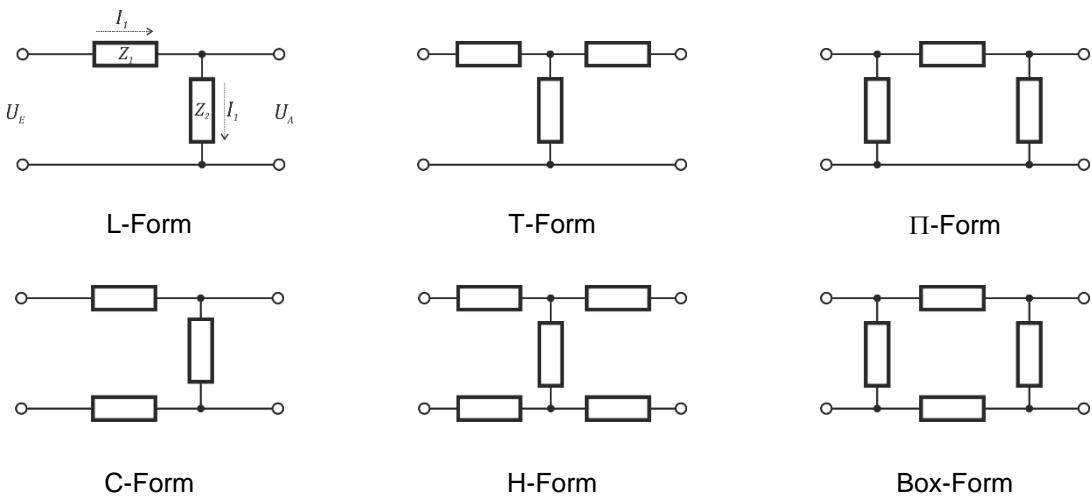


Abbildung 7.6: Topologien passiver Analogfilter.

Tabelle 7.3 zeigt typische einstufige analogelektronische L-Filter und deren Übertragungsfunktionen. Zu deren Berechnung greifen wir auf die Impedanzen aus Tabelle 7.2 und die Spannungsteilerregel zurück. Diese ergeben die Frequenzraumübertragungsfunktion:

$$H(\omega) = \frac{U_A(\omega)}{U_E(\omega)} = \frac{Z_2}{Z_1 + Z_2}. \quad (7.10)$$

**Herleitung:** Es sei auf die Bezeichnungen in Abbildung 7.6 (L-Form) verwiesen. Wir nehmen an, dass in den Ausgang des Filters kein Strom abfließt (hochohmiger Eingang der nachfolgenden Stufe). Dann fließt durch die beiden Bauteile ein Wechselstrom

$$I(\omega) = \frac{U_E(\omega)}{Z_1 + Z_2}$$

Durch die zweite Impedanz fließt der Wechselstrom

$$I_2(\omega) = \frac{U_A(\omega)}{Z_2}$$

Da  $I(\omega) = I_1(\omega) = I_2(\omega)$  gilt, folgt aus

$$\frac{U_E(\omega)}{Z_1 + Z_2} = \frac{U_A(\omega)}{Z_2}$$

Gl. (7.10).

Ein passives RC-Tiefpassfilter (Tabelle 7.3) hat demnach die Übertragungsfunktion

$$H(\omega) = \frac{Z_C}{R + Z_C} = \frac{\frac{1}{j\omega C}}{R + \frac{1}{j\omega C}} = \frac{1}{1 + j\omega RC}. \quad (7.11)$$

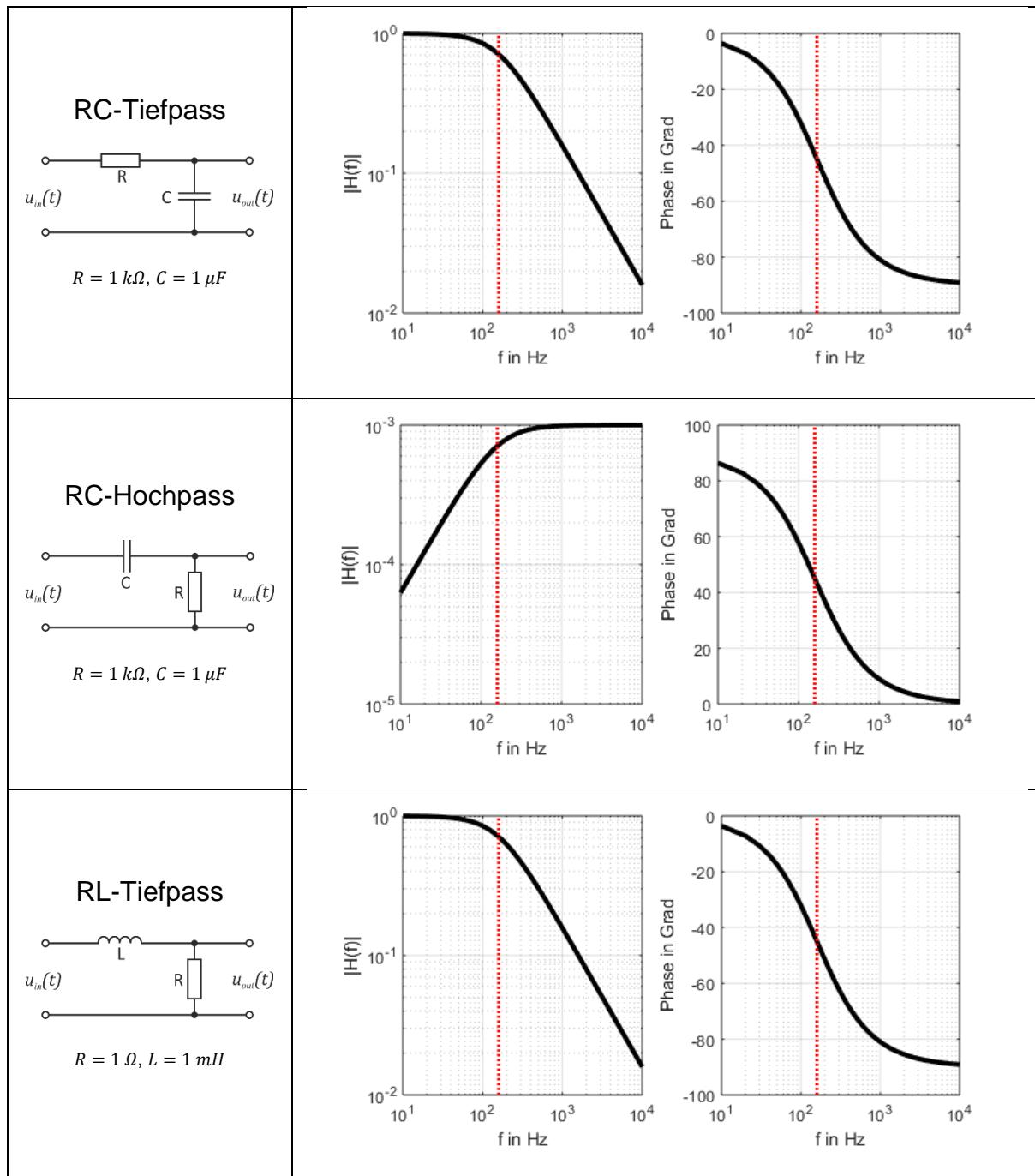
Die anderen Übertragungsfunktionen ergeben sich analog.

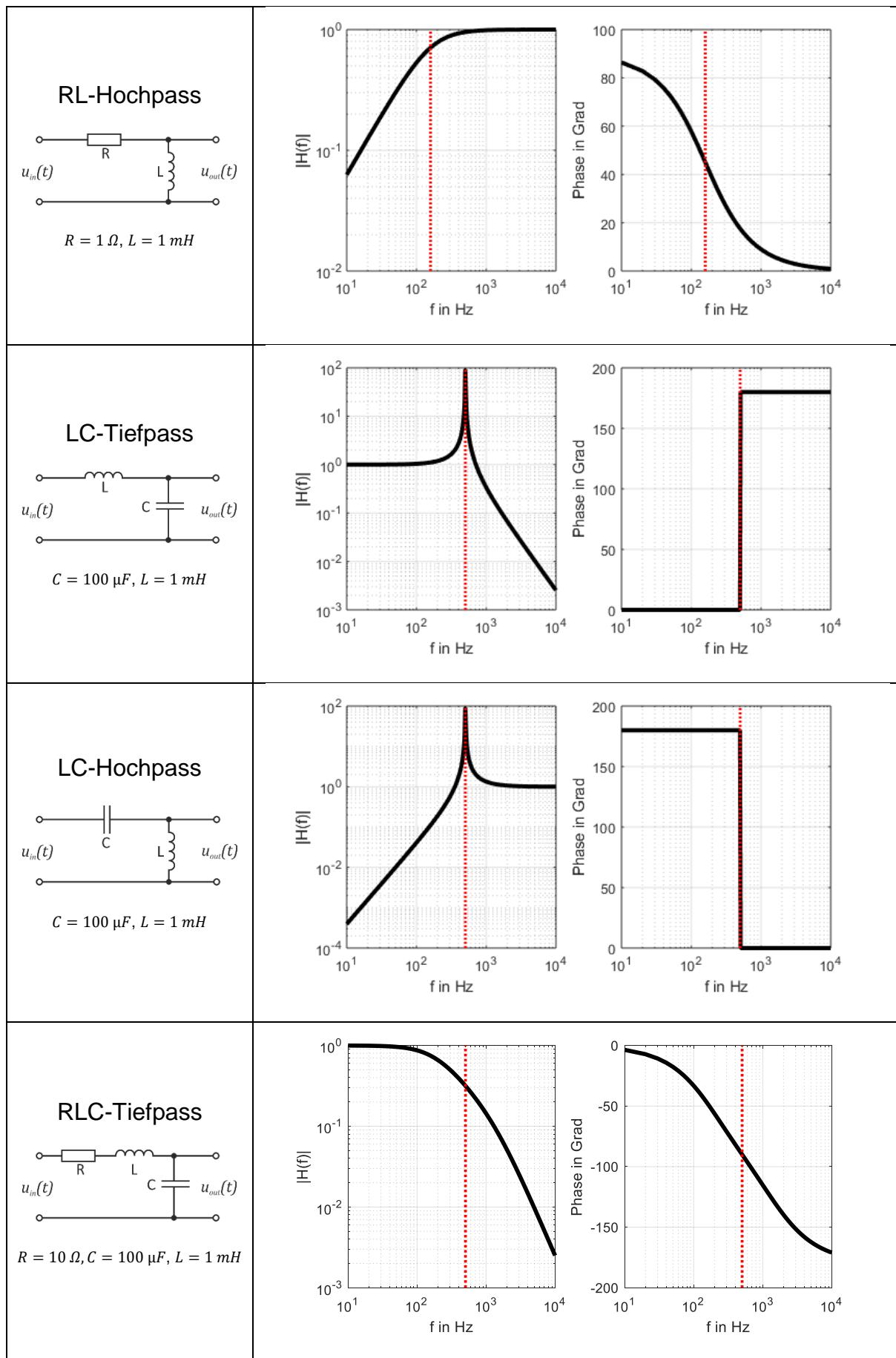
Tabelle 7.3: Einstufige passive analogelektronische Filter.

RC-Tiefpass		$H(\omega) = \frac{1}{1 + j\omega RC}$
RC-Hochpass		$H(\omega) = \frac{j\omega RC}{1 + j\omega RC}$
RL-Tiefpass		$H(\omega) = \frac{R}{R + j\omega L}$
RL-Hochpass		$H(\omega) = \frac{j\omega L}{R + j\omega L}$
LC-Tiefpass		$H(\omega) = \frac{\omega^2 LC}{\omega^2 LC - 1}$
LC-Hochpass		$H(\omega) = \frac{1}{1 - \omega^2 LC}$

Nachfolgend wird das Amplituden- und Phasenverhalten analoger Filter betrachtet.

Tabelle 7.4: Übertragungsfunktion (Amplitudenbetrag und Phase) passiver Filter.





Das RC-Tiefpassfilter erreicht als einstufiges Filter eine Frequenzunterdrückung von 20 dB pro Dekade. Die als rote Linie eingezeichnete Grenzfrequenz  $f_G$  ergibt sich aus der Definition, dass die Signalleistung auf 50% abfällt:

$$H^2(\omega) \stackrel{\text{def}}{=} \frac{1}{2} \quad (7.12)$$

Daraus folgt für den Tiefpass:

$$\frac{1}{2} = \left| \frac{1}{1 + j\omega_G RC} \right|^2 = \frac{1}{|1 + j2\pi f_G RC|^2}$$

$$1 + (2\pi f_G RC)^2 = 2 \quad (7.13)$$

$$f_G = \frac{1}{\sqrt{2\pi RC}}$$

Werden R und C vertauscht, ergibt sich ein Hochpassfilter mit gleichen Parametern. RL-Filter haben zu den RC-Filtern vergleichbare Charakteristiken. Da aber besonders im Niederfrequenzbereich große Induktivitäten benötigt werden, sind Schaltungen mit Induktivitäten eher im Hochfrequenzbereich zu finden.

Das LC-Filter stellt dagegen mit der Übertragungsfunktion

$$H(\omega) = \frac{j\omega L}{j\omega L + \frac{1}{j\omega C}} = \frac{\omega^2 LC}{\omega^2 LC - 1}. \quad (7.14)$$

ein Filter 2. Ordnung mit einer Dämpfung von 40 dB pro Dekade dar. Dies kommt durch die zwei Blindwiderstände, die den frequenzabhängigen Term quadrieren. Nachteilig ist die Resonanz bei der Grenzfrequenz  $f_G = \frac{1}{\sqrt{2\pi LC}}$ . Will man die Resonanz loswerden, muss diese mit Hilfe eines Serienwiderstandes bedämpft werden (RLC-Glied, Tabelle 7.4). Dies verbreitert aber auch wieder den Übergangsbereich des Filters.

Reicht die Steilheit des Übergangsbereichs eines einstufigen passiven Filters nicht aus, können passive Filter mehrstufig ausgeführt werden (Abbildung 7.7). Durch die Hintereinanderschaltung lässt sich die Übertragungsfunktion unmittelbar als Produkt ableiten:

$$H(\omega) = H_1(\omega) \cdot H_2(\omega) \cdot \dots \quad (7.15)$$

Die Modulationsübertragungsfunktion (Abbildung 7.7) zeigt dann eine Vergrößerung der Frequenzunterdrückung um 20 dB/Dekade pro Filterstufe in den Sperrbereichen, eine ebenso lineare Zunahme des Phasengangs sowie schärfere Übergangsbereiche für gleichartige Filterstufen. Mit verschiedenen dimensionierten Filterstufen lassen sich entsprechend Bandpass- oder Bandsperrenfilter realisieren.

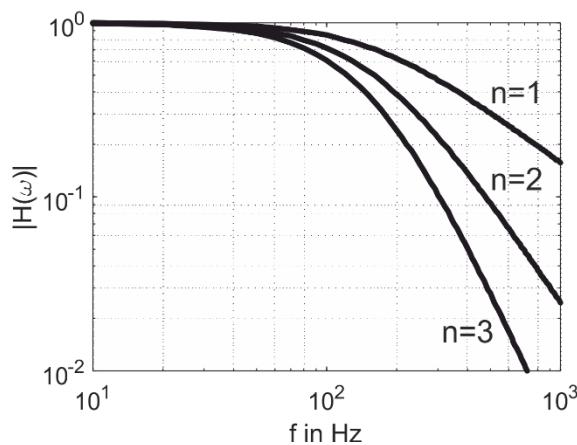
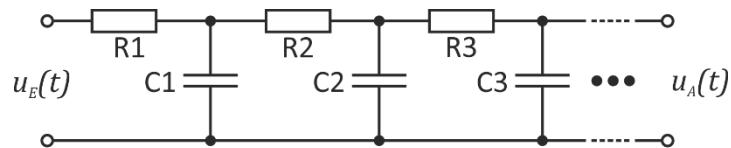


Abbildung 7.7: Mehrstufige RC-Filter und deren Amplitudenspektren.

## 7.7 Aktive analoge Filter

Passive Filter haben verschiedene Nachteile. Einmal ist ihre Genauigkeit durch die Bauelementetoleranz eingeschränkt. Weiterhin nimmt die Signalamplitude durch Ohmsche Verluste stets ab. Da in elektronischen Signalverarbeitungsstufen aber Signale ohnehin fast immer verstärkt werden, bietet es sich an, Operationsverstärker so zu beschalten, dass sie gleichzeitig der Amplitudenanpassung und der Filterung im Frequenzbereich dienen. Auch hier wollen wir das Prinzip nur schematisch erläutern und nicht in Details elektronischer Schaltungsentwicklung verfallen. Daher diskutieren wir aktive analoge Filter nur im Rahmen des Konzeptes eines idealen Operationsverstärkers.

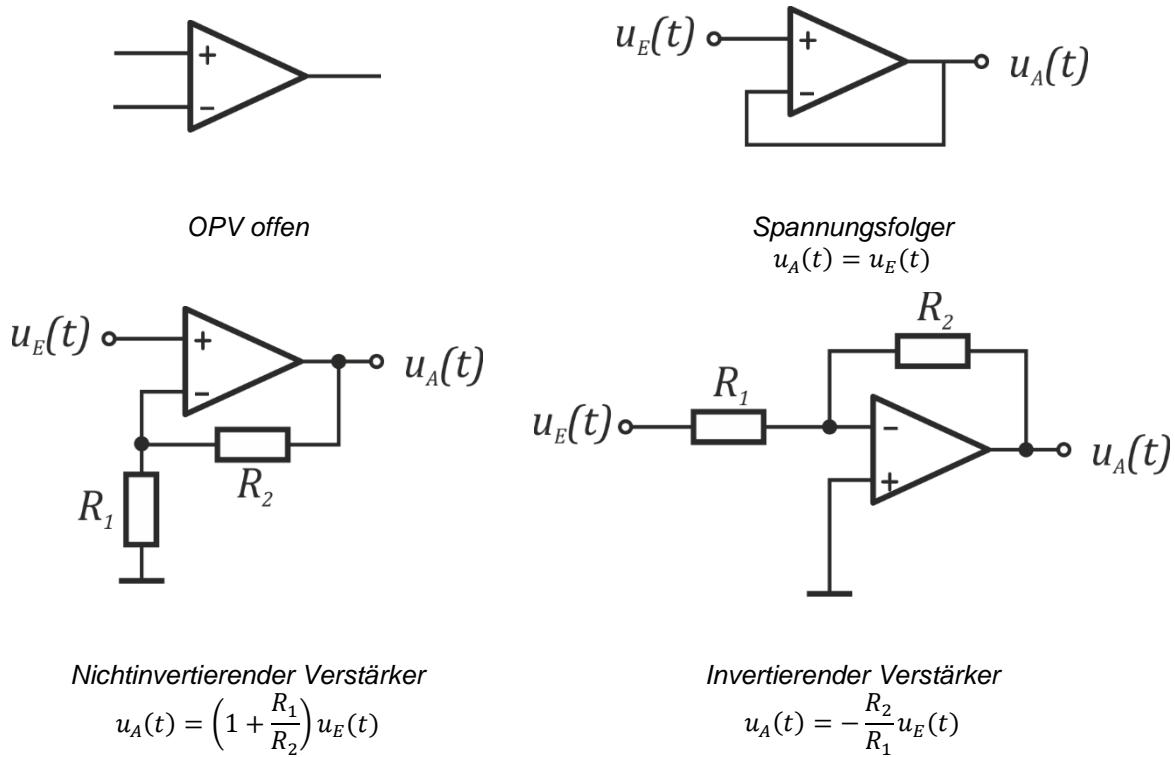


Abbildung 7.8: Unbeschalteter und beschalteter Operationsverstärker

Der ideale Operationsverstärker ist ein hypothetisches aktives Bauteil mit zwei Eingängen (nichtinvertierender Eingang +, invertierender Eingang -) mit folgenden Eigenschaften: Eine Spannungsdifferenz  $\Delta u = u_+ - u_-$  an den beiden Eingängen wird

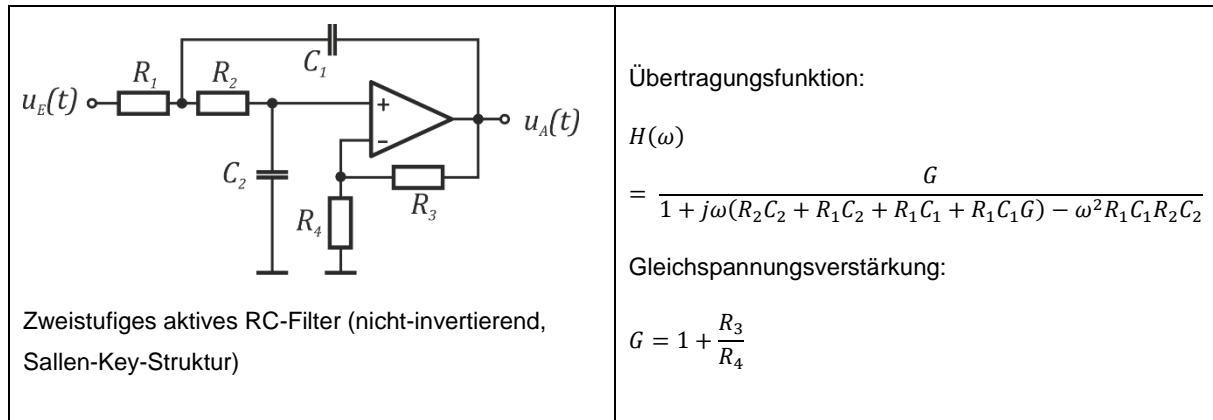
1. sofort und
2. mit unendlicher Verstärkung

auf den Ausgang übertragen. Der offene Operationsverstärker ist so praktisch nutzlos. Der Ausgangsspannung muss rückgekoppelt werden, um ein sinnvolles Verhalten zu produzieren. Dies ist in Abbildung 7.8 illustriert. Beim Spannungsfolger ist der Ausgang direkt auf den invertierenden Eingang zurückgekoppelt. Ändert sich die Eingangsspannung, so steigt die Ausgangsspannung ebenfalls sehr schnell an, aber nur so lange, bis der invertierende Eingang auf gleichem Potenzial, wie der nichtinvertierende Eingang ( $u_- = u_E$ ) liegt. Die Ausgangsspannung folgt der Eingangsspannung. Beim nichtinvertierenden Verstärker sorgt der Widerstandsteiler dafür, dass die Ausgangsspannung nur im Verhältnis  $u_- = R_2/(R_1 + R_2)u_A$  zurückgeführt wird. Der Ausgang des Operationsverstärkers muss also auf das  $(R_1 + R_2)/R_2$ -fache ansteigen, um die Eingangsspannung zu kompensieren. Entsprechend

wird also  $u_E$  um das  $(R_1 + R_2)/R_2$ -fache verstärkt. Der invertierende Verstärker ist so beschalten, dass die Spannung  $u_-$  am invertierenden Eingang Null sein muss. Das wird genau dann erreicht, wenn  $u_A = -(R_2/R_1)u_E$ . Die Ausgangsspannung ist also von umgekehrter Polarität und der Verstärkungsfaktor ist  $R_2/R_1$ . Auf elektrotechnische Details eines praktisch immer nichtidealen OPV (endliche Verstärkung, endliche Antwortzeit, Spannungsbegrenzung) soll hier nicht weiter eingegangen werden.

Tabelle 7.5: Aktive RC-Tiefpassfilter.

<p>Einstufiges aktives verstärkendes RC-Filter mit Filterglied im Eingangskreis</p>	<p>Übertragungsfunktion:</p> $H(\omega) = \left( \frac{1}{1 + j\omega R_1 C} \right) \left( 1 + \frac{R_3}{R_2} \right)$ <p>Spannungsverstärkung:</p> $G = 1 + \frac{R_3}{R_2}$
<p>Einstufiges aktives RC-Filter mit Filterglied im Rückkoppelkreis (nichtinvertierend)</p>	<p>Übertragungsfunktion:</p> $H(\omega) = 1 + \frac{R_3}{R_2} + j\omega R_1 C$ <p>Spannungsverstärkung:</p> $G = 1 + \frac{R_3}{R_2}$
<p>Einstufiges aktives RC-Filter mit Filterglied im Rückkoppelkreis (invertierend)</p>	<p>Übertragungsfunktion:</p> $H(\omega) = -\frac{R_2}{R_1} \frac{1}{1 + j\omega R_2 C}$ <p>Spannungsverstärkung:</p> $G = -\frac{R_2}{R_1}$



Analog zu den oben erläuterten passiven Filtern lässt sich nun ebenfalls die aktive Schaltung mit komplexen Impedanzen im Frequenzraum beschreiben. Wir wollen das beispielhaft für das erste Filter in Tabelle 7.5 zeigen.

**Beispielrechnung aktives RC-Filter 1. Ordnung mit nichtinvertierender Verstärkung:**

Zunächst berechnen wir die komplexe Eingangsspannung am nichtinvertierenden Verstärkereingang mit der Spannungsteilerregel:

$$U_+(\omega) = \frac{Z_C}{Z_{R1} + Z_C} U_E(\omega) = \frac{\frac{1}{j\omega C}}{R_1 + \frac{1}{j\omega C}} U_E(\omega) = \frac{1}{1 + j\omega R_1 C}$$

Die Ausgangsspannung des Verstärkers ergibt sich aus

$$U_A(\omega) = G U_+(\omega)$$

mit der Gleichspannungsverstärkung  $G = \left(1 + \frac{R_3}{R_2}\right)$ . Daraus folgt direkt die Übertragungsfunktion

$$H(\omega) = \frac{U_A(\omega)}{U_E(\omega)} = \left(1 + \frac{R_3}{R_2}\right) \frac{1}{1 + j\omega R_1 C}$$

## 8 Digitale Signalfilter

### 8.1 Nichtrekursive digitale Filter

Mit dem Fortschreiten der digitalen Signalverarbeitung werden digitale Filter immer beliebter und werden häufiger genutzt. Ein digitales Filter ist im Prinzip eine mathematische Berechnungsvorschrift, die eine Zahlenreihe (diskretes Signal) verarbeitet. Dabei ist die Verarbeitungsvorschrift einem digitalen Faltungsoperator nachempfunden. Digitale Filter werden in schnellen Digitalelektronikeinheiten verwendet. Beispiele sind Field Programmable Arrays (FPGA), Signalprozessoren und Mikrocontroller. Hier sind die Zahlenreihen Werte eines mit der Abtastfrequenz  $f_A = T_A^{-1}$  abgetasteten analogen Signals.

Es sei das Signal als Zahlenreihe mit Werten  $x_n$  gegeben. Ein nichtrekursives digitales Filter bildet daraus die Zahlenreihe

$$y_n = \sum_{k=0}^N a_k x_{n-k}. \quad (8.1)$$

Es werden also Zahlenwerte des Ausgangssignals durch gewichtete Summation der Zahlenwerte des aktuellen Wertes  $x_n$  sowie der zurückliegenden  $N - 1$  Werte des Eingangssignals gebildet. Die Gewichte sind durch die Filterkoeffizienten  $a_k$  festgelegt. Die Gleichung (8.1) stellt eine diskrete Faltungsoperation dar. Um die nachfolgende Behandlung des Themas formell etwas zu vereinfachen, verwenden wir anstelle von (8.1) fortan eine symmetrische Notation der Filterkoeffizienten

$$y_n = \sum_{k=-N}^N a_k x_{n-k}. \quad (8.2)$$

Regen wir das digitale Filter mit einem Einheitsimpuls der Form  $x_n = \{1 \text{ für } N = 0; \text{ sonst } 0\}$  an, so erhalten wir die Impulsantwort  $y_n = a_k x_{n-k} = a_k$ . Da die Filterlänge  $N$  endlich ist, ist auch die Filterantwort endlich. Wir sprechen von einem **Finite Impulse Response Filter (FIR-Filter)**.

Regt man das digitale Filter mit einer Harmonischen der Form  $x(nT_A) = e^{j\omega nT_A}$  an, so ergibt sich wegen der Linearität und Zeitinvarianz des Filters (Gl. (8.1)) ist eine Faltungsoperation!) die Ausgangsfunktion

$$y(nT_A) = \sum_{k=-N}^N a_k e^{j\omega(n-k)T_A}. \quad (8.3)$$

Daraus lässt sich die Modulationsübertragungsfunktion angeben:

$$H_D(\omega) = \frac{y(nT_A)}{x(nT_A)} = \frac{\sum_{k=-N}^N a_k e^{j\omega(n-k)T_A}}{e^{j\omega nT_A}} = \sum_{k=-N}^N a_k e^{-j\omega kT_A}. \quad (8.4)$$

Der Subskript steht D hier für „digital“. Wie man sieht, ist die Übertragungsfunktion identisch mit der Diskreten Fourier-Transformation der Filterkoeffizienten, wenn man formell eine unendliche Serie mit  $a_k = 0$  für  $|k| > N$  ansetzt:

$$H_D(\omega) = DFT(a_k) = \sum_{k=-\infty}^{\infty} a_k e^{-j\omega kT_A}. \quad (8.5)$$

Mit Gl. (8.5) folgt sofort eine Anleitung zur Konstruktion eines FIR-Filters. In der Regel wird man mit einem digitalen Filter die Modulationsübertragungsfunktion eines analogen Filters  $H_A(\omega)$  nachbilden wollen. Also ist es sinnvoll zu fordern, dass die Modulationsübertragungsfunktionen des analogen und des digitalen Filters im Sinne der kleinsten Fehlerquadrate möglichst minimal voneinander abweichen:

$$\frac{T_A}{2\pi} \int_{-\pi/T_A}^{+\pi/T_A} [H_D(\omega) - H_A(\omega)]^2 d\omega \stackrel{!}{\rightarrow} \text{Min.} \quad (8.6)$$

Wir wissen aus der Theorie der Fourier-Transformation, dass die Fourier-Reihe einer Funktion genau dies tut. Das heißt, die Fourier-Reihe einer Funktion approximiert diese im Sinne der kleinsten Fehlerquadrate. Für uns heißt das hier, dass die Fourier-Reihe aus Gl. (8.5) genau dies erfüllt. Mit anderen Worten, die Filterkoeffizienten sind als Fourier-Koeffizienten der abgetasteten Funktion  $H_A(\omega)$  zu bestimmen:

$$a_k = \frac{T_A}{2\pi} \int_{-\pi/T_A}^{+\pi/T_A} H_A(\omega) e^{j\omega k T_A} d\omega. \quad (8.7)$$

**Beispiel:** Wir möchten ein FIR-Filter konstruieren, dessen Übertragungsfunktion der eines idealen Tiefpass bis zur Grenzfrequenz  $\omega_G$  entspricht.

Der ideale Tiefpass ist durch  $H_A(\omega) = \text{rect}(\omega; \omega_G)$  gegeben. Wir finden

$$a_k = \frac{T_A}{2\pi} \int_{-\frac{\pi}{T_A}}^{+\frac{\pi}{T_A}} \text{rect}(\omega; \omega_G) e^{j\omega k T_A} d\omega.$$

Da es sich bei  $\text{rect}(\omega; \omega_G)$  um eine gerade symmetrische Funktion handelt, gilt weiter

$$\begin{aligned} a_{|k|} &= \frac{T_A}{\pi} \int_0^{+\frac{\pi}{T_A}} \text{rect}(\omega; \omega_G) \cos(\omega k T_A) d\omega \\ a_{|k|} &= \frac{T_A}{\pi} \int_0^{\omega_G} \cos(\omega k T_A) d\omega \\ a_{|k|} &= \frac{T_A}{\pi} \frac{\sin(\omega_G k T_A)}{\omega_G k T_A} = 2 \frac{f_G}{f_A} \text{sinc}\left(2\pi \frac{f_G}{f_A} k\right). \end{aligned}$$

Erwartungsgemäß sind die Filterkoeffizienten die Funktionswerte der sinc-Funktion (Fourier-Transformierte der Rechteckfunktion) an den Vielfachen der Frequenzabtastpunkte  $f_G/f_A$ .

## 8.2 Die z-Transformation

Für die Diskussion rekursiver Filter im Folgenden ist es passend, kurz auf die z-Transformation einzugehen. Mit dieser lassen sich rekursive Filter besser beschreiben und auslegen. Die z-Transformation ist der Fourier- und der Laplace-Transformation sehr ähnlich. Mit ihr wird eine Zeitreihe in den Raum der komplexen Variablen

$$z = e^{j\omega T_A} \quad (8.8)$$

überführt. Wozu das gut ist, werden wir sofort sehen. Prägen Sie sich hier aber auch den Unterschied zur Fourier-Transformation und zur Laplace-Transformation ein. Die Fourier-Transformation transformiert eine Funktion in den Raum der Frequenzvariablen  $j\omega$ , die Laplace-Transformation in den Raum der gedämpften Frequenzvariablen  $s = \sigma + j\omega$ .

Der Ausdruck (8.8) steht für eine Zeitverschiebung eines Signals von der Größe eines Abtastintervalls  $\Delta T = T_A$  im Fourier-Raum:

$$f(t + T_A) \circledast e^{j\omega T_A} F(\omega) = zF(\omega). \quad (8.9)$$

Entsprechend steht

$$f(t + nT_A) \circledast e^{jn\omega T_A} F(\omega) = z^n F(\omega). \quad (8.10)$$

für eine  $n$ -fache Verschiebung. Die z-Transformation transformiert nun mit Hilfe dieser Verschiebungsterme das diskrete Signal  $x_n = x(nT_A)$  in den Raum der komplexen Variablen  $z$ :

$$\hat{X}(z) = \sum_{n=-\infty}^{+\infty} x_n z^n \quad (8.11)$$

Obwohl es sich um eine andere Transformationsvariable handelt, ist die z-Transformation zur Fourier-Transformation sehr ähnlich. Insbesondere gelten alle Korrespondenzen in analoger Form. Abbildung 8.1 zeigt die Verhältnisse in der komplexen Zahlenebene.

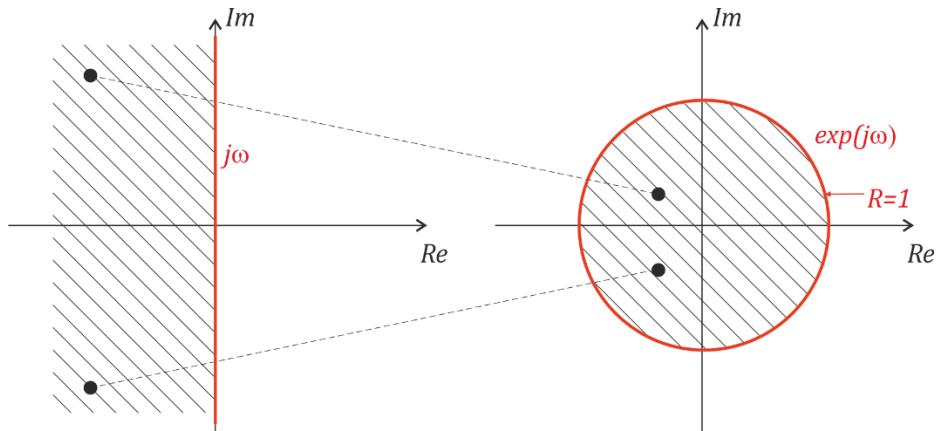


Abbildung 8.1: Komplexe Zahlenebene mit dem Stabilitätsgebiet der Laplace-Transformation (links) und der z-Transformation (rechts) sowie der Trajektorie der Transformationsvariablen  $j\omega$  der Fourier-Transformation (links) und  $e^{j\omega T_A}$  der z-Transformation. Das schraffierte Gebiet der linken komplexen Zahlenebene ist das Stabilitätsgebiet der analogen Übertragungsfunktionen. Dieses kann durch eine affine Transformation in den Einheitskreis der z-Transformation übertragen werden. Das heißt, übertragen werden die Pole und Nullstellen, hier durch die zwei schwarzen Pole angedeutet. Auf diesem Wege werden analoge Übertragungsfunktionen diskretisiert.

Die z-Transformation mit  $z = e^{j\omega T_A}$  transformiert die Zeitreihe  $x_n$  in eine Funktion  $\hat{X}(z)$  auf dem Einheitskreis, so wie die Fourier-Transformation mit ihrer Transformationsvariablen  $j\omega$  die Funktion  $x(t)$  in eine Funktion entlang der imaginären

Achse der komplexen Zahlenebene transformiert. Daraus wird sofort ersichtlich, dass die  $z$ -Transformation periodische Spektren beschreibt, denn die Spektralwerte wiederholen sich mit  $2\pi f_A$  auf dem Einheitskreis. Übertragungsfunktionen, die mit der Laplace-Transformation ermittelt werden, sind vollständig durch ihre Nullstellen und Polstellen in der komplexen Zahlenebene bestimmt. Aus Stabilitätsgründen müssen die Polstellen links der imaginären Achse (schraffiertes Gebiet links) liegen. Durch die bilineare Transformation

$$s = \frac{2}{T_A} \frac{1 - z^{-1}}{1 + z^{-1}} \Leftrightarrow z = \frac{1 + \frac{sT_A}{2}}{1 - \frac{sT_A}{2}} \quad (8.12)$$

können Polstellen und Nullstellen aus dem Gebiet der komplexen Zahlenebene für analoge Funktionen in das Gebiet der komplexen Zahlenebenen für diskrete Funktionen übertragen werden. Auf diesem Weg lassen sich Polynomkoeffizienten für rekursive Filter bestimmen.

### 8.3 Rekursive digitale Filter

Ein rekursives digitales Filter kann als ein nichtrekursives digitales Filter betrachtet werden, welches um ein „Gedächtnis“ erweitert wurde:

$$y_n = \sum_{k=0}^N a_k x_{n-k} - \sum_{k=1}^M b_k y_{n-k}. \quad (8.13)$$

Der „Gedächtnisterm“ ist steht ganz rechts. Das Ausgangssignal wird nunmehr nicht allein direkt aus dem Eingangssignal berechnet sondern auch aus den zurückliegenden Werten  $y_{n-k}$  der Ausgangsfunktion. Wegen dieses Rückbezugs („alte“ Ausgangswerte gehen in die Berechnung des aktuellen Ausgangswertes ein) wird das Filter rekursiv genannt. Führt man formell in Gleichung (8.13) noch  $y_n = b_0 y_n$  mit  $b_0 = 1$  ein, so wird die Filtergleichung symmetrisch bezgl. der  $x_n$  und  $y_n$ :

$$\sum_{k=0}^M b_k y_{n-k} = \sum_{k=0}^N a_k x_{n-k}. \quad (8.14)$$

Die Impulsantwort des rekursiven digitalen Filters lässt sich wie folgt angeben. Nur zum ersten Abtastzeitpunkt  $n = 0$  ist  $x_n = 1$ , sonst 0. Weiterhin sei für alle  $n < 0$   $y_n = 0$ . Daraus ergibt sich

$$\begin{aligned} y_0 &= a_0 \\ y_1 &= a_1 - b_1 y_0 = a_1 - b_1 a_0 \\ y_2 &= a_2 - b_2 y_1 - b_1 y_0 = a_2 - b_2(a_1 - b_1 a_0) - b_1 a_0 \\ &= a_2 - b_2 a_1 + b_2 b_1 a_0 - b_1 a_0 \end{aligned} \quad (8.15)$$

$$y_2 = \dots$$

Man erkennt, dass die Impulsantwort durch den Rückbezug auf den Ausgang niemals auf Null zurückgeht. Das Filter hat eine unendliche lange Antwort und wird deshalb als **Infinite Impulse Response Filter (IIR Filter)** bezeichnet.

Zur Übertragungsfunktion des IIR-Filters gelangen wir durch z-Transformation von Gleichung (8.14):

$$\sum_{n=-\infty}^{\infty} \left( \sum_{k=0}^M b_k y_{n-k} \right) z^{-n} = \sum_{n=-\infty}^{\infty} \left( \sum_{k=0}^N a_k x_{n-k} \right) z^{-n}. \quad (8.16)$$

Nun werden wieder die Summationsgrenzen durch Nullsetzen der übrigen Koeffizienten erweitert

$$\sum_{n=-\infty}^{\infty} \left( \sum_{k=-\infty}^{\infty} b_k y_{n-k} \right) z^{-n} = \sum_{n=-\infty}^{\infty} \left( \sum_{k=-\infty}^{\infty} a_k x_{n-k} \right) z^{-n}. \quad (8.17)$$

und die Summanden nach Potenzen von  $z$  sortiert:

$$\sum_{k=-\infty}^{\infty} b_k z^{-k} \sum_{m=-\infty}^{\infty} y_m z^{-m} = \sum_{k=-\infty}^{\infty} a_k z^{-k} \sum_{m=-\infty}^{\infty} x_m z^{-m}. \quad (8.18)$$

Dies ergibt zusammengefasst

$$\hat{B}(z)\hat{Y}(z) = \hat{A}(z)\hat{X}(z) \quad (8.19)$$

und damit die Übertragungsfunktion

$$\hat{H}(z) = \frac{\hat{X}(z)}{\hat{Y}(z)} = \frac{\hat{A}(z)}{\hat{B}(z)}. \quad (8.20)$$

Damit ist das rekursive Filter beschrieben. Beim Design eines Filters wählt man die gewünschte analoge Übertragungsfunktion und wandelt deren Zähler- und Nennerpolynome mittels der bilinearen Transformation (Gleichung (8.12)) in die Polynome  $\hat{A}(z)$  und  $\hat{B}(z)$  um.

#### 8.4 Blockschaltbilder nichtrekursiver und rekursiver Filter

Die algorithmische Vorgehensweise bei der Berechnung von Filterantworten aus diskreten Eingangssignalen lässt sich anschaulich mit dem in Abb. gezeigten Blockdiagramm darstellen. Die mit  $z^{-1}$  gekennzeichneten Elemente stellen Verzögerungsglieder der Dauer  $T_A$  dar, die Dreiecke die Filterkoeffizienten. Der obere Teil des Blockschaltbildes steht für ein nichtrekursives (FIR) Filter, der obere und untere Teil gemeinsam bilden ein rekursives (IIR) Filter.

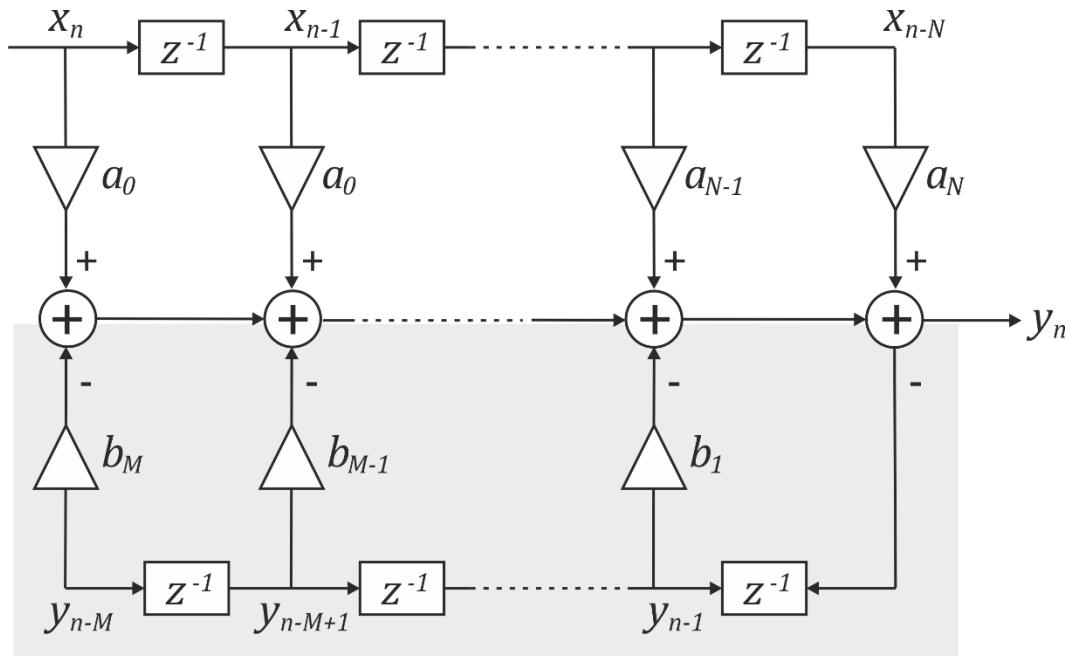


Abbildung 8.2: Blockschaltbild nichtrekursiver und rekursiver Filter. Beim nichtrekursiven Filter entfällt der grau hinterlegte Teil.