

Proseminar
„Algorithmen, Schnittstellen und Werkzeuge zur Audiobearbeitung“

Digitale Filter

Oskar Klett
Matthias Schwegler

Technische Universität München
Fakultät für Informatik

Sommersemester 2003

Inhaltsverzeichnis

Kapitel 1	Signale im Frequenzspektrum	2
Kapitel 2	Filter	5
2.1	Was ist ein Filter?	5
2.2	Anwendungsgebiete	6
2.3	Filterarten	6
2.4	Grundlagen	8
2.4.1	Impulsantwort	8
2.4.2	Dirac-Impuls	8
2.4.3	Faltung	8
2.4.4	Filterkriterien	10
2.5	Ideales Filter	11
2.6	Reales Filter	12
Kapitel 3	Digitale Filter	13
3.1	Digitale Filtertypen	13
3.2	Bestandteile digitaler Filter	13
3.3	Aufbau des FIR-Filters	14
3.4	Aufbau des IIR-Filters	15
Kapitel 4	Filterentwurf	16
4.1	Vorgabe der Filtercharakteristik	16
4.2	Fourier-Transformation	17
4.3	Realisierung	17
4.3.1	FIR-Filter	18
4.3.2	IIR-Filter	19
Kapitel 5	Nicht-Echtzeitfilter	20
	Literaturverzeichnis	22

Kapitel 1 Signale im Frequenzspektrum

Periodische Signale können durch Darstellung (mindestens) einer Schwingung im Zeitbereich komplett beschrieben werden. Für das Thema Filter ist aber eine andere, gleichwertige Darstellung des Signals nützlich, die Darstellung im Frequenzbereich. Dieses Kapitel soll helfen eine Vorstellung vom Frequenzbereich zu bekommen. Betrachten wir mal ein einfaches periodisches Signal, einen Sinus, im Zeitbereich:

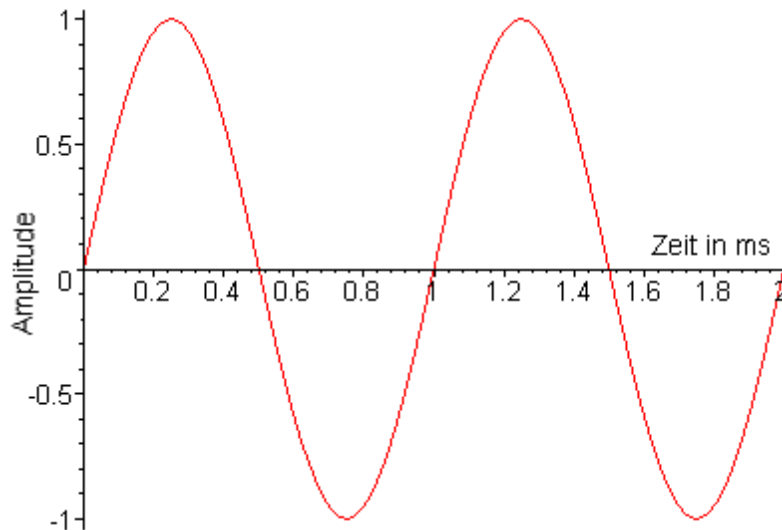


Bild 1.1: Sinusschwingung im Zeitbereich

Sofort ersichtlich: dieses Signal hat eine Amplitude von eins (zum Beispiel Volt, hier nicht weiter wichtig) und die Periodendauer 1 ms (Millisekunde = 1/1000 Sekunde). Mit der bekannten Formel $f = 1/T$ (Frequenz entspricht dem Kehrwert der Periodendauer) erhalten wir die Frequenz des Signals. Diese beträgt 1 kHz (Kilohertz = 1000 Schwingungen pro Sekunde).

Deswegen sieht die Darstellung im Frequenzbereich so aus:

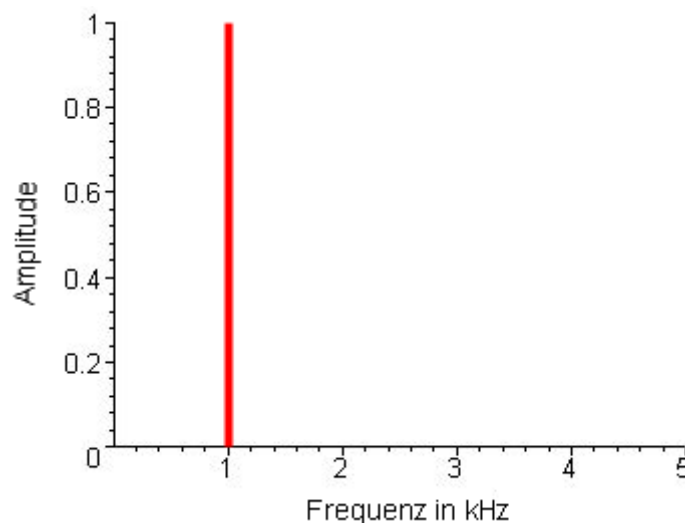


Bild 1.2: Sinusschwingung im Frequenzbereich

Betrachten wir nun ein Sägezahnsignal:

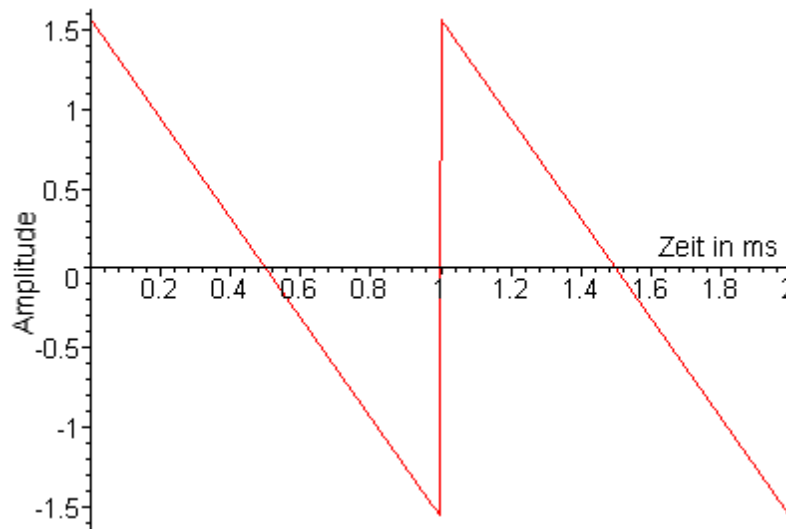


Bild 1.3: Sägezahn im Zeitbereich

Das Signal hat eine grössere Amplitude, aber die gleiche Periodendauer, also die gleiche Grundfrequenz. Dennoch kann es nicht das gleiche Frequenzspektrum haben wie der Sinus. Tatsächlich sieht der Sägezahn im Frequenzbereich so aus:

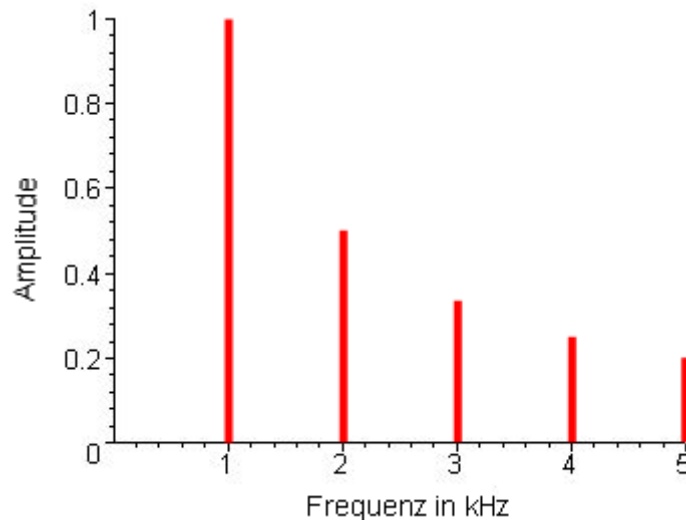


Bild 1.4: Sägezahn im Frequenzbereich (Ausschnitt)

Es enthält neben der Grundschwingung zusätzlich noch alle Harmonischen mit sinkender Amplitude. Das obige Bild 1.4 zeigt nur einen Ausschnitt aus dem Spektrum. Die einzelnen Linien symbolisieren die jeweiligen Sinusanteile des Signals. Dieses Frequenzspektrum berechnet man mit Hilfe der Fourier-Reihen.

Formel für Sägezahn:
$$\text{sägezahn}(t) =: \sum_{n=1}^{\infty} \frac{1}{n} \cdot \sin(n \cdot 2 \cdot \pi \cdot f \cdot t)$$

Andere Signalformen (z. B. Rechteck) können durch Weglassen bestimmter Harmonischen erzeugt werden.

Man kann Signale approximieren, indem man mehrere Sinusschwingungen aufaddiert. Je mehr es sind, desto mehr nähert sich das erzeugte Signal der Vorgabe. Das folgende Bild 1.5 zeigt eine Approximation des Sägezahns durch 8 Sinusschwingungen.

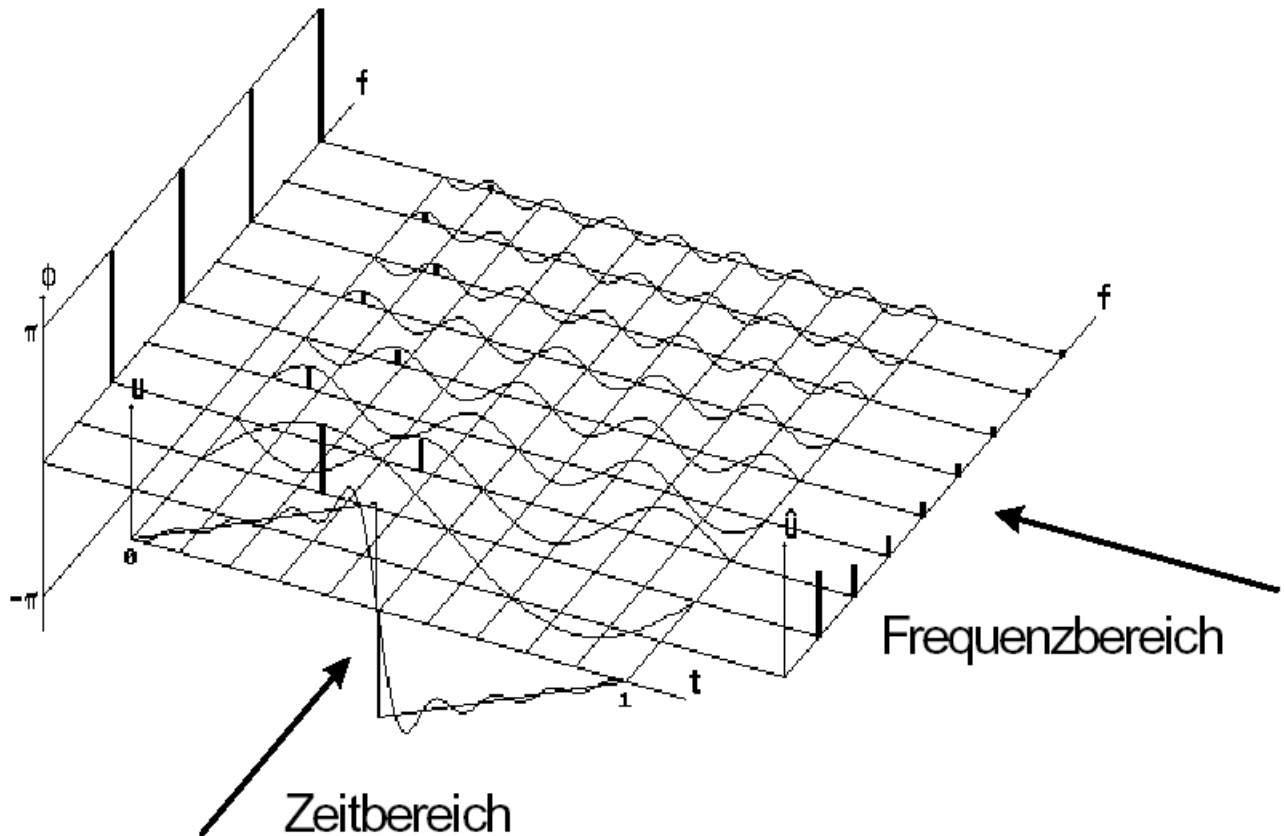


Bild 1.5: Sägezahn dargestellt im Zeit- und Frequenzbereich

Im alltäglichen Leben hat man es selten mit periodischen Signalen zu tun. Stattdessen müssen Filter sogenannte stochastische Signale verarbeiten. Sprache, Musik oder auch Rauschen gehören zu diesen Signalen.

Stochastische Signale haben unvorhersagbaren Verlauf und können aus beliebigen Frequenzanteilen bestehen. Ihr Frequenzspektrum ist im Gegensatz zu den periodischen Signalen zu jedem Zeitpunkt unterschiedlich.

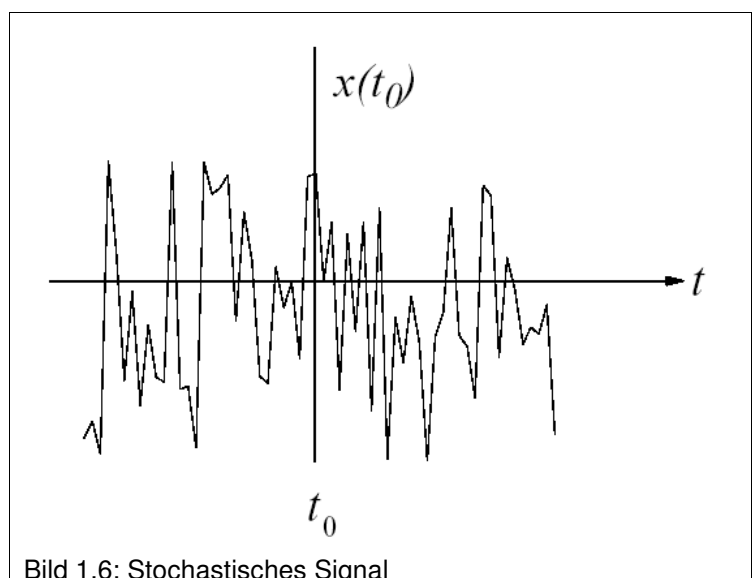


Bild 1.6: Stochastisches Signal

In allen weiteren Beispielen werden periodische Signale betrachtet, da sie wesentlich einfacher zu handhaben und bestimmte Effekte an ihnen deutlicher zu erkennen sind. Alle vorgestellten Prinzipien gelten aber ebenso für stochastische Signale.

Kapitel 2 Filter

2.1 Was ist ein Filter?

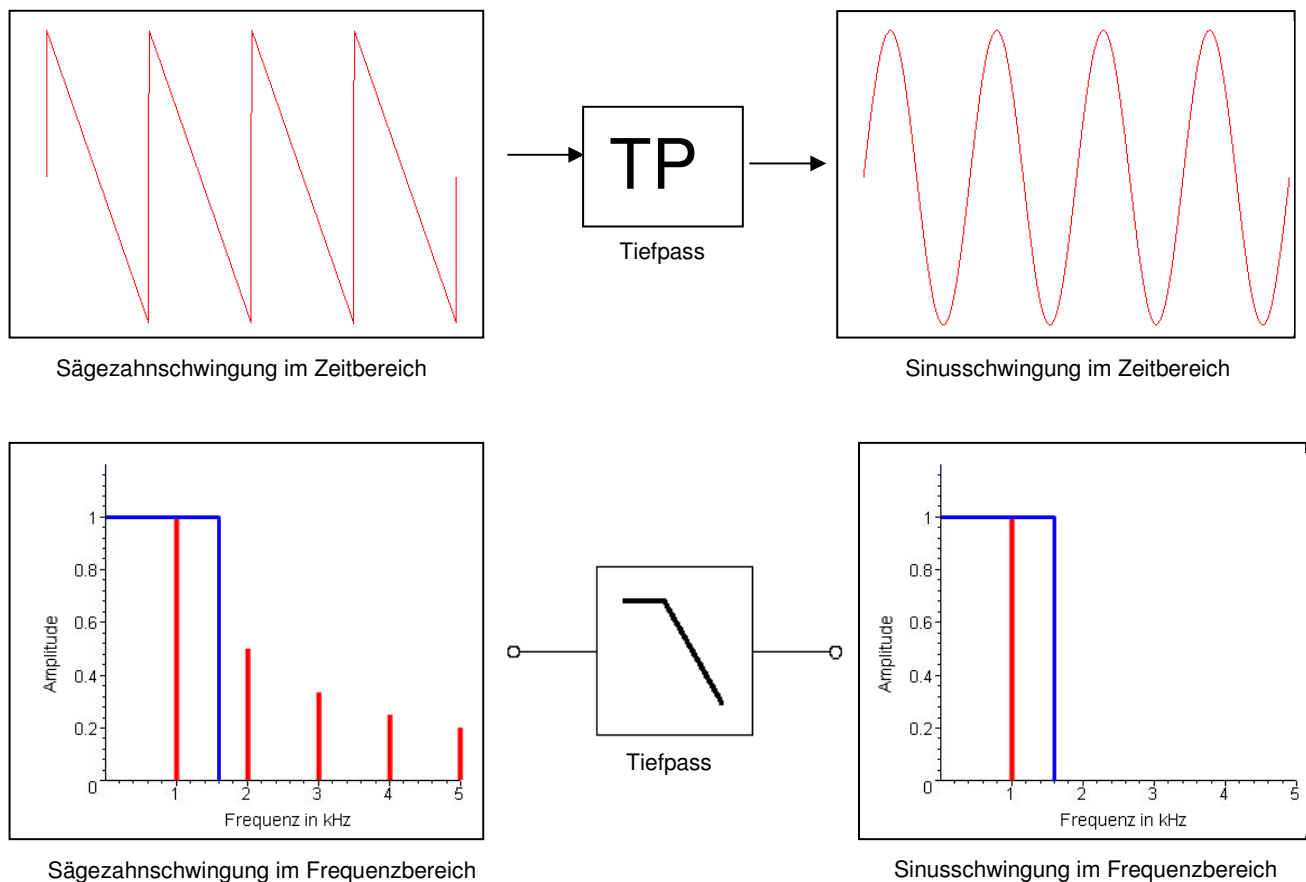


Bild 2.1: Auswirkung eines Filters auf ein Signal (oben Zeitbereich / unten Frequenzbereich)

Die Antwort auf diese Frage lässt sich leicht ermitteln, indem man die Auswirkung eines Filters betrachtet. Hilfreich ist hierbei eine geeignete Darstellungsform zu verwenden, wie man anhand des Beispiels in Bild 2.1 erkennt. So wird die Betrachtung im Zeitbereich kaum Rückschlüsse zulassen, ganz im Gegensatz zur Betrachtung der Signale im Frequenzbereich. Man kann im Beispiel sehr schön sehen wie die höheren Frequenzanteile des Eingangssignals gesperrt werden und nur die tiefen (hier nur eine 1 kHz Schwingung) Frequenzen passieren können. Das Filter arbeitet im Frequenzbereich. Im Rahmen dieser Ausarbeitung beschränken wir uns auf die Behandlung sogenannter frequenzselektiver Filter. Das heisst solche Filter sperren Frequenzanteile eines Signals. Anders formuliert: Es werden nur bestimmte (vom Filter abhängige) Frequenzen "durchgelassen", störende Anteile werden gefiltert. Anhand des Sperrverhaltens werden Filter in verschiedenen Arten gegliedert. Das Filter im obigen Einführungsbeispiel ist ein Tiefpassfilter, näheres auf Seite 6 in 2.3 Filterarten.

Grundsätzlich werden Filter in analog und digital unterschieden. Ein analoges Filter ist eine Schaltung, die aus Kondensatoren, Spulen und Widerständen aufgebaut sein kann. Das digitale Filter besteht lediglich aus Rechenoperationen, die überwiegend auf optimierten Digital-Signal-Prozessoren, sogenannten DSP's, ausgeführt werden.

Innerhalb dieses Kapitels werden grundlegende Kenntnisse vermittelt die sowohl für analoge als auch digitale Filter ihre Gültigkeit haben. Im dritten Kapitel wird speziell auf das digitale Filter eingegangen.

Übrigens: es heisst das Filter, denn in der Technik ist ein Filter sachlich.

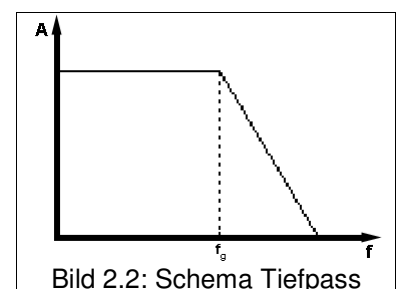
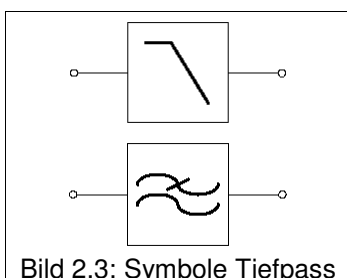
Als nächstes folgt ein kleiner Einblick in das Umfeld in dem Filter eingesetzt werden.

2.2 Anwendungsgebiete [6]

- Messwerterfassung
Bereits im Messgerät sind Filterfunktionen anzutreffen, um das Signal so aufzubereiten, dass nachfolgende Störeinflüsse wieder leichter vom Nutzsignal getrennt werden können. Ein digitales Filter kann allerdings nie als Ersatz für ein analoges Anti-Aliasing-Filter eingesetzt werden.
- Messdatenverarbeitung
Filter dienen hier einerseits dazu, erwünschte von nicht erwünschten Signalanteilen zu trennen und andererseits dazu, die im Signal steckende Information leichter zugänglich zu machen.
- Regelungs- und Prozesstechnik
Klassische Funktionen von Reglern und analog nicht realisierbare Reglertypen werden digital simuliert. Störende Eigenschaften von Prozesskomponenten (z.B. Trägheit von Sensoren) können im Nachhinein aus einem Signal herausgerechnet werden.
- Medizintechnik
Signale in Medizin und Biologie sind oft äußerst schwach und störanfällig oder haben ihre Information in periodischen Anteilen oder in der Abweichung von der Periodizität versteckt.
- Bildverarbeitung
In der Bildverarbeitung werden die Filterverfahren auf zwei- und dreidimensionale Objekte erweitert. Auch die zur Bildkompression benutzten Verfahren (Wavelets) basieren auf Digitalfiltern.
- Audiotbearbeitung
Typische Tonstudio-Effekte, die sich mit digitalen Filtern erzeugen lassen sind, neben den Equalizer-Funktionen, Raumklang, Echo, Stereo-Basisverbreiterung, Ausblenden von Stimmen und Frequenzverschiebungen (Vocoder).

2.3 Filterarten

Filter werden in vier verschiedene Arten eingeteilt. Als erstes behandeln wir das schon anfangs kennengelernte Tiefpassfilter (Bild 2.2). Wie der Name schon sagt, lässt es tiefe Frequenzen passieren und sperrt oberhalb seiner Grenzfrequenz alle hohen Frequenzen. Die beiden Symbol-Varianten (Bild 2.3) für das Tiefpass (engl. low-pass) spiegeln den Charakter des Filters wieder. Einmal wird die typische Kennlinie verwendet, das anderemal werden durch zwei übereinander angeordnete Wellen hohe und tiefe Frequenzen angedeutet. Wobei die Sperrwirkung durch eine durchgestrichene obere "Welle" symbolisiert wird.



Ganz analog ist das Hochpassfilter (engl. high-pass filter) das exakte Gegenstück zum Tiefpassfilter. Es blockiert die tiefen Frequenzen unterhalb seiner Grenzfrequenz und lässt die hohen Frequenzen ungehindert durch.

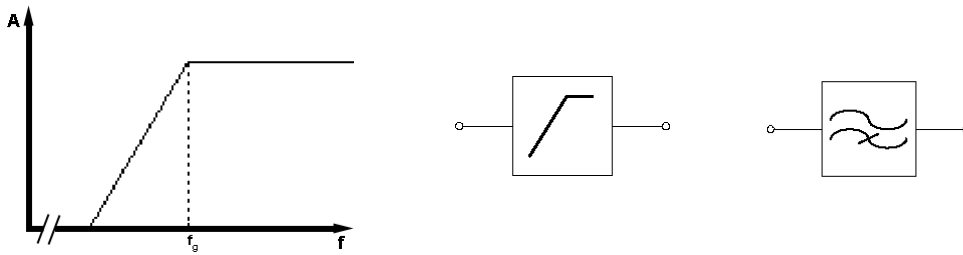


Bild 2.4: Schema eines Hochpassfilters mit beiden Symbolvarianten

Die Kombination von Tief- und Hochpassfiltern ermöglichen zwei weitere Filterarten. Schaltet man Tief- und Hochpass in Reihe (seriell), dann entsteht ein sogenannter Bandpass (engl. bandpass). Zu beachten ist, dass sich die Sperrbereiche der Filter nicht überschneiden, sonst kommt es zu einer Auslöschung des Eingangssignals. Das Prinzip ist ganz einfach: Das Eingangssignal durchläuft, hier im Beispiel (Bild 2.5), zuerst das Hochpassfilter → tiefe Frequenzen werden gesperrt. Das gefilterte Signal liegt nun am Eingang des Tiefpasses an. Dieses sperrt nun hohe Frequenzen und ausschliesslich ein "Band von Frequenzen", die sowohl Hoch- als auch Tiefpassfilter passieren ließen, sind im Ausgangssignal enthalten.

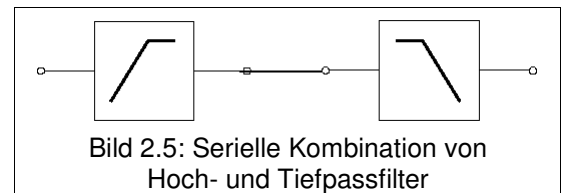


Bild 2.5: Serielle Kombination von Hoch- und Tiefpassfilter

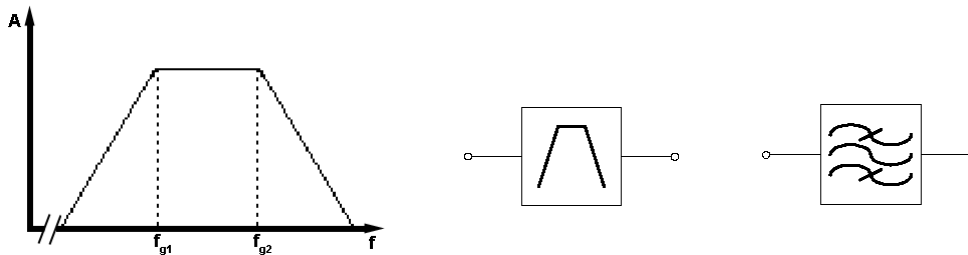


Bild 2.6: Schema eines Bandpassfilters mit beiden Symbolvarianten

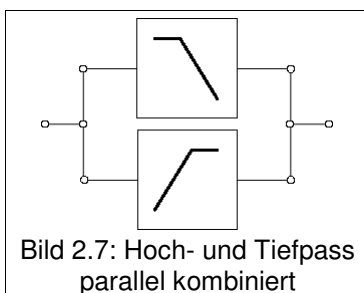


Bild 2.7: Hoch- und Tiefpass parallel kombiniert

Die andere Möglichkeit ist Hoch- und Tiefpass parallel zu kombinieren, wie zum Beispiel im rechten Bild 2.7 zu sehen ist. Hier ist es wichtig, dass sich die Sperrbereiche überlappen. Denn das Eingangssignal liegt gleichzeitig an beiden Filtern an und deren Ausgänge werden zu einem Ausgangssignal "addiert". Somit wird ein Band von Frequenzen gesperrt, welche beide Filter nicht passieren konnten. Daher auch die Bezeichnung Bandsperre (engl. notch).

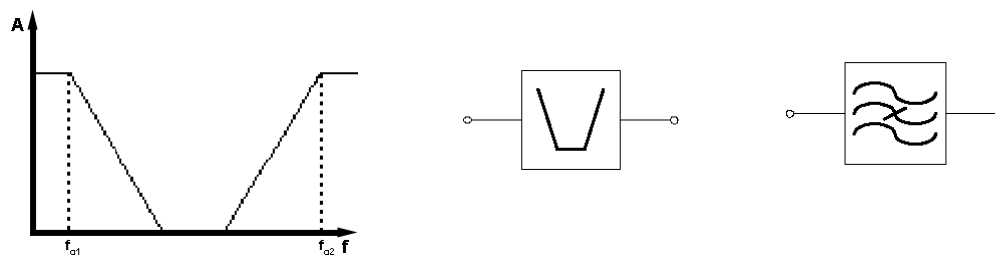


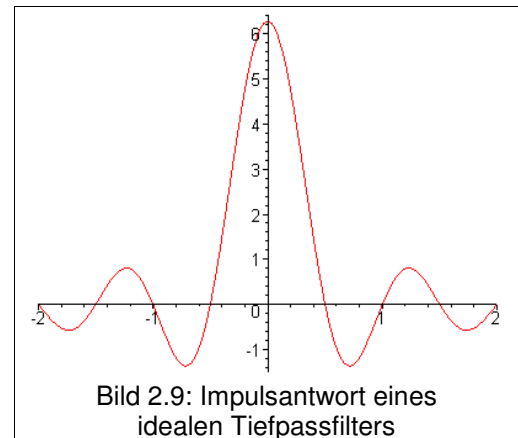
Bild 2.8: Schema einer Bandsperre mit beiden Symbolvarianten

2.4 Grundlagen

Ein Filter (allgemein LTI-System, mehr in 2.4.4 Filterkriterien) ist durch seine Impulsantwort vollständig charakterisiert.

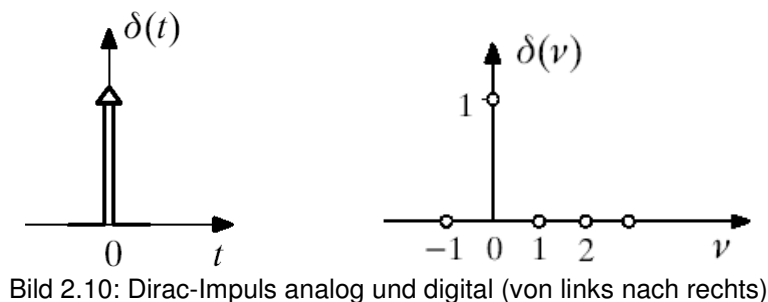
2.4.1 Impulsantwort

Unter der Impulsantwort versteht man die Reaktion eines LTI-Systems auf einen speziellen Impuls (siehe 2.4.2 Dirac-Impuls). Mithilfe der Impulsantwort kann durch Faltung (siehe 2.4.3 Faltung) mit dem Eingangssignal, das Ausgangssignal berechnet werden.



2.4.2 Dirac-Impuls

Der sogenannte Einheitsimpuls wurde nach dem britischen Physiker Paul Adrien Maurice Dirac (1902 – 1984) benannt. Im Zeitkontinuierlichen hat der Impuls zum Zeitpunkt $t = 0$ eine Amplitude mit unendlicher Höhe, ansonsten null. Die Fläche des Impulses ist eins. Im Zeitdiskreten ist, zum Abtastwert $n = 0$, die Höhe des Impulses gleich eins, ansonsten gleich null. Die Faltung eines Signals mit dem Dirac-Impuls ergibt wiederum das Signal (Identität der Faltung).



2.4.3 Faltung (S. 31 [3])

Die Faltung (engl. convolution) für zwei Signale $x_1(t)$ und $x_2(t)$ ist definiert als:

$$y(t) = x_1(t) * x_2(t) = \int_{-\infty}^{\infty} x_1(\tau) \cdot x_2(t - \tau) d\tau$$

Das Faltungsintegral (auch Duhamel-Integral genannt) wird so häufig gebraucht, dass dafür die abgekürzte Schreibweise mit dem Sternsymbol eingeführt wurde.

Schema für die Durchführung der Faltung:

1. Zeitvariable t durch τ ersetzen
2. $x_2(\tau)$ an der y -Achse spiegeln (falten!), dies ergibt $x_2(-\tau)$
3. $x_2(-\tau)$ um t_1 nach rechts verschieben, dies ergibt $x_2(t_1 - \tau)$
4. $x_1(\tau)$ und $x_2(t_1 - \tau)$ miteinander multiplizieren

5. Dieses Produkt über alle τ integrieren, ergibt $y(t_1)$
6. t_1 von $-\infty$ bis $+\infty$ variieren und jeweils Schritte 1-6 ausführen, dies ergibt das Signal $y(t)$

Durch das Einsetzen obiger Formel in die Fouriertransformierte von $y(t)$, betrachtet man den Faltungsvorgang im Frequenzbereich.

$$Y(\omega) = \int_{-\infty}^{+\infty} y(t) \cdot e^{-j\omega t} dt$$

Das Ergebnis ist, dass eine Faltung im Zeitbereich einer Multiplikation im Frequenzbereich (und umgekehrt) entspricht. Die Faltung ist kommutativ.

Zur Verdeutlichung des Faltungsvorgangs betrachten wir ein Beispiel.

Beispiel: Die Faltung eines kausalen Rechteckimpulses mit sich selbst (S.62 [4])

$$y(t) = x(t) * x(t) = \int_{-\infty}^{+\infty} x(\tau) \cdot x(t - \tau) d\tau \quad \text{mit } x(t) = x_T \left(t - \frac{T}{2}\right)$$

Die Faltung lässt sich wie in Bild 2.11 (Seite 10) veranschaulicht bestimmen:

Für einen beliebigen aber fest vorgegebenen Wert t ist im Integranden das Produkt aus $x(\tau) \cdot x_T(t - \tau)$ zu bilden. Hierzu ist in Bild 2.11 das Signal $x(\tau)$ und $x(t - \tau)$ für die Werte $t=0, T/2, T, 3T/2$ und $2T$ gezeigt. Für $t = 0$ erhält man für $x(-\tau)$ wegen des Minuszeichens im Argument den an der Ordinate gespiegelten (gefalteten) Rechteckimpuls. Die Spiegelung wird im Bild durch ein kleines Dreieck verdeutlicht.

Für $t = T/2$ verschiebt sich der Rechteckimpuls um $T/2$ nach rechts, usw.

Im Integranden ist das Produkt aus dem Rechteckimpuls und seiner gespiegelten und verschobenen Replik zu bilden. Da im Beispiel $x(\tau) = 0$ für $\tau < 0$ oder $\tau > T$ gilt, werden vier Fälle unterschieden:

(i) $t < 0$

Der Integrand ist für $t < 0$ identisch null, da sich $x(\tau)$ und $x(t - \tau)$ nicht überlappen. Demzufolge liefert die Faltung den Wert null, wie in der rechten Hälfte von Bild 2.11 gezeigt wird.

(ii) $0 \leq t < T$

Für wachsendes t schiebt sich der Rechteckimpuls $x(t - \tau)$ mehr und mehr unter den Rechteckimpuls $x(\tau)$. Da es sich um Rechteckimpulse handelt nimmt die vom Produkt der beiden Rechteckimpulse eingeschlossene Fläche linear zu.

(iii) $T \leq t < 2T$

Für $t = T$ decken sich die Rechteckimpulse vollständig. Die Faltung liefert den Maximalwert. Mit wachsendem t schiebt sich $x(t - \tau)$ aus dem Abszissenbereich des Rechteckimpulses $x(\tau)$. Die von den beiden Rechteckimpulsen gemeinsam eingeschlossene Fläche nimmt linear ab.

(iv) $2T \leq t$

Für $t \geq 2T$ ist der Integrand und damit das Ergebnis der Faltung identisch null.

Das gesuchte Faltungsergebnis ist in Bild 2.11 unten rechts gezeigt. Es resultiert ein Dreiecksimpuls mit der Breite $2T$ und der Höhe T .

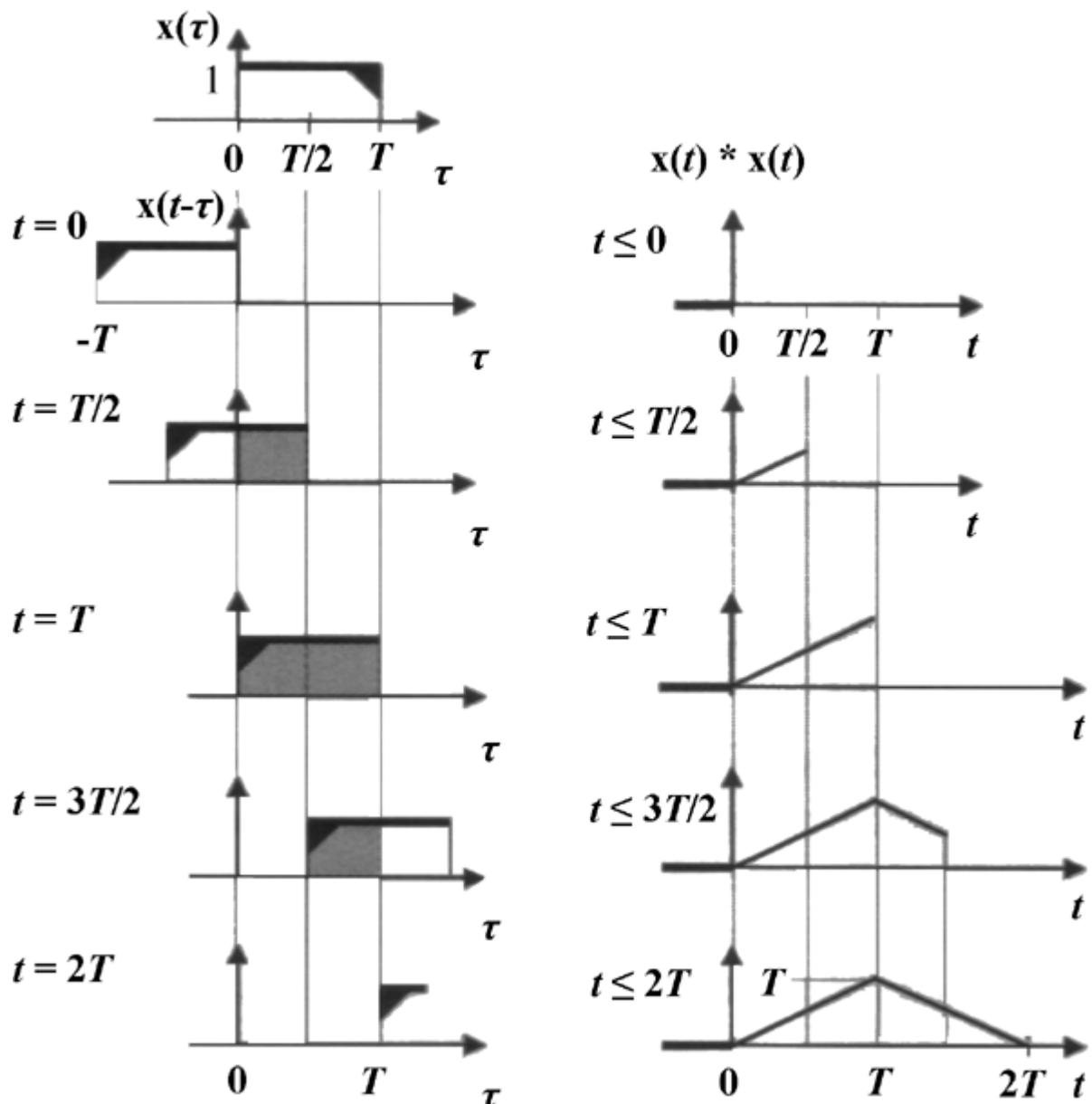


Bild 2.11: Faltung eines kausalen Rechteckimpulses mit sich selbst

Weitere Beispiele mit Animationen zur Veranschaulichung des Faltungsvorgangs, sind auf den Folien 15 – 17 der Powerpoint-Präsentation zu sehen.

2.4.4 Filterkriterien

Ein Filter gehört zu den sogenannten LTI-Systemen (Linear TimeInvariant) und muss insgesamt vier Kriterien erfüllen: Linearität, Zeitinvarianz, Kausalität und Stabilität.

Bei linearen Systemen ist die Abbildung einer Linearkombination von Eingangssignalen identisch mit der gleichen Linearkombination der zugehörigen Ausgangssignale.

Unter Zeitinvarianz versteht man, dass jedes Ausgangssignal nur vom Eingangssignal abhängig ist und nicht von einem bestimmten Zeitpunkt. Analoge Filter sind streng genommen zeitvariant, denn durch Erwärmung der Bauteile (Kondensator, Spule) verändert sich ihre Impedanz und mit ihr die Filtereigenschaft.

Ein kausales System reagiert erst dann mit einem Ausgangssignal, wenn ein Eingangssignal anliegt. Die Impulsantwort eines kausalen Systems verschwindet für $t < 0$. Ist ein System nicht kausal so ist es technisch auch nicht realisierbar.

Ein System ist stabil, falls beschränkte Eingangswerte beschränkte Ausgangswerte liefern, das heißt die Impulsantwort ist absolut summierbar.

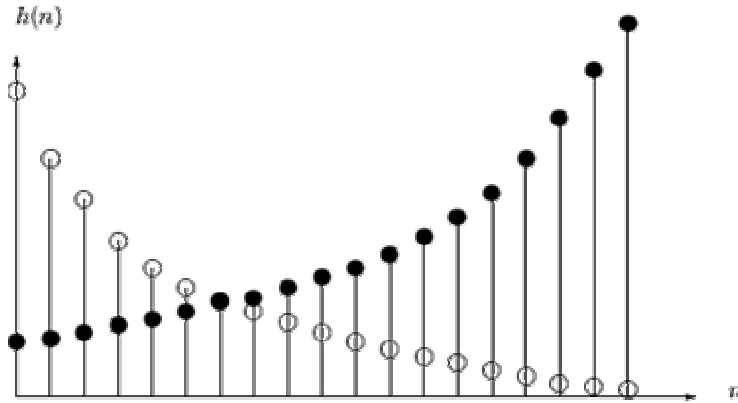


Bild 2.12: Ausgangsfolge eines stabilen (leere Kreise) und instabilen (ausgefüllte Kreise) Systems als Reaktion auf einen Einheitsimpuls

2.5 Ideales Filter

Diese Filter nennt man ideal, nicht weil sie alle Filterkriterien erfüllen, das tun sie nämlich nicht. Sie sind nicht kausal, also nicht realisierbar. Allerdings ist die Sperrwirkung des Filters ideal. Das heißt Frequenzen, die den Filter passieren, werden in ihrer Amplitude nicht verändert ("Multiplikation mit der Kennlinie": Durchlassbereich = 1, Sperrbereich = 0). Hingegen Frequenzen oberhalb der Grenzfrequenz (im Beispiel des idealen Tiefpasses, Bild 2.13) werden komplett gesperrt (Amplitude = 0). Der Übergang von Durchlass- und Sperrbereich, auch Flanke genannt, ist also unendlich steil. Im Bild 2.9, wurde die Impulsantwort eines idealen Tiefpassfilters gezeigt. Hier erkennt man, dass diese für die Zeit $t < 0$ nicht verschwindet (Widerspruch zur Kausalität).

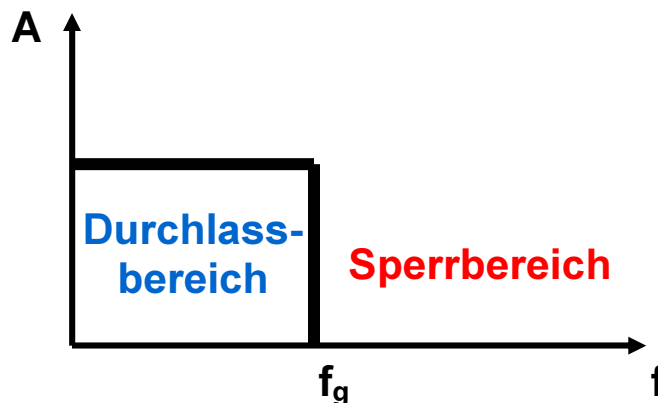


Bild 2.13: Ideales Tiefpassfilter

2.6 Reales Filter

Das reale Filter ist im Gegensatz zum idealen Filter kausal.

Vergleicht man die Kennlinie des idealen (Bild 2.13) mit der des realen Tiefpasses (Bild 2.14), so fallen zwei Dinge auf: Die Kennlinie ist im Sperr- und Durchlassbereich "wellig" und die Flanke ist nicht mehr unendlich steil. Durch die Welligkeit wird die Amplitude der Frequenzen verändert. Da die Flanke sich nun über ein Band von Frequenzen erstreckt, ist eine Definition der Grenzfrequenz nicht mehr trivial. Die Grenze von Durchlass- und Übergangsbereich nennt man Eckfrequenz. Frequenzen im sogenannten Übergangsbereich (engl. roll-off) werden ebenso im Ausgangssignal auftreten, wenn auch mit abgeschwächter Amplitude. Die Grenzfrequenz wurde deshalb definiert als Frequenz, deren Amplitude so groß ist wie dem 0,707-fachen der Eingangsspannung. Dies entspricht einer 3dB Abschwächung der Amplitude der Eingangsspannung. Für den Grad der Flankensteilheit hat man den Begriff der Ordnung eingeführt. Man definiert eine 6dB Abschwächung pro Oktave als erste Ordnung. Jede weitere Ordnung entspricht dem Vielfachen der ersten. Zum Beispiel hat ein Filter der vierten Ordnung eine Flankensteilheit von 24dB pro Oktave.

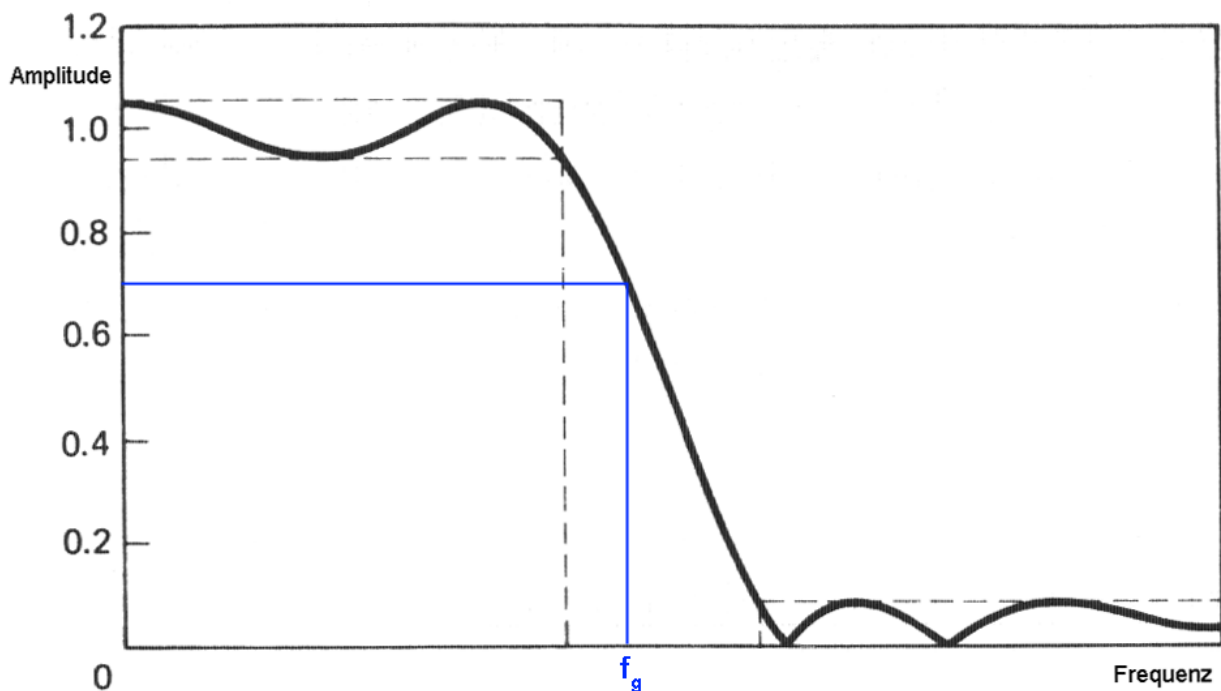


Bild 2.14: Beispiel für ein reales Tiefpassfilter

Am Ende des zweiten Kapitels sei noch erwähnt, dass Filter deren Ausgangssignale eine höhere Amplitude besitzen als ihre Eingangssignale, also die Amplituden verstärkt wurden, auch als aktive Filter bezeichnet werden. Wird die Amplitude abgeschwächt (oder bleibt im Idealfall identisch) dann spricht man von einem passiven Filter.

Kapitel 3 Digitale Filter

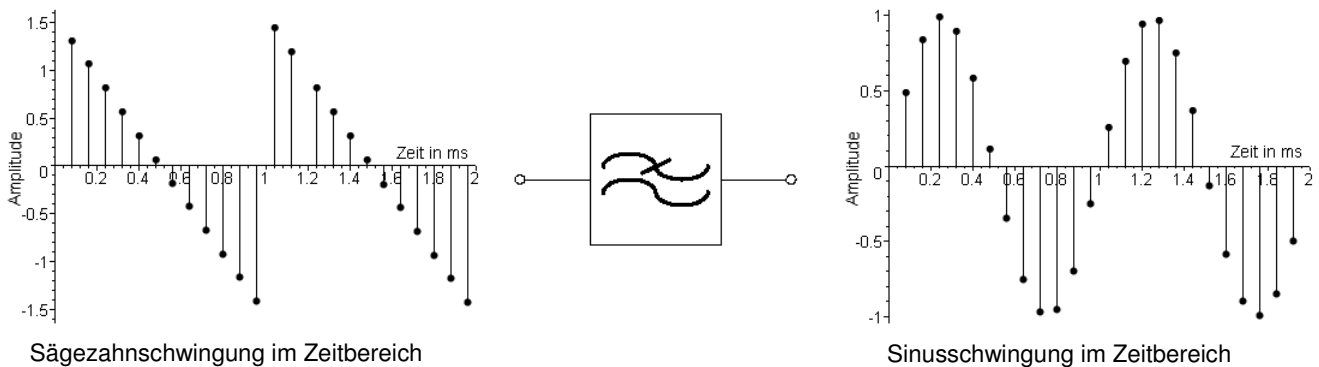


Bild 3.1: Digitales Tiefpass mit Ein- und Ausgangssignal

Bild 3.1 verdeutlicht die Besonderheit der digitalen Filter. Deren Eingabe besteht aus Abtastwerten eines Signals (in diesem Fall Sägezahn) und als Ausgabe müssen sie die Abtastwerte des gefilterten Signals liefern. Dieses Beispiel entspricht dem Beispiel aus Bild 2.1, das digitale Tiefpass hat also die gleiche Charakteristik. Deswegen ist die Ausgabe wieder die Grundschwingung des Sägezahns.

3.1 Digitale Filtertypen

Es gibt zwei Arten von Digitalfiltern: die FIR- und die IIR-Filter. Sie unterscheiden sich in der Länge ihrer Impulsantworten.

FIR steht für Finite Impulse Response, also endliche Impulsdauer. Das bedeutet, dass wenn man einen Dirac-Impuls an den Eingang dieses Filters gibt, seine Impulsantwort nach N Takten gleich 0 ist. N gibt dabei gleichzeitig die Ordnung dieses Filters an. FIR-Filter sind relativ einfach zu entwerfen und werden deshalb bevorzugt eingesetzt. Wegen der endlichen Impulsantwort sind sie immer stabil.

IIR steht für Infinite Impulse Response. Sie sind rekursiv aufgebaut, das Ausgangssignal wird also an den Eingang zurückgeführt. Die Impulsantwort ist deshalb unendlich. Da man das Kriterium der Stabilität beim Entwurf einhalten muss, ist der Entwurf an sich komplizierter als bei den FIR-Filtern. Andererseits erreicht man die gleichen Kennlinieneigenschaften mit IIR-Filtern geringerer Ordnung, man braucht also weniger Bauteile um dieses Filter zu realisieren.

3.2 Bestandteile digitaler Filter

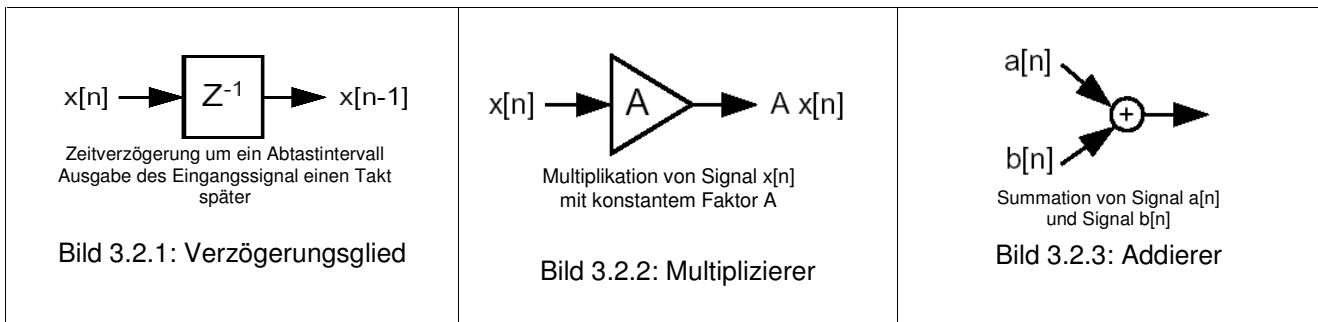
Beide Filtertypen sind aus lediglich 3 verschiedenen Komponenten aufgebaut.

Das Verzögerungsglied verzögert das Signal um einen Takt. Es speichert den aktuellen Signalwert ab und gibt gleichzeitig den Signalwert vom vorherigen Takt aus. Die Stückzahl dieser Komponente gibt die Ordnung des Filters an.

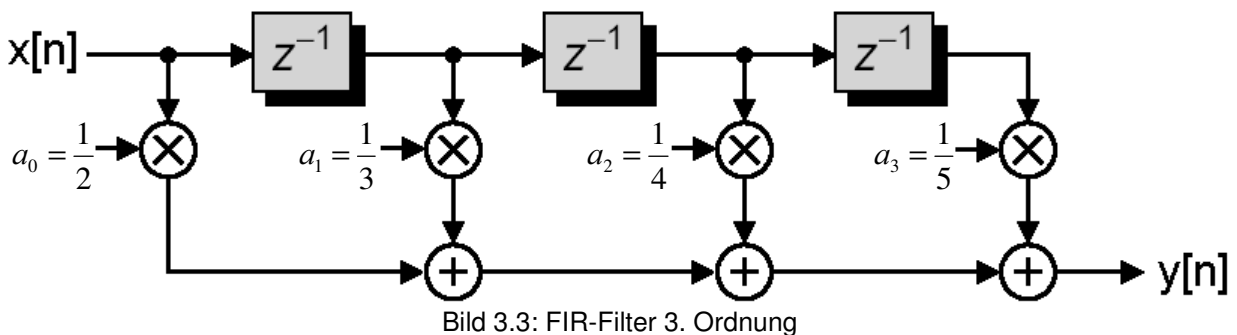
Der Multiplizierer skaliert jeden Signalwert mit einem konstanten Faktor, verändert also seine Amplitude.

Der Addierer summiert zwei Signalwerte zu einem auf.

Mit diesen Komponenten können wir jeden FIR- und IIR-Filter zusammenbauen.



3.3 Aufbau des FIR-Filters



Das Bild 3.3 zeigt den Aufbau eines Filters 3. Ordnung. Durch den Aufbau ist ersichtlich, warum die Impulsantwort des Filters endlich ist. Der Signalwert zum Zeitpunkt n wird in jedem Takt skaliert ausgegeben und gleichzeitig verzögert. Man kann die Ordnung des Filters erhöhen, indem man rechts zusätzliche Komponenten (Bestehend aus Verzögerungsglied, Summierer, Addierer) hinzufügt.

Berechnen wir mal die Impulsantwort dieses Filters. Wiederholung: Impulsantwort ist das Ausgangssignal bei einem diskreten Dirac-Impuls am Eingang zum Zeitpunkt 0. Das Ausgangssignal ist 0 für $n < 0$.

-
- Takt 0: Der Dirac wird abgespeichert und gleichzeitig halbiert und ausgegeben. Also der Wert 0,5 als Ausgabe.
- Takt 1: Der Dirac verlässt das erste Verzögerungsglied, wird vom zweiten gespeichert und gleichzeitig skaliert ausgegeben. Der Wert 0,33 als Ausgabe.
- Takt 2: Der Wert 0,25 wird ausgegeben.
- Takt 3: 0,2 wird ausgegeben.

Für alle späteren Takte gibt der Filter den Wert 0 aus.

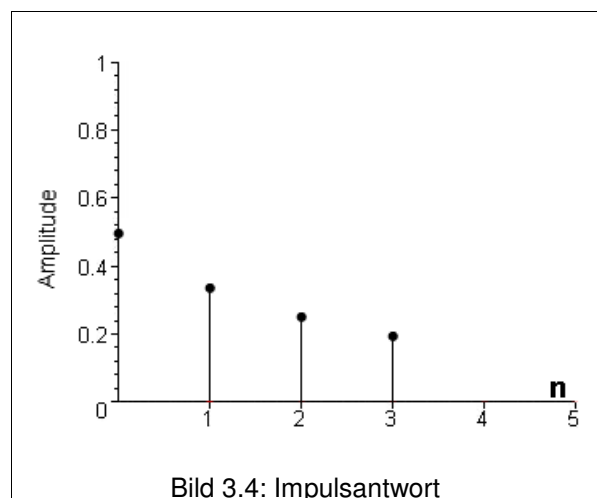


Bild 3.4 zeigt die resultierende Impulsantwort unseres Beispiels aus Bild 3.3.

3.4 Aufbau des IIR-Filters

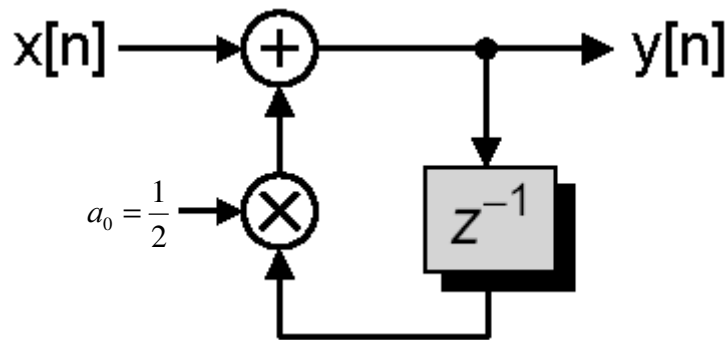


Bild 3.5: IIR-Filter 1. Ordnung

Bild 3.5 zeigt den einfachsten IIR-Filter (Ordnung 1). Die Rekursivität des Filter ist aus der Schaltung sofort einsichtig. Das einfache Beispiel wurde gewählt, um die Impulsantwort leicht berechnen zu können. Schon bei Filtern zweiter Ordnung wird dies nämlich kompliziert.

In unserer Schaltung wird der Signalwert am Eingang sofort ausgegeben und gleichzeitig vom Verzögerungsglied gespeichert. Im nächsten Takt wird dieser Wert skaliert, ausgegeben und gleichzeitig gespeichert. Im nächsten Takt dann wieder, und wieder, und wieder...

Betrachten wir wieder die Impulsantwort:

- Takt 1: Der Impuls wird gespeichert und ausgegeben. Ausgabewert: 1.
- Takt 2: Der zwischengespeicherte Impuls wird mit 0,5 skaliert, ausgegeben und gespeichert. Ausgabe: 0,5.
- Takt 3: Ausgabe: 0,25.
- Im nächsten Takt 0,125 usw.

Aufsummiert konvergiert diese Reihe gegen 2. Dieses Filter ist also stabil. Das folgende Bild zeigt einen Ausschnitt der Impulsantwort:

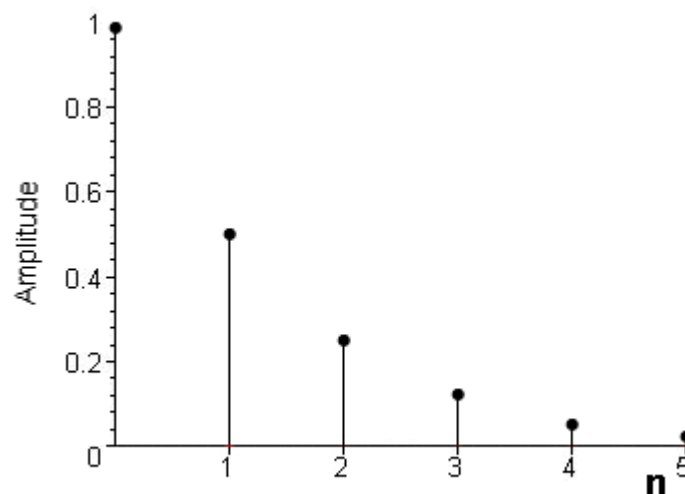


Bild 3.6: Impulsantwort (Ausschnitt)

Kapitel 4 Filterentwurf

Den Entwurf digitaler Filter (FIR/IIR) kann man grob in 3 Schritte unterteilen. Die Grenzen zwischen den einzelnen Schritten verfließen jedoch bei näherem Hinsehen. Deswegen beginnt der später vorgestellte Algorithmus bei Schritt 1 und nicht erst bei Schritt 3. Die Unterteilung dient eher dem systematischeren Verständnis der Vorgehensweise.

4.1 Vorgabe der Filtercharakteristik

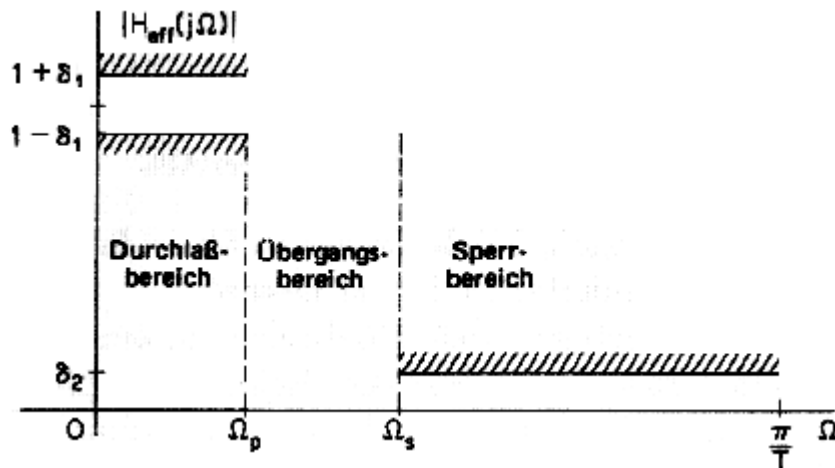


Bild 4.1: Entwurfsvorlage

Für die Charakteristik muss man insgesamt 4 Parameter vorgeben.

Man legt fest, welche Frequenzanteile durchgelassen und welche gesperrt werden sollen. Dazwischen müssen Übergangsbereiche vorgesehen werden, die nicht die Breite 0 haben dürfen. Ausserdem muss man festlegen, in welchem Bereich die Amplitude im Durchlass- und im Sperrbereich schwanken darf. Man gibt also keine Kennlinie vor, sondern ein Toleranzschema, in dem die spätere Kennlinie verlaufen soll.

Die Kennlinie selbst benötigen wir aber nicht als Bild, sondern als mathematischen Ausdruck (sprich: Formel). Da es viele Möglichkeiten für das Aussehen der Kennlinie gibt und auch unterschiedliche wichtige Aspekte für das spätere Filter vorgegeben sein könnten (z. B. maximal glatter Verlauf im Sperrbereich) gibt es mehrere verschiedene Verfahren, um diesen Ausdruck zu erzeugen.

Betrachten wir mal kurz 3 dieser Verfahren:

- Das Butterworth-Filter (im Bild 4.2 oben links) hat eine maximal glatte Kennlinie als Vorgabe.
- Oben rechts sieht man die Kennlinie des Chebyschew-Filters. Dieses Filter ist wellig im Durchlassbereich und glatt im Sperrbereich. Seine Flanke ist steiler als beim Butterworth bei gleicher Ordnung. Bei seinem Abgleich wird der Betrag der maximalen Abweichung minimiert.
- Nutzt man die Toleranzen im Durchlass- und Sperrbereich voll aus, so erreicht man die maximale Steilheit der Flanke im Übergangsbereich. Dies geschieht beim sog. elliptischen oder Cauer-Filter (im Bild 4.2 unten).

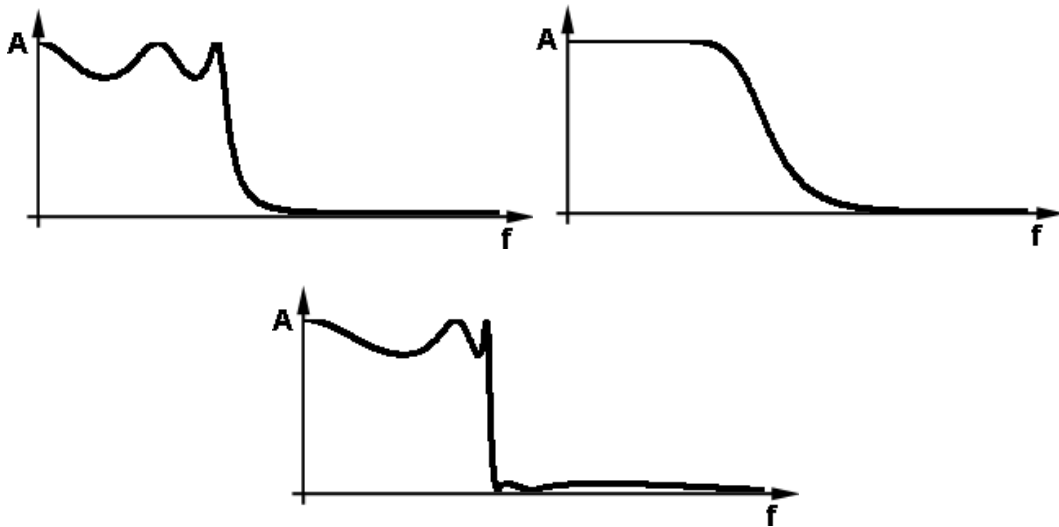


Bild 4.2: Filterkennlinien (von links nach rechts und oben nach unten) Butterworth-, Chebyschew- und Cauer-Filter (alle 5. Ordnung)

4.2 Fourier-Transformation

Nachdem die Formel für die Kennlinie vorliegt, berechnet man die Impulsantwort des Filters mit Hilfe der ZDFT, der Zeitdiskreten-Fourier-Transformation.

Die Fourier-Transformation ist ein Verfahren mit dem man Signale aus der Zeitdarstellung in die Frequenzdarstellung bringen kann (Analyse) und zurück (Synthese).

Für verschiedene Signalarten gibt es verschiedene Verfahren, die alle nach Jean Baptiste Joseph Fourier (1768-1830) benannt sind. Es handelt sich dabei immer um zwei Formeln (Analyse/Synthese).

Mit Hilfe der Fourier-Reihen können periodische kontinuierliche Signale durch unendliche Summen von Sinusschwingungen approximiert werden. Dieses Verfahren war bereits vor der Zeit Fouriers bekannt, wurde aber in seiner Bedeutung unterschätzt.

Fourier-Transformation wurde von Fourier aus den Fourier-Reihen entwickelt. Seine Idee war ein nichtperiodisches Signal als Grenzfall eines periodischen Signals mit unendlich langer Periodendauer zu betrachten. Die Summen werden dadurch zu Integralen.

Für diskrete Signale gibt es die ZDFT (Zeitdiskrete-Fourier-Transformation) und die DFT (Diskrete-Fourier-Transformation). Die DFT hat im Gegensatz zur ZDFT auch ein diskretes Frequenzspektrum.

$$F(k) = \sum_{n=0}^{N-1} f(n) \cdot e^{-i \cdot 2 \cdot \pi \cdot k \cdot \frac{n}{N}} \qquad f(n) = \frac{1}{N} \cdot \sum_{k=0}^{N-1} F(k) \cdot e^{i \cdot 2 \cdot \pi \cdot k \cdot \frac{n}{N}}$$

Formeln der DFT

Als Beispiel sieht man hier die Synthese-/Analyseformeln der DFT. Die Formeln der anderen Verfahren sehen ähnlich aus.

Gesondert zu erwähnen ist die Fast-Fourier-Transform (FFT). Dieses Verfahren ist keine Eigenständige Transformation, sondern eine schnelle Variante der DFT. Es wird im Bereich der Nichtezeitfilter eingesetzt (beim filtern von aufgezeichneten Daten).

4.3 Realisierung

Die berechnete Impulsantwort bildet man in Hardware nach. Man muss sich zwischen FIR und IIR entscheiden. Beide Filterarten haben ihre Vor- und Nachteile.

4.3.1 FIR-Filter

Da sie einfacher zu entwerfen sind, entscheidet man sich meistens für die FIR-Filter. Bereits in den 60er Jahren wurden fertige Verfahren entwickelt, die Toleranzgrenzen als Eingabe verlangen und eine fertige Schaltung ausgeben. Je nach Algorithmus müssen die beiden Frequenzen des Übergangsbereichs, die Toleranzen für den Durchlass- oder Sperrbereich vorgegeben werden, aber nicht alles gleichzeitig!

Die Verfahren sind nach ihren Entwicklern benannt:

- Hermann und Schüssler
- Hofstetter
- Rabiner
- Parks-McLellan

Von diesen Verfahren ist der Parks-McLellan-Algorithmus der wichtigste.

Schauen wir am Beispiel der "Fenstermethode" ein mögliches Vorgehen eines solchen Algorithmus mal näher an:

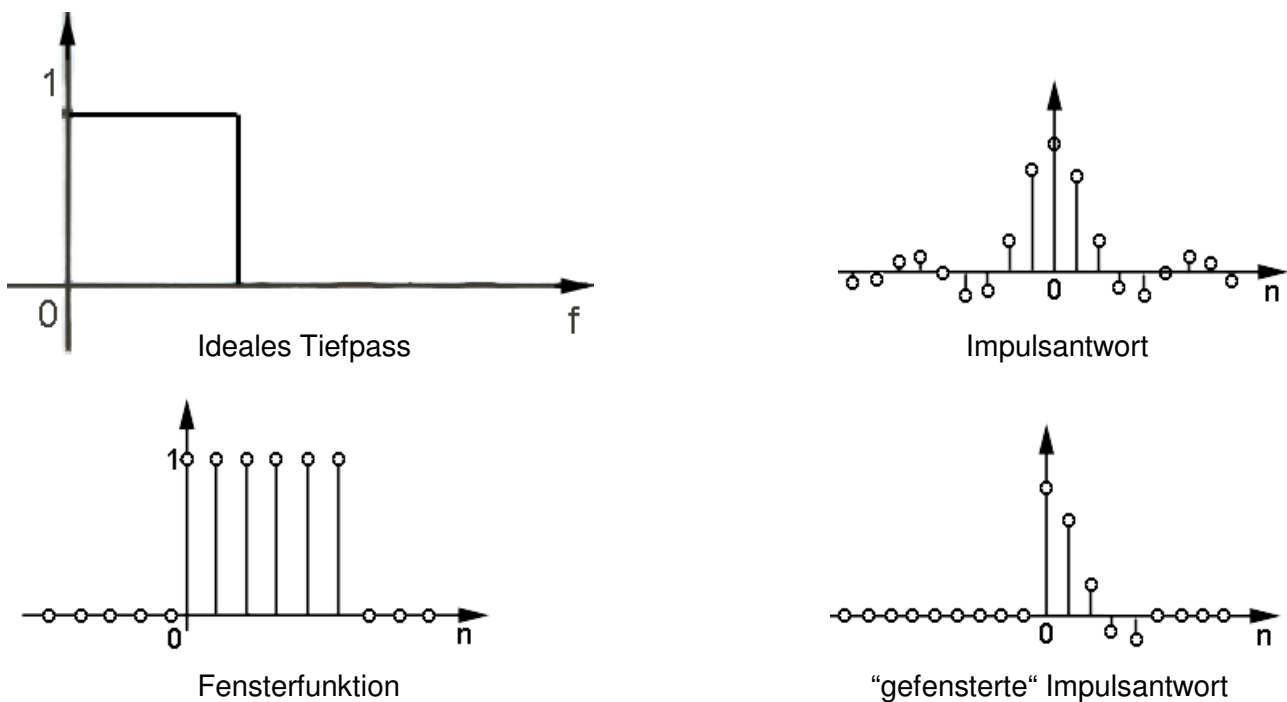


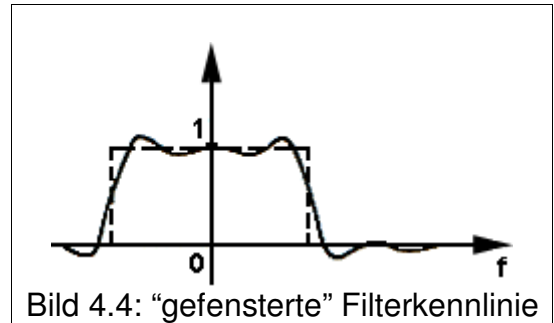
Bild 4.3: Schritte der "Fenstermethode"

Bei dieser Methode gibt man als erstes die Kennlinie vor. Wie aus dem Bild 4.3 oben links ersichtlich, kann es sich dabei auch um ein ideales Filter (in diesem Fall Tiefpass) handeln. Mit Hilfe der ZeitdiskretenFT berechnet man die Impulsantwort dieses Filters (oben rechts). Eine solche Impulsantwort ist nicht realisierbar, da sie bereits vor dem Zeitpunkt $n = 0$ Werte ungleich 0 besitzt. Deswegen kausalisieren wir die Impulsantwort mit der "Fensterfunktion" (unten links). Je nach Ordnung des späteren Filters hat diese Funktion N Werte, die gleich 1 sind. Die restlichen sind 0. Man multipliziert die Impulsantwort mit der Fensterfunktion und erhält die "gefensterte" Impulsantwort (unten rechts). Dies ist die Impulsantwort des späteren Filters, falls die "gefensterte" Filterkennlinie in unser Toleranzschema passt. Die Werte der Impulsantwort können dann als Multiplikatoren in die einzelnen Stufen des FIR übernommen werden.

Doch zuerst müssen wir die "gefensterte" Kennlinie berechnen. Multiplikation im Zeitbereich bedeutet Faltung im Frequenzbereich. Wir müssen also die Fensterfunktion

mit der ZDFT transformieren und mit der vorgegebenen Kennlinie falten. Bild 4.4 zeigt das mögliche Ergebnis.

Falls die Kennlinie nicht ins Toleranzschema passt, müssen wir die Fensterfunktion ändern (z. B. Ordnung erhöhen) und dann die Kennlinie neu berechnen.



4.3.2 IIR-Filter

Wie der Name schon sagt, haben die IIR-Filter eine unendlich lange Impulsantwort. Das macht es schwierig, ein stabiles digitales Filter zu entwerfen.

Im Gegensatz zu den FIR-Filtern gibt es für die IIR keine geschlossenen Entwurfsverfahren.

Deswegen geht man beim Entwurf anders vor:

Man gibt die Filtercharakteristik vor. Als nächstes entwirft man ein passendes analoges Filter. Dieses Filter wird dann transformiert mit der bilinearen Transformation oder einem anderen Verfahren.

Da wir analoge Filter nicht behandelt haben und die bilineare Transformation ein rein mathematisches Verfahren ist, wird hier nicht weiter darauf eingegangen.

Kapitel 5 Nicht-Echtzeitfilter

Wenn das Signal nicht in Echtzeit gefiltert werden muss, sondern als Aufzeichnung vorliegt, können wir die Zeitkoordinate als Ortskoordinate betrachten. Da wir das gesamte Signal zur Verfügung haben, müssen die darauf angewendeten Filter nicht mehr kausal sein. Ausserdem kann man mit der FFT die Frequenztransformierte berechnen und direkt darauf arbeiten. Es ist möglich beliebige Kennlinien zu realisieren, da man einfach einzelne Frequenzanteile vergrössern oder direkt auf 0 setzen kann. Zusätzlich wird es möglich mehrdimensionale Daten wie Bilder oder Bildsequenzen zu filtern.

Beim ersten Anblick ist es ungewohnt, ein Bild als ein Signal zu betrachten. Greift man aber eine Pixelreihe heraus, ist es nichts anderes als ein Vektor von 8-Bit-Zahlen (sw-Bild), also nichts anderes als ein entsprechendes Tonsignal. Die Amplitude entspricht der Helligkeit. Frequenzanteile erkennt man als langsame oder schnelle Änderungen der Helligkeit. Da man die Pixelreihe entlang jeder Geraden herausgreifen kann, ist das Bild ein 2-Dimensionaler Ortsraum und seine Fourier-Transformierte ein 2D-Frequenzraum.

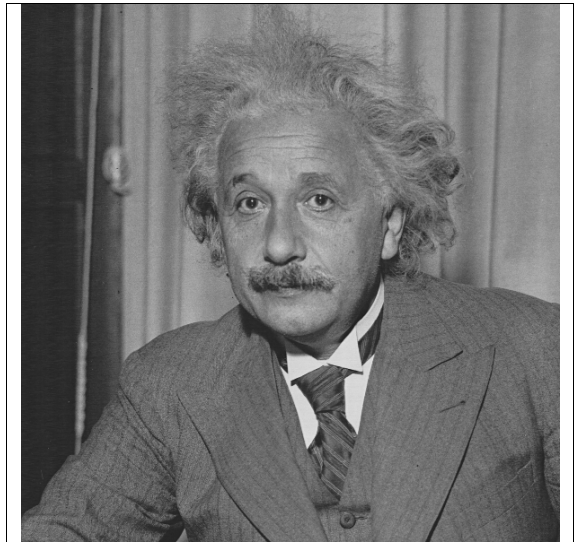
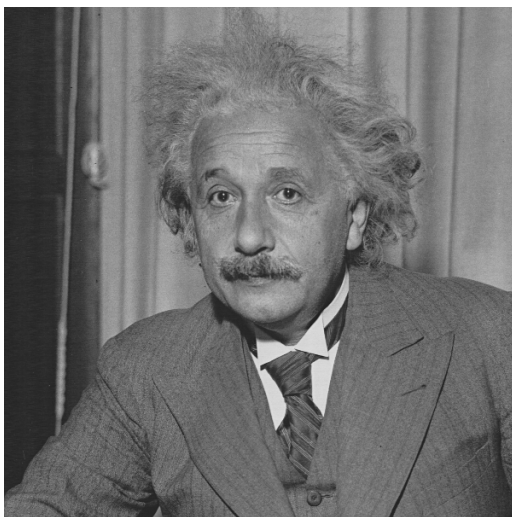


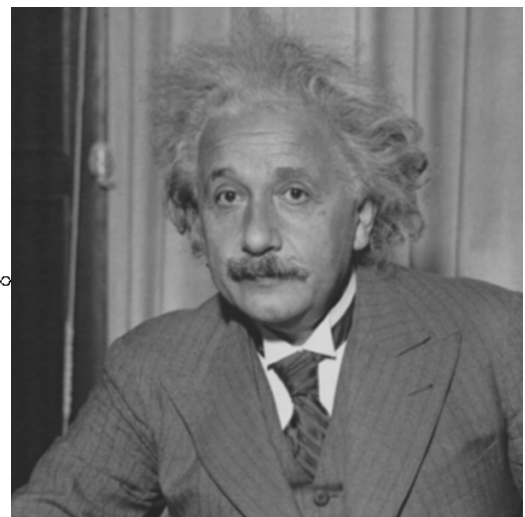
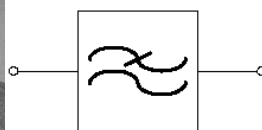
Bild 5.1: Schwarz-Weiss Foto

Betrachten wir mal zwei Beispiele für 2D-Filter:

Beispiel 1: Weichzeichnen



Original Einstein Bild



Weichgezeichnetes Einstein Bild

Bild 5.2: Anwendung eines 2D-Tiefpasses auf ein Bild

Das erste Filter hat folgendes Aussehen:

$\begin{pmatrix} 1/9 & 1/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \end{pmatrix}$ Eine 3x3 Matrix, mit der man das vorgegebene Bild weichzeichnet, also ein Tiefpass. Scharfe Kanten werden glatter oder mit anderen Worten, die Amplitude der hohen Frequenzanteile wird abgesenkt. Wenn man den Punkt (m,n) im gefilterten Bild berechnen möchte, muss man die

Matrix mit der mittleren Zelle an die entsprechende Stelle im Eingangsbild anlegen, dann die Werte der Matrix mit den "darunterliegenden" Pixelwerten multiplizieren und anschliessend aufsummieren. Das macht man nacheinander mit jedem Bildpunkt und erhält so das gefilterte Bild.

Wichtig ist, dass die maximale Summe der Matrixelemente 1 ergibt, da sich sonst im gefilterten Bild Werte über 255 (bei 8 Bit pro Pixel) ergeben könnten. Diese würden dann auf 255 abgesenkt werden, es würde also zu Verzerrungen kommen.

Aus der Sicht der Signalverarbeitung stellt die Matrix nichts anderes als die Frequenztransformierte des Filter dar. Da wir die Kennlinie des Filters mit der Kennlinie der Bilddatei multiplizieren, müssen wir deren Frequenztransformierte im Ortsraum miteinander falten, was wir durch das Anlegen der Matrix auch tun.

Beispiel 2: Kantendetektion

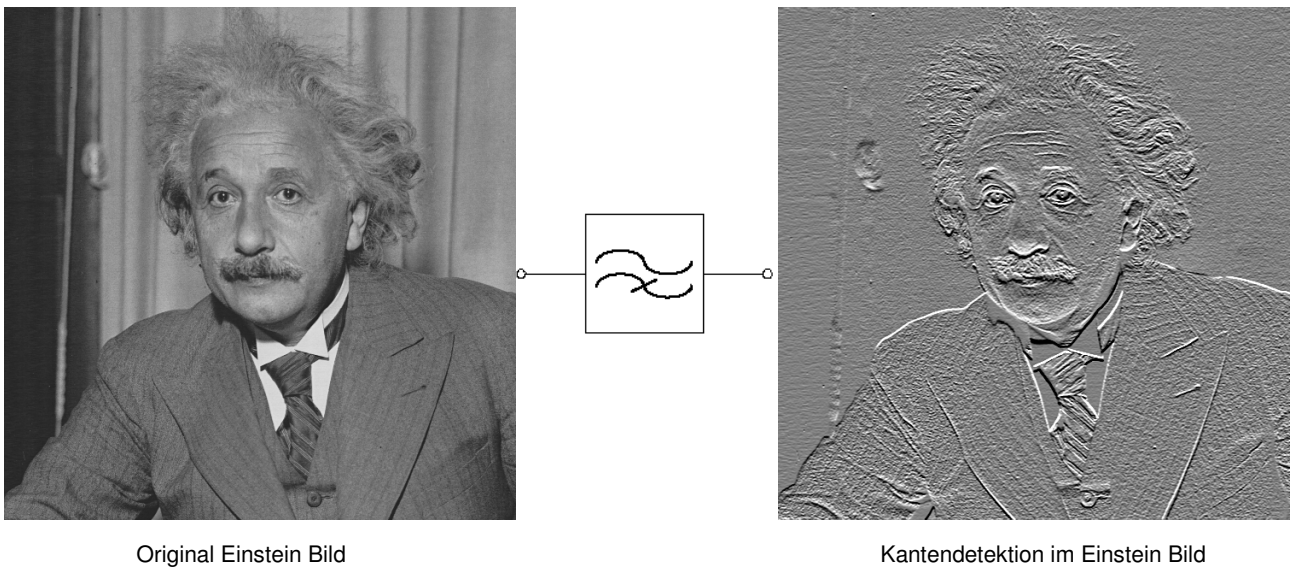


Bild 5.3: Anwendung eines 2D-Hochpasses auf ein Bild

Das zweite Filter ist ein Hochpass und hat folgende Transformierte:

$$\begin{pmatrix} 1/4 & 1/2 & 1/4 \\ 0 & 0 & 0 \\ -1/4 & -1/2 & -1/4 \end{pmatrix}$$

Zu beachten ist der Faktor 4, denn obwohl die Summe der Elemente 0 ergibt, ergibt der Betrag der maximalen Summe 1.

Literaturverzeichnis

[1] Oppenheim/Schafer, Zeitdiskrete Signalverarbeitung, 1995

[2] Texas Instruments, DSP Teaching Kit, 1996

[3] Martin Meyer, Signalverarbeitung, 2000²

[4] Martin Werner, Nachrichtentechnik, 2002³

[5] http://disi.eit.uni-kl.de/difi_n.html (Vorlesung Digitale Filter)

[6] <http://www.ces.ka.bw.schule.de/lehrer/culm/mathematik/filterung/anwendungen.htm>

[7] (Artikelsammlung aus Elektronik)

<http://digital.ni.com/worldwide/germany.nsf/3528e4c55a5f22d4852568e100532f33/5f0bb6c216c0df648025680100348c48?OpenDocument>