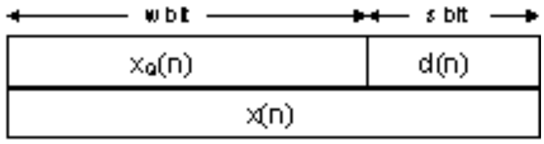
Prof. Dr. Stefan Weinzierl 21.11.2013

Musterlösung: 18. November 2013, 17:04

1 Erwartungswert nach Requantisierung

Ein 12-bit-Signal x(n) wird mit w = 6 bit requantisiert. Vor der Requantisierung wird ihm ein gleichverteiltes Dithersignal mit der Amplitude -0,5 LSB bis 0,5 LSB beigemischt. Das Dithersignal wird mit s = 6 bit quantisiert.



a) Stellen Sie mit Hilfe von Matlab den Erwartungswert des requantisierten Ausgangssignals über der Eingangsamplitude innerhalb eines Quantisierungsintervalls dar. Die Berechnungsvorschrift des Erwartungswerts findet sich im Abschnitt „Dither-Techniken" bei Zölzer (2005)[1].

Lösung: Es wird nur ein Ausschnitt aus der Quantisierungskennlinie betrachtet (Abb. 1) und untersucht, wie sich die Addition des Dither zum Eingangssignal auf das Quantisierungsergebnis, bzw. Ausgangssignal auswirkt. Dabei lässt sich der Erwartungswert eines requantisierten Signals wie folgt

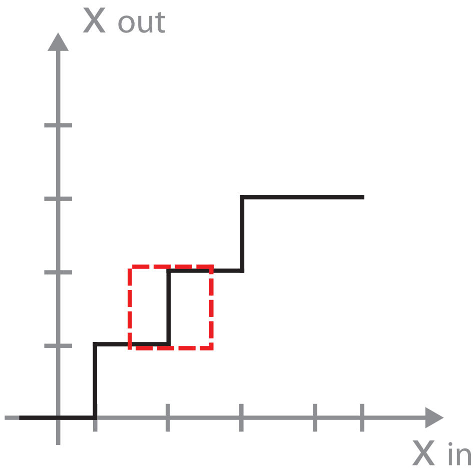


Abbildung 1: Für die Berechnung des Erwartungswertes berücksichtiger Ausschnitt aus der Quantisierungskennlinie

berechnen:

$$\bar{g}(V) = \sum_k g(V + d_k)P(d_k) \quad , mit \; -2^{s-1} \leq k \leq 2^{s-1} - 1$$

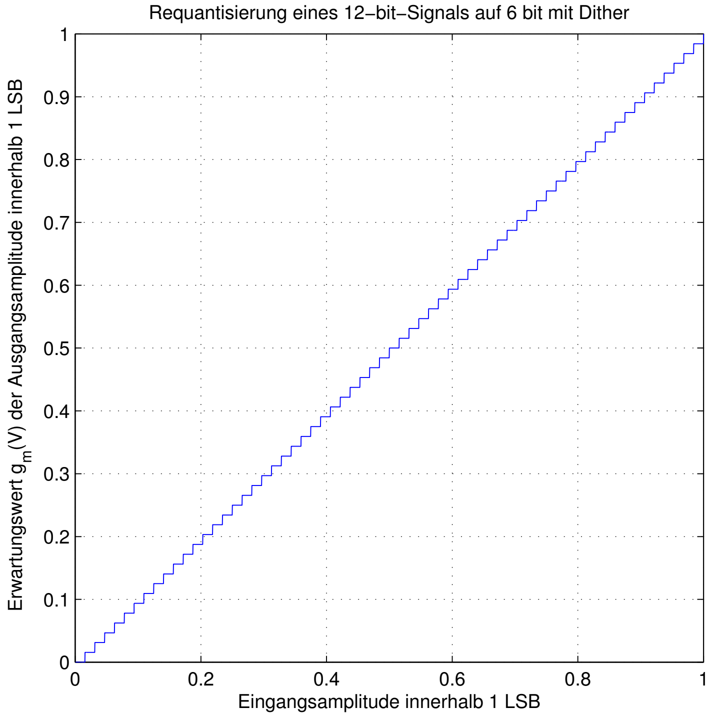
Dabei stellt $g(V + d_k) = Q\lfloor \frac{V+d_k}{Q} + 0.5 \rfloor$ die Quantisierung mit mid-tread Kennlinie nach Dither-Addition dar (Es könnte auch round$\left(\frac{V+d_k}{Q}\right)$ geschrieben werden). Es wird also zu der Eingangsamplitude V eine der 2^s Ditheramplituden d_k addiert und das Ergebnis quantisiert. Der Erwartungswert

[1] Zölzer U (2005) Digitale Audiosignalverarbeitung. 3. Auflage, Teubner, Stuttgart. Dieses Buch finden Sie im Downloadbereich auf der Webseite zu dieser Lehrveranstaltung.

einer bestimmten Eingangsamplitude V ist die Summe dieser k Quantisierungen gewichtet mit der Wahrscheinlichkeit der Ditheramplituden. Wird dies für jede Eingangsamplitude V berechnet, ergibt sich die gesuchte Requantisierungskennlinie.

```
%% a)
clear all
clc

n = 12;                      % Wortbreite der ersten Quantisierung
s = 6;                       % Wortbreite der Ditherquantisierung

V = 0:2^-n:1;                % Eingangsvektor
amp_dither = .5;            % Ditheramplitude

q_s = 2*amp_dither/2^s;      % Quantisierungsintervall des Dithers

d_k = -amp_dither:q_s:amp_dither-q_s;    % Dithersignal
P_d = 1/2^s;% Wahrscheinlichkeit der Ditheramplituden

% Zu jedem Amplitudenwert von V innerhalb eines LSB werden alle moeglichen
% Amplitudenwerte des Dithers addiert, einzeln quantisiert und der
% Erwartungswert berechnet. Es resultiert ein Vektor mit den
% Erwartungswerten fuer alle Eingangsamplituden von V.

for i = 1:length(V);
     g_m(i) = sum(round((V(i)+d_k))*P_d);
end

figure
plot(V,g_m)
axis ([0 1 0 1])
title('Requantisierung eines 12-bit-Signals auf 6 bit mit Dither')
xlabel('Eingangsamplitude innerhalb 1 LSB')
ylabel('Erwartungswert g_m(V) der Ausgangsamplitude innerhalb 1 LSB')
grid on
axis square
```

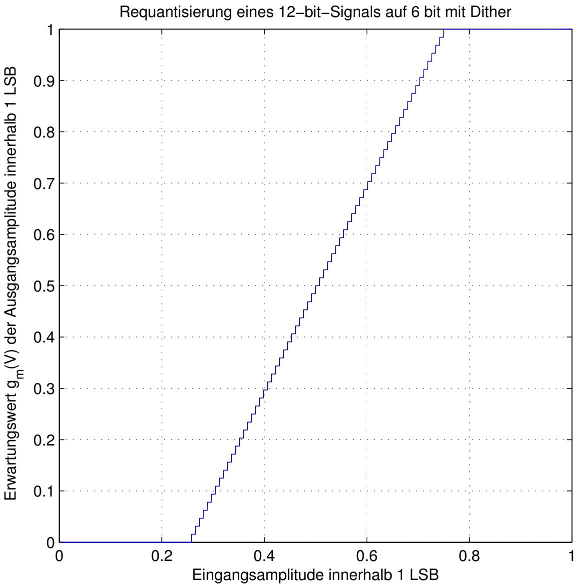
Durch die Addition des Dithers ergibt sich eine linear ansteigende Kennlinie. Die mittleren Ausgangswerte einer mit Dither durchgeführten Requantisierung entsprechen den Eingangswerten, obwohl diese tatsächlich ja nicht dargestellt werden können. Durch das Dithering kann also eine höhere Auflösung erzielt werden, als von der Wortbreite bei der Requantisierung vorgegeben. Sie entspricht w bit + s bit.



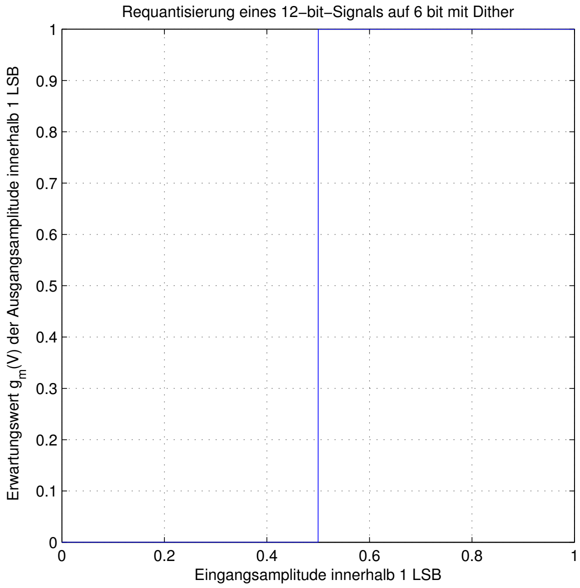
b) Wiederholen Sie dies für die Dither-Aussteuerungen $[-0.25;\ 0.25]$ LSB und $[-0.6;\ 0.6]$ LSB sowie für eine Requantisierung ohne Dither und interpretieren die Ergebnisse

Lösung: m-Code siehe Aufgabenteil a). Bei geringer ausgesteuertem Dithersignal nähert sich die Kennlinie immer mehr einer einzigen Treppenstufe, wie sie sich bei einer Requantisierung ohne Dither ergibt (Abb. oben). Da die Ditheramplitude zu gering ist, „landen" weniger Werte über einem halben LSB, sodass der mittlere Ausganswert länger auf der Stufe Null bleibt. Bei einer Quantisierung ohne Dither wird die Hälfte der Eingangswerte auf Null LSB, die andere Hälfte auf 1 LSB quantisiert.

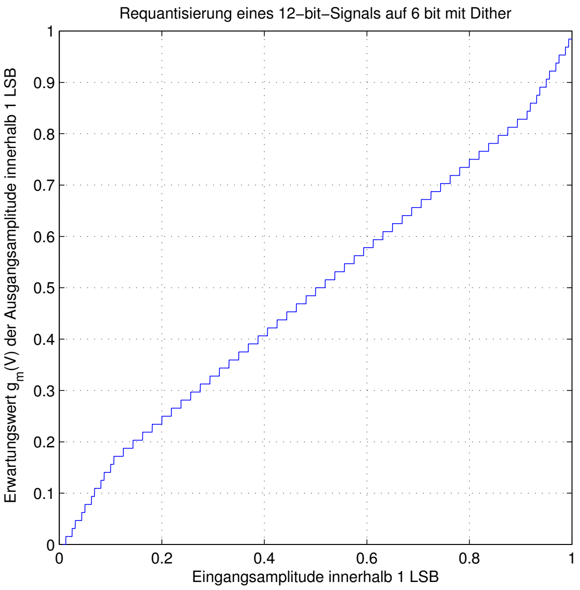
Wird das Dithersignal höher ausgesteuert als $\pm 0,5$ LSB, wird die Kennlinie bis zu einem bestimmten Punkt steiler ansteigen, da im Mittel zu viele Werte auf 1 LSB quantisiert werden.



$[-0.25;\ 0.25]$ LSB

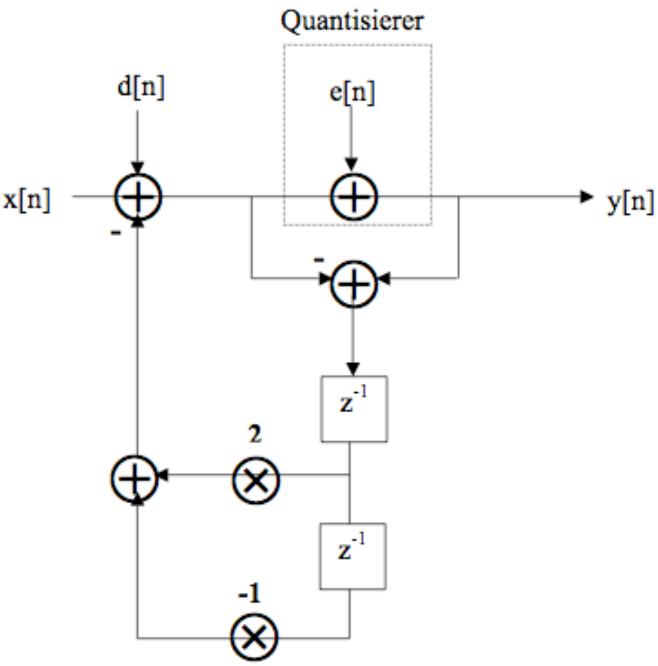


$[-0.001;\ 0.001]$ LSB

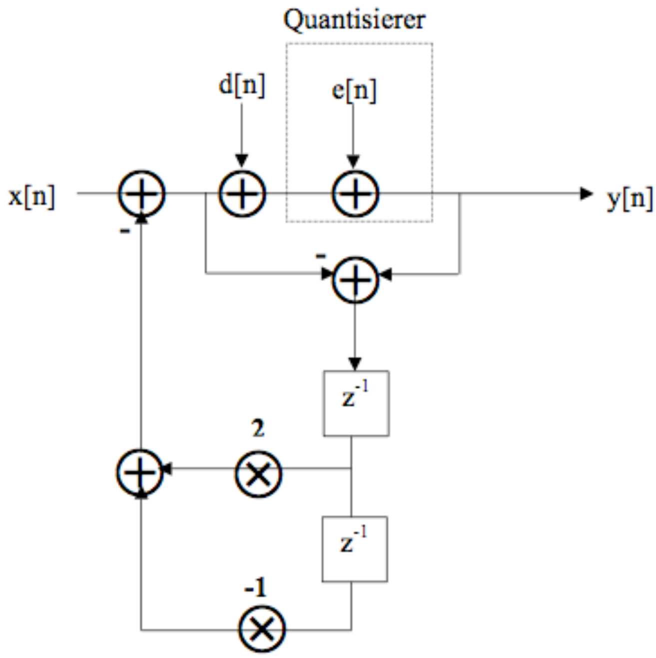


$[-0.6;\ 0.6]$ LSB

2 Noiseshaping



(A)



(B)

a) Stellen Sie die Differenzengleichungen für die beiden dargestellten Noiseshaping-Systeme auf und berechnen Sie jeweils die Signal- und die Rauschübertragungsfunktion im Z-Bereich.
Lösung: System A:

$$\begin{aligned}
y[n] &= x[n] + d[n] + e[n] - 2e_0[n-1] + e_0[n-2] \\
e_0 &= x[n] + d[n] + e[n] - (x[n] + d[n]) \\
&= e[n] \\
\Rightarrow y[n] &= x[n] + d[n] + e[n] - 2e[n-1] + e[n-2]
\end{aligned}$$

Transformiert in den z-Bereich

$$\begin{aligned}
Y(z) &= X(z) + D(z) + E(z) - 2E(z)z^{-1} + E(z)z^{-2} \\
&= X(z) + \underbrace{D(z) + E(z)}_{Rauschen} \underbrace{\left(1 - 2z^{-1} + z^{-2}\right)}_{spektraleFormung}
\end{aligned}$$

ergeben sich Signal und Rauschübertragungsfunktion.

$$\begin{aligned}
H_X(z) &= \frac{Y(z)}{X(z)} = H_D(z) = \frac{Y(z)}{D(z)} = 1 \\
H_E(z) &= \frac{Y(z)}{E(z)} = 1 - 2z^{-1} + z^{-2} \quad \Rightarrow \text{Noiseshaping 2. Ordnung}
\end{aligned}$$

System B:

$$\begin{aligned}
y[n] &= x[n] + d[n] + e[n] - 2e_0[n-1] + e_0[n-2] \\
e_0 &= x[n] + d[n] + e[n] - x[n] \\
&= e[n] + d[n] \\
\Rightarrow y[n] &= x[n] + d[n] - 2d[n-1] + d[n-2] + e[n] - 2e[n-1] + e[n-2]
\end{aligned}$$

Transformiert in den z-Bereich

$$\begin{aligned} Y(z) &= X(z) + D(z) - 2D(z)z^{-1} + D(z)z^{-2} + E(z) - 2E(z)z^{-1} + E(z)z^{-2} \\ &= X(z) + \underbrace{[D(z) + E(z)]}_{Rauschen} \underbrace{\left(1 - 2z^{-1} + z^{-2}\right)}_{spektraleFormung} \end{aligned}$$

ergeben sich Signal und Rauschübertragungsfunktion.

$$\begin{aligned} H_X(z) &= \frac{Y(z)}{X(z)} = 1 \\ H_E(z) &= \frac{Y(z)}{E(z)} = H_D(z) = \frac{Y(z)}{D(z)} = 1 - 2z^{-1} + z^{-2} \qquad \Rightarrow \text{Noiseshaping 2. Ordnung} \end{aligned}$$

b) Berechnen Sie Null- und Polstellen von der in a) berechneten Rauschübertragungsfunktion und stellen Sie diese im Pol-Nullstellen-Diagramm dar. Was für eine Art Übertragungsfunktion vermuten Sie?

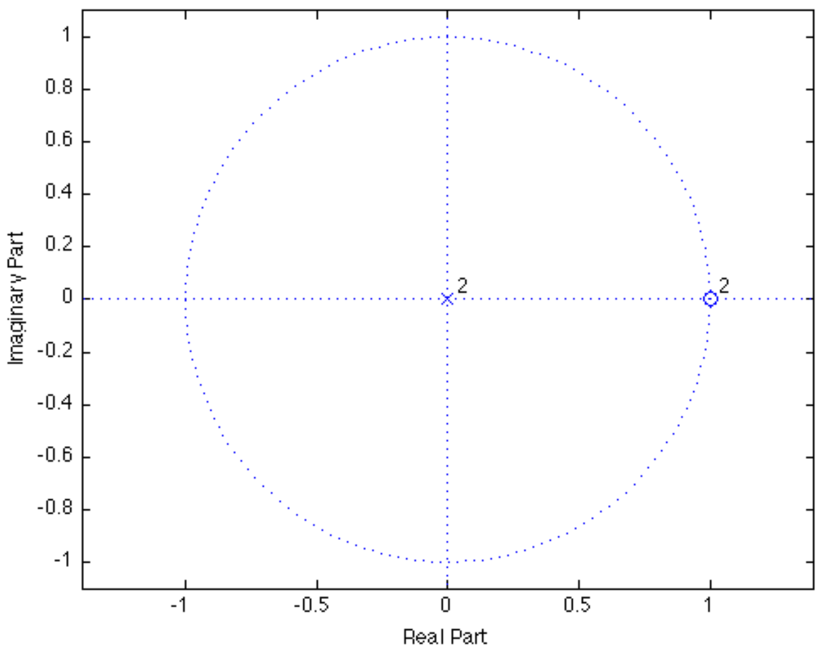
Lösung:

$$1 - 2z^{-1} + z^{-2} = \frac{z^2 - 2z + 1}{z^2}$$

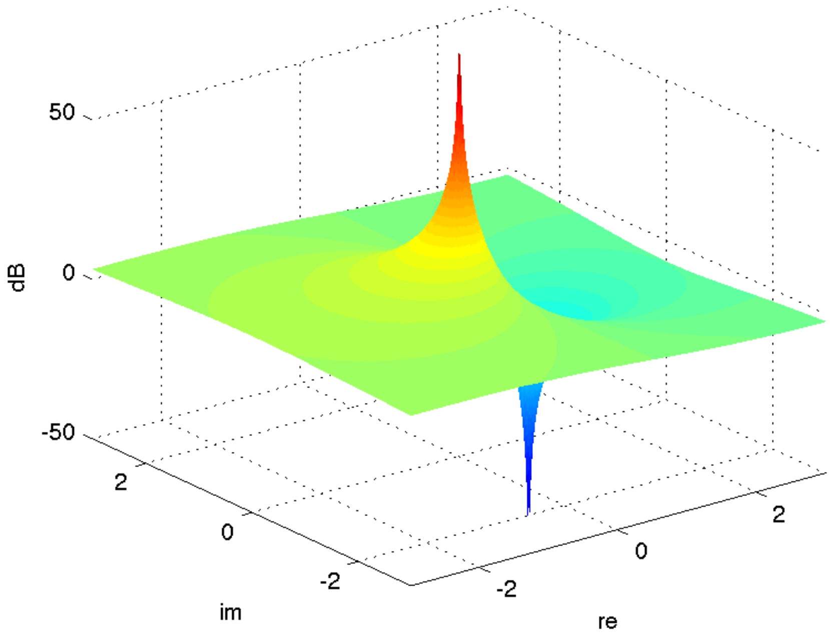
Null- und Polstellen ergeben sich über die Berechnung der Nullstellen von Zähler- bzw. Nennerpolynom:

$$\begin{aligned} n_{1,2} &= -\frac{p}{2} \pm \sqrt{\left(\frac{p}{2}\right)^2 - q} = 1 \pm \sqrt{1-1} = 1 \\ p_{1,2} &= 0 \end{aligned}$$

Bei $z = 1$ befindet sich also eine doppelte Nullstelle und bei $z = 0$ eine doppelte Polstelle. Aufgrund der Nullstellen bei 1, was der Frequenz f=0 Hz entspricht, wird eine Hochpass-übertragungsfunktion erwartet.



c) Stellen Sie das Betragsspektrum der Rauschübertragungsfunktionen in Matlab dar.
Lösung: Zunächst wird die Z-Transformierte zur Anschauung über der Gesamten z-Ebene ausgewertet, so dass die Auswirkungen von Pol- und Nullstellen sichtbar werden:



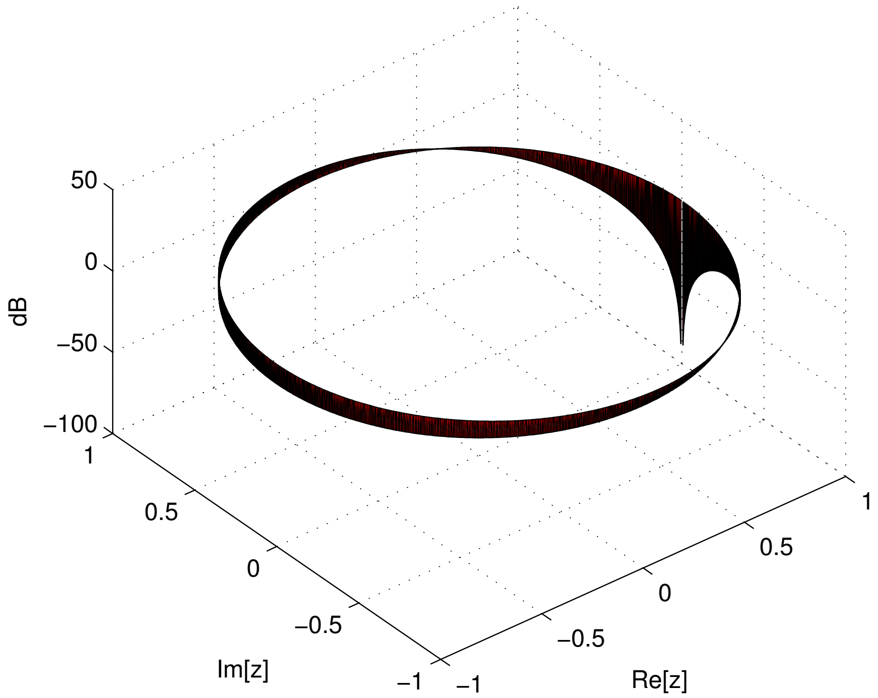
Die Fouriertransformierte ergibt sich durch Auswertung der Z-Transformierten über dem Einheitskreis, da dann die Z-Tranformation in die Fouriertransformation übergeht:

$$\underbrace{X(z) = \sum_{n=-\infty}^{\infty} x[n] z^{-n}}_{Z-Transformation} \quad , mit\ z = r \cdot e^{j\omega}$$

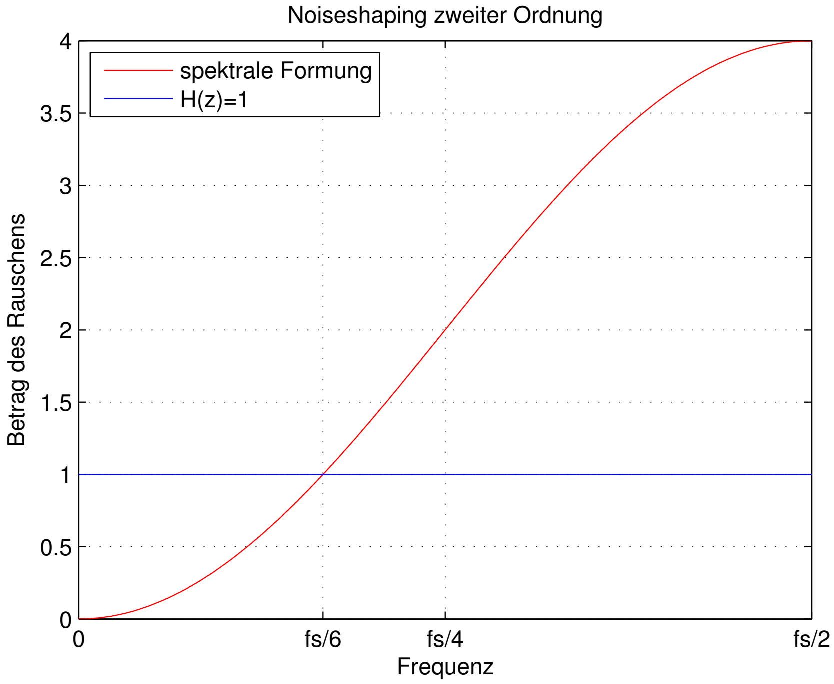
$$X(z) = \sum_{n=-\infty}^{\infty} x[n] \left(r \cdot e^{j\omega}\right)^{-n} = \sum_{n=-\infty}^{\infty} x[n] \cdot r^{-n} e^{-j\omega n} \quad , für\ r = 1\ folgt$$

$$X(z = e^{j\omega}) = \underbrace{X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x[n] \cdot e^{-j\omega n}}_{Fourier-Transformation}$$

Die Übertragungsfunktion über dem Einheitskreis kann auch mit Hilfe von Matlab betrachtet werden:



Die Vermutung, dass es sich bei dem System um einen Hochpass handelt wird also bestätigt. Genaue Werte lassen sich aber aus den dreidimensionalen Darstellungen schlecht bestimmen. Dafür muss der Frequenzgang zweidimensional dargestellt werden:



Das Noise-shaping verändert also die spektrale Verteilung der Rauschleistung. Ein Teil der Leistung wird zu hohen Frequenzen verschoben, bei denen das Gehör weniger empfindlich ist. Der Effekt kann durch Noise-Shaping höherer Ordnung in Kombination mit Überabtastung noch verstärkt werden.

```
%% c)

% Zusatz: plot in der Z-Ebene

% grid erzeugen
[re, im] = meshgrid(-3:.0125:3);
% z-Ebene erzeugen
z = (re+1j*im);
% Zaehlerpolynom berrechnen
num = polyval([1 -2 1], z);
% Nennerpolynom brechnen
denum = polyval([1 0 0], z);
% H(Z) berechnen
Z = num./denum;

% plot
figure
surf(re, im, 20*log10(abs(Z)), 'EdgeAlpha', 0)
xlabel('re')
ylabel('im')
zlabel('dB')
axis([-3 3 -3 3 -50 50])


%% plot ueber Einheitskreis in der Z-Ebene
N = 2^10;
[re, im] = cylinder(1, N-1);                 % Kreiskoordinaten generieren
H = freqz([1 -2 1], [1 0 0], N/2);           % Freqenzgang bis pi
H(length(H)+1:2*length(H)) = flipud(H);      % Frequenzgang bis 2*pi durch Spiegelung
H(:, 2) = H;                                 % Fuer surf/mesh modifizieren
```

```
H(:, 1) = 0;
H(:, 2) = 20*log10(abs(H(:, 2)));           % Betrag
figure
hold on
surf(re, im, H', 'EdgeAlpha', 1)
xlabel Re[z]; ylabel Im[z]; zlabel dB
grid on

%% Betragsspektrum plotten
% frequenzvektor
w = linspace(0, pi, 512);
% Da Fouriertransformation = H(z=e^jw), und H(z) = 1-2z^-1+z^-2
H = 1-2*exp(-1j*w)+exp(-2*1j*w);

figure
plot(w,abs(H),'r')
title('Noiseshaping zweiter Ordnung')
xlabel('Frequenz'), ylabel('Betrag des Rauschens')
xlim([0 pi])
set(gca,'XTick',[0 pi/3 pi/2 pi])
set(gca,'XTickLabel',{'0';'fs/6';'fs/4';'fs/2'})
grid on
hold on
plot(w(1:512),ones(1,512))
legend('spektrale Formung','H(z)=1','Location','NW')
```

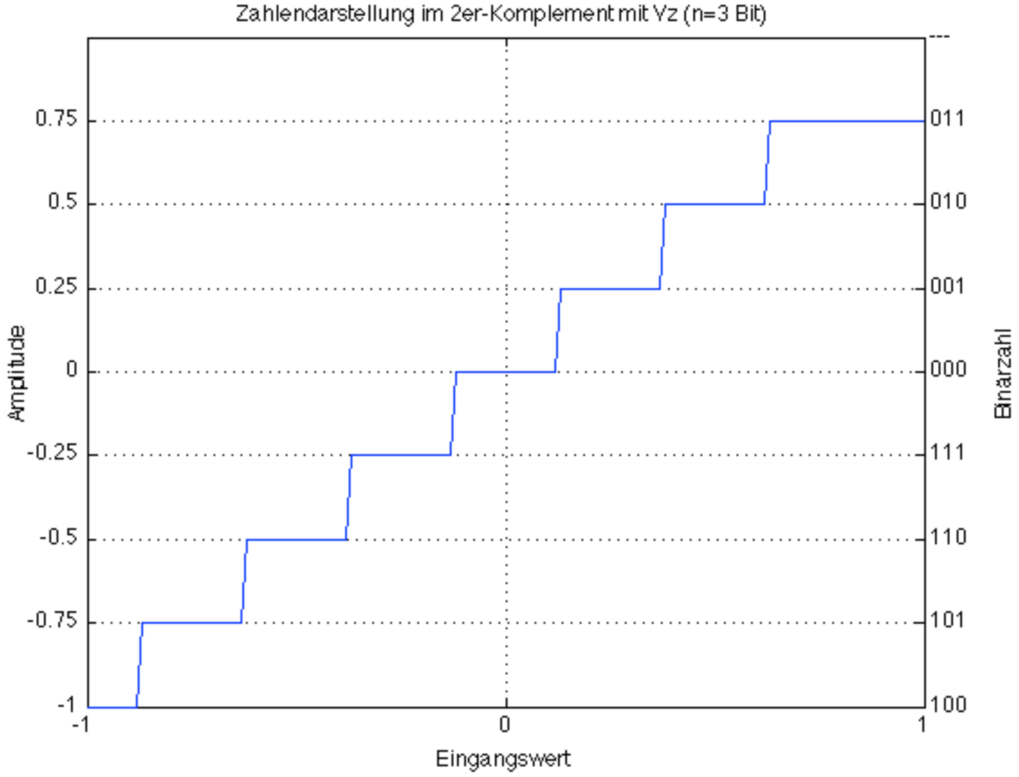
d) Welches der beiden Systeme aus a) würden Sie für eine Requantisierung eines digitalen Audiosignals wählen. Begründen Sie Ihre Entscheidung.
Lösung: Im System A wird zwar der Quantisierungsfehler spektral geformt, nicht aber das Dithersignal. Dieses wird genau wie das Nutzsignal mit $H(z) = 1$ übertragen. Im System B hingegen wird auch das Dithersignal bei der Spektralformung berücksichtigt. Dieser Algorithmus wird deswegen bevorzugt.

3 Zahlenformate: Fixed-Point Darstellung

a) Skizzieren Sie eine 3 bit Mid-Tread Quantisierungskennlinie für einen Ein- und Ausgangswertebereich von -1 bis 1. Beschriften sie die Ausgangswerte mit den Ampituden, sowie den entsprechenden Binärzahlen im 2er-Komplementär mit Vorzeichen.
Lösung: Der Zusammenhang von Zahlenwerten zur Binärsarstellung ist

$$x_Q = -b_0 + \sum_{i=1}^{w-1} b_{-i} 2^{-i} \quad , mit \ -1 \leq x_Q \leq 1 - 2^{-(w-1)}$$

wobei w die Wortbreite und i, das i-te bit ist. Daraus folgt die Skizze:



b) Erläutern sie die Unterschiede zwischen Fest- und Gleitkommadarstellung, sowie deren Vor- und Nachteile.
Lösung: Durch die Gleitkommadarstellung wird eine Variable Stufengröße für die Quantisierung erreicht, wodurch für kleine Eingangswerte ein geringerer Quantisierungsfehler als für große resultiert. Dies führt für kleine Eingangswerte im Vergleich zur Festkommadarstellung mit fester Stufengröße zu einem besseren SNR. Sie findet u.a. in Audio-Workstations mit leistungsfähigen Gleitkommaeinheiten moderner Prozessoren Anwendung. Die Festkommadarstellung wird aus Kostengründen in vielen DSP-Chips, wie sie z.B. in MP3 Playern verbaut sind verwendet, zudem hat sie eine geringere Leistungsaufnahme.