

III. Erfassen physikalischer Signale

Nachdem wir bisher die Eigenschaften von Signalen und von Rauschen bzw. Störungen in idealisierter Form behandelt haben, soll im folgenden der „reale“ physikalische Meßprozeß im Vordergrund stehen. Bei einem derartigen Meßprozeß kann man die im folgenden skizzierten grundsätzlichen Elemente zumeist abgrenzen:

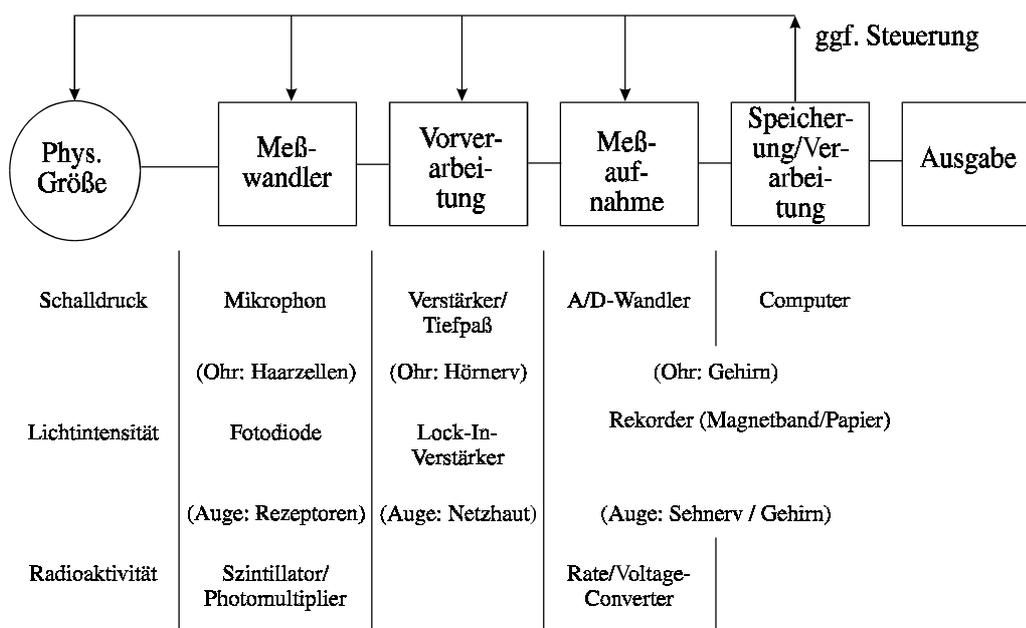


Abb. 3.1: Typischer Meßaufbau

Die zu bestimmende physikalische Größe kann beispielsweise der Schalldruck, die Lichtintensität, der Zerfallprozeß einer radioaktiven Substanz oder andere mikroskopische oder makroskopische Größen sein. Diese Größe wird mit einem geeigneten Sensor (Meßwandler bzw. Meßaufnehmer) zumeist in eine elektrische Spannung umgesetzt, die dann weiterverarbeitet bzw. aufbereitet wird. Zur Auswertung muß dieses elektrische Signal dann aufgenommen werden, um dann gespeichert, weiterverarbeitet oder z. B. graphisch ausgegeben zu werden. Diese Weiterverarbeitung kann auch die vorhergehenden Elemente der Meßanordnung beeinflussen, z. B. zur Einstellung eines geeigneten Meßbereichs oder zur Einstellung der geeigneten Tiefpaßfrequenz eines Filters.

In oben stehender Tabelle sind neben den technischen Systemen auch in einem groben Vergleich die Sinnesorgane aufgeführt, die beim Menschen zu einer ähnlichen Erfassung von physikalischen Größen benutzt werden. Bei dem Schalldruck sind neben dem technischen System des Mikrophons die im Innenohr angesiedelten Haarzellen zu betrachten, die mechanische Bewegungen in Rezeptorpotentiale umwandeln. Analog zu

dem einem Mikrophon nachgeschaltetem Verstärker und Tiefpaßfilter (zur Unterdrückung des Rauschens) kann beim Ohr als Vorverarbeitungsstufe der Hörnerv angesehen werden, bei dem ein Großteil der im akustischen Signal enthaltenden Schallinformation bereits entfernt worden ist, so daß bereits eine Informationsreduktion stattgefunden hat. Den bei der Schallaufnahme vom Computer mit AD-Wandler übernommenen Teil würde im Fall des Ohrs das Gehirn spielen.

Bei der Messung der Lichtintensität wäre das der (technischen) Fotodiode analoge System die Fotorezeptoren des Auges, die Lichtenergie in neuronale Impulse umwandeln. Die Vorverarbeitung, die im Fall der Fotodiode durch einen Lock-In Verstärker (s. u.) zur Störgeräuschreduktion durchgeführt wird, würde beim Auge durch die neuronale Verschaltung der Netzhaut bewerkstelligt werden, während die Einheit aus Sehnerv und Gehirn wieder dem technischen System aus AD-Wandler und Computer bzw. Magnetband-Rekorder oder Papierstreifen-Rekorder analog wäre.

Aus der Analogie-Betrachtung zwischen einem physikalischen Meßsystem und einem physiologischen Sinnesorgan kann zum einen die Parallele zwischen dem physikalischen Meßprozeß und dem für das Überleben des Individuums notwendigen ständigen Erfassens der Umwelt gezogen werden. Andererseits kann bei ausgewählten physikalischen Meßvorgängen (z. B. Erfassung des Schalldrucks und der Lichtintensität) anhand des Vergleichs mit dem Auge bzw. dem Ohr festgestellt werden, wie sehr unsere zur Verfügung stehenden Techniken noch hinter der Leistungsfähigkeit unserer Sinnesorgane hinterherhinken. Im folgenden sollen die Elemente dieser prinzipiellen Meßanordnung anhand einiger exemplarischer Beispiele näher behandelt werden.

III.1 Beispiel: Mikrophon

Als erstes Beispiel für einen physikalischen Meßwandler sei das dynamische Mikrophon betrachtet. Bei ihm wird durch den Schalldruck eine Membran in Schwingungen versetzt, die mit einer konzentrisch in einem homogenen Magnetfeld aufgehängten Schwingspule verbunden ist. Durch Bewegung des Leiters in dem Magnetfeld wird eine Spannung induziert, die der Magnetfeldstärke B und der Länge der Leiterschleife l proportional ist. Diese Spannung wird abgegriffen und als Ausgangsspannung des Mikrophons verwendet, die ein Maß für den Schalldruck bzw. die Schallschnelle darstellen soll.

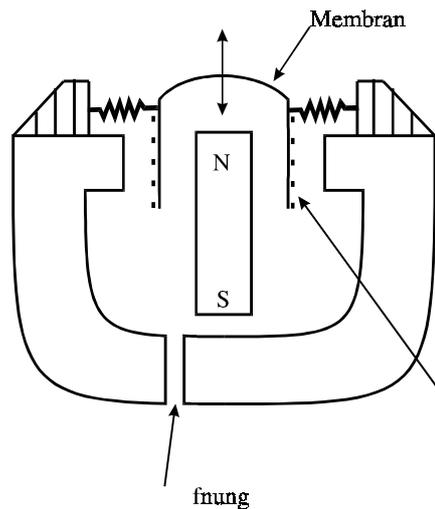


Abb. 3.2: Dynamisches Mikrophon

Unter Verwendung der zweiten elektromechanischen Analogie (d. h. Proportionalitätsbeziehung zwischen Spannung und Schallschnelle bzw. Strom und Schalldruck) kann ein elektrisches Ersatzschaltbild des dynamischen Mikrophons wie folgt aufgestellt werden:

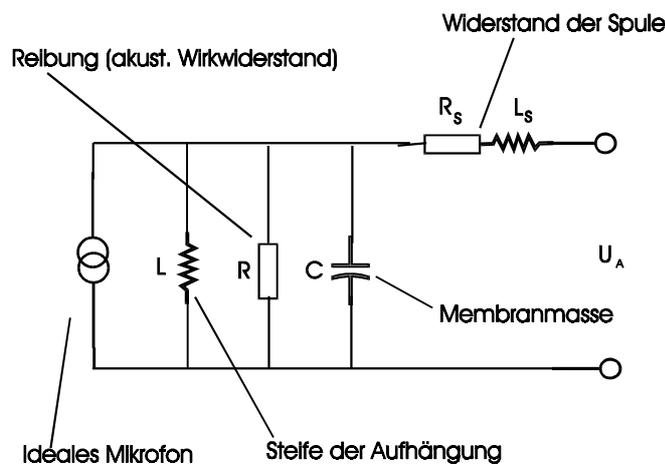


Abb. 3.3: Elektrisches Ersatzschaltbild des Mikrophons

Die ideale Strom- bzw. Spannungsquelle bezeichnet das „ideale“ Mikrophon, bei dem die Spannung proportional zur anliegenden Schallschnelle und der Strom proportional zum anliegenden Schalldruck ist. Die Parallelschaltung aus Induktivität L , Widerstand R und Kapazität C bezeichnet die Steifigkeit der Aufhängung der Membran, die Reibungsverluste (einschließlich der verrichteten, akustischen Wirkleistung) und die Membranmasse. Die Induktivität und der ohmsche Widerstand der Schwingspule wird dagegen durch die Längs-Wechselstrom-Widerstände R_s und L_s gekennzeichnet. Aufgrund dieses Ersatzschaltbildes ergibt sich als Übertragungsfunktion des dynamischen Mikrophons (d. h. Verhältnis zwischen

Ausgangs-Spannung- und Eingangs-Schalldruck) eine Charakteristik, die einen Abfall zu hohen und niedrigen Frequenzen beinhaltet:

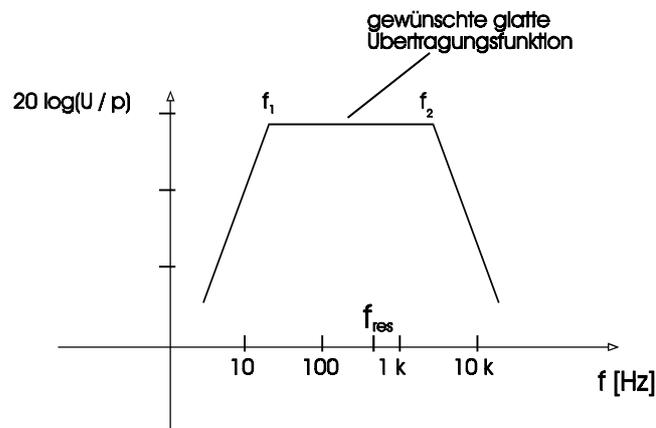


Abb. 3.4: Übertragungsfunktion des Mikrophons (schematisch)

Die untere Grenzfrequenz f_1 wird dabei durch die Nachgiebigkeit der Aufhängung bestimmt, da die Induktivität L bei niedrigen Frequenzen als Kurzschluß wirkt. In der Nähe der Resonanzfrequenz

$$f_{\text{res}} = \frac{1}{2\pi\sqrt{L \cdot C}} \quad (\text{III.1})$$

ist die Übertragungsfunktion maximal und nimmt in der Nähe dieser Resonanzfrequenz einen relativ konstanten Verlauf an, während die obere Grenzfrequenz f_2 durch die Membranmasse bedingt ist, die im elektrischen Ersatzschaltbild einen Kurzschluß durch die Kapazität C verursacht.

Von praktischer Wichtigkeit ist nun die Größe der am Mikrophon meßbaren Rauschspannung

$$\tilde{U}_R = \sqrt{4RkT \cdot \Delta f} \quad (\text{III.2})$$

Diese Rauschspannung kann als thermisches Rauschen an dem Innenwiderstand R angesehen werden, wobei im Unterschied zu Kapitel II der Faktor 4 in der effektiven Rauschleistung zum einen durch die Zusammenfassung von negativen und positiven Frequenzen (Δf wird nur für den Bereich positiver Frequenzen angegeben) und zum anderen durch die Darstellung als Spitzenwert (und nicht als Effektivwert) zustande kommt. Diese Angabe entspricht jedoch der allgemein üblichen Konvention. Es ist dabei zu beachten, daß in dem Ersatzwiderstand R kein elektrischer,

sondern ein akustischer Wirkwiderstand zu verstehen ist, so daß hier die Fluktuationen der Luftteilchen einen ähnlichen Effekt haben, wie die Fluktuationen der Ladungsträger beim reinen ohmschen Widerstandsrauschen.

Bei einem typischen dynamischen Mikrofon mit $R = 200 \text{ Ohm}$, $\Delta f = 20 \text{ kHz}$ und $T = 300 \text{ K}$ beträgt U_r etwa $0,3 \text{ } \mu\text{V}$, so daß der Pegel des Rauschens L_r bei ungefähr 16 dB liegt (Für eine typische Mikrofonempfindlichkeit von 2 mV/Pa). Damit entspricht ein reales Mikrofon einem rauschfreien idealen Mikrofon und einem Rausch-Schallfeld von 16 dB . Selbst unter Verwendung eines der rauschärmsten Mikrophone, einem Kondensatormikrofon mit einem Durchmesser von $2,54 \text{ cm}$, beträgt der äquivalente Rausch-Pegel etwa 13 dB(A) , (d. h. bei niedrigerer Gewichtung von niedrigen und hohen Frequenzen, die als A-Gewichtung bezeichnet wird und die in Anlehnung an die Lautstärkenwahrnehmung im Gehör bei niedrigen Pegeln vorgenommen wird). Es ist dagegen die Leistungsfähigkeit des Ohres hervorzuheben, das im wesentlichen bei gesunden Probanden kein Rauschen oberhalb der Hörschwelle wahrnehmen läßt, so daß der äquivalente Pegel des Rausch-Schallfeldes 0 dB(A) wäre.

Die Empfindlichkeit und die Rauscharmut des Ohres wird daher nur mit einem sehr hohen technischen Aufwand von derzeitigen technischen Systemen erreicht. Dieses zunächst erstaunlich klingende Resultat wird durch die massive parallele Verarbeitung von Informationen im Gehör ermöglicht, bei der eine unübersehbare Vielzahl von Nervenfasern die akustische Information in aufbereiteter Form dem Gehirn zuträgt, so daß durch Parallelverarbeitung eine höhere Mittelung und damit eine bessere Rauschunterdrückung als mit nur einem als Sensor dienenden technischen Mikrofon hervorgerufen wird.

III.2 Fotodiode

Als zweites Beispiel für einen Meßwandler soll die Fotodiode betrachtet werden, die für die Messung der Lichtintensität eingesetzt wird, d. h. der in der Fotodiode induzierte Fotostrom ist proportional zu der Intensität des einfallenden Lichtes. Das grundlegende Problem eines derartigen Fotodetektors ist dabei die Quantisierung der elektromagnetischen Strahlung: Bei einer vorgegebenen Wellenlänge bzw. Frequenz des zu detektierenden Lichts, d. h. bei vorgegebener Energie $E=h \cdot \nu$ des Lichtquants, ist die Schwankung (Standardabweichung) der Anzahl der Photonen genauso groß wie die erwartete Anzahl der Photonen. Diese zumindestens für kleine Wellenlängen zu beobachtende Proportionalität ist bedingt durch den zugrundeliegenden Poisson-Prozeß. Für kleine Energien bzw. große Wel-

lenlängen geht die erwartete Schwankung der Photonenzahl (und damit die Leistung des Quantenrauschens) allerdings nicht gegen 0, sondern gegen einen Grenzwert, der durch die **Bose-Einstein-Statistik** bestimmt wird. Die spektrale Leistungsdichte des Quantenrauschens $W(\nu)$ beträgt demnach:

$$W(\nu) = h \cdot \nu + \frac{h\nu}{e^{h\nu/kT} - 1} \quad (\text{Bose-Einstein-Statistik})$$

$$\rightarrow h\nu \quad \text{für} \quad T \gg \frac{h\nu}{k} \quad (\text{III.3})$$

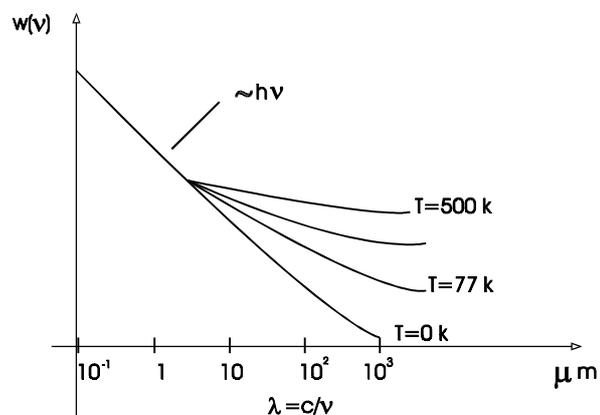


Abb. 3.5: Leistungsdichte des Quantenrauschens $W(\nu)$

Für sichtbares Licht ist $W(\nu)$ proportional $h\nu$, d. h. das aufgrund der Quantelung der Lichtenergie im Fotodetektor auftretende Quantenrauschen ist proportional zu der Frequenz des beobachteten Lichts. Für die Fotodiode als speziellen Lichtdetektor sieht der Aufbau schematisch folgendermaßen aus:

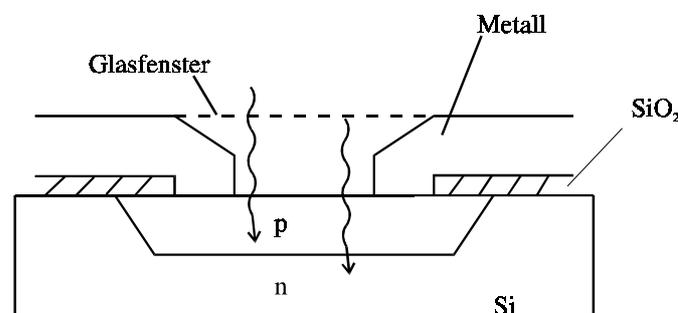


Abb. 3.6: pn-Fotodiode

Es handelt sich dabei um eine in Sperrichtung betriebene Halbleiter-Diode, bei der aufgrund der angelegten Spannung sämtliche Ladungs-

träger im Grenzbereich zwischen der p-dotierten und der n-dotierten Schicht „abgesaugt“ werden. Jedes in diese Sperrschicht eindringende Photon kann nun ein Ladungsträgerpaar erzeugen, das zu einem Stromfluß führt, bei dem das erzeugte Elektron zum Plus-Pol und die erzeugte Fehlstelle zum Minus-Pol wandern. Der resultierende Photostrom I hängt dabei von der Elementarladung e , der optischen Eingangsleistung P und dem Quantenwirkungsgrad η ab:

$$I = e \cdot \frac{P}{h\nu} \cdot \eta = s \cdot P \quad (\text{III.4})$$

Der Quotient $\frac{P}{h\nu}$ gibt damit die Zahl der Photonen pro Zeiteinheit an, die auf die Fotodiode eintreffen. Der Quantenwirkungsgrad η gibt dabei an, wieviel Prozent der einfallenden Photonen tatsächlich zu dem Stromfluß beitragen. Die Größen e , $h\nu$ und η lassen sich zu der Empfindlichkeit des Detektors s zusammenfassen, für die gilt:

$$s = \frac{e}{h\nu} \eta = \frac{\lambda}{1,24} \cdot \eta \quad (\text{III.5})$$

Die Empfindlichkeit s des Detektors hängt dabei von der Wellenlänge des zu empfangenden Lichts ab und ist für verschiedene Detektorarten in folgender Abbildung aufgetragen:

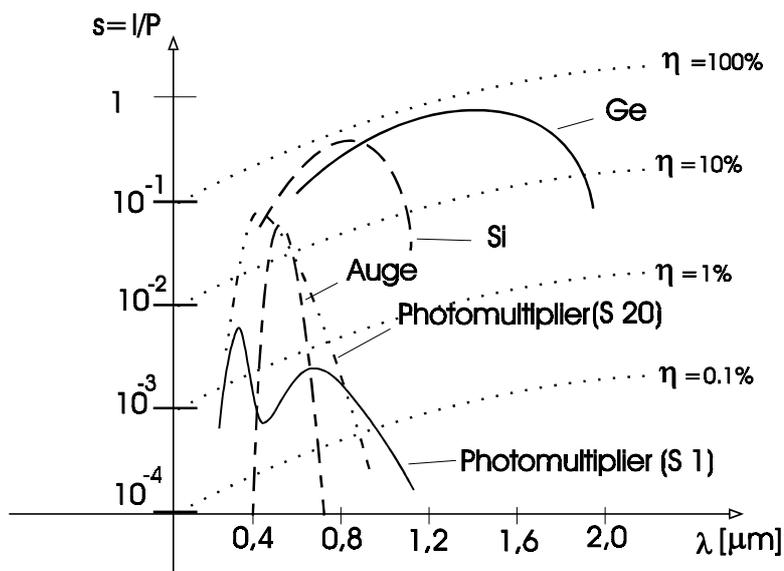


Abb. 3.7: Empfindlichkeit $s=I/P$

In der Abbildung ist neben der Abhängigkeit der Empfindlichkeit des Detektors s von der Wellenlänge λ für unterschiedliche Quantenwirkungsgrade η die theoretisch maximal erreichbare Empfindlichkeit für Silizium und Germanium als gängige Halbleiter-Materialien angegeben. Ebenfalls ist die Empfindlichkeit für einen Photomultiplier mit einer bzw. zwanzig Kathoden (S1 bzw. S20) angegeben. Die Empfindlichkeit des Auges liegt im „optimalen“ Spektralbereich über 10 % und liegt damit sehr dicht an dem theoretisch erreichbaren Optimum. Das Auge erweist sich damit als ein sehr leistungsfähiger und empfindlicher Fotodetektor, in dem fast jedes einlaufende Photon in dem empfindlichsten Frequenzbereich auch in eine Wahrnehmung umgesetzt werden kann. Für den Bereich der Fotodetektoren schneidet das Auge als physiologisches System im Vergleich mit dem technisch realisierbaren Systemen insgesamt daher sehr gut ab.

III. 3 Vorverarbeitung

Nachdem mit den bisher behandelten Wandlern die physikalische Meßgröße in ein (zumeist) elektrisches Signal gewandelt wurde, findet eine Vorverarbeitung statt, die z. B. auf Verminderung von störenden Anteilen im Signal (z. B. Verminderung von Rauschen) abzielt. Bei der **Korrelationsanalyse**, die in der Meßtechnik dem **Lock-In-Verstärker** zugrunde liegt, geht man davon aus, daß das eigentliche Nutzsignal von einem unkorrelierten, statistischen Rauschprozeß überlagert ist. Das Nutzsignal ist zudem periodisch mit einer Grundfrequenz f_s . Falls diese Bedingung für das Nutzsignal nicht erfüllt ist, wird es in der Meßtechnik zu einem periodischen Signal gemacht, z. B. durch Verwendung eines „Choppers“, der den Lichtstrahl periodisch unterbricht. Die Aufgabe ist es nun, die Leistung des Nutzsignals $s(t)$ und seine Phasenlage bzgl. eines Referenzsignals mit der Periode f_s aus dem vorliegenden Gemisch (d. h. der Summe aus $n(t)$ und $s(t)$) zu extrahieren. Der prinzipielle Aufbau eines Lock-In-Verstärkers ist dabei folgendermaßen:

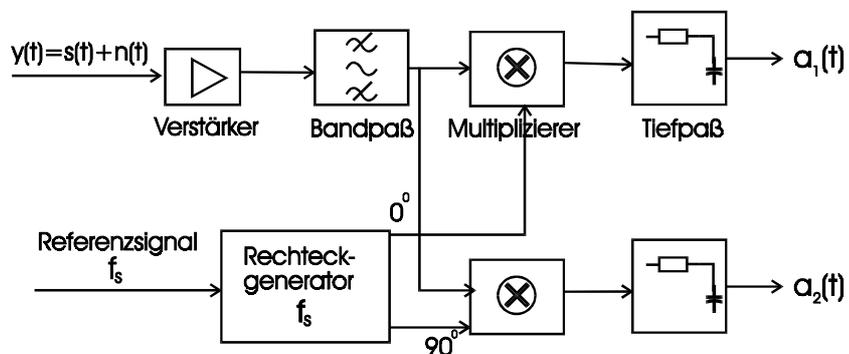


Abb. 3.8: Lock-In-Verstärker

Das Summensignal $y(t)=n(t)+s(t)$ wird zunächst verstärkt, bandpassgefil-

tert (zur Beschränkung auf den Frequenzbereich, der ohnehin nur von Interesse ist, die Bandpassfilterung kann auch entfallen), multipliziert und tiefpassgefiltert. Als zu multiplizierendes Signal wird ein rechteckförmiges Signal mit der Grundfrequenz f_s verwendet, das von einem Referenzsignal mit derselben Frequenz abgeleitet wird. Dieses Referenzsignal ist also mit dem Nutzsignal korreliert, bzw. wird bei der Periodisierung des Nutzsignals gewonnen. Aus diesem Referenzsignal werden zwei Rechtecksignale gewonnen, die jeweils um 90 Grad gegeneinander phasenverschoben sind, und mit dem gefilterten Nutzsignal multipliziert werden. Der anschließende Tiefpaß erzeugt den zeitlichen Mittelwert des Ausgangs dieser Multiplikatoren, so daß durch die Kombination aus Multiplikation und Mittelung das Korrelationsintegral approximiert wird. An den Ausgängen $a_1(t)$ und $a_2(t)$ liegt dann die „effektive“ Amplitude des Nutzsignals an, das 0° bzw. um 90° phasenverschoben mit dem Referenzsignal in dem Summensignal $Y(t)$ enthalten ist. Auf diese Art und Weise kann die Amplitude und Phasenlage des Nutzsignals selbst bei relativ ungünstigen Signal-Rauschverhältnissen im Summensignal genau bestimmt werden. Die Multiplikation mit dem Rechteck und anschließende Tiefpaßfilterung läßt sich im Zeit- und Frequenzbereich folgendermaßen beschreiben:

Ausgangssignal:

$$y(t) \cdot \left(\text{rect}_{\frac{1}{2f_s}}(t) * \text{rect}_{\frac{1}{f_s}}(t) \right) * h(t)$$


$$H(f) \cdot \left[(N(f) + S(f)) * \sum_{k=-\infty}^{\infty} \frac{\delta(f - (2k-1)f_s)}{(2k-1)f_s \cdot \pi} \right]$$

(III.6)

Im Frequenzbereich kann die Funktionsweise des Lock-In-Verstärkers daher als Verschiebung des Original-Spektrums um ein jeweils ungeradzahliges Vielfaches der Frequenz f_s aufgefaßt werden, von dem die anschließende Tiefpaßfilterung gerade den Gleichanteil heraussiebt. Dies bedeutet ein Heraussieben der bei $(2K-1) \cdot f_s$ liegenden Komponenten von $Y(f)$, d. h. sämtlicher ungeradzahlig harmonischen Komponenten des periodischen Nutzsignals.

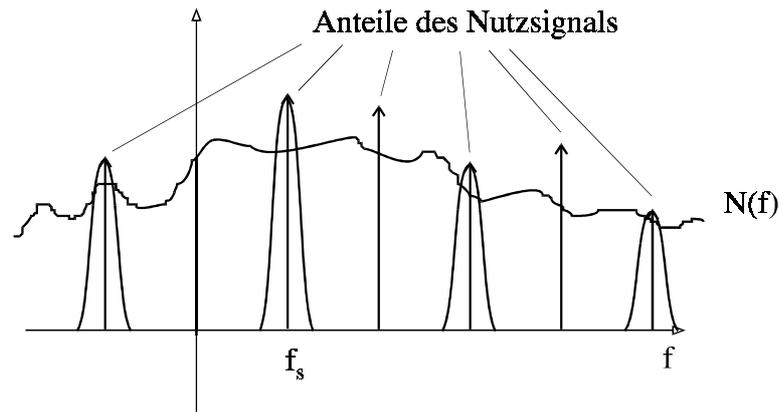


Abb. 3.9: Verdeutlichung des Vorgangs im Frequenzbereich

Je niedriger die Eckfrequenz des nachgeschalteten Tiefpaßfilters (und dementsprechend je länger die „effektive“ Integrationszeit), desto schärfer werden die periodischen Komponenten des Nutzsignals von dem kontinuierlichen Spektrum des Störsignals getrennt. Die Phasenlage von $s(t)$ wird gerade durch den \arctan des Verhältnisses von $a_1(t)/a_2(t)$ angegeben. Aufgrund dieser beim Lock-In-Verstärker erfolgenden Phasendetektion und Herausfilterung des Gleichanteils wird das Verfahren auch als **phasenempfindliche Gleichrichtung** bezeichnet. Durch die Umwandlung des Referenzsignals in ein Rechtecksignal mit zwei unterschiedlichen Phasenlagen wird beim Lock-In-Verstärker das Nutzsignal allerdings nicht direkt mit dem Referenzsignal korreliert, sondern mit einer generalisierten Form des Referenzsignals, das sämtliche ungradzahligen Harmonischen beinhaltet. Diese Art der Korrelationsanalyse ist sinnvoll, wenn man außer der Periodizität des Nutzsignals wenig Information über das zu erwartende Nutzsignal hat. Wenn allerdings die Zeitfunktion des Nutzsignals (bis auf einen Amplitudenfaktor und eine Zeitverzögerung) bekannt sind, kann der Rechteckgenerator entfallen und das Schaltbild des Lock-In-Verstärkers geht in die bereits oben behandelte Korrelationsanalyse über. Somit stellt sich der Lock-In-Verstärker als eine weit verbreitete, in der Meßtechnik gebräuchliche Form der Korrelationsanalyse dar.

Ein wesentlicher Nachteil des Lock-In-Verstärkers ist jedoch die Tatsache, daß die Zeitfunktion verlorengeht und nur die Amplitude und die Phasenlage extrahiert wird. Um dieses Problem zu umgehen und dennoch eine wirksame Rauschbefeuerung durchzuführen, wird ein **Boxcar-Verstärker** eingesetzt. Die grundlegende Idee besteht darin, daß das Signal zu einem festen Zeitpunkt der Periode abgetastet wird und dieser Abtastwert über mehrere Perioden gemittelt wird. Dabei nimmt das Nutzsignal immer denselben Wert an, während sich der statistisch fluktuierende Anteil des Störsignals herausmitteln soll, so daß mit zunehmender Anzahl von Mittelungen eine Verbesserung des Signal- Rauschabstandes zu erzielen ist.

Der grundlegende Aufbau sieht wie folgt aus:

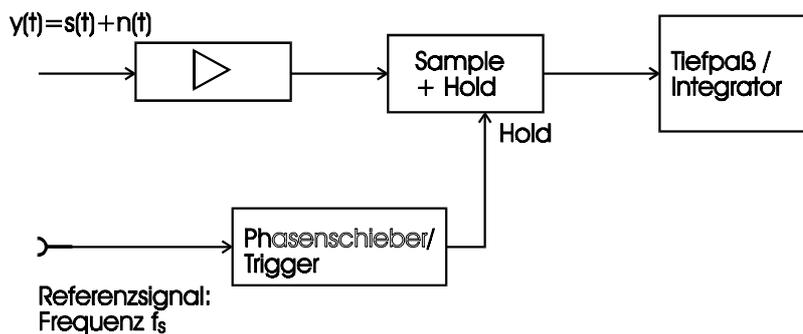


Abb. 3.10: Schaltbild des Boxcar-Verstärkers

Das Summensignal $y(t) = n(t) + s(t)$ wird nach Verstärkung einem „Sample and Hold“ zugeführt, d. h. einer Schaltung, die bei Triggerung durch ein externes Signal die gerade anliegende Eingangsspannung am Ausgang konstant hält („hold“). Dieses Triggersignal wird vom Referenzsignal mit der Frequenz f_s abgeleitet, wobei jeweils ein bestimmter Phasenwinkel eingehalten wird. Der Ausgang des „Sample and Hold“ wird tiefpassgefiltert, so daß über mehrere Perioden gemittelt wird und sich der Rauschanteil herausmitteln kann. Nach N Mittelungsperioden addiert sich der Nutzsignalanteil zu einer Amplitude, die proportional zu N ist, da die Amplituden des Nutzsignals kohärent aufaddiert werden. Die gemittelte Amplitude des statistischen Rausch-Anteils steigt dagegen nur mit \sqrt{N} , weil sich die Rausch-Signale inkohärent aufaddieren und damit nur die Leistung aufaddiert wird. Ein weiterer „Trick“ beim Boxcar-Verstärker ist das langsame Verschieben des Trigger-Zeitpunktes, so daß sich die gesamte gesuchte Wellenform als Funktion der Zeit darstellen läßt. Dieser Trick wird auch beim Sampling-Oszilloskop angewandt. Anstelle des Verschieben des Trigger-Zeitpunktes können auch mehrere Boxcar-Verstärker mit zeitversetztem Trigger-Zeitpunkt parallel eingesetzt werden, so daß die gesamte Wellenform zeitgleich in unterschiedlichen Ausgängen der jeweiligen Verstärker abgebildet (abgetastet) wird. Diese teure Alternative zum Boxcar-Verstärker wird als **Transientenrekorder** bezeichnet. Er wird überall dort eingesetzt, wo Zeitvorgänge mit sehr hoher Geschwindigkeit abgetastet werden müssen.

III. 4 Digitalisierung

Bis jetzt wurde die Vorverarbeitung von Signalen besprochen, die zum größten Teil mit analoger Technik durchgeführt werden kann. Zur exakten Speicherung und Bearbeitung von Signalen hat sich jedoch in letzter Zeit zunehmend die Digitaltechnik etabliert, bei der die analogen Eingangsgrößen in digitale, mit dem Computer darstellbare Werte umgewandelt

und weiterverarbeitet werden. Diese Digitalisierung setzt dabei eine Umsetzung der zeit- und wertkontinuierlichen analogen Signale in zeit- und wertdiskrete Zahlen folgen voraus, die an unterschiedlichen Speicherstellen im Computer abgespeichert werden können. In der Praxis wird diese Diskretisierung zumeist mit einem Analog-Digital-Wandler (**A/D-Wandler**) durchgeführt.

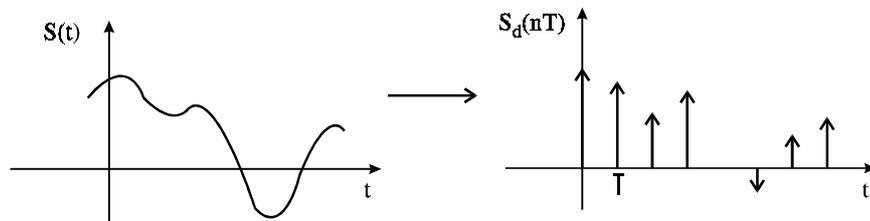


Abb. 3.11: A/D Wandlung

Je nach Anwendungsbereich des A/D-Wandlers gibt es verschiedene praktische Realisierungen, die sich in deren Präzision und ihrer Wandlungsgeschwindigkeit unterscheiden. Als Beispiel sei hier ein nach dem Wägeverfahren bzw. der sukzessiven Approximation funktionierender A/D-Wandler aufgeführt:

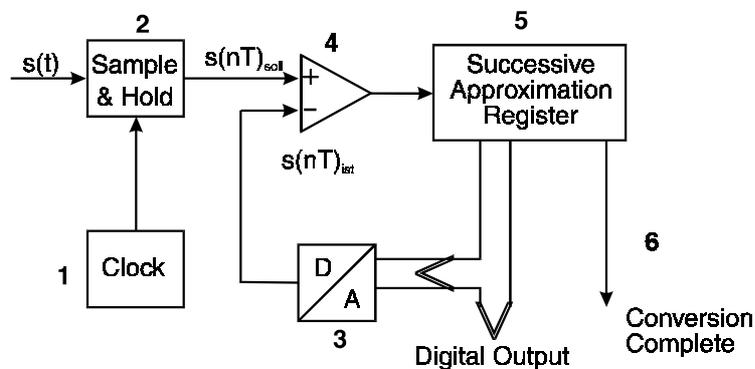


Abb. 3.12: A/D-Wandler

Mit Hilfe eines Taktgebers (Clock, (1)) wird der Abtastzeitpunkt vorgegeben, der typischerweise bei einer Flanke eines rechteckförmigen Trigger-Signals stattfindet. Der analoge Signalwert $s(t)$ wird mit einem Sample-and-Hold-Schaltkreis (2) auf einen festen Wert festgehalten, so daß während der nachfolgenden Umwandlung der Eingangswert sich nicht mehr zeitlich ändert. Dieser analoge „Soll“-Wert $s(nT)_{\text{soll}}$ wird nun mit dem digitalen Ist-Wert verglichen, indem mit Hilfe eines Digital-Analog-Wandlers (D/A-Wandler, (3)) der aktuelle digitale Wert wieder in einen analogen Wert verwandelt wird. Mit Hilfe eines Komparators (4), d. h. Operationsverstärker mit invertierendem und nichtinvertierendem Eingang, wird die

Differenz zwischen dem Soll- und dem Ist-Wert gebildet. Je nachdem, welches Vorzeichen diese Differenz annimmt, wird der digitale Ist-Wert im „Successive-Approximation-Register“ (5) verändert. Dabei wird zunächst das höchstwertigste Bit (d. h. das Vorzeichen-Bit) verändert und sukzessiv die nachfolgenden Bits. Sobald der Ist- und der Soll-Wert bis auf einen Fehler übereinstimmen, der durch die Anzahl der Bits des A/D-Wandlers bestimmt ist, wird das Ausgangssignal „Conversion Complete“ (6) vom A/D-Wandler ausgegeben, das die erfolgreiche Wandlung des Analog-Werts in eine digitale Zahl anzeigt.

Das hier vorgestellte Verfahren ist relativ langsam (d. h. die höchsten praktisch damit erreichbaren Sampling-Frequenzen liegen bei etwa 1 Mhz). A/D-Wandler für höhere Abtastfrequenzen arbeiten dagegen parallel und erfordern damit einen größeren Schaltungsaufwand. A/D-Wandler mit besonders hohen Genauigkeitsanforderungen, aber geringeren Geschwindigkeits-Anforderungen approximieren den erwünschten Wert dagegen nicht sukzessive Bit für Bit (d. h. logarithmisch), sondern durch sukzessives Durchfahren aller möglicher Zustände. Der Quantisierungsfehler des hier vorgestellten A/D-Wandlers ist immer größer oder gleich dem Fehler des dabei verwendeten D/A-Wandlers und insgesamt ist der Schaltungsaufwand für A/D-Wandler wesentlich größer als für vergleichbare D/A-Wandler.

Sowohl bei der Diskretisierung im Wertebereich („Quantisierung“) als auch im Zeitbereich („Abtastung“) können Fehler auftreten, die im folgenden betrachtet werden sollen. Der **Quantisierungsfehler** ε , den man bei der Abbildung eines wertekontinuierlichen, analogen Eingangssignals s in ein wertediskretes Ausgangssignal Z macht, ist in folgender Abbildung dargestellt:

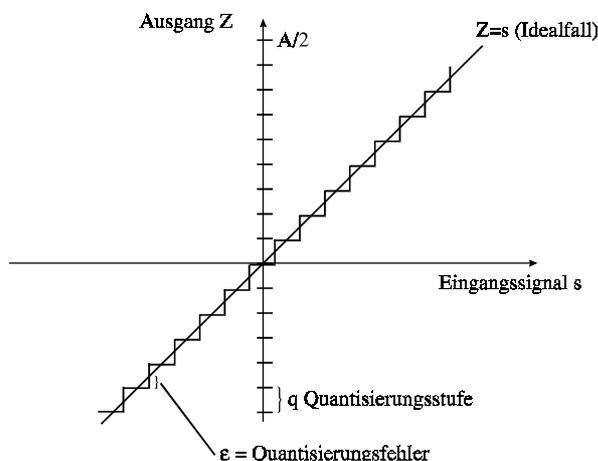


Abb. 3.13: Quantisierungsfehler

Wenn durch q die Quantisierungsstufe bezeichnet wird, ist der Quantisierungsfehler ε auf dem Bereich von $-\frac{q}{2}$ bis $+\frac{q}{2}$ beschränkt:

$$-\frac{q}{2} \leq \varepsilon \leq \frac{q}{2} \quad (\text{III.7})$$

Unter bestimmten Voraussetzungen können wir annehmen, daß der Quantisierungsfehler ε eine gleichverteilte Zufallsgröße ist, deren zeitlicher Verlauf das sogenannte **Quantisierungsrauschen** bildet. Die Wahrscheinlichkeitsdichtefunktion $p(\varepsilon)$ nimmt dann folgende Form an:

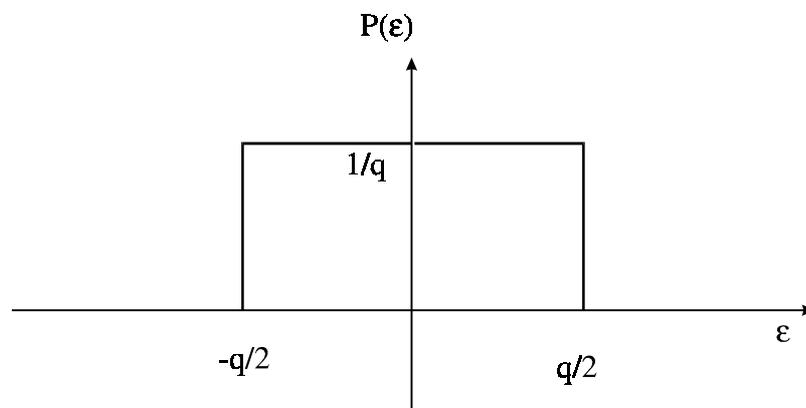


Abb. 3.14: Wahrscheinlichkeitsdichte des Quantisierungsrauschens

Die Annahme eines gleichverteilten Quantisierungsfehlers gilt, wenn die Signalamplitude sehr viel größer ist als die Quantisierungsstufe q und wenn der Effekt der zeitlichen Diskretisierung (Abtastung) vernachlässigt werden kann (d. h. wenn die Abtastrate wesentlich größer als die höchste im Signal vorkommende Frequenz ist). Die Leistung des Quantisierungsrauschens ist proportional zur Varianz σ_ε^2 der Zufallsgröße ε , die sich unter der Annahme obiger Gleichverteilung berechnet zu:

$$\sigma_\varepsilon^2 = \int_{-q/2}^{q/2} \varepsilon^2 \cdot p(\varepsilon) d\varepsilon = \frac{1}{12} q^2 \quad (\text{III.8})$$

Wenn wir ferner annehmen, daß der Gesamt-Wertebereich des A/D-Wandlers mit A bezeichnet wird (z. B. ± 5 V oder 10 V) und mit W die Wortbreite des A/D-Wandlers (z. B. 12 Bit), dann ist die Zahl der mit dem A/D-Wandler darstellbaren Zahlen 2^W (z. B. $2^{12} = 4096$), so daß sich die

Quantisierungsstufe q berechnet zu:

$$q = 2^{-W} \cdot A \quad (\text{III.9})$$

(Mit obigem Beispielswerten ergibt sich q zu 2,44 mV)

Um nun den Effekt des Quantisierungsrauschens abschätzen zu können, muß seine Leistung in das Verhältnis zur Leistung des Eingangssignals gesetzt werden. Unter der Annahme, daß das Eingangssignal ein voll ausgesteuerter Sinus mit der Amplitude $A/2$ ist, berechnet sich die Leistung dieses Eingangssignals σ_s^2 zu:

$$\sigma_s^2 = \frac{1}{2} \left(\frac{A}{2} \right)^2 = \frac{1}{8} q^2 2^{2W} \quad (\text{III.10})$$

Das Verhältnis aus maximaler Signalleistung und Leistung des Quantisierungsrauschens (Signal-Rauschverhältnis oder Signal-to-Noise-Ratio, SNR) errechnet sich damit zu:

$$\frac{\sigma_s^2}{\sigma_\varepsilon^2} = \frac{12}{8} \cdot 2^{2W} \quad (\text{III.11})$$

Das Signal-Rauschverhältnis wird meistens in Dezibel (dB) angegeben, also

$$\text{SNR (dB)} = 10 \log \left(\frac{\sigma_s^2}{\sigma_\varepsilon^2} \right) = W \cdot 6 \text{ dB} + 1,8 \text{ dB} \quad (\text{III.12})$$

Bei der oben angegebenen Rauschleistung des 12-Bit-Wandlers ergibt sich damit ein max. SNR von 74 dB. In der Praxis ist das Signal-Rausch-Verhältnis jedoch wesentlich geringer, weil die Leistung des Quantisierungsrauschens annähernd konstant ist, die Leistung des Eingangssignals jedoch stark variiert. Dabei wird in der Regel der A/D-Wandler nicht vollständig ausgesteuert, um eine Reserve für größere Eingangssignale zu haben, die im Fall einer vollständigen Aussteuerung nicht mehr richtig übertragen werden könnten. In der Audiotechnik liegt die Aussteuerungsreserve bei etwa 10 bis 20 dB. Außerdem ist bei realen A/D-Wandlern die Wandlungsgenauigkeit schlechter als die kleinste Quantisierungsstufe, weil die Quantisierungsstufen nicht für alle möglichen Zahlenwerte die

gleiche Größe aufweisen. Durch diese Abweichung der Quantisierungskennlinie von der idealen Form ist das Quantisierungsrauschen in der Praxis größer als hier berechnet, so daß bei einem 12-Bit-Wandler in der Praxis ein nutzbarer Dynamikbereich von maximal etwa 50 dB resultiert. Für hochwertige Audio-Übertragungsanlagen (z. B. CD-Spieler, Digital-Audio-Tape(DAT)) wird daher eine Quantisierung mit 16 Bit vorgenommen, wobei die A/D-Wandlung z. T. mit einer noch höheren Auflösung geschieht.

Während wir bisher nur die Diskretisierung im Wertebereich betrachtet haben, soll nun die **Diskretisierung im Zeitbereich** betrachtet werden. Um ein analoges Signal im Computer darstellen zu können, wird es zu diskreten, äquidistant aufeinanderfolgenden Zeitpunkten abgetastet, so daß das Signal $s(t)$ in ein Signal

$$s_d(nT) = s(t) \cdot \frac{1}{T} \text{II}\left(\frac{t}{T}\right) \quad (\text{III.13})$$

umgewandelt wird.

Wenn mit $S(f)$ das Spektrum des Signals $s(t)$ bezeichnet wird, berechnet sich das Spektrum $S_d(f)$ des abgetasteten Signals zu:

$$S_d(f) = S(f) * \text{II}(T \cdot f) = \sum_{n=-\infty}^{\infty} S\left(f - \frac{n}{T}\right) \quad (\text{III.14})$$

Das Spektrum des abgetasteten Signals ist damit periodisch.

Falls das Original-Spektrum $S(f)$ nun Anteile oberhalb von $f_s/2$ aufweist, überlappen sich diese Bereiche mit der periodischen Wiederholung des Spektrums, so daß sogenannte „**Spiegelfrequenzen**“ auftreten. Ein derartiger Überlapp verhindert es, durch ein ideales rechteckförmiges Tiefpaßfilter das Original-Spektrum aus dem periodischen Spektrum rekonstruieren zu können, da in dem rekonstruierten Spektrum noch „umgekehrt“ angeordnete Bereiche der periodischen Wiederholung des Spektrums vorliegen.

Signal $s(t)$

Signal $s_d(t)$

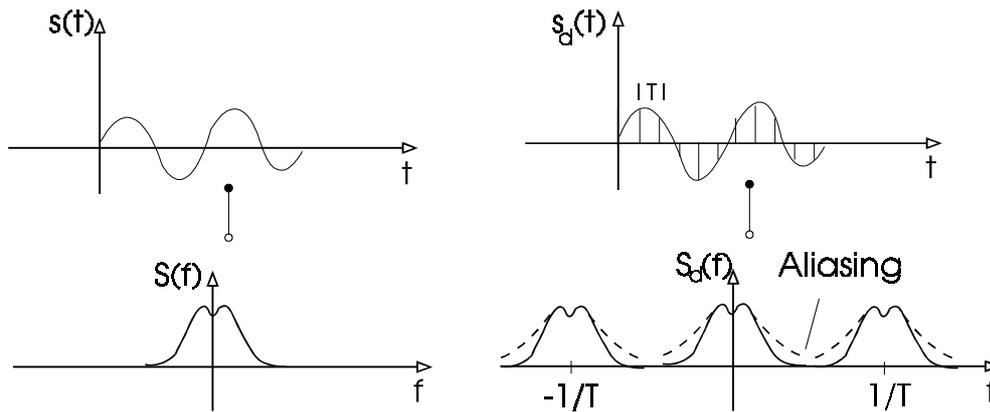


Abb. 3.15: Veranschaulichung des Abtasttheorems

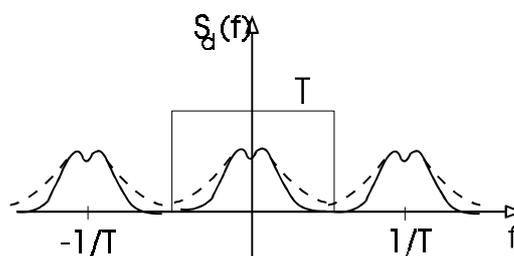


Abb. 3.16: Rekonstruktion mit idealem Tiefpaßfilter

Die Rekonstruktion des Original-Spektrums ist daher nur möglich, falls

$$S(f) = 0 \text{ für } f \geq f_s/2 \quad (\text{III.15})$$

Aus dieser als **Abtasttheorem** bezeichneten Beziehungen folgt, daß ein Zeitsignal mit einer Abtastfrequenz abgetastet werden muß, die mindestens dem doppelten der höchsten im Spektrum vorkommenden Frequenz entspricht. In der Praxis wird die Einhaltung dieses Abtasttheorems zum einen dadurch sichergestellt, daß vor dem Abtasten das Signal durch einen analogen Tiefpaßfilter mit einer Eckfrequenz gefiltert wird, die unterhalb der halben Abtastfrequenz liegt. Diese niedrigere Wahl der Eckfrequenz ist notwendig, weil in der Praxis keine „ideale“ Tiefpaßfilterung bei genau der halben Abtastfrequenz möglich ist und aufgrund der endlichen Filterflankensteilheit ein genügender „Sicherheitsabstand“ vorgesehen werden muß. Alternativ kann die Abtastfrequenz so hoch gesetzt werden, daß in den zu digitalisierenden Signalen keine spektralen Anteile oberhalb der Abtastfrequenz mehr vorhanden sind. Dies ist z. B. bei Audiosignalen der Fall, die für Compact-Disc-Aufnahmen mit einer Abtastfrequenz von 44,1 kHz digitalisiert werden. Da das Leistungsspektrum von natürlichen Musikinstrumenten und anderen Klangquellen (z. B. auch Sprache) oberhalb von 10 kHz stark abfällt, ist oberhalb von 22 kHz nur noch ein so geringer Anteil der Signalleistung vorhanden, daß eine Tiefpaßfilterung in

der Regel nicht notwendig ist. Allerdings sind Extremfälle möglich, bei denen das Leistungsspektrum erst langsam zu hohen Frequenzen hin abfällt (z. B. elektronische Musikinstrumente, Blech-Schlaginstrumente), so daß ohne Tiefpaßfilterung vor der Digitalisierung störende Spiegelfrequenzen im hörbarem Frequenzbereich auftreten könnten. Um dies zu verhindern, wird generell vor der Digitalisierung auch für Audio-CD-Aufnahmen tiefpaßgefiltert.

III.5 Rekonstruktion eines digitalisierten Signals

Für den Fall, daß man ohne Fehler das analoge Signal digitalisiert hat, ist das nächste Problem die „ideale“ Rekonstruktion des analogen Signals aus dem digitalisierten Signal. Im Idealfall wird dies durch ein „ideales“ Tiefpaßfilter durchgeführt, daß rechteckförmig bei der halben Abtastfrequenz die höheren Frequenzen abschneidet und so die δ -Pulsfolge des digitalen Signals in ein kontinuierliches Zeitsignal zurückverwandelt. In der Praxis wird die Rekonstruktion durch einen Digital-Analog-Wandler (**D/A-Wandler**) durchgeführt, der ein zeit- und wertediskretes Signal $s_d(nT)$ in ein zeitkontinuierliches und wertediskretes Signal $s^*(t)$ umwandelt. Das in der Schaltungstechnik für D/A-Wandlung am meisten verbreitete Prinzip des D/A-Wandlers nach dem Wägeverfahren ist im folgenden Blockschaltbild schematisch dargestellt:

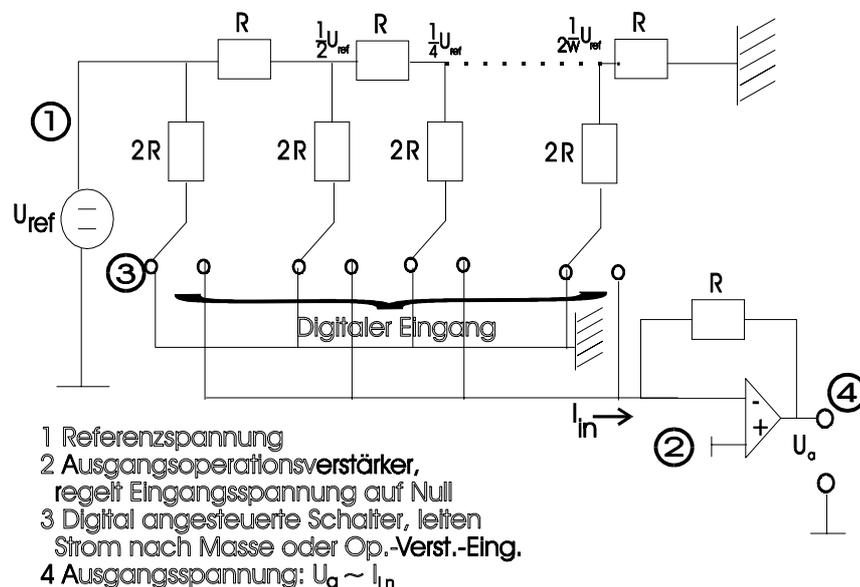


Abb. 3.17: D/A-Wandler

Bei diesem D/A-Wandler wird eine Referenz-Spannung, die entweder von einer internen Referenz-Spannungsquelle erzeugt wird oder von außen angelegt wird („multiplizierender D/A-Wandler“) sukzessive im Verhältnis 1:2 geteilt und je nach anliegender Bit-Kombination zu einer resultieren-

den Ausgangsspannung zusammen addiert. Die Spannungsteilung wird durch ein $R/2R$ -Widerstandsnetzwerk durchgeführt, bei dem jeder Zweig den Längs-Widerstand $2R$ aufweist. An jedem Knotenpunkt dieses Widerstandsnetzwerks liegt also ein Widerstand der Größe $2R$, der entweder direkt nach Masse oder über den Eingang des nachfolgenden Operationsverstärkers auf Masse gelegt wird, parallel mit dem Rest des Widerstands-Netzwerks. Dieser Rest stellt sich als Hintereinander-Schaltung aus einem Widerstand der Größe R und einem Rest-Widerstand der Größe R dar (d. h. Gesamtwiderstand $2R$). Diese Parallel-Schaltung aus zwei Widerständen der Größe $2R$ ergibt die sukzessive Herunter-Teilung der Referenz-Spannung im Verhältnis 1:2. Der am Ausgang liegende Operationsverstärker regelt die am invertierenden Eingang („minus“-Eingang) anliegende Eingangsspannung auf denselben Wert wie die am nichtinvertierenden Eingang („plus“-Eingang) liegende Spannung von 0 V . Dazu regelt der Operationsverstärker seine Ausgangsspannung auf einen so großen (negativen) Spannungswert, daß der einlaufende Strom I_{in} über den Gegenkopplungswiderstand R gerade abfließt, so daß am invertierenden Eingang 0 V anliegt. Auf diese Weise ist die Ausgangsspannung U_a proportional zu dem Eingangsstrom I_{in} und es gilt:

$$U_a = -U_{ref} \frac{Z}{Z_{max} + 1} \quad (\text{III.16})$$

Z bezeichnet dabei die digital anliegende Zahlenkombination. Das höchstwertige Bit steuert dabei den in der Zeichnung am weitestens links liegenden Schalter und das am niedrigsten wertige Bit den am weitesten in der Zeichnung rechts liegenden Schalter. Z_{max} bezeichnet die größte darstellbare Zahl (d. h. sämtliche Schalter-Stellungen auf logisch 1). Die Ausgangsspannung U_a bleibt so lange konstant, bis eine neue digitale Zahl Z an die Schalter angelegt wird, so daß aus dem zeit- und wertediskreten Signal $s_d(nT)$ eine Treppenfunktion resultiert, bei der die zu den Abtastzeitpunkten anliegenden Zahlen-Werte als analoge Spannungswerte für jeweils eine Abtastperiode festgehalten werden.

In der Realität geht die Ausgangsspannung des D/A-Wandler von einem Ausgabewert bis zum nächsten Wert nicht instantan über und nimmt auch noch nicht sofort konstante Werte an. Dies ist insbesondere ein Problem, wenn ein höherwertiges Bit z. B. auf Null geschaltet wird und gleichzeitig sämtliche niederwertigen Bits auf 1 geschaltet werden, um eine Zahl zu realisieren, die nur um einen Zählwert niedriger liegt, trotzdem aber viele Schaltvorgänge gleichzeitig erforderlich macht. Durch diese simultan erfolgenden Schaltvorgänge können Ausgleichsströme innerhalb des D/A-Wandlers fließen, die zu sogenannten „**Glitches**“ führen. Diese z. T. stö-

renden hochfrequenten Schaltartefakte können durch Tiefpaßfilterung oder durch „Deglitchen“ z. B. mit einem Sample and Hold-Schaltkreis kompensiert werden.

Die mit dem D/A-Wandler durchgeführte Rekonstruktion des Zeitsignals kann also wie folgt beschrieben werden:

$$s_d(nT) = \begin{cases} s(t) & \text{für } t = nT \\ 0 & \text{sonst} \end{cases} \rightarrow s^*(t) = s_d(nT) \text{ für } nT \leq t \leq (n+1)T$$

$$\Rightarrow s^*(t) = s_d(nT) * \text{rect}_T(t) = (s(t) \cdot \text{rect}_T(t)) * \text{rect}_T(t)$$

$$S^*(f) = S_d(f) \cdot T \cdot \text{sinc}(f \cdot T) = (S(f) * \text{rect}_{\frac{1}{T}}(f)) T \text{sinc}(fT) \quad (\text{III.17})$$

Die mit dem D/A-Wandler durchgeführte Approximation durch eine Treppenfunktion kann also durch die Faltung des ursprünglichen digitalisierten Signals mit einem Rechteck der Länge T beschrieben werden und im Frequenzbereich durch die Multiplikation mit einer Spaltfunktion, die die erste Nullstelle bei $1/T$ aufweist. Diese Rekonstruktion des D/A-Wandlers erweist sich damit nicht als eine ideale Tiefpaßfilterung mit einer steilen Flanke bei der halben Abtastfrequenz, sondern als eine sehr ineffiziente Tiefpaßfilterung, die bei der halben Abtastfrequenz bereits eine gewisse Absenkung des Spektrums und erst bei der Abtastfrequenz die vollständige Abschwächung bewirkt. Im weiteren Verlauf fällt diese Tiefpaßfilterung relativ langsam mit $1/f$ ab.

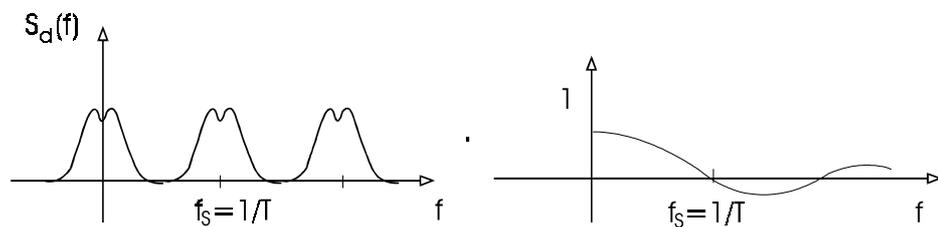


Abb. 3.18: Multiplikation des periodischen diskreten Spektrums mit der sinc-Funktion

Bei der halben Abtastfrequenz $f_s/2$ ist das Spektrum des rekonstruierten Signals schon um

$$\operatorname{sinc}\left(\frac{1}{2}\right) = \frac{\sin \pi / 2}{\pi / 2} = \frac{2}{\pi} = 0,636 \approx -3,9\text{dB} \quad (\text{III.18})$$

abgesunken. Dieser Effekt der spektralen Verfärbung durch den D/A-Wandler kann beispielsweise durch einen nachgeschalteten analogen Tiefpaß wieder kompensiert werden, der bei der halben Abtastfrequenz eine Überhöhung durchführen kann. Eine bessere Rekonstruktion des digitalisierten Signals kann durch einen D/A-Wandler erzeugt werden, der das digitalisierte Signal nicht durch eine Treppenfunktion, sondern durch eine Abfolge kurzer, rechteckförmiger Pulse jeweils bei den Abtastzeitpunkten rekonstruiert, wobei in den Pausen zwischen diesen rechteckförmigen Pulsen der Ausgangs-Spannungswert 0 angenommen wird. Diese Art der Rekonstruktionen entspricht einer Verringerung der Dauer des Rechtecks, mit dem das digitale Signal gefaltet wird und somit einer Verlagerung der ersten Nullstelle der sinc-Funktion auf höhere Frequenzen. Dadurch wird zwar die Abschwächung des Spektrums bei der halben Abtastfrequenz geringer, dennoch muß ein analoges Tiefpaßfilter bei der halben Abtastfrequenz nachgeschaltet werden.

Das benötigte scharfe analoge Tiefpaßfilter kann durch ein digitales Filter weitgehend ersetzt werden, wenn man bei der Technik des sogenannten **Oversamplings** die Abtastfrequenz vor der Rekonstruktion künstlich erhöht (N-faches Oversampling) und dann bei dieser höheren Frequenz einen digitalen Tiefpaßfilter erzeugt, der gerade bei der halben ursprünglichen Abtastfrequenz eine sehr scharfe Filterung durchführt. Wenn nun dieses resultierende Signal mit der neuen, wesentlich höheren Abtastfrequenz rekonstruiert wird, liegen die ersten Spiegelfrequenzen erst etwas unterhalb der neuen, wesentlich höheren Abtastfrequenz. Daher reicht ein einfacheres analoges Filter mit relativ geringer Flankensteilheit aus, um diese höhere Frequenzanteile vollständig zu eliminieren. Aufgrund der Einsparung von analogem Filteraufwand durch digitale Signalverarbeitung ist diese Art der Rekonstruktion weit verbreitet (z. B. in Compact-Disc-Spielern).

III.6 Anpassung von Daten an eine Modellfunktion

Mit den bisher behandelten Elementen einer physikalischen Meßwert-Aufnahme ist es uns gelungen, physikalische Meßwerte in elektrische Signale umzuwandeln, sie vorzuverarbeiten und über die Digitalisierung z. B. in einen Computer oder ein anderes Speichermedium einzulesen. Ein fundamentales Problem stellt sich nun für die Interpretation dieser Daten, bei der zwischen dem (unerwünschten) Rauschanteil und dem (erwünschten) Signalanteil unterschieden werden sollte. Ohne den Analyse-

Verfahren der digitalen Signalverarbeitung (Kap. IV) vorzugreifen, soll im folgenden eine allgemeine Methode zur Rauschbefreiung bzw. Datenanalyse betrachtet werden: Über den erwünschten Teil des Signals gibt es zumeist Erwartungswerte oder Modellvorstellungen, so daß das Problem der Dateninterpretation auf die Anpassung einer vorgegebenen Modellfunktion an die aufgenommenen Daten reduziert werden kann. Dieses grundlegende Problem der **Anpassungsrechnung** für eine beliebige Modellfunktion (die im allereinfachsten Fall der „linearen Regression“ eine Gerade ist) soll etwas ausführlicher behandelt werden. Dabei bezeichnen $y_i, i = 1 \dots N$ die gemessenen, abhängigen Datenwerte bei Vorgabe von (unabhängigen) Parameterwerten $x_i, i = 1 \dots N$. Diese gemessenen Daten können z. B. die Werte einer Zeitfunktion sein, die an N Abtastwerten gewonnen wurden, wobei die x_i gerade die Abtast-Zeitpunkte beschreiben würden. Die „gesuchte“ Beziehung oder Modellfunktion hat dabei die allgemeine Form:

$$y = F(x, \underline{\alpha}) \quad (\text{III.19})$$

F bezeichnet dabei die Modellfunktion, die von den m Parameterwerten $\alpha_1 - \alpha_m$ abhängt, die als Parameter-Vektor $\underline{\alpha}$ zusammengefaßt werden. Beispielsweise kann als Modellfunktion eine sinusförmige Zeitfunktion angenommen werden, für die die Amplitude, die Frequenz und die Phasenlage die drei anzupassenden Parameter wären:

$$y = \alpha_1 \cdot \sin(\alpha_2 \cdot x + \alpha_3) \quad (\text{III.20})$$

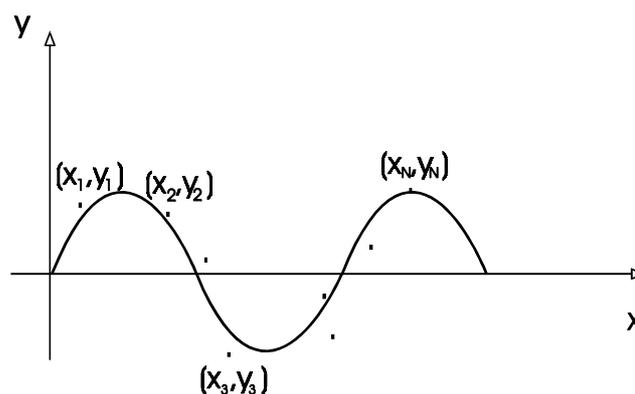


Abb. 3.19: Anpassung einer sinus-förmigen Funktion $y(x)$

Das Problem der Anpassungsrechnung besteht nun darin, daß der Parameter-Vektor $\underline{\alpha}$ so bestimmt werden muß, daß die Modellfunktion am besten die Daten erklärt. Diese als Anpassungs-Rechnung oder „Fit“ bezeichnete Aufgabe kann formal damit gelöst werden, daß man die sog-

nannte Likelihood L maximiert, d. h. die Wahrscheinlichkeit für das Auftreten der beobachteten Daten unter Annahme der Modellfunktion mit dem Parameter-Vektor $\underline{\alpha}$. Diese Likelihood-Funktion ist unter der Voraussetzung von N statistisch unabhängigen Beobachtungswerten y gerade das Produkt aus den Wahrscheinlichkeiten, daß der jeweilige Wert y_i auftritt als Funktionswert zu dem Eingangswert x_i bei Zugrundelegung des Parameter-Vektors $\underline{\alpha}$:

$$L(\underline{\alpha}) = \prod_{i=1}^N P(y_i = F(x_i, \underline{\alpha}) | \underline{\alpha}) \quad (\text{III.21})$$

Diese Funktion $L(\alpha)$ wird als Likelihood und nicht als Probability (Wahrscheinlichkeit) bezeichnet, weil ja nicht die Wahrscheinlichkeit für das Vorliegen des Parameter-Vektors $\underline{\alpha}$ als Funktion y_i ausgedrückt wird, sondern umgekehrt bei vorgegebenen $\underline{\alpha}$ die Wahrscheinlichkeit, daß die beobachteten Werte y_i tatsächlich hätten auftreten können. Daher ist der Begriff der Wahrscheinlichkeit (Probability) von dem der Likelihood deutlich zu trennen. Die nun resultierende Aufgabe, für eine vorgegebene beliebig geformte Modellfunktion die Likelihood zu maximieren, kann auf verschiedene Arten gelöst werden:

- Methode der kleinsten Quadrate

Die klassische von Gauß bereits entwickelte Methode der kleinsten Fehlerquadrate geht davon aus, daß der Meßfehler normalverteilt ist und daß die Varianz der Meßfehler für sämtliche N Meßpunkte gleich ist und den Wert σ^2 annimmt. Damit errechnet sich die Likelihood zu

$$L = \prod_{i=1}^N \left\{ e^{-\frac{(y_i - F(x_i, \underline{\alpha}))^2}{2\sigma^2}} \cdot \Delta y \right\} \quad (\text{III.22})$$

Der konstante Faktor Δy tritt hier auf, weil Wahrscheinlichkeiten miteinander multipliziert werden, die Normalverteilung in geschlossener Form jedoch nur als Wahrscheinlichkeitsdichte angegeben werden kann, die man durch Multiplikation mit dem Differential dy bzw. der endlichen Differenz Δy in eine Wahrscheinlichkeit umrechnen kann. Der Faktor Δy spielt aber im folgenden keine Rolle, weil eine Maximierung der Likelihood äquivalent ist mit der Minimierung des negativen Logarithmus von L :

$$-\log L = \left[\sum_{i=1}^N \frac{(y_i - F(x_i, \underline{\alpha}))^2}{2\sigma^2} \right] - N \log \Delta y \quad (\text{III.23})$$

Dabei muß α so gewählt werden, daß der erste Summand minimal wird. Der zweite Summand ist konstant und kann daher vernachlässigt werden. Durch die Bildung des Logarithmus geht das Produkt der Wahrscheinlichkeiten in die Summe der Abweichungsquadrate über und die Minimierung dieser Abweichungsquadrate entspricht gerade der von Gauß eingeführten Methode der kleinsten Fehlerquadrate. Zu ihrer numerischen Berechnung werden N-dimensionale Optimierungsalgorithmen eingesetzt (s. u.).

- χ^2 - Fit

Eine etwas andere Vorgehensweise beim sogenannten χ^2 (Chi-Quadrat-Fit) geht von der Annahme aus, daß der Meßfehler wiederum normalverteilt ist, die Varianzen σ_i^2 jedoch für die einzelnen Meßwerte verschieden groß sind. Diese Annahme ist in der Praxis realistischer, wenn die Meßwerte einen großen Wertebereich überdecken und der Meßfehler abhängig von der Größe des jeweiligen Meßwertes ist. Entsprechend der o. a. Minimierungsaufgabe ergibt sich für diesen Fall die Minimierung der Größe χ^2 mit:

$$\chi^2 = \sum_{i=1}^N \frac{(y_i - F(x_i, \alpha))^2}{2\sigma_i^2} \quad (\text{III.24})$$

- Robuste, lokale Schätzung für beliebige Fehler-Verteilung

In der Praxis ist die Annahme eines normalverteilten Meßfehlers oft eine zu starke Einschränkung, da „Ausreißer“ häufiger vorkommen, als es durch eine Normalverteilung vorhergesagt wird. Um den Einfluß dieser Ausreißer auf die Gesamt-Anpassung gering zu halten, ist es daher sinnvoll, eine andere, nicht notwendigerweise normalverteilte Wahrscheinlichkeitsdichte für den Meßfehler anzunehmen. Diese Verteilungsdichte kann man jedoch insofern einschränken, als daß die Verteilung jeweils nur von

$$z_i = \frac{y_i - F(x_i, \alpha)}{\sigma_i} \quad (\text{III.25})$$

abhängt. Die Wahrscheinlichkeit, daß ein Meßfehler auftaucht, hängt damit nicht explizit von y_i und $F(x_i, \alpha)$ und σ_i ab, sondern nur von der „normierten“, lokalen Abweichung z_i . „Lokal“ bedeutet dabei, daß die Wahrscheinlichkeit eines Fehlers nicht z. B. von einer bestimmten Kombination der y_i abhängt, sondern von jedem z_i getrennt. Das Problem der Anpassung einer optimalen Vielfunktion beschränkt sich damit auf das Minimie-

rungsproblem von

$$-\log L = \sum_{i=1}^N \rho(z_i) - N \log \Delta y \quad (\text{III.26})$$

Dabei bezeichnet $\rho(z_i)$ den negativen Logarithmus der Wahrscheinlichkeitsdichte für den normierten Fehler z_i . Der zweite Term $N \log \Delta y$ ist genau wie oben konstant, so daß er im weiteren vernachlässigt wird. Wenn wir mit $\psi(z) = d\rho(z)/dz$ die Ableitung der Wahrscheinlichkeitsdichte bezeichnen, so ist das Minimierungsproblem identisch mit der Nullstellensuche von:

$$\sum_{i=1}^N \frac{1}{\sigma_i} \psi\left(\frac{y_i - F(x_i, \underline{\alpha})}{\sigma_i}\right) \cdot \frac{\partial F(x_i, \underline{\alpha})}{\partial \alpha_k} = 0 \quad (\text{III.27})$$

Ein absolutes Minimum der Likelihood liegt also vor, wenn die Summen für jeden Wert von k den Wert 0 annehmen. In diesen Summen ist einerseits die partielle Ableitung der Zielfunktion $F(x_i, \underline{\alpha})$ nach der Parameterkomponente α_k enthalten, die von der jeweils angenommenen Fehlerverteilungsfunktion unabhängig ist. Sie wird zum einen mit der Varianz $1/\sigma_i$ gewichtet, die genau wie beim χ^2 -Fit ein Maß für die zu erwartende Abweichung bildet und zum anderen mit dem Fehlermaß $\psi(z)$. Dieser Funktion $\psi(z)$ kommt dabei eine große Bedeutung zu, da sie direkt von der angenommenen Wahrscheinlichkeitsdichtefunktion des Meßfehlers abhängt. Beispielsweise können die folgenden Verteilungen auftreten:

i) Normalverteilung:

$$P \approx e^{-\frac{z^2}{2}}, \quad \rho(z) = \frac{1}{2} z^2, \quad \psi(z) = z \quad (\text{III.28})$$

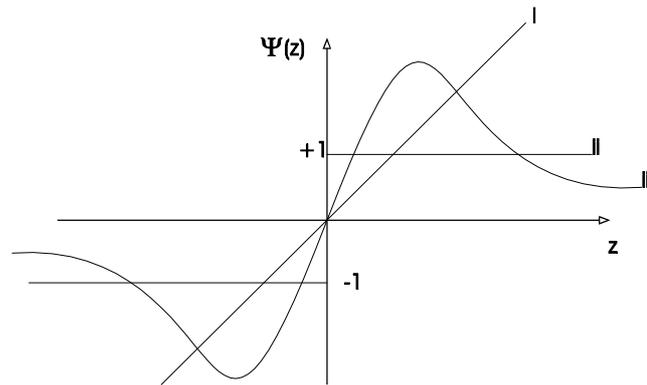
ii) Zweiseitig exponentiell

$$P \approx e^{-|z|}, \quad \rho(z) = |z|, \quad \psi(z) = \text{sgn}(z) \quad (\text{III.29})$$

iii) Cauchy (oder Lorentz)

$$P \approx \frac{1}{1 + \frac{1}{2} z^2}, \quad \rho(z) = \log\left(1 + \frac{1}{2} z^2\right), \quad \psi(z) = \frac{z}{1 + \frac{1}{2} z^2} \quad (\text{III.30})$$

Die Abhängigkeit von $\psi(z)$ ist für diese drei Verteilungsfunktionen im folgenden Bild aufgetragen:

Abb. 3.16: Fehlermaß $\psi(z)$

Während die Normalverteilung „Ausreißer“ sehr stark gewichtet, indem der Wichtungswert $\psi(z)$ mit zunehmenden Abstand vom Erwartungswert linear ansteigt, bietet die zweiseitige Exponentialverteilung den Vorteil, daß sämtliche Abweichungen gleich behandelt werden. Dies hat jedoch den Nachteil, daß bereits kleine Abweichungen vom Erwartungswert zu großen Schwankungen in der Fehlerfunktion führen. Für praktische Anwendungen sehr günstig ist eine Art „Kompromiß“ zwischen beiden Funktionen in Form der Cauchy-Verteilung, die für kleine Werte von z einen linearen Anstieg der Gewichtung mit z aufweist, für große Werte von z jedoch eine (kleine) Konstante annimmt, so daß diese Verteilung relativ robust gegenüber „Ausreißern“ ist. Die Annahme einer derartigen Fehlerverteilungsfunktion kann daher für die Praxis empfohlen werden.

III. 7 Minimierung in M Dimensionen

Bis jetzt wurde behandelt, wie das Fit-Problem in ein Minimierungsproblem von $-\log L(\alpha_k)$ bei vorgegebenen Meßwerten x_i umgewandelt werden kann. Dabei handelt es sich um ein m -dimensionales Minimierungsproblem in den Parametern α_k , $k = 1 \dots m$. Dieses Minimierungsproblem ist äquivalent mit der Suche nach Nullstellen der partiellen Ableitungen $\delta F(x_i, \underline{\alpha}) / \delta \alpha_k$. Je nachdem, ob diese partiellen Ableitungen der Zielfunktion überhaupt bekannt sind, (d. h. analytisch ausrechenbar), kann man zwei Fälle unterscheiden:

- Partielle Ableitung bekannt

In diesem Fall ist das Fit-Problem durch eine Nullstellensuche im m -dimensionalen Raum \mathbb{R}^m zu lösen. Diese Nullstellensuche kann z. B. mit dem m -dimensionalen Newton-Verfahren gelöst werden, bei dem die Funktion, deren Nullstelle gesucht wird, iterativ durch eine lineare Abbildung bei einer vorhergehenden Approximation der Nullstelle angenähert wird. Bei der Nullstelle dieser linearen Abbildung wird die Funktion erneut

geschätzt. Durch die an dieser Stelle berechnete lineare Approximation wird die Nullstelle weiter gesucht. Im ein-dimensionalen Fall entspricht dieses Verfahren einem Tangenten-Verfahren, das relativ schnell eine Nullstelle in der Praxis finden läßt.

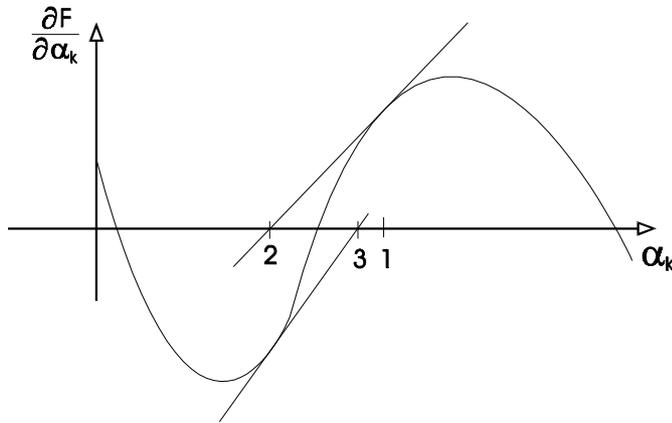


Abb. 3.21: Ein-dimensionales Newton-Verfahren

Für das m -dimensionale Problem wird die Funktion durch die m -dimensionale lineare Abbildung approximiert, die aus der Jacobi-Matrix hervorgeht, so daß dieses Problem mit einer Matrix-Inversion einhergeht. Dieses Verfahren ist daher programmieretechnisch aufwendig, führt jedoch mit wenigen Iterationsschritten zu einer sehr guten Annäherung der Nullstelle.

- **Partielle Ableitungen unbekannt**

Falls die partiellen Ableitungen und die aus den partiellen Ableitungen errechnete Jacobi-Matrix nicht bekannt ist bzw. wegen des algebraischen Aufwandes nicht berechnet werden soll, eignen sich Optimierungsverfahren aus der numerischen Mathematik, die eher empirischen Regeln gehorchen. Am verbreitetesten ist das sogenannte Simplex-Verfahren, bei dem im m -dimensionalen Raum ein Simplex mit $m + 1$ Eckpunkten aufgespannt wird (z. B. im \mathbb{R}^2 ein Dreieck bzw. im \mathbb{R}^3 ein Tetraeder).

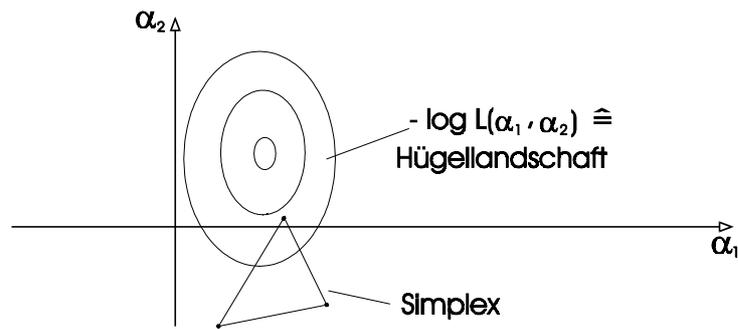


Abb. 3.22: Simplex-Verfahren

An den Eckpunkten dieses Simplex wird die zu minimierende Funktion (z. B. $-\log L(\alpha_1, \alpha_2)$) errechnet. Dann wird die Ecke mit dem größten Funktionswert ausgewählt und das Simplex an der Kante gespiegelt, die dieser Ecke gegenüberliegt. Auf diese Weise wird angestrebt, daß das Simplex sich langsam in die Richtung bewegt, in der die Funktion minimal ist. Gleichzeitig wird bei jeder Spiegelung eine Verkleinerung des Simplex durchgeführt (bzw. eine Vergrößerung, wenn trotz Suche kein Minimum gefunden werden konnte oder eine Reihe von Zuständen zyklisch angenommen wird). Dann wird mit dem ersten Schritt fortgefahren (d. h. erneut Berechnen der Funktionswerte an allen Ecken des Simplex usw.), so daß sich das Simplex im Laufe der Zeit auf das lokale Minimum zusammenzieht und nach einem bestimmten Kriterium die Suche abbricht. Für dieses und andere numerische Verfahren gibt es in den gängigen Hochsprachen zumeist Standard-Algorithmen (z. B. Press et al., 1992). Außerdem sei auf die Literatur der numerischen Mathematik verwiesen.