5.5 Theorie und Praxis der Signalabtastung

Wie gut ist eigentlich "digital"?
Von der digitalen Speicherung und Verarbeitung eigentlich analoger Signale werden Wunder erwartet - und es werden nicht selten auch Wunder versprochen. Tatsächlich gilt aber: "digital" bedeutet keineswegs "100% fehlerfrei".

Es gibt zwei grundsätzliche Fehlerquellen:

1. die Wandlung des einzelnen analogen Signalwertes in eine digitale Darstellung (Digitalisierung),
2. die Wiedergabe des zeitlichen Signalverlaufs durch einzelne aufeinanderfolgende digitale Signalwerte (Signalabtastung).

Worin liegen die Vorteile der Digitalisierung?
Digitale Signale können, einmal gewandelt, infolge des extremen Störabstandes der digitalen Verarbeitung, praktisch immer wieder 100%-ig identisch (also fehlerfrei) *reproduziert* werden. (Im Innern digitaler Systeme kann man systematische Fehler beim Transportieren und Speichern praktisch ausschließen.)

Hinzu kommen die allgemeinen Vorzüge der digitalen Schaltungstechnik:

- digitale Information kann man in praktisch unbegrenztem Umfang unbegrenzt lange speichern*),
- digitale Information kann man an sich beliebigen Wandlungen unterziehen (z. B. Rechen- und Durchmusterungsalgorithmen),
- die binäre Codierung ist eine universelle und einheitliche Form der Informationsdarstellung (für Zahlenwerte, Zeichenketten, Steuerungsangaben (Programme/Befehle), Sprache, Musik, Fernsehbilder usw.),
- Fehlererkennung und Fehlerkorrektur sind mit akzeptablen Aufwand durchführbar,
- digitale Systeme lassen sich leichter in hoch integrierten Schaltungen verwirklichen und somit kostengünstig fertigen.

*): es bereitet beachtliche Schwierigkeiten, analoge Signale zu speichern (technische Mittel hierfür sind u. a. Magnetbänder und Ultraschall-Verzögerungsleitungen).

Signalabtastung
Das Ziel besteht in einer möglichst getreuen und exakten Wiedergabe des tatsächlichen Signalverlaufs. Dem stehen folgende Tatsachen entgegen:

- die Analog-Digital-Wandlung ist (siehe oben) grundsätzlich mit einem Quantisierungsfehler behaftet,
- abgetastet wird nur zu festen ("diskreten" Zeitpunkten); den Signalverlauf zwischen den Abtastzeitpunkten können wir nicht beobachten.

Abtastfrequenz und Wiedergabetreue

Abbildung 5.5.1 veranschaulicht die einleuchtende Tatsache, daß Einzelheiten des Signalverlaufs verlorengehen, wenn wir ihn nicht oft genug abtasten. Die Dichte der Abtastungen wird durch die Abtastfrequenz bzw. Abtastrate (Sampling Rate) f_s gemessen (Impulsfolgefrequenz der Abtastimpulse).

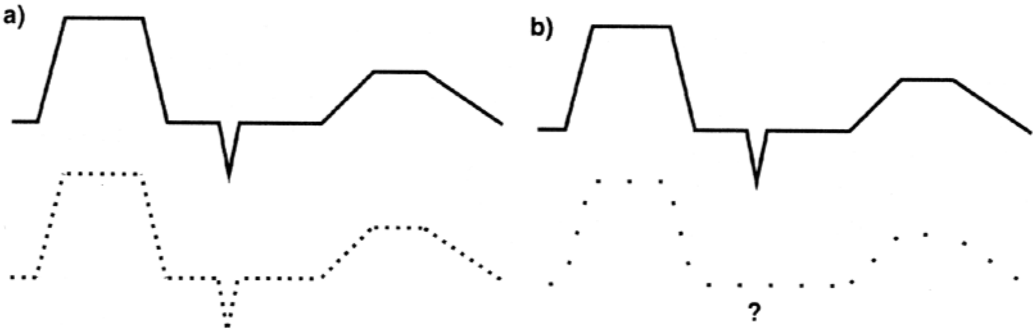

Abbildung 5.5.1 Signalwiedergabe durch Abtastung. a) bei hoher, b) bei niedriger Abtastfrequenz. Oben jeweils der tatsächliche Signalverlauf, darunter die Darstellung aus abgetasteten Punkten

Das Abtasttheorem (Nyquist-Kriterium)

Diese einfache Formel beschreibt den Zusammenhang zwischen der höchsten Frequenz (Grenzfrequenz f_g) eines bandbreitenbeschränkten periodischen Signals und der Abtastfrequenz (Abtastrate) f_s, die mindestens nötig ist, um den Signalverlauf exakt rekonstruieren zu können:

$$f_s > 2\, f_g$$

Das heißt, es ist mit mehr als dem dem Doppelten der oberen Grenzfrequenz abzutasten. Dieser Zusammenhang ist weithin bekannt, wird aber nicht immer zutreffend interpretiert. Die Abbildungen 5.5.2 bis 5.5.12 sollen die Zusammenhänge veranschaulichen.

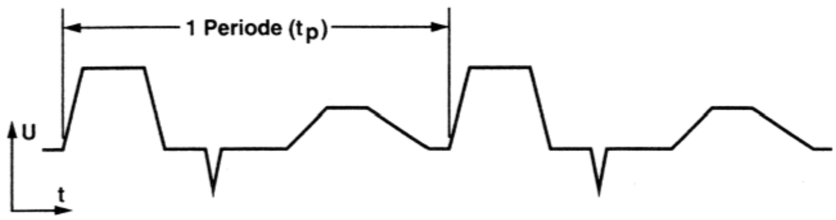

Abbildung 5.5.2 Ein periodischer Signalverlauf

Erklärung:

Der Signalverlauf von Abbildung 5.5.1 ist hier als periodische Folge dargestellt. Aus der Periodendauer t_p ergibt sich eine Grund- bzw. Folgefrequenz $f_0 = 1/t_p$.

Jeder periodische Ablauf kann grundsätzlich durch eine Überlagerung von Sinusschwingungen nachgebildet werden, deren Frequenz ein ganzzahliges Vielfaches der Grundfrequenz f_0 ist (also f_0, 2 f_0, 3 f_0 usw. (Fourier-Zerlegung)). In mathematisch exakter Form hat eine solche Fourier-Reihe unendlich viele Glieder.

Da es keine unendlich hohen Frequenzen gibt, muß man die Reihe nach einer gewissen Anzahl von Gliedern abbrechen. Der Signalverlauf wird also durch eine endliche Anzahl von Sinusschwingungen näherungsweise wiedergegeben (Abbildung 5.5.3):

Folgefrequenz des periodischen Signalverlaufs: $f_0 = \frac{1}{t_p}$ = Grundwelle (1. Harmonische).

Der Gesamt-Signalverlauf ergibt sich als unendliche Summe von Sinusschwingungen (Fourier-Reihe):

$$u(t) = a_1 \cdot \sin \omega t + a_2 \cdot \sin 2\omega t + a_3 \cdot \sin 3\omega t + \ldots = \sum_{i=1}^{\infty} a_i \cdot \sin i\omega t$$

$$\text{mit } \omega = 2\pi f_0 = \frac{2\pi}{t_p}$$

Abbildung 5.5.3 Fourier-Zerlegung am Beispiel einer Rechteckschwingung

Erklärung:
Bei einfachen Signalverläufen können bereits wenige Sinusschwingungen gute Näherungen ergeben. (Hinweis: Die Fourierreihe einer symmetrischen Rechteckschwingung enthält nur die ungeraden Harmonischen.) Allein die Hinzunahme der 3. Harmonischen gibt bereits brauchbare Impulse, die man durchaus einer üblichen Logikschaltung anbieten könnte. Gehen wir bis zur 11. Harmonischen, so entstehen wirklich annehmbare Impulse mit steilen Flanken und nur noch geringfügig welligen Dächern.

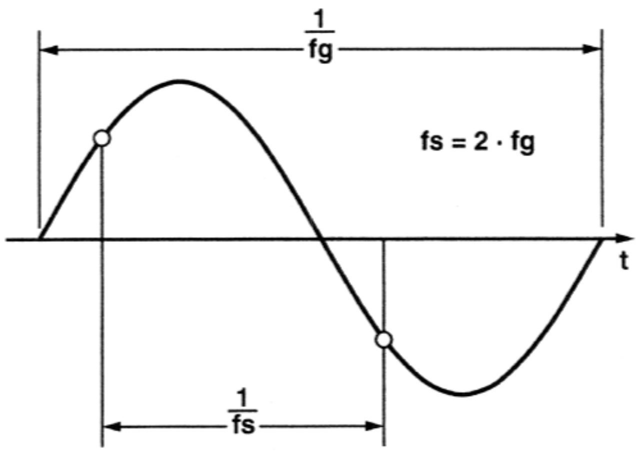

Abbildung 5.5.4 Die volkstümliche Interpretation des Abtasttheorems

Erklärung zu Abbildung 5.5.4:
„Es genügt, mit dem Doppelten der höchsten Signalfrequenz abzutasten. Eine Sinusschwingung kann man aus 2 Abtastwerten, die im Abstand einer halben Periodendauer vorliegen, genau rekonstruieren.“ Diese Vorstellung ist weit verbreitet. Sie trifft aber leider nur sehr näherungsweise zu.

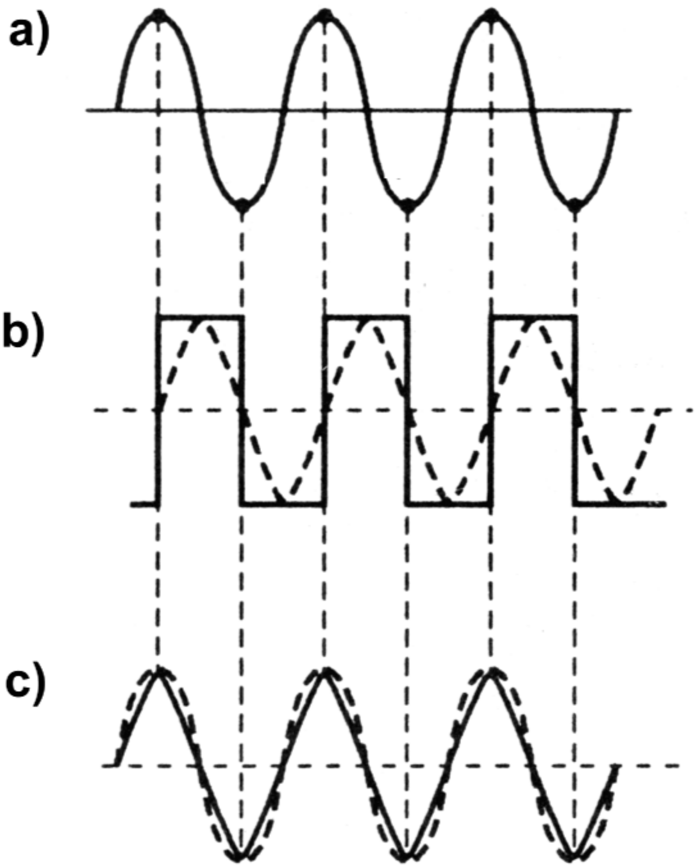

Abbildung 5.5.5 Auch im günstigsten Fall gibt es schon Fehler (nach: Burr-Brown)

Erklärung:

a) Abtastung einer Sinusschwingung mit $f_s = 2\ f_g$. Die Abtastzeitpunkte liegen bei 90 und 270°, so daß in jeder Periode die beiden Spitzenwerte erfaßt werden (besser geht‘s nicht).

b) Rekonstruktion 0. Ordnung. Auf der digitalen Seite sehen wir je Signalperiode nicht mehr als zwei Abtastwerte. Wenn wir diese Werte auf einen Digital-Analog-Wandler geben, entsteht eine Rechteckschwingung mit der Frequenz f_g.

c) Rekonstruktion 1. Ordnung. Die einzelnen Werte werden durch Geraden untereinander verbunden. Es ergibt siche eine Dreieckschwingung. Ein ähnlicher Signalverlauf entsteht, wenn wir die Rechteckschwingung auf einen einfachen Tiefpaß geben (Verschleifen der Flanken).

Fehlerbetrachtung:
Wie groß ist die Abweichung des näherungsweise rekonstruierten Signals vom Original? - Vergleichen wir hierzu die vom Kurvenverlauf eingeschlossenen Flächen in einer Halbperiode, bezogen auf die Dauer der Halbperiode (mit anderen Worten: die Gleichrichtwerte):

Sinus: Fläche = 2 U_S; $|U| = \frac{2}{\pi} U_S$ => 0,64

Rechteck: Fläche = π U_S; $|U| = U_S$ => 1

Dreieck: Fläche π /2 U_S; $|U| = \frac{U_S}{2}$ = > 0,5

- Abweichung bei Rekonstruktion 0. Ordnung = Abweichung zwischen Rechteck und Sinus: 1 - 0,64 = 0,36 = 36%.
- Abweichung bei Rekonstruktion 1. Ordnung = Abweichung zwischen Dreieck und Sinus: 0,64 - 0,5 = 0,14 = 14%.

Die Genauigkeit der Signalverlaufswiedergabe steigt an, wenn wir die Abtastfrequenz erhöhen (Abbildung 5.5.6).

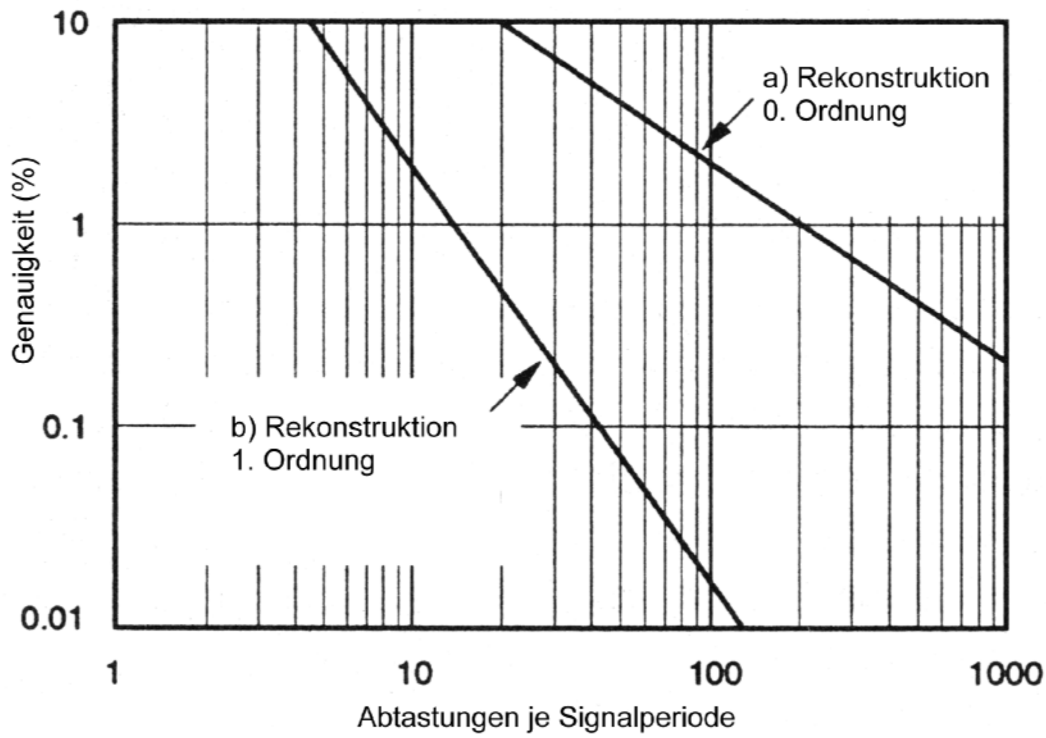

Abbildung 5.5.6 Die Wiedergabegenauigkeit in Abhängigkeit von den Abtastungen je Signalperiode (nach: Burr-Brown)

Wir haben bisher idealisierte Verhältnisse angenommen. In der Praxis treten die aber nur selten auf (Abbildung 5.5.7).

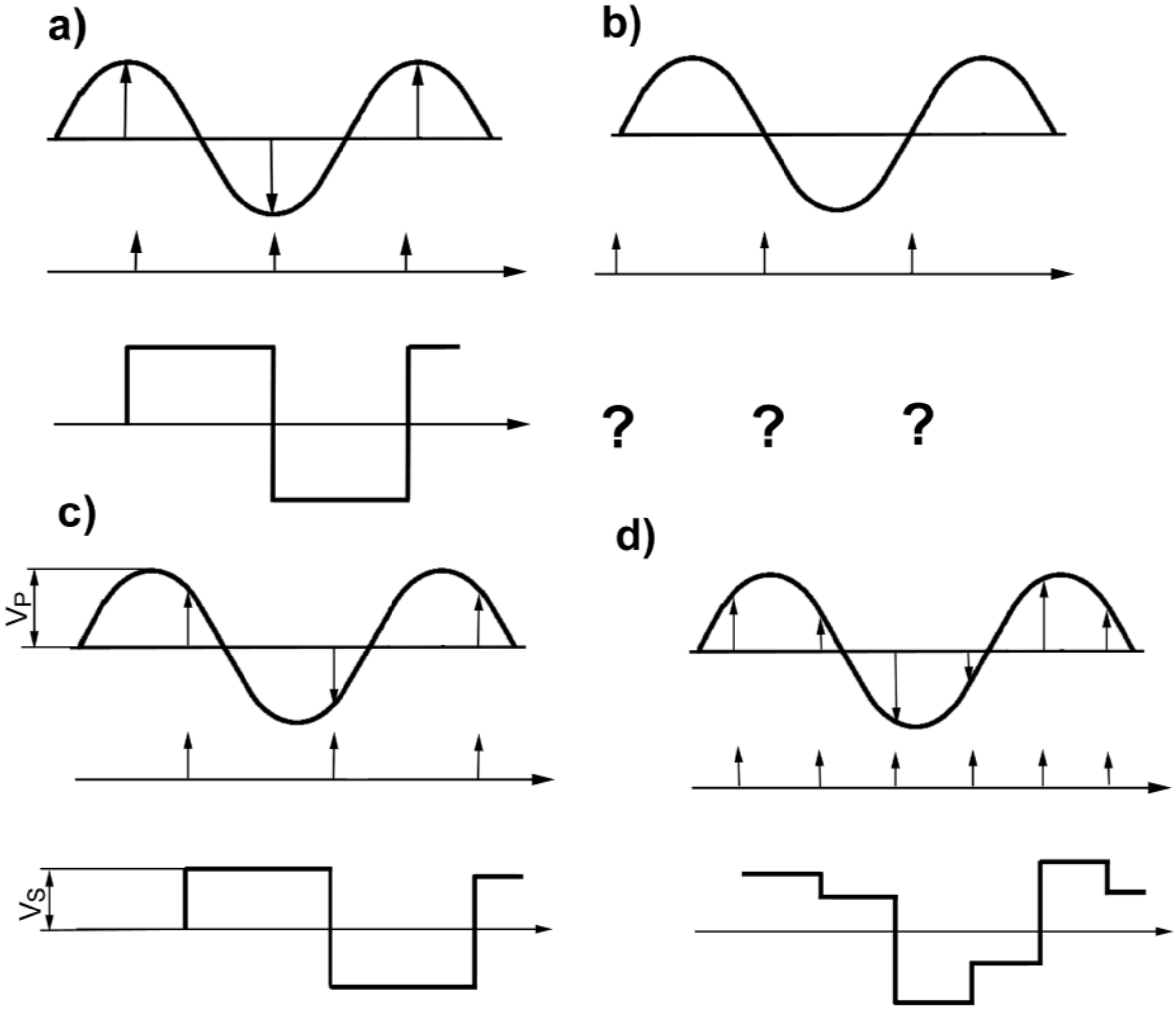

Abbildung 5.5.7 Genügt es, mit $f_s = 2\ f_g$ abzutasten?

Erklärung:

a) bilderbuchmäßige Abtastung mit $f_s = 2\ f_g$. Die Abtastzeitpunkte treffen mit den Spitzenwerten des Sinussignals zusammen. Ein so abgetastetes Signal läßt sich ohne weiteres vollständig rekonstruieren ((Frequenz und Amplitude).

b) Abtastung mit $f_s = 2\ f_g$. Die Abtastzeitpunkte treffen mit den Nulldurchgängen des Sinussignals zusammen. Wir tasten also gar nichts ab.

c) Abtastung mit $f_s = 2\ f_g$. Die Abtastzeitpunkte treffen irgendwo auf den Signalverlauf. Ein Signal mit gleicher Frequenz läßt sich rekonstruieren, aber woher sollen wir - auf der digitalen Seite - wissen, daß der Spitzenwert des Sinus eigentlich V_P beträgt und nicht etwa nur V_S?

d) Abtastung mit $f_s > 2\ f_g$. Das Sinussignal läuft gleichsam unter den Abtastzeitpunkten durch, so daß - wenn wir die gesamte Erfassungszeit gegen Unendlich gehen lassen - jeder Punkt der Signalperiode von der Abtastung erfaßt werden wird.

Wie oft abtasten?
Auf jeden Fall öfter als zweimal in der Signalperiode. Richtwert: wenigstens 7...10 Abtastungen.

Theorie und Praxis
Das Abtasttheorem wird unter Nutzung der Fourier-Transformation hergeleitet, wobei vorzusgweise im Frequenzbereich gerechnet wird. Die verwendete Mathematik kennt nur Frequenzen, weiß aber nichts davon, daß in der Praxis die Zeit eine wichtige Rolle spielt: das Abtasten muß schnell gehen (Erfassungszeit), die Signale werden beim Durchlaufen durch Schaltmitttel verzögert (Phasenverschiebung) usw.

Was das Abtasttheorem wirklich besagt:

- wenn wir eine Sinusschwingung (also keinen x-beliebigen Signalverlauf!) der Frequenz f_g mehr als zweimal je Periodendauer abtasten (Abtastfrequenz $f_s > 2\ f_g$), können wir den ursprünglichen Signalverlauf auch dann wieder rekonstruieren, wenn f_s nur geringfügig größer als 2 f_g ist - vorausgesetzt, daß wir uns genügend Zeit lassen, die Abtastwerte heranzuschaffen (Erfassungszeit gegen Unendlich).
- wenn wir mit einer Abtastfrequenz f_s arbeiten, dürfen im abzutastenden Signal nur Sinusschwingungen enthalten sein, deren Frequenz < $f_s/2$ ist (Bandbreitenbeschränkung). Abbildung 5.5.8 veranschaulicht die technischen Voraussetzungen.

Hinweis:
Die volkstümliche Auffassung, es genüge, mit 2 f_g abzutasten, ist genaugenommen unzutreffend. Je mehr sich eine Frequenzkomponente der Grenzfrequenz f_g nähert, desto mehr Fehler handelt man sich ein. (Bleibt die Frage, ob man sich diese Fehler leisten kann. Beispiele: Digitalisierung im Telefonnetz, Audio-CD.)

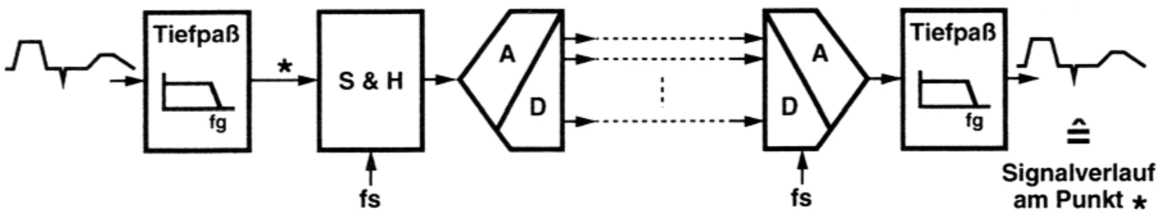

Abbildung 5.5.8 System mit Signalabtastung und -wiederherstellung

Erklärung:
Die Abbildung zeigt das technische Umfeld, in dem das Abtasttheorem gilt: wir begnügen uns damit, den Signalverlauf durch eine endliche Zahl von Sinusschwingungen (Harmonische) näherungsweise nachzubilden. Die Frequenz der letzten Harmonischen, die wir mitnehmen wollen, wird zur Grenzfrequenz f_g. Demzufolge müssen wir mit einer Abtastfrequenz $f_s > 2\ f_g$ abtasten. Höhere Signalfrequenzen als f_g dürfen wir nicht verarbeiten. Deshalb muß das Signal bandbreitenbeschränkt werden. Dazu schicken wir es durch einen Tiefpaß mit der Grenzfrequenz f_g, der dem Abtast- und Halteglied (S & H) vorgeschaltet ist. Die Abtastwerte können wir dann auf beliebige Weise speichern, verarbeiten usw. Um das Signal wieder zu rekonstruieren, führen wir die Abtastwerte im Rhythmus der Abtastfrequenz f_s einem

Digital-Analog-Wandler zu. Dessen Ausgang ist wiederum - im Sinne der Bandbreitenbeschränkung - ein Tiefpaß mit der Grenzfrequenz f_g nachgeschaltet. An dessen Ausgang wird dann der ursprüngliche Signalverlauf wieder erscheinen, allerdings in der Form, die sich am Ausgang des eingangsseitigen Tiefpasses ergeben hatte. Nehmen wir an, wir wollten einen Signalverlauf mit f_0 = 1 MHz abtasten und dabei noch die 11. Harmonische berücksichtigen. Das heißt, f_g = 11 MHz. Unser Tiefpaß müßte also nach 11 MHz dichtmachen, und wir müßten mit f_s = 2 f_g, also 22 MHz*), abtasten. Am Ausgang würde dann der rekonstruierte Signalverlauf erscheinen.

*): genügt das? - Es kommt darauf an. Kann sein, daß Fehler in der Erfassung der 11. Harmonischen sich praktisch kaum auswirken.

Bandbreitenbeschränkung - die Theorie

Das Abtasttheorem gilt an sich nur dann, wenn das abzutastende Signal einer exakten Bandbreitenbeschränkung unterworfen wird, wenn es also vor der Abtastung über einen idealen Tiefpaß der Grenzfrequenz $f_g = f_s/2$ geführt wurde. Ein solcher Tiefpaß läßt alle Sinusschwingungen mit Frequenzen < f_g unbeeinträchigt durch und sperrt alle Sinusschwingungen mit Frequenzen > f_g.

Bandbreitenbeschränkung in der Praxis:

- einen idealen Tiefpaß kann man nicht bauen,
- ein näherungsweise idealer Tiefpaß verfälscht aber alle Signale, die nicht sinusförmig sind (Abbildung 5.5.9),
- ein nicht idealer Tiefpaß läßt auch Sinusschwingungen oberhalb f_g durch - und das hat Folgen (Alias-Frequenzen).

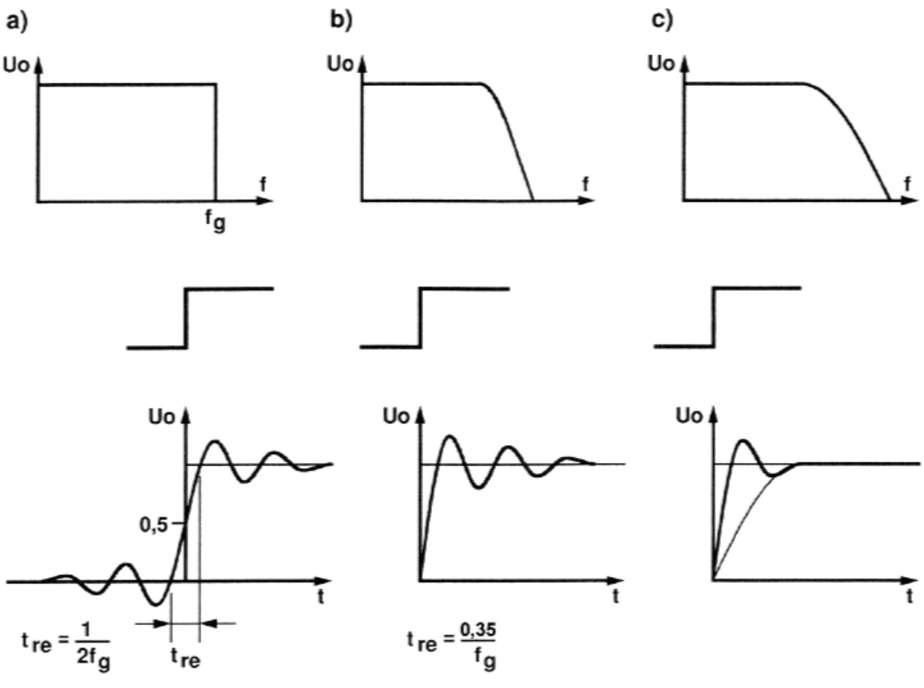

Abbildung 5.5.9 Zur Auslegung von Tiefpässen. Oben: Frequenzgänge, Mitte: die eingangsseitige Signalflanke, unten: das Ausgangssignal im zeitbereich (Sprungantwort)

Erklärung zu Abbildung 5.5.9:

a) ein idealer Impuls am Eingang eines idealen Tiefpasses führt zu Überschwingern am Ausgang. Die mathematische Behandlung ergibt sogar ausgangsseitige Schwingungen *vor* der Impulsflanke - eine physikalische Unmöglichkeit, die sich daraus erklärt, daß hier zwei Idealisierungen (Grenzübergänge gegen ∞) zusammenfallen. Es ist ersichtlich, daß auch der ideale Tiefpaß auf eine ideale Flanke (Anstiegszeit 0) mit einer Flanke antwortet, die eine endliche Anstiegszeit (Eigenanstiegszeit t_{re}) hat.

b) ein realer Frequenzgang mit weitgehender Annäherung an das ideale Tiefpaßverhalten. Auch diese Auslegung führt zu Überschwingern.

c) Frequenzgang mit flacherem Abfall. Je nachdem, wie die Kurve im einzelnen aussieht, gibt es entweder nur ein geringes Überschwingen oder gar keines. Bei zu flachem Abfall wird aber die Eigenanstiegszeit zu groß.

Ein beliebiger periodischer Signalverlauf kann exakt aus einer unendlichen, näherungsweise aus einer endlichen Summe von Sinusschwingungen rekonstruiert werden (Fourier-Zerlegung). Das ist aber eine mathematische Idealisierung, die nur unter der Bedingung gilt, daß die einzelnen Sinusschwingungen nicht zueinander phasenverschoben sind (vgl. Abbildung 5.5.3: wenn die Grundschwingung durch Null geht, haben auch alle Harmonischen einen Nulldurchgang). Die Sinusschwingungen in Abbildung 5.5.3 stehen auf dem Papier und lassen sich leicht zur resultierenden näherungsweisen Rechteckwelle addieren. Jedes wirkliche Signal braucht aber *Zeit*, um wirkliche Signalwege (Leitungen und Schaltungen) zu durchlaufen. Eine Zeitverzögerung ist aber nichts anderes als eine Phasenverschiebung zwischen ein- und ausgangsseitigem Signalverlauf (Abbildung 5.5.10).

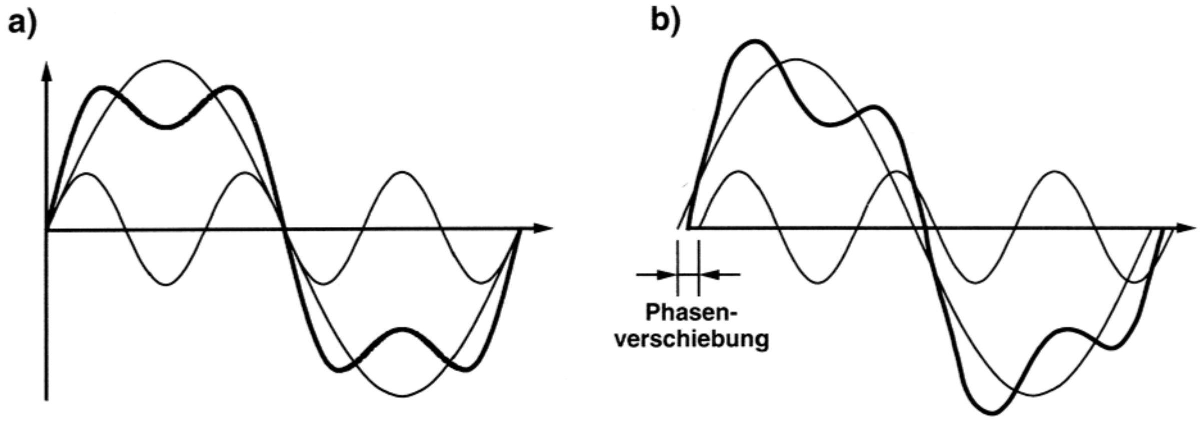

Abbildung 5.5.10 Signalverzerrung durch Phasenverschiebung (Beispiel)

Erklärung:

a) näherungsweise Darstellung einer Rechteckschwingung durch Überlagerung zweier Sinusschwingungen. Keine Phasenverschiebung = korrekte Darstellung (vgl. auch Abbildung 5.5.3).

b) die Sinusschwingungen sind gegeneinander phasenverschoben. Das führt zu einer stark verzerrten Darstellung.

Damit solche Verzerrungen nicht eintreten, ist zu gewährleisten, daß Sinusschwingungen aller Frequenzen beim Passieren des Tiefpasses die gleiche zeitliche Verzögerung Δt erfahren. Phasenwinkel φ, Frequenz f und zeitliche Verzögerung Δt hängen folgendermaßen zusammen: $\varphi = 2\pi\ f\ \Delta t$. Mit $\Delta t = const$ ergibt sich somit die Abhängigkeit $\varphi = const \cdot f$. Das bedeutet, die Phasenverschiebung muß linear mit der Frequenz zunehmen. Ein solcher Tiefpaß hat aber einen vergleichsweise flach abfallenden (sog. Gaußschen) Frequenzgang.

Oversampling
Dies ist der gängige Begriff, um auszudrücken, daß ein bandbreitenbeschränktes Signal (obere Grenzfrequenz f_g) mit einer Abtastrate abgetastet wird, die über der unbedingt notwendigen liegt: f_s deutlich über $2\ f_g$. Die Oversampling-Rate ist dann das Verhältnis von tatsächlicher zu notwendiger Abtastrate, also $f_s : 2\ f_g$. Beispiel: Ein Signal mit einer Grenzfrequenz $f_g = 100$ MHz wird mit einer Abtastrate $f_s = 500$ MSamples/s abgetastet. Dann spricht man von einem 2,5-fachen Oversampling ($500 : 2 \cdot 100$).

Undersampling
Damit bezeichnet man ein Abtasten mit einer an sich zu geringen Abtastrate ($f_s < 2\ f_g$). Es gibt Anwendungen, in denen diese Betriebsweise zulässig ist.

Was geschieht, wenn der Tiefpaß Signale oberhalb f_g durchläßt?
Das gleiche wie beim absichtlichen Undersampling - der ursprüngliche Signalverlauf von Sinussignalen mit Frequenzen > 0,5 f_s läßt sich nicht rekonstruieren (Abbildungen 5.5.11, 5.5.12).

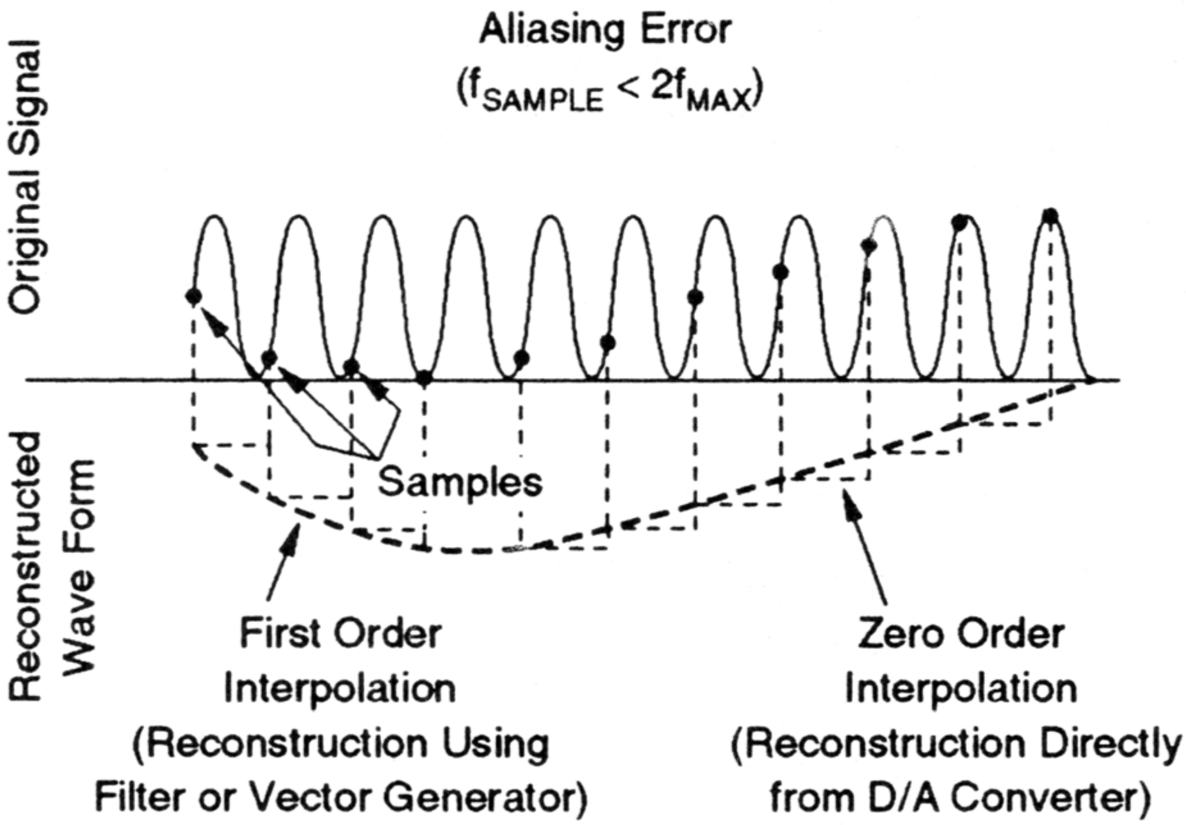

Abbildung 5.5.11 Undersampling (1). Nach: Burr-Brown

Erklärung zu Abbildung 5.5.11:
Es sind offensichtlich zuwenig Abtastzeitpunkte. Aus den einzelnen Abtastwerten kann man beim besten Willen kein Sinussignal mit der ursprünglichen Frequenz rekonstruieren; es ergibt sich vielmehr nur eine Art Hüllkurve.

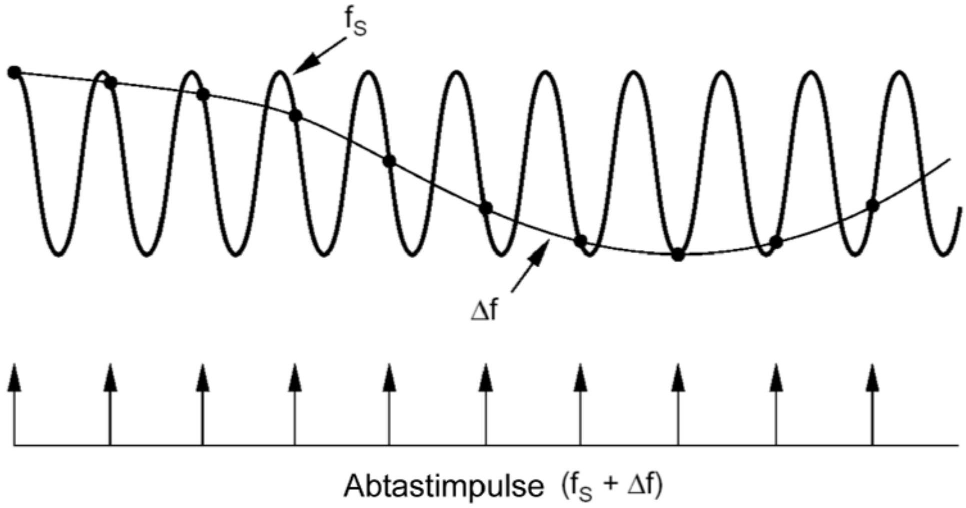

Abbildung 5.5.12 Undersampling (2). Die Alias-Frequenz (nach: Texas Instruments)

Wenn wir eine Sinusschwingung der Frequenz f_g mit einer Abtastfrequenz $f_s < 2\ f_g$ abtasten (oder - in anderer Sichtweise - die Bandbreite des abzutastenden Signals nicht auf $f_g \leq 0{,}5\ f_s$ beschränken), so erscheint eine Sinusschwingung mit der sogenannten Alias-Frequenz $f_s - f_g$. (Und die schafft auch kein ausgangsseitiger Tiefpaß aus der Welt...)

Bei der Auslegung von Abtastsystemen gilt es, einen Kompromiß zu finden:

- damit das Abtasttheorem gilt, ist ein idealer Tiefpaß erforderlich,
- verwirklicht man aber diese Forderung näherungsweise (um Alias-Schwingungen auszuschließen), so verfälscht der Tiefpaß den Signalverlauf (Überschwinger, Phasenverschiebung).
- ein Tiefpaß mit Gaußschem Frequenzverhalten hingegen, der das Signal nahezu unverfälscht weitergibt, läßt auch Frequenzen oberhalb $f_g = f_s/2$ durch, so daß sich ausgangsseitig Alias-Schwingungen ergeben.