Digital Signal Processing for Music

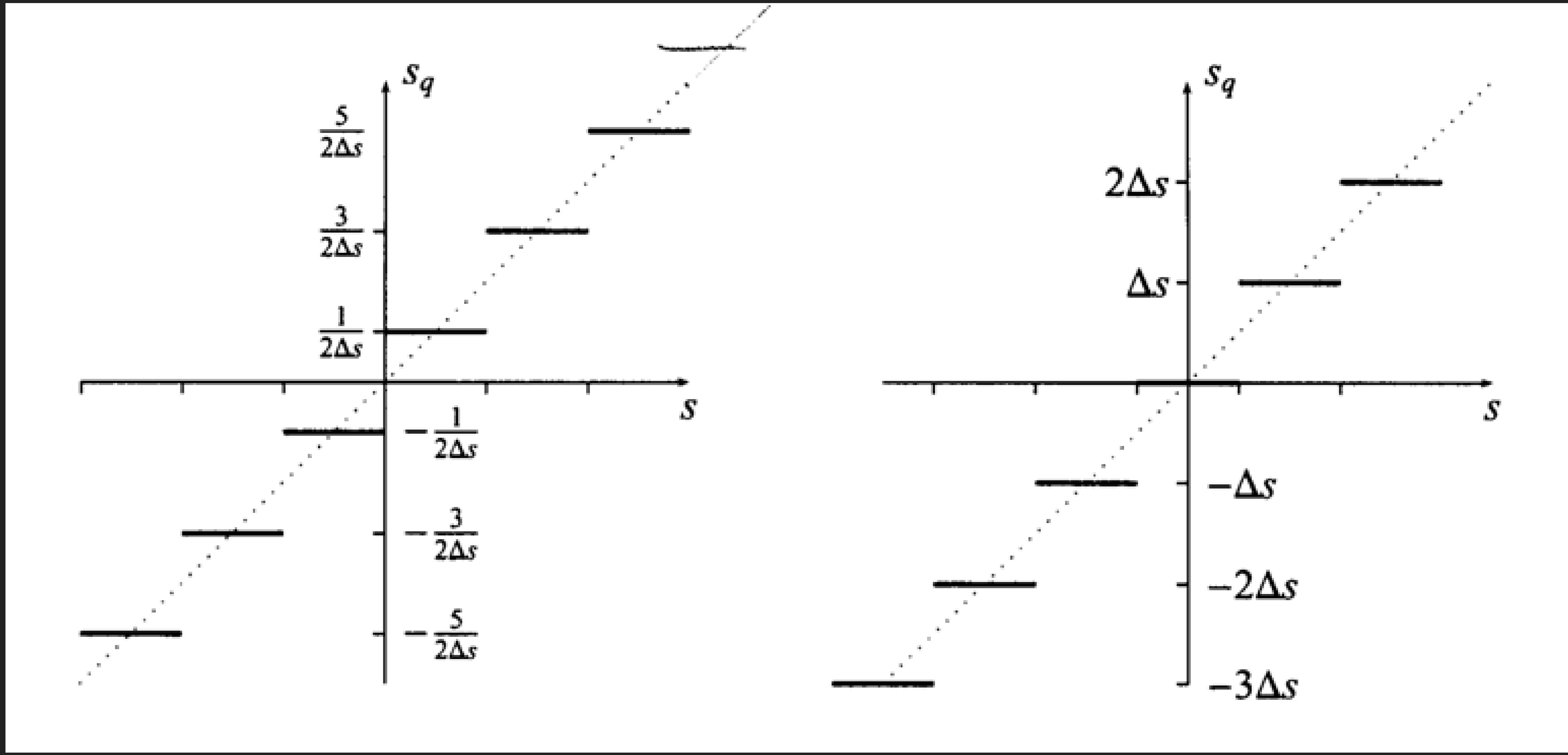
Part 9: Discretization, Part 2 - Quantization

Andrew Beck

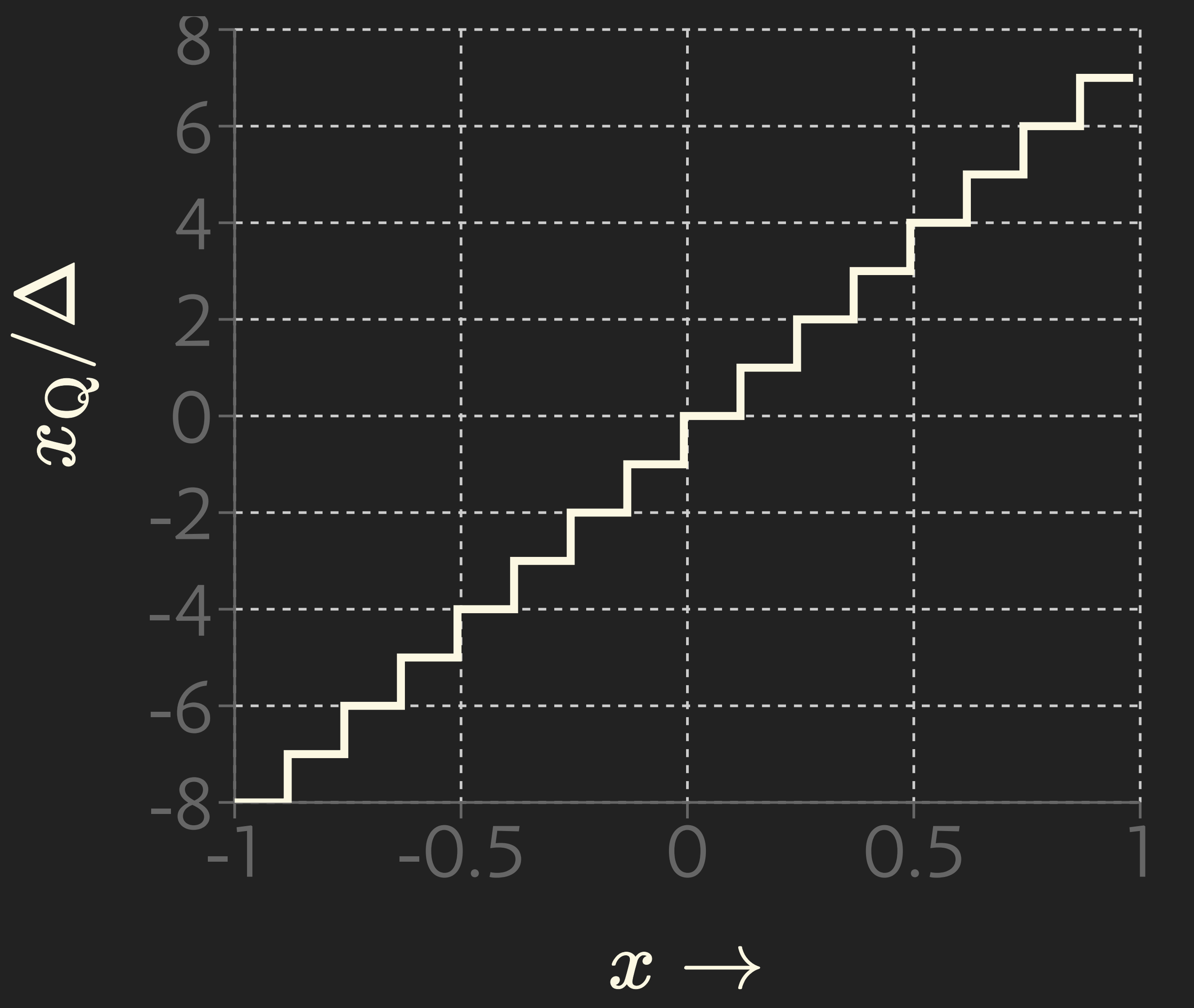
Quantizer:

Continuous $\mapsto$ Discrete (pre-defined set of allowed values)

- Quantization is **non-linear**
- Quanitzation is **irreversible**



Mid-Rise   Mid-Tread

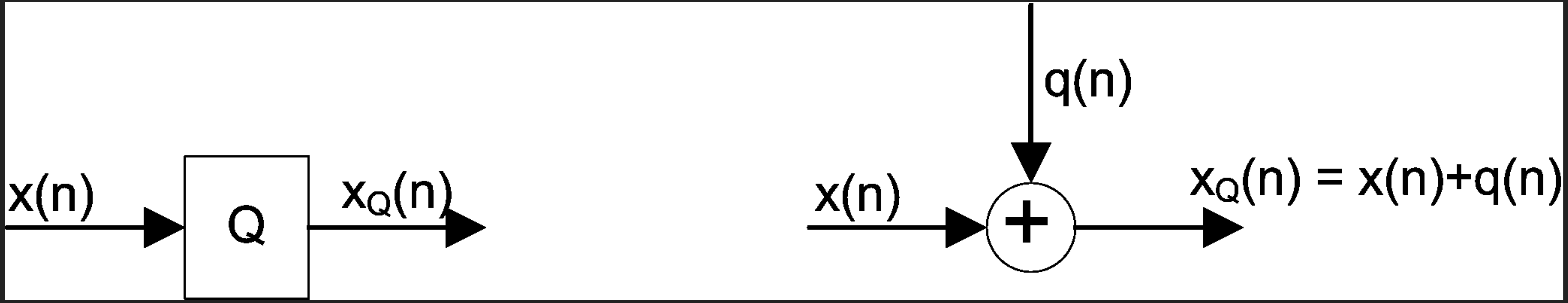


- **Number of quantization steps:** $\mathcal{M} = 16$
- **Word Length (bits):** $w = \log_2(\mathcal{M}) = 4\text{bit}$

Quanitzation: Word Length & Number of Steps

w	$\mathcal{M} = 2^w$
1	2
2	4
4	16
8	256
12	4096
16	65536
20	1048576
24	16777216

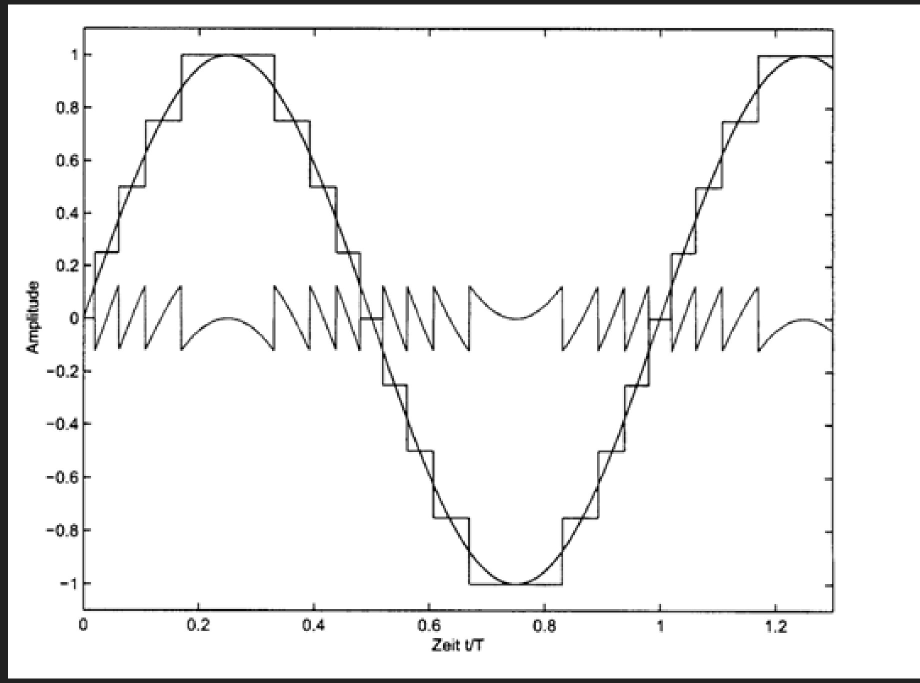
Quantization Error: Definition

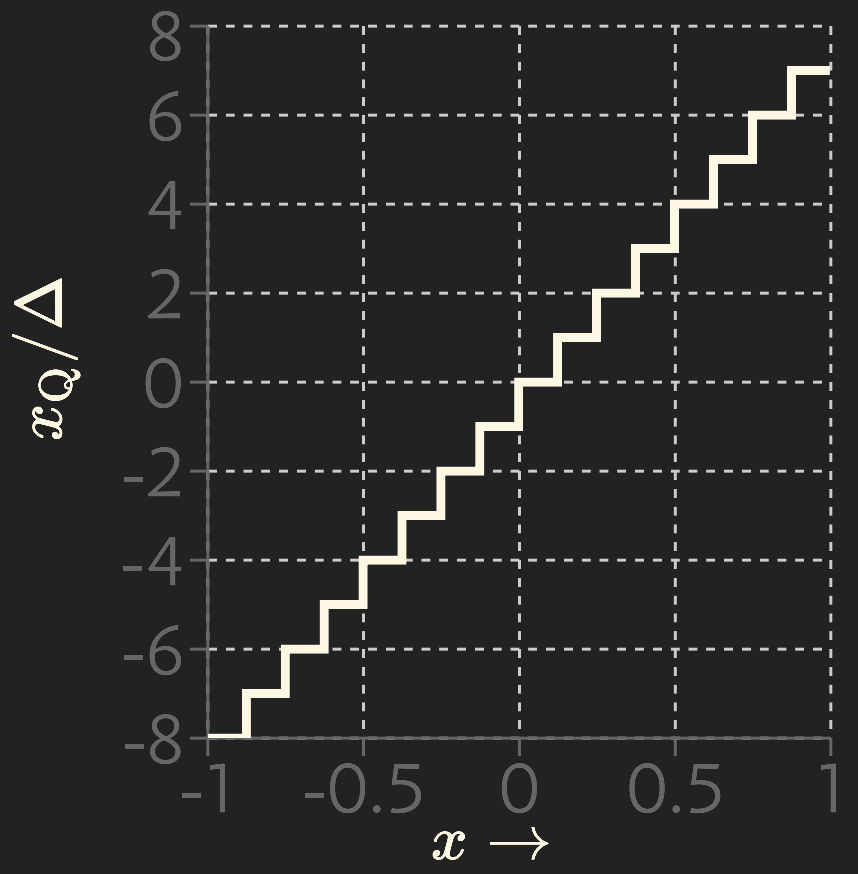


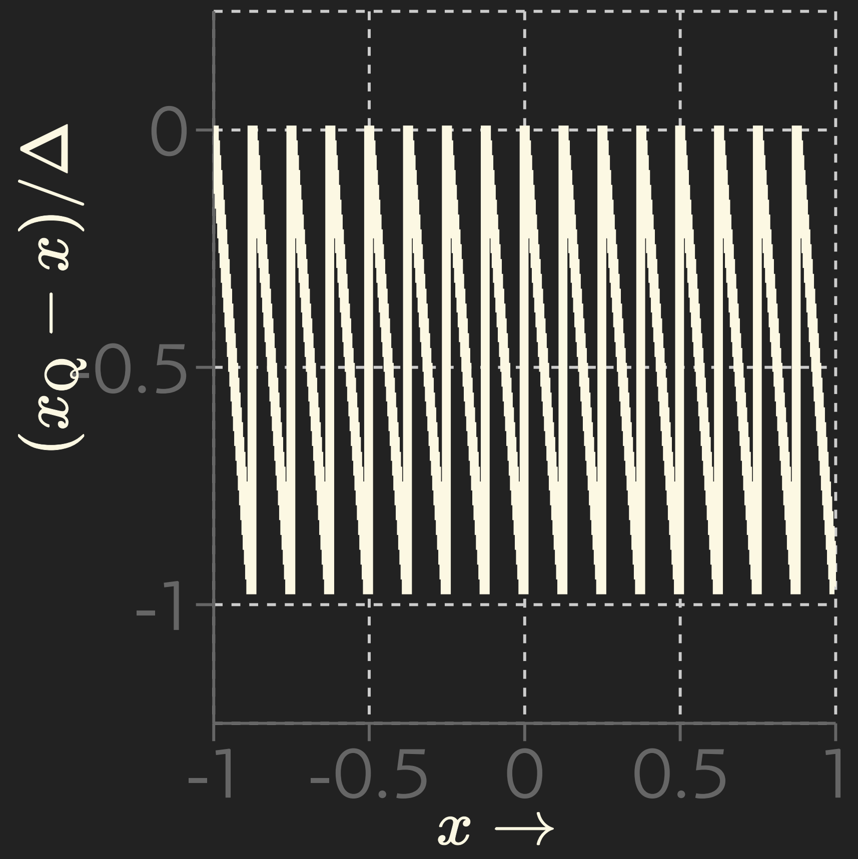
$$q(i) = x_{\mathrm{Q}}(i) - x(i)$$

What is the maximum amplitude of the quantization error?

Maximum Amplitude of Quantization Error





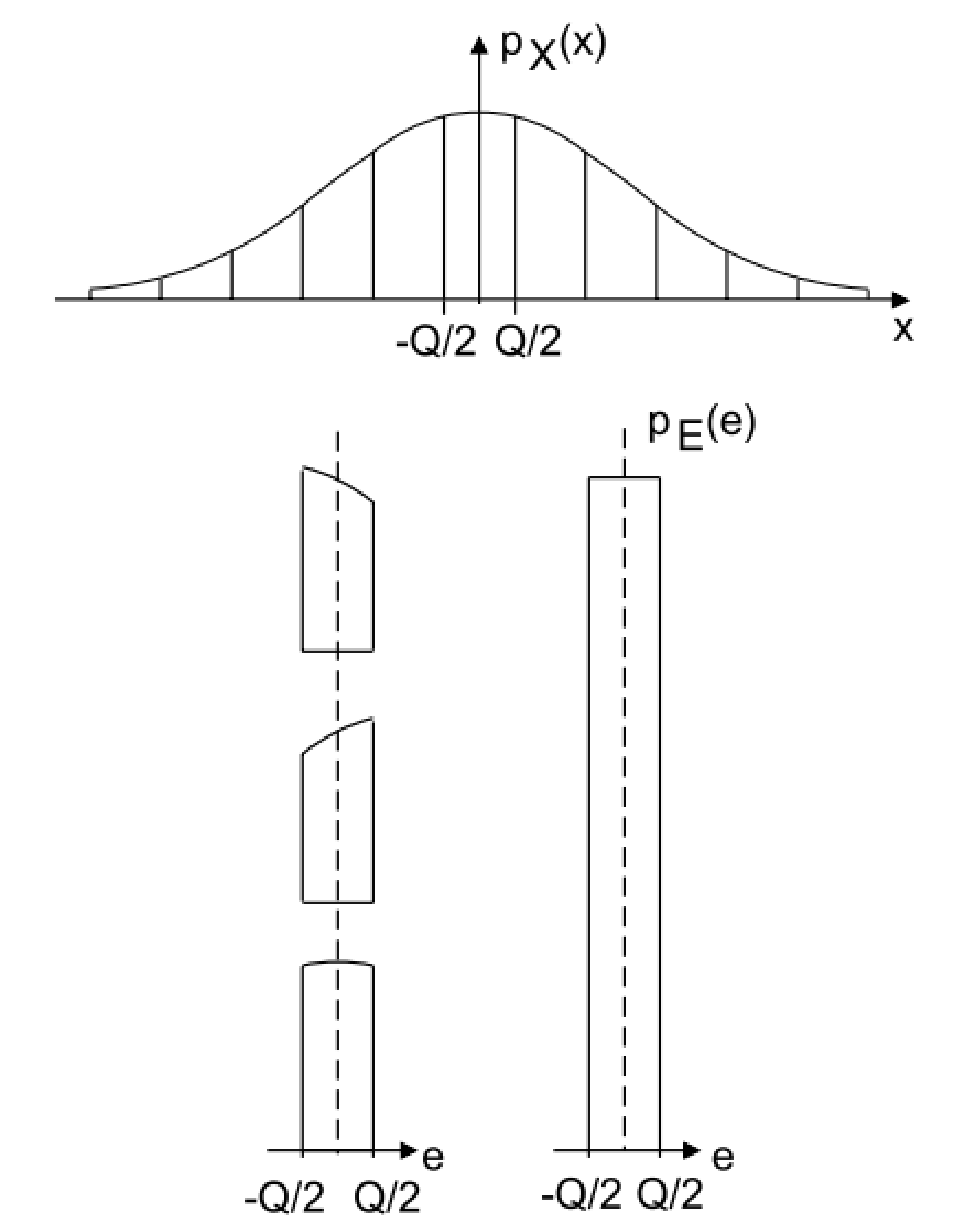


$$|q(i)| \leq \frac{\Delta}{2}$$

What is the PDF of the quantization error?

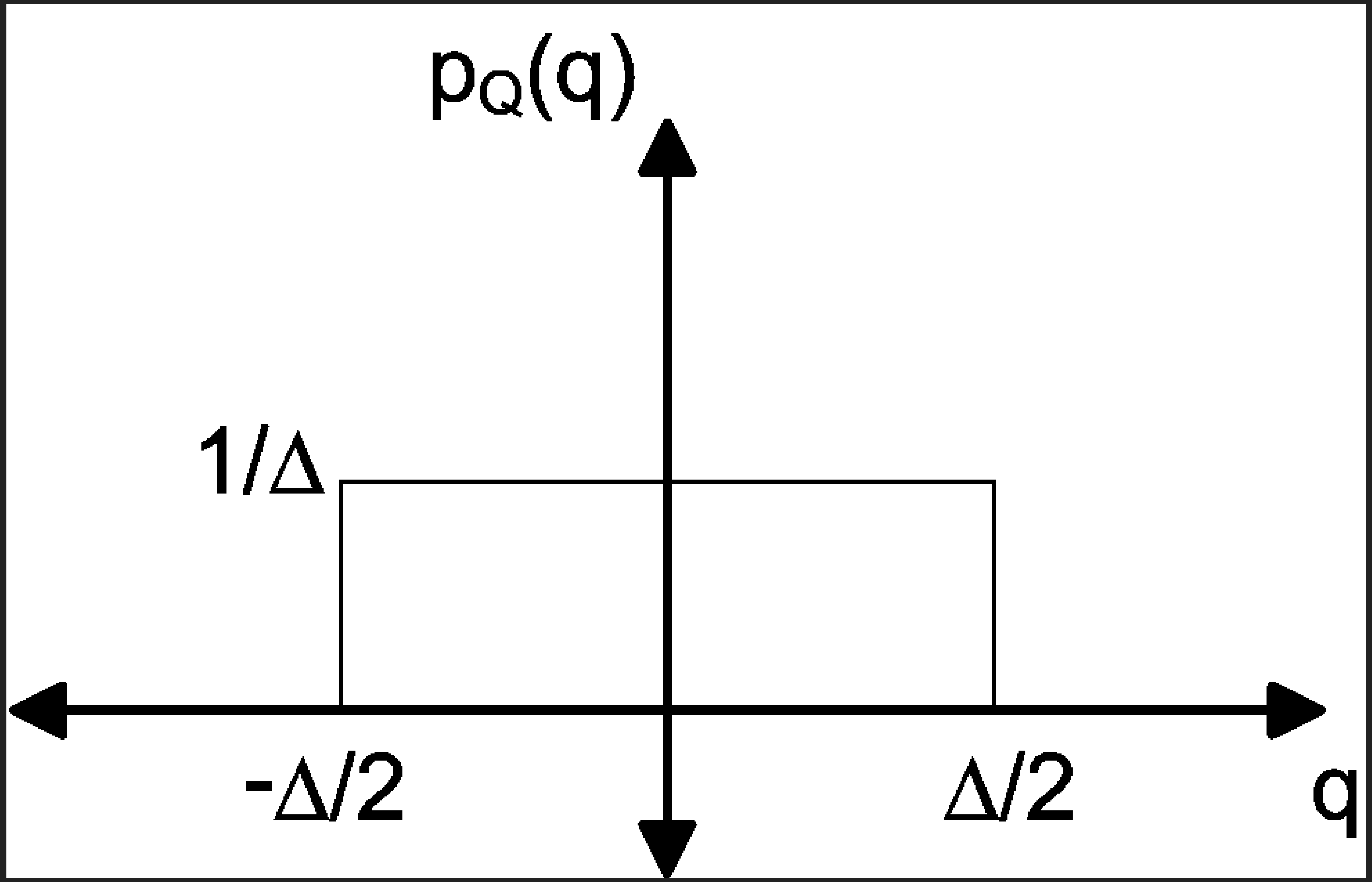
PDF of Quantization Error

Assuming $\Delta \ll max(|x(i)|)$



PDF of Quantization Error

Assuming $\Delta \ll max(|x(i)|)$



It can be shown that the PDF of the quanitzation error depends (without derivation)

- on the **variance of the input** signal in relation to the step size
- on the **pdf of the input** signal

→ will be **uniform for large values of $\frac{\sigma_X}{\Delta}$**

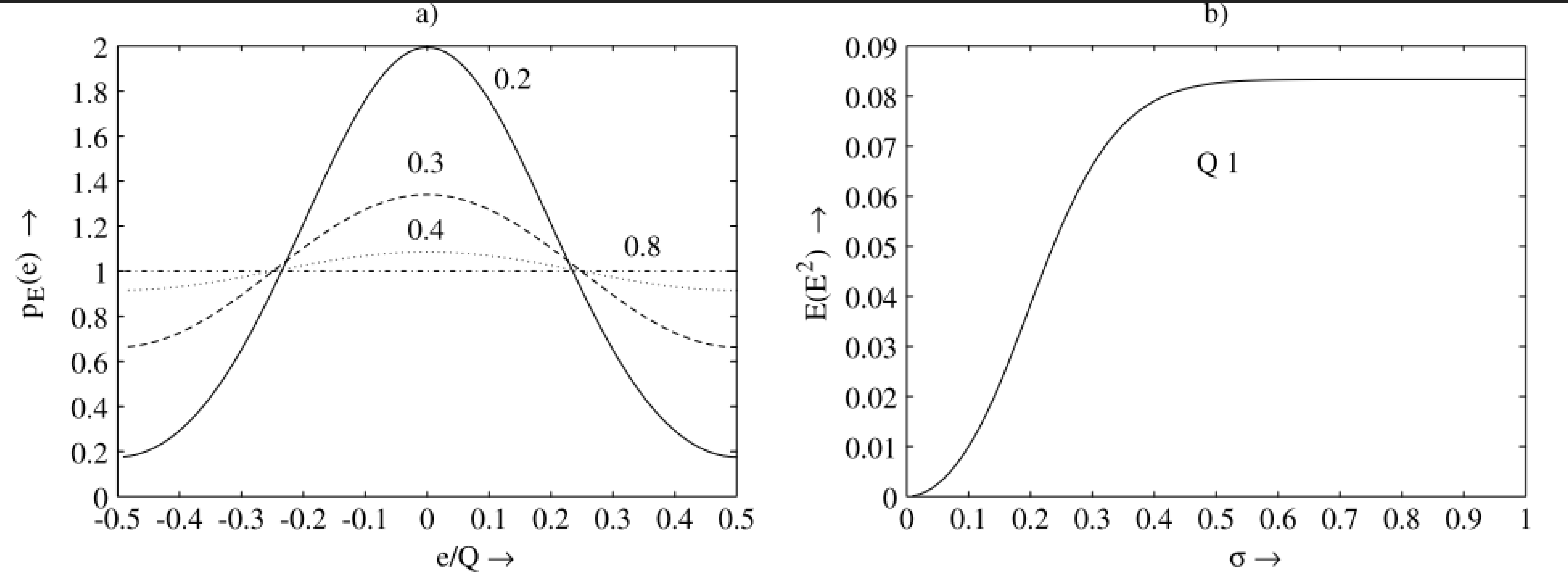


Figure 2.16 (a) PDF of quantization error for different standard deviations of a Gaussian PDF input. (b) Variance of quantization error for different standard deviations of a Gaussian PDF input.

How to computer the power W_Q of Quantization Error?

Computing power W_{Q} of Quantization Error

From PDF:

$$W_{\mathrm{Q}} = \int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} q^2 \cdot \underbrace{p_{\mathrm{Q}}(q)}_{\frac{1}{\Delta}} \, dq$$

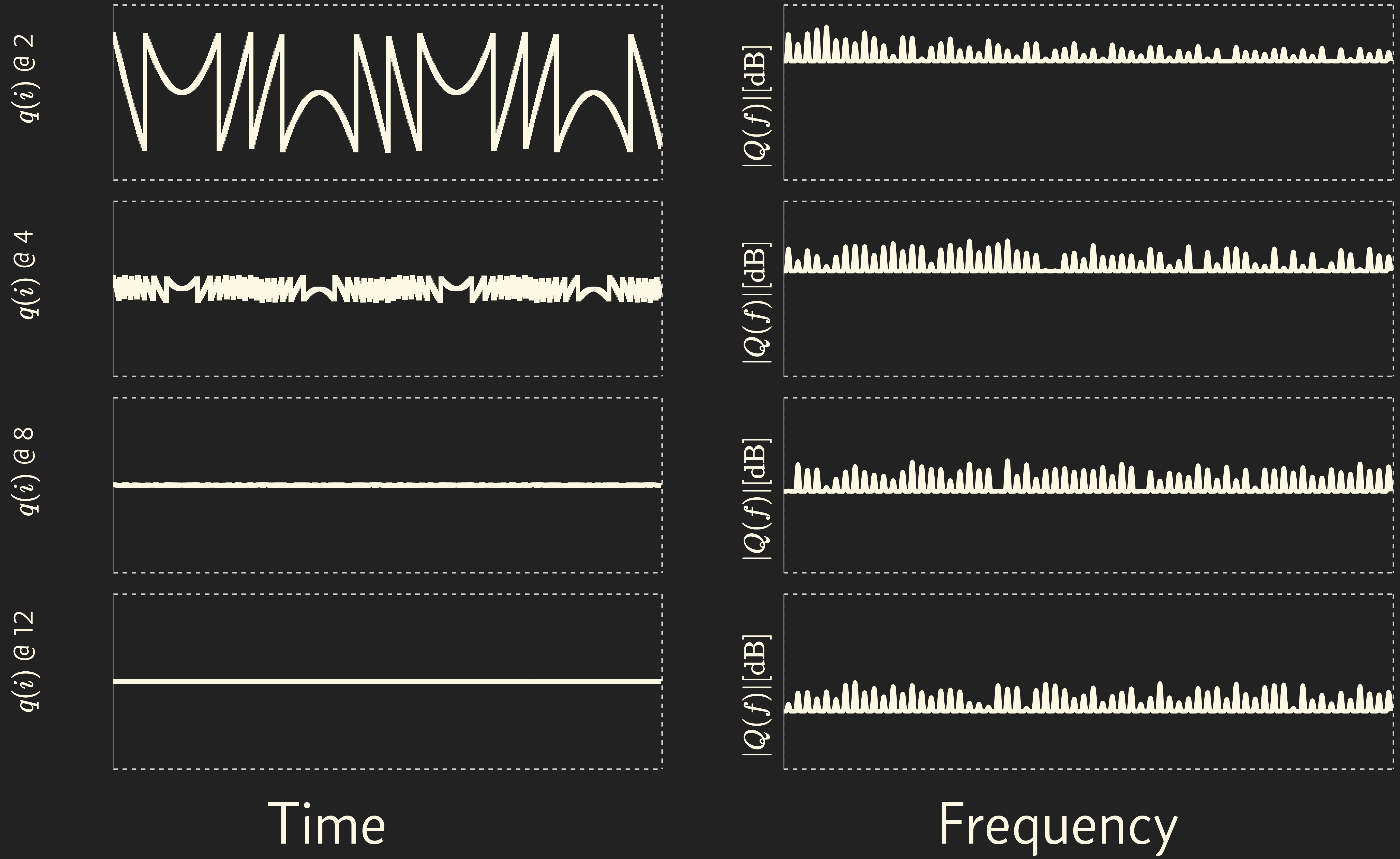
$$= \frac{1}{\Delta} \int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} q^2 \, dq$$

$$= \frac{1}{\Delta} \left[\frac{1}{3} q^3 \right]_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}}$$

$$= \frac{1}{3\Delta} \left(\frac{\Delta^3}{8} + \frac{\Delta^3}{8} \right)$$

$$= \frac{\Delta^2}{12}$$

Quantization Error of Full-Scale Sinusoidal



Audio Examples

w	$x_{\mathrm{Q,sine}}(i)$	$q_{\mathrm{sine}}(i)$	$x_{\mathrm{Q,speech}}(i)$	$q_{\mathrm{speech}}(i)$	$x_{\mathrm{Q,music}}(i)$	$q_{\mathrm{music}}(i)$
16						
12						
8						
6						
4						
2						

Quality Assessment of a Quantizer: Signal-to-Noise Ratio (SNR)

- Power of the signal in relation to power of the (quantization) noise.

$$SNR' = \frac{\text{signal energy}}{\text{noise energy}} = \frac{W_{\mathrm{S}}}{W_{\mathrm{Q}}}$$

- Often in decibel

$$SNR = 10 \cdot \log_{10}\left(\frac{W_{\mathrm{S}}}{W_{\mathrm{Q}}}\right) \text{ [dB]}$$

- SNR grows by:
  - Reducing the noise power
  - Increasing the signal power

Derive the SNR of quantized full-scale sinusoidal

$$SNR = 10 \cdot \log_{10}\left(\frac{W_\mathrm{S}}{W_\mathrm{Q}}\right)\ [\mathrm{dB}]$$

Use $\sin^2(t) = \frac{1-\cos(2t)}{2}$

$$W_\mathrm{S} = \frac{A^2}{2} \ \xrightarrow{\text{full-scale}} \ W_\mathrm{S} = \frac{(\Delta \cdot 2^{w-1})^2}{2}$$

$$W_\mathrm{Q} = \frac{\Delta^2}{12}$$

$$\frac{W_\mathrm{S}}{W_\mathrm{Q}} = \frac{3}{2} \cdot 2^{2w}$$

$$SNR = w \cdot 20 \log_{10}(2) + 10 \cdot \log_{10}\left(\frac{3}{2}\right)\ [\mathrm{dB}]$$

Derive the SNR of full-scale square wave

$$SNR = 10 \cdot \log_{10}\left(\frac{W_\mathrm{S}}{W_\mathrm{Q}}\right) \; [\mathrm{dB}]$$

$$W_\mathrm{S} = A^2 \;\; \overrightarrow{\text{full-scale}} \; W_\mathrm{S} = (\Delta \cdot 2^{w-1})^2$$

$$W_\mathrm{Q} = \frac{\Delta^2}{12}$$

$$\frac{W_\mathrm{S}}{W_\mathrm{Q}} = 3 \cdot 2^{2w}$$

$$SNR = w \cdot 20 \log_{10}(2) + 10 \cdot \log_{10}(3) \; [\mathrm{dB}]$$

Signal-to-Noise Ratio

$$SNR = 6.02 \cdot w + c_{\mathrm{S}} \quad [\mathrm{dB}]$$

- Every addtional bit adds ~6 dB SNR
- Constant C_{S} depends on signal (scaling and PDF shape)

SNR for different input signal examples:

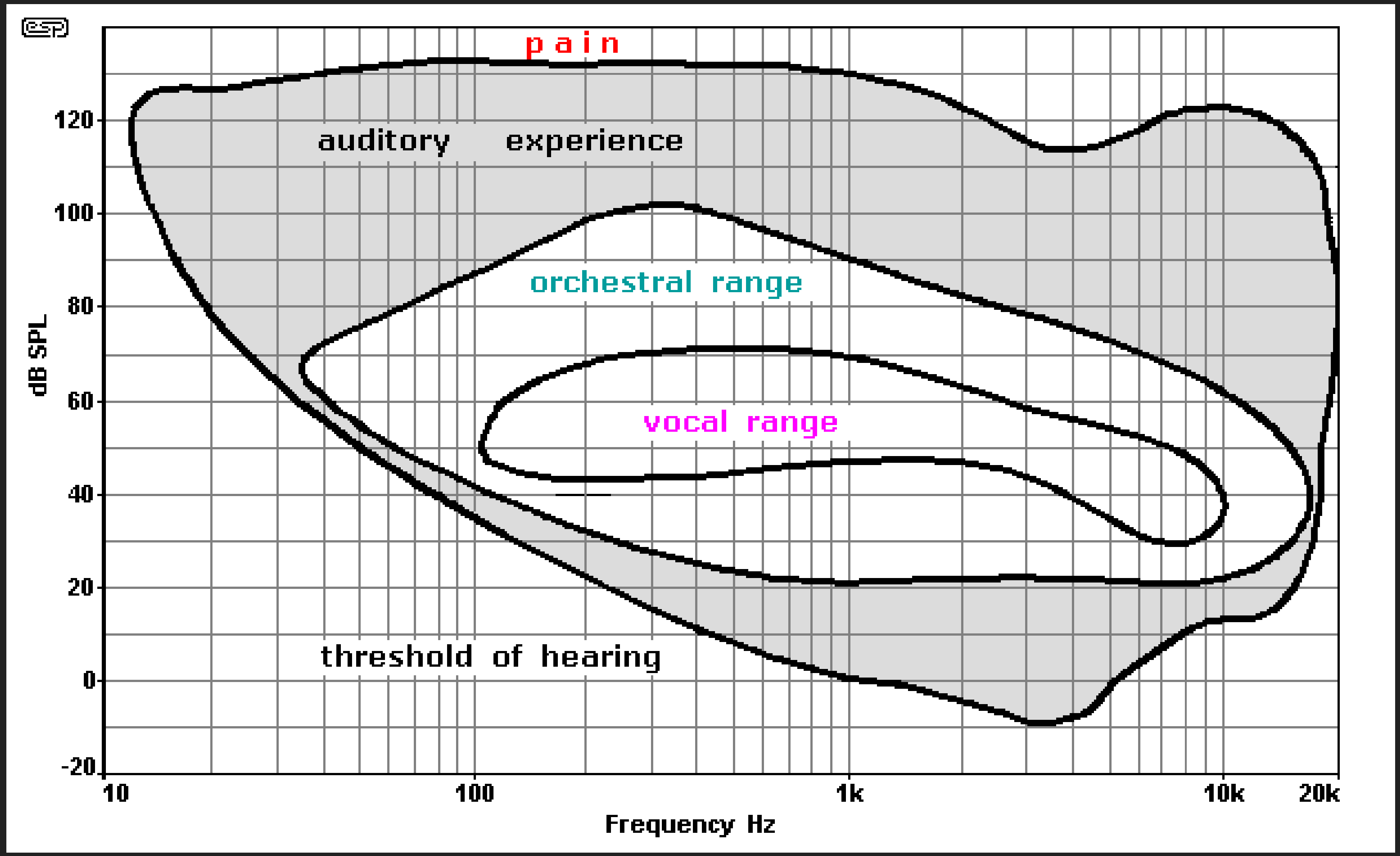
- Square wave (full scale): $C_{\mathrm{S}} = 4.77\mathrm{dB}$
- Sinusoidal wave (full scale): $C_{\mathrm{S}} = 1.76\mathrm{dB}$
- Rectangular PDF (full scale): $C_{\mathrm{S}} = 0\mathrm{dB}$
- Gaussian PDF (full scale $= 4\sigma_g$): $C_{\mathrm{S}} = -7.27\mathrm{dB}$

Quantization: Word Length and SNR

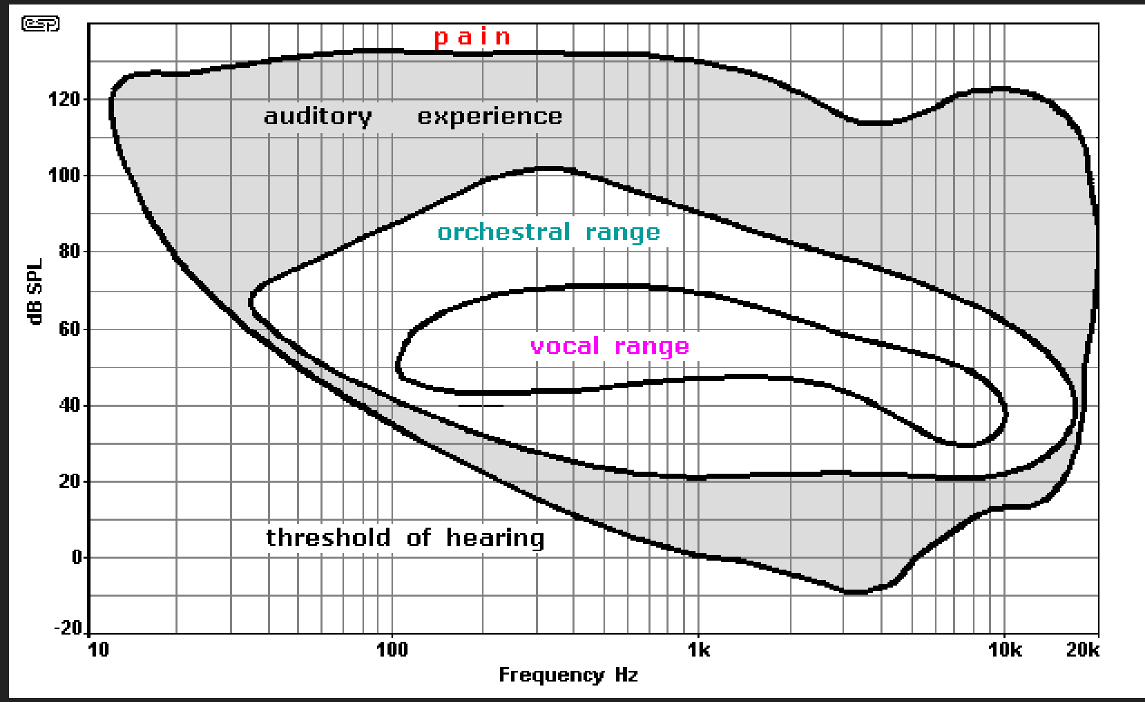
w	Δ	Max. Amp	theo. SNR
8 (Int)	± 1	$0 \ldots 255$	$\approx 48\,\mathrm{dB}$
16 (Int)	± 1	$-32768 \ldots 32767$	$\approx 96\,\mathrm{dB}$
20 (Int)	± 1	$-524288 \ldots 524287$	$\approx 120\,\mathrm{dB}$
24 (Int)	± 1	$-16777216 \ldots 16777215$	$\approx 144\,\mathrm{dB}$
32 (Float)	$\pm 1.175 \cdot 10^{-38}$	$\pm 3.403 \cdot 10^{1038}$	$1529\,\mathrm{dB}$
64 (Float)	$\pm 2.225 \cdot 10^{-308}$	$\pm 1.798 \cdot 10^{10308}$	$12318\,\mathrm{dB}$

SNR and Auditory Sensation Area

How many bits do we need?

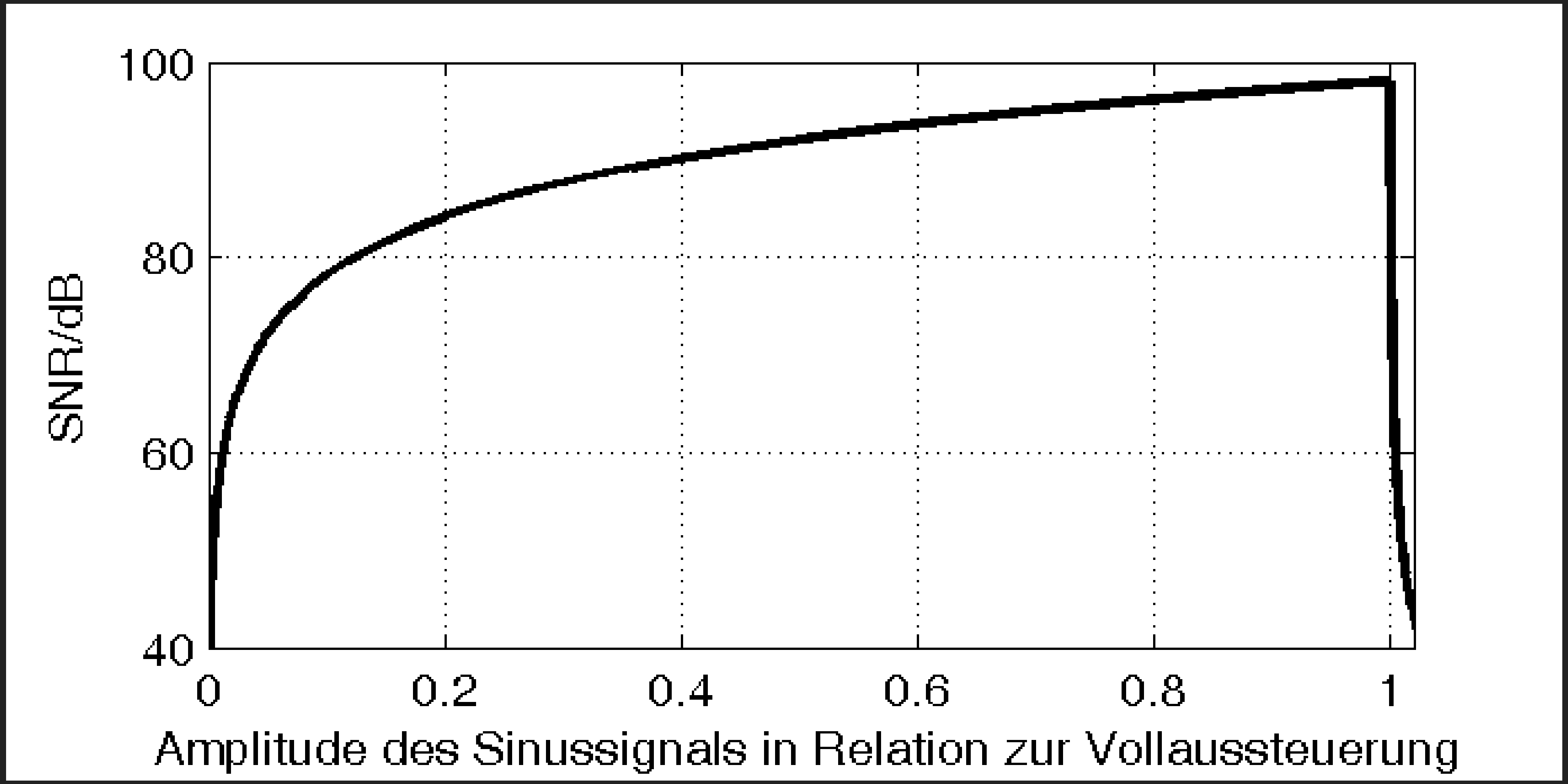


How many bits do we need?



- To cover the whole range of hearing: 20-24 bit
- Practically, a lower range is sufficientas the dynamic range of recordings has to be much lower
- In production with many processing and possible requantization steps, high resolution (if possible floating point) is recommended

SNR and Signal Scaling



Full Scale:

- Absolute maximum before clipping
- Usually 1 (in floating point systems)
- Marks 0 dbFS

- Quantization is **non-linear** & **irreversible**
  - Information is lost
  - Error is introduced
- Quantization **error**
  - Power is determined by number of bits (word length)
  - Is approxiamately white noise (float spectrum and uncorrelated to signal) when the signal power is much higher than the quantization step size
  - Special severe case: clipping
- **SNR** is used to assess quantizer quality
  - Depends on both signal power and quant error power (ratio)
  - Each additional bit gains 6 dB SNR
  - Different signals with identical maximum amplitude yield different SNRs
- **Typical word lengths** include
  - 8 bit: Phone
  - 16 bit: Consumer audio
  - 24 bit and higher: Production audio