Kommunikationstechnik II – Wintersemester 08/09

Prof. Dr. Stefan Weinzierl

Musterlösung: 6. Aufgabenblatt

Lösung in der Rechenübung am 15.1.2009

1. Aufgabe: Dynamik und SNR der Zahlenformate

Generieren Sie eine Periode eines Sinussignals mit 1024 samples/Periode und quantisieren Sie es auf eine Wortbreite von 16 bit im Festkomma-Format sowie im Fließkomma-Format. Bei letzterem betrage nach IEEE 754r die Wortbreite der Mantisse 11 bit (davon 1 bit für das Vorzeichen), und die Wortbreite des Exponenten 5 bit.

Gleitkomma-Quantisierung[1]

Eine Gleitkomma-Zahl wird dargestellt in der Form:

$x_Q = M \cdot 2^E$

Die Mantisse *M* wird dabei zur eindeutigen Darstellung normalisiert und nimmt nur Werte zwischen 0,5 und 1 (ohne 1!) an. Sie wird im Festkomma-Format dargestellt mit einer Wortbreite von w_m bit, wobei davon 1 bit für das Vorzeichen benutzt wird.

Für den Exponenten gilt $E = e - bias$. e ist eine Zahl zwischen 1 und $2^{w_e} - 2$ (hier werden die Sonderfälle 0 und Infinity/NaN berücksichtigt). Um auch negative Exponenten zu ermöglichen wird der so genannte *bias* abgezogen, der sich wie folgt berechnen lässt (und auch in der IEEE-Norm definiert ist): $bias = e_{max}/2 = (2^{w_e} - 2)/2 = 2^{w_e-1} - 1$.

Die Werte für 0 und unendlich sind als Sonderfälle definiert (siehe KT 2-Skript), hierbei ist zu beachten, dass sich ein positiver und ein negativer Wert für 0 ergeben.

```
close all; clear all; clc

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Musterloesung zum 6. Aufgabenblatt                                          %
%                                                                             %
%-----------------------------------------------------------------------------%

% Aufgabe 1: SNR und Dynamik der Zahlenformate

% Eingangssignal mit N samples
N = 1024;
n = 0:1/N:1-1/N;
y = sin(2*pi*n);


%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% fixed point %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
w = 16;
q = 2/2^w;
y_fixed = quant(y,q);
```

[1] siehe auch: Zölzer (2005), Abschnitt „Zahlendarstellung“

```
% die hoechste Quantisierungsstufe (1) wird weggelassen, um genau 2^w
% Stufen zu erreichen
maximum = max(y);
for i = 1:length(y_fixed);
    if y_fixed(i) > maximum - q;
    y_fixed(i) = maximum - q;
    end
end


%%%%%%%%%%%%%%%%%%%%%%%% floating point %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
w_e = 5;                % Wortbreite des Exponenten
w_m = 11;               % Wortbreite der Mantisse
e = 1 : 2^w_e - 2;      % Vektor mit allen Werten von e
bias = max(e)/2;        % bias, wird spaeter von e abgezogen
q_m = 2/2^w_m;          % Quantisierungsstufe der Mantisse (diese wird im fixed
point Format dargestellt)
m = 0.5:q_m:1-q_m;      % Vektor mit allen Werten der Mantisse
s = [1 0];              % Vorzeichen

% Vorbereitung fuer die for-Schleife
l_s = length(s);
l_e = length(e);
l_m = length(m);

% Es wird eine Vektor mit allen Werten des floating point Formats erzeugt
for i = 1:l_s;
    for j = 1:l_e;
        floatingpoint_(1+(j-1)*l_m:l_m+(j-1)*l_m) = (-1)^s(i) * m .* 2^(e(j)-
bias);
    end
    floatingpoint(1+(i-1)*(l_e*l_m) : l_e*l_m+(i-1)*(l_e*l_m)) = floatingpoint_;
end

floatingpoint = [floatingpoint 0 -0];    % es werden 0 und -0 eingefuehrt
floatingpoint = sort(floatingpoint);     % die Werte werden nach ihrer Groesse
sortiert

% Vor der Quantisierung wird das Eingangssignal wird auf das Maximum des
% floating point Vektors skaliert
y_floating = y*max(floatingpoint);

% Um zu ermitteln, welcher Wert des floating point Vektors den
% Eingangssignalwerten am naechsten ist (also auf welche floating point Zahl
% die Eingangswerte gerundet werden), wird die Differenz zwischen
% floating point Zahlen und Eingangssignalwerten gebildet (differenz). Die
% Position der kleinsten Differenz wird ermittelt (index_min), um den
% Eingangssignalwert auf eben diese Stufe im floating point Vektor zu runden.
for i = 1:length(y);
    differenz = floatingpoint - y_floating(i);
    [min_differenz index_min] = min(abs(differenz));
    y_floating(i) = floatingpoint(index_min);
end

% Das quantisierte Signal wird wieder auf die urspruengliche Skalierung
% gebracht.
y_floating = y_floating./max(floatingpoint);
```

a) Wie groß ist die Dynamik für beide Signale in dB?

Die Dynamik lässt sich berechnen, in dem man das Verhältnis zwischen maximal möglicher Amplitude und minimal möglicher Amplitude des Signals bildet:

$$\mathrm{DR} = 20\log\left(\frac{x_{\max}}{x_{\min}}\right)$$

Im Falle der Festkomma-Quantisierung (mit midtread-Kennlinie) ist ein Zahlenbereich von -2^{15} bis $2^{15} - 1$, bzw. [-32.768...32.767], darstellbar. Die minimale Auslenkung liegt zwischen ±1.
Es ist oberste Quantisierungsstufe
$x_{\max} = 1 - q = 1 - 2^{1-w}$
und die kleinste Quantisierungsstufe
$x_{\min} = q = 2^{1-w}$
Dies resultiert in einer Dynamik von

$$\mathrm{DR} = 20\log\left(\frac{1-q}{q}\right) = 20\log\left(\frac{1-2^{-15}}{2^{-15}}\right) = 20\log\left(\frac{2^{15}-1}{1}\right) = 20\log\left(\frac{32.767}{1}\right) = 90{,}31\ \mathrm{dB}$$

Im Falle der Fließkomma-Quantisierung gilt:
$x_{\max} = M_{\max}\cdot 2^{E\_\max}$ und $x_{\min} = M_{\min}\cdot 2^{E\_\min}$

Dabei ist die maximale Mantisse: $M_{\max} = 1 - q_m = 1 - 2^{1-w\_m}$
Die minimale Mantisse $M_{\min}$ ist wegen der Normalisierung gleich 0,5.

Weiterhin ist
$E_{\max} = (2^{w\_e} - 2) - bias = (2^{w\_e} - 2) - (2^{w\_e-1} - 1) = 2^{w\_e-1}\cdot(2 -1) -1 = 2^{w\_e-1} - 1.$
$E_{\min} = 1 - bias = 1 - (2^{w\_e-1} - 1) = -2^{w\_e-1} + 2$
Der Zahlenbereich des Exponenten liegt also zwischen [-14...15].

Dies führt auf eine maximale Auslenkung zwischen $x_{\max} = M_{\max}\cdot 2^{E\_\max} = \pm$ 32.736 Die minimale Auslenkung liegt zwischen $x_{\min} = M_{\min}\cdot 2^{E\_\min} = \pm 3{,}052\cdot 10^{-5}$.

$$\mathrm{DR} = 20\log\left(\frac{M_{\max}\cdot 2^{E_{\max}}}{M_{\min}\cdot 2^{E_{\min}}}\right) = 20\log\left(\frac{(1-q_m)\cdot 2^{2^{w_e}-bias}}{0{,}5\cdot 2^{1-bias}}\right) = 20\log\left(\frac{(1-2^{-10})\cdot 2^{15}}{0{,}5\cdot 2^{-14}}\right) = 180{,}61\ \mathrm{dB}$$

```
%%%%%%%%  a) %%%%%%%%

DR_fixed = 20*log10((1-q)/q)
DR_floating = 20*log10((max(m)*2^(max(e)-bias))/(min(m)*2^(min(e)-bias)))
```

b) Bestimmen Sie den theoretischen SNR des Festkomma-kodierten Signals. Warum ist diese Berechnung bei der Fließkomma-Kodierung nicht möglich?

Der theoretische SNR lässt sich mit Hilfe der folgenden Formel berechnen:
SNR = 6.02·w + 1.76 [dB]
Im vorliegenden Fall also: SNR = 6.02·16 + 1.76 [dB] = 98,08 dB

Bei der Fließkomma-Quantisierung wird der Zahlenbereich durch den Exponenten in Abschnitte von Zweierpotenzen geteilt, die mit steigendem Exponenten einen immer größeren Bereich abdecken. Diese Bereiche werden jeweils mit der linearen Auflösung der Mantisse quantisiert, die Quantisierungsstufen werden also bei steigender Amplitude größer. Eine theoretische SNR-Berechnung müsste diese Unregelmäßigkeit berücksichtigen. Man bevorzugt daher die numerische Berechnung des SNR im Frequenzbereich (s. Aufgabe c)).

c) Transformieren Sie beide Signale mittels FFT in den Frequenzbereich und berechnen Sie jeweils den SNR, indem Sie die Signalleistung ins Verhältnis zur Fehlerleistung setzen. Wie interpretieren Sie die Ergebnisse im Hinblick auf die in a) berechnete Dynamik?

Der SNR ist das Verhältnis zwischen Leistung des Nutzsignals und Leistung des Fehlers:

$$\mathrm{SNR} = 10\log\left(\frac{P_\mathrm{S}}{P_\mathrm{N}}\right)$$

Eine numerische Berechnung des SNR basierend auf der FFT-Technik ist dann auf einfache Weise möglich, wenn Signallängen verwendet werden, die mit einer ganzzahligen Anzahl von Perioden in das FFT-Analyse-Intervall passen. Genau diese Voraussetzung wurde mit den Rahmenbedingungen in dieser Aufgabe geschaffen (das Signal passt mit genau einer Periode in einen Block der FFT). In diesem Fall sind alle Rauschanteile in den diskreten Harmonischen des FFT-Spektrums enthalten, sodass zur Berechnung des SNR folgende Formel verwendet werden kann (s. Handbuch der Audiotechnik, Kap. 17):

$$\mathrm{SNR} = 10\log\left(\frac{Y_1^2}{Y_2^2 + Y_3^2 + Y_4^2 + \ldots + Y^2_{N/2-1}}\right)$$, mit Y_i = Spektralkomponenten

Hierbei wird die parseval'sche Beziehung ausgenutzt.

```
%%%%%%%%  b) %%%%%%%%

% FFT der quantisierten Signale
Y_fixed = fft(y_fixed);
Y_floating = fft(y_floating);

% Betragsbildung
Y_fixed = abs(Y_fixed);
Y_floating = abs(Y_floating);


%-------- SNR-Berechnung ---------%

% fixed point
P_grund_fixed = Y_fixed(2)^2;              % Leistung der Grundschwingung
P_teil_fixed = sum(Y_fixed(3:(N/2-1)).^2); % Gesamtleistung der Rauschanteile
                                           % und Oberschwinungen ohne DC
                                           % und Grundschwingung


% Der SNR ist das Verhaeltnis der Leistung der Grundschwingung zur Leistung
```

```
% des Fehlers
snr_fixed = 10*log10(P_grund_fixed/P_teil_fixed)


% floating point
P_grund_floating = Y_floating(2)^2;                % Grundschwingung
P_teil_floating = sum(Y_floating(3:(N/2-1)).^2); % Gesamtleistung der
Oberschwingungen
                                                   % und Oberschwinungen ohne DC
                                                   % und Grundschwingung

% Der SNR ist das Verhaeltnis der Leistung der Grundschwingung zur Leistung
% des Fehlers
snr_floating = 10*log10(P_grund_floating/P_teil_floating)
```

Die Berechnung liefert einen SNR von 98.67 dB für die Festkommazahl und 69,72 dB für die Fließkommazahl, obwohl beide mit derselben Wortbreite quantisiert sind. Dies zeigt, dass obwohl im Fließkommaformat eine deutlich größere Dynamik erreicht werden kann, der SNR sogar kleiner ist als bei der Festkomma-Kodierung. Dies liegt daran, dass die Quantisierungsintervalle bei großen Amplituden ebenfalls immer größer werden, sodass natürlich auch der relative Fehler größer wird. Allerdings werden bei Audiosignalen große Amplituden viel seltener erreicht als niedrige. Letztere werden bei der Fließkomma-Kodierung deutlich feiner quantisiert, während bei Festkomma-Kodierung der SNR mit abnehmender Aussteuerung sinkt.

c) Welche Frequenz müsste das Signal bei einer Abtastfrequenz von 48 kHz haben?

$$\text{Signalfrequenz} = \frac{\text{Abtastfrequenz}}{\text{samples in einer Periode}} = \frac{48000\ \text{Hz}}{1024} = 46{,}875\ \text{Hz}$$

Diese Frequenz müsste für das Testsignal gewählt werden, um bei einer FFT-basierten Messung des SNR eines A/D-Wandlers ein exaktes Ergebnis zu erhalten.

2. Aufgabe: A/D- und D/A-Wandler-Daten

a) Erläutern Sie das Messverfahren für die Größen THD+N und Dynamic Range.

THD+N :
THD+N (total harmonic distortion plus noise) ist das Verhältnis aller Oberschwingungen *inkl.* Grundrauschen zu den Oberschwingungen, Grundrauschen und dem voll ausgesteuerten Messsignal (Effektivwerte).[2] Der THD+N ist also einfacher zu messen als der THD allein, da bei letzterem auch Aliasinganteile der Obertöne an völlig anderen Positionen als der Obertonreihe auftreten können und berücksichtigt werden müssten. Diese aufwändige Spektralanalyse kann entfallen, da beim THD+N einfach „alles“ im

[2] Die Festlegung der Fullscale-Amplitude erfolgt dabei während der Messung selbst, da diese nach der AES-17 Norm wie folgt definiert ist: „… wenn das digitale Signal nicht zugänglich ist, wird Input-Fullscale 0,5 dB unterhalb des Amplitudenwertes eines 997-Hz-Sinustons gesetzt, bei dem 1% THD+N oder 0,3dB Kompression erreicht werden…“.

hörbaren Audiobereich in den Messwert einfließt, was sich vom Messton unterscheidet. Die Messung berücksichtigt somit nicht nur harmonische Oberwellen, sondern das gesamte Störspektrum einschließlich unharmonischer Anteile, Einstreuungen, Brummen, Rauschanteile u.ä.

Ablauf der Messung:
- rege mit f_0 = 997Hz Ton bei -0,5 dB FS / -1 dB FS an
- messe Effektivwert von Rauschen und Verzerrungsprodukten im gesamten Audioband bei Unterdrückung von f_0 durch Notch-Filter
- bilde Verhältnis zu ungefiltertem Signal und berechne Pegel

Notation z.B.: THD+N (997 Hz, -1 dB FS) = -85 dB FS

Für das Grundrauschen wird üblicherweise kein Gewichtungsfilter verwendet, da von voll ausgesteuerten also „lauten“ Nutzsignalen ausgegangen wird (ungefähr linearer Bereich der Hörkurve).

Dynamic Range:
Dynamic Range bezeichnet das Verhältnis der Vollaussteuerung des Wandlers zum (frequenzbewerteten) Grundrauschen in Anwesenheit eines (leisen) Signals (in [dB FS]). Es erscheint in der AES-17-Norm als „*Signal-to-Noise Ratio (SNR) or noise in the presence of a signal*“.

Ablauf der Messung:
- zur Angabe des DR-Wertes muss vorher eine FS-Messung stattgefunden haben
- dann wird ein Messton f_0 = 997 Hz bei -60 dB FS benutzt, damit möglichst keine Nichtlinearitäten angeregt werden, der Wandler aber sicher aktiv ist (keine Stummschaltung)
- messe Effektivwert von Rauschen und Verzerrungsprodukten im gesamten Audioband bei Unterdrückung von f_0 durch Notch-Filter
- bilde Verhältnis zur FS-Spannung und berechne Pegel

Notation z.B.: DR = 85 dB FS A

- Gewichtungsfilter z.B.: A, CCIR

b) Wieso unterscheiden sich bei Messungen oftmals die Werte für THD+N und DR?

- bei Systemen mit wenig Verzerrungsprodukten sind THD+N und DR annähernd gleich, es messen dann beide hauptsächlich den Rauschpegel
- THD+N ist aber meist schlechter, da sich durch den hohen Messsignalpegel (z.B. -0.5 dB FS) deutliche Verzerrungsprodukte ergeben
- Frequenzbewertung bei der DR-Messung