EE4C5 Digital Signal Processing

Lecture 6 – Quantisation

This lecture

- Based on Chapter 4 (Section 8) of O&S
- All images from O&S book unless otherwise stated.

Working with discrete signals

Ideal scenario

Real-world

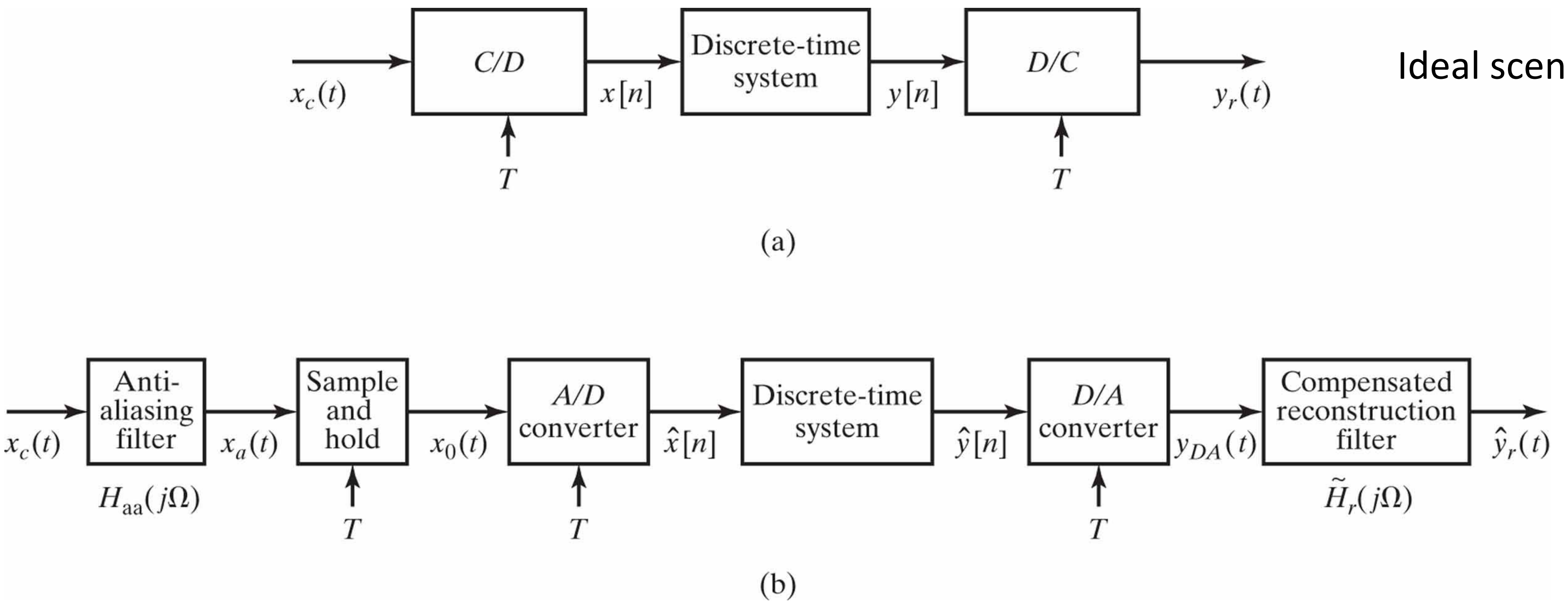


Figure 4.47 (a) Discrete-time filtering of continuous-time signals. (b) Digital processing of analog signals.

Finite Precision

- Ideally, the system parameters along with the signal variables have infinite precision taking any value between $-\infty$ and ∞
- In practice, they can take only discrete values within a specified range since the registers of the computer/electronic device where they are stored are of finite length.
- However, if the quantization amounts are small compared to the values of signal variables and filter parameters, can still have useful system.
- Useful to model the effects/impacts on your system.

Digital Signals

- Digital signal discrete in time and amplitude.
- Can only represent sample amplitudes with finite number of values
  - E.g. 16-bit audio.
- Introduce errors due to rounding and truncation effects.
  - Propagate through the system.

Analysis of Finite Wordlength Effects

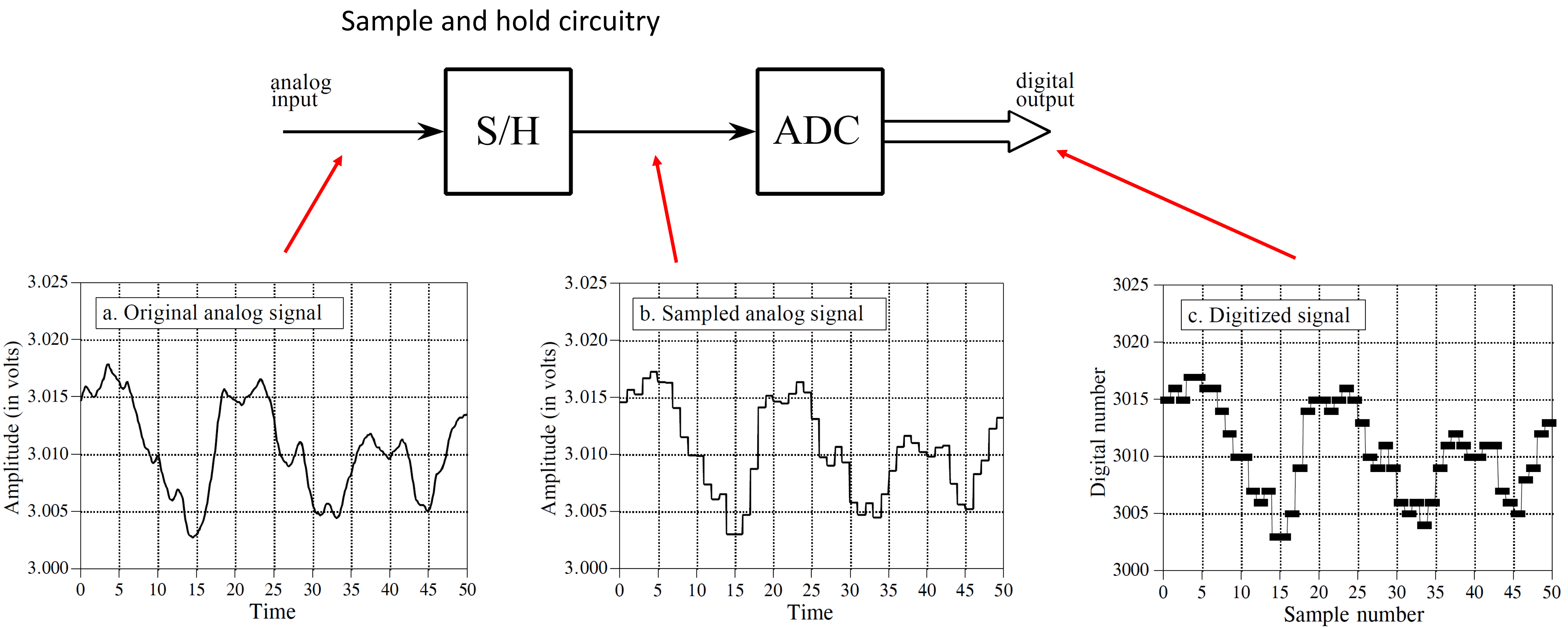
- Sources where we will encounter this sort of error in 4C5
  - Filter coefficient quantization.
  - Quantization of arithmetic operations.
  - A/D conversion.

A/D Conversion

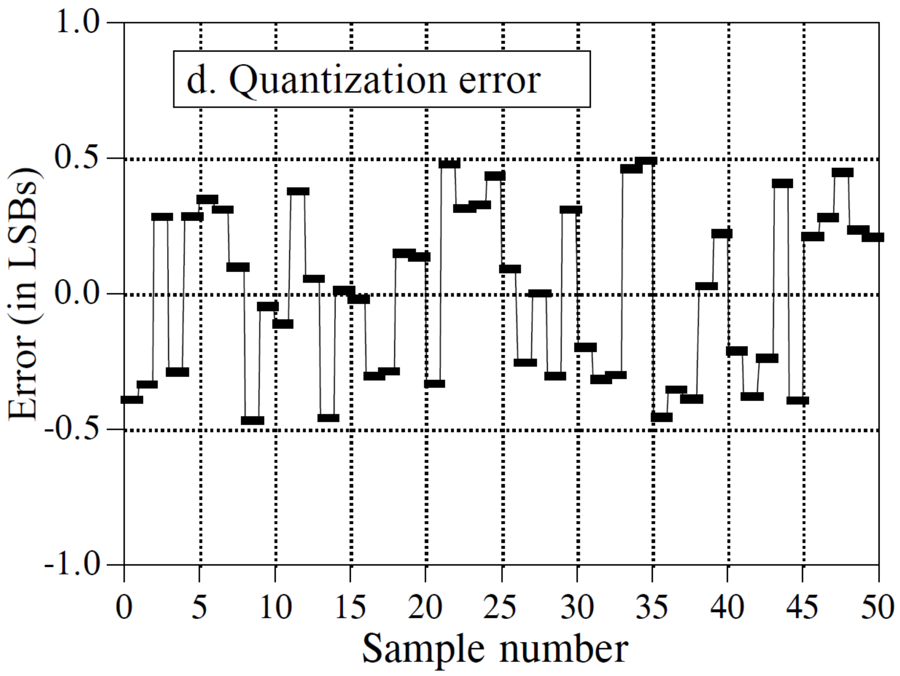
Analog to digital (A/D) conversion

- ADC
- Physical device that converts voltage or current amplitude at the input into a binary code.
- The binary code represents a quantised amplitude value closest to the amplitude of the input.
  - Or “best” in some sense.

Digitising a signal



Resulting error



Uniform Quantisation

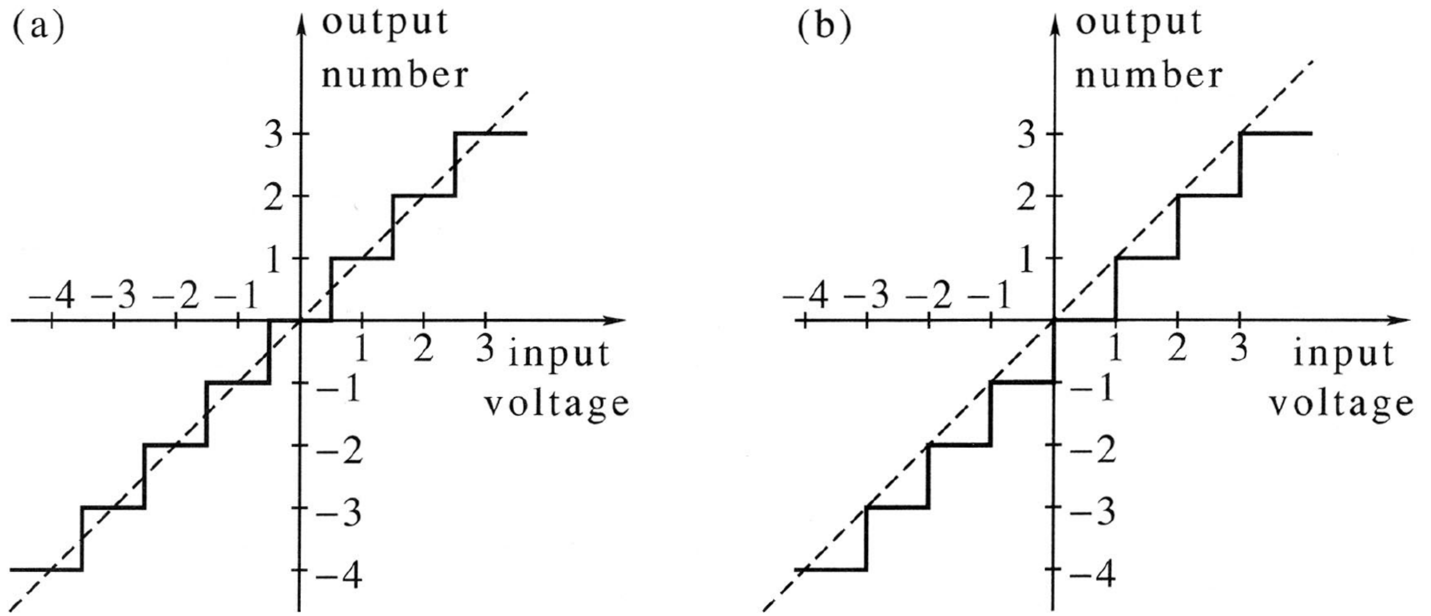


Figure 3.20 Quantization in analog-to-digital converter: (a) rounding; (b) truncation. Staircase lines show the actual responses; dashed lines show the ideal responses.

Source: Porat "A Course in Digital Signal Processing

A 3-bit uniform quantiser

- Quantisation step Δ - an LSB.
- Note positive and negative values.
- Recall 2's complement format?
- 2^{B+1} levels with $(B+1)$-bit binary code.
- $\Delta = \frac{2X_m}{2^{B+1}} = \frac{X_m}{2^B}$

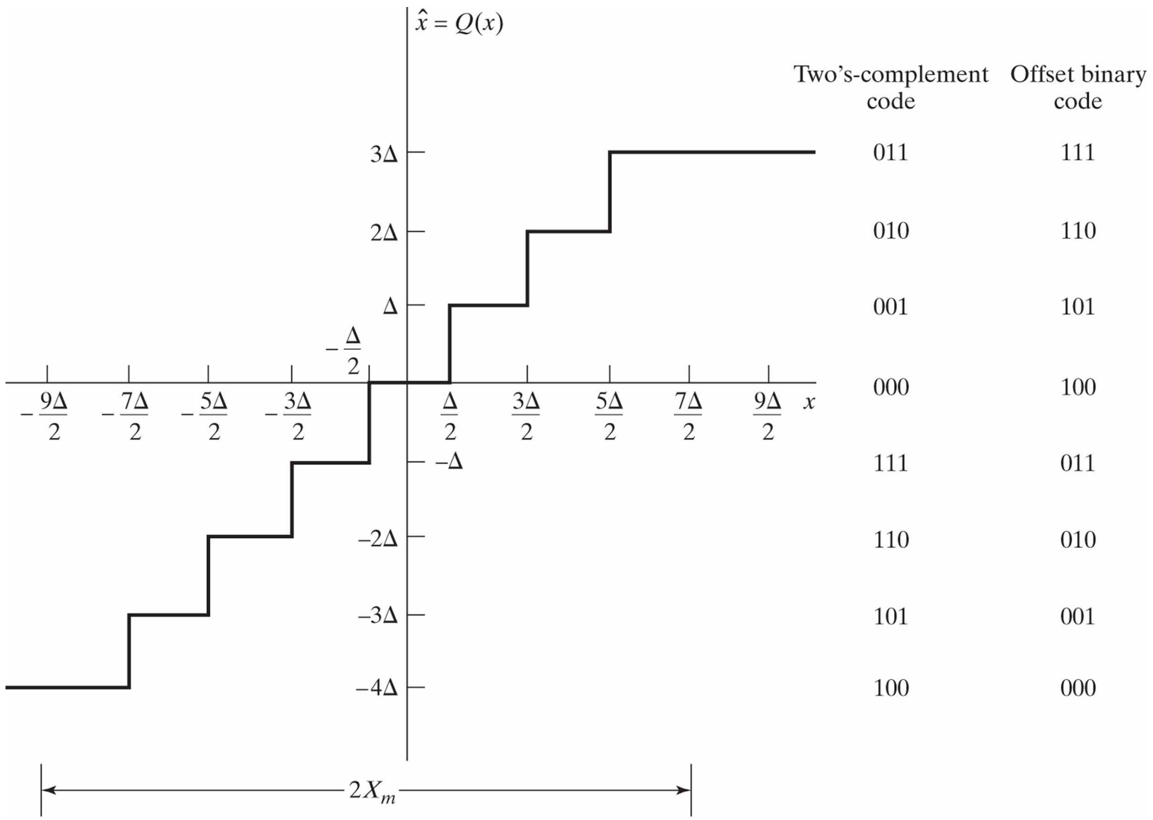


Figure 4.54 Typical quantizer for A/D conversion.

Quantisation Errors

- Quantised sample $\hat{x}[n]$ is the approximation of the "true" sample $x[n]$.
- Then can define quantisation error as:
  - $e[n] = \hat{x}[n] - x[n]$
- Errors due to:
  - Quantisation noise – accumulates due to rounding and truncation. Sufficiently small step size.
  - Saturation – input exceeds max or min value you can represent (clipping). Plan for range.

Errors

- Where do the two types of errors occur opposite?

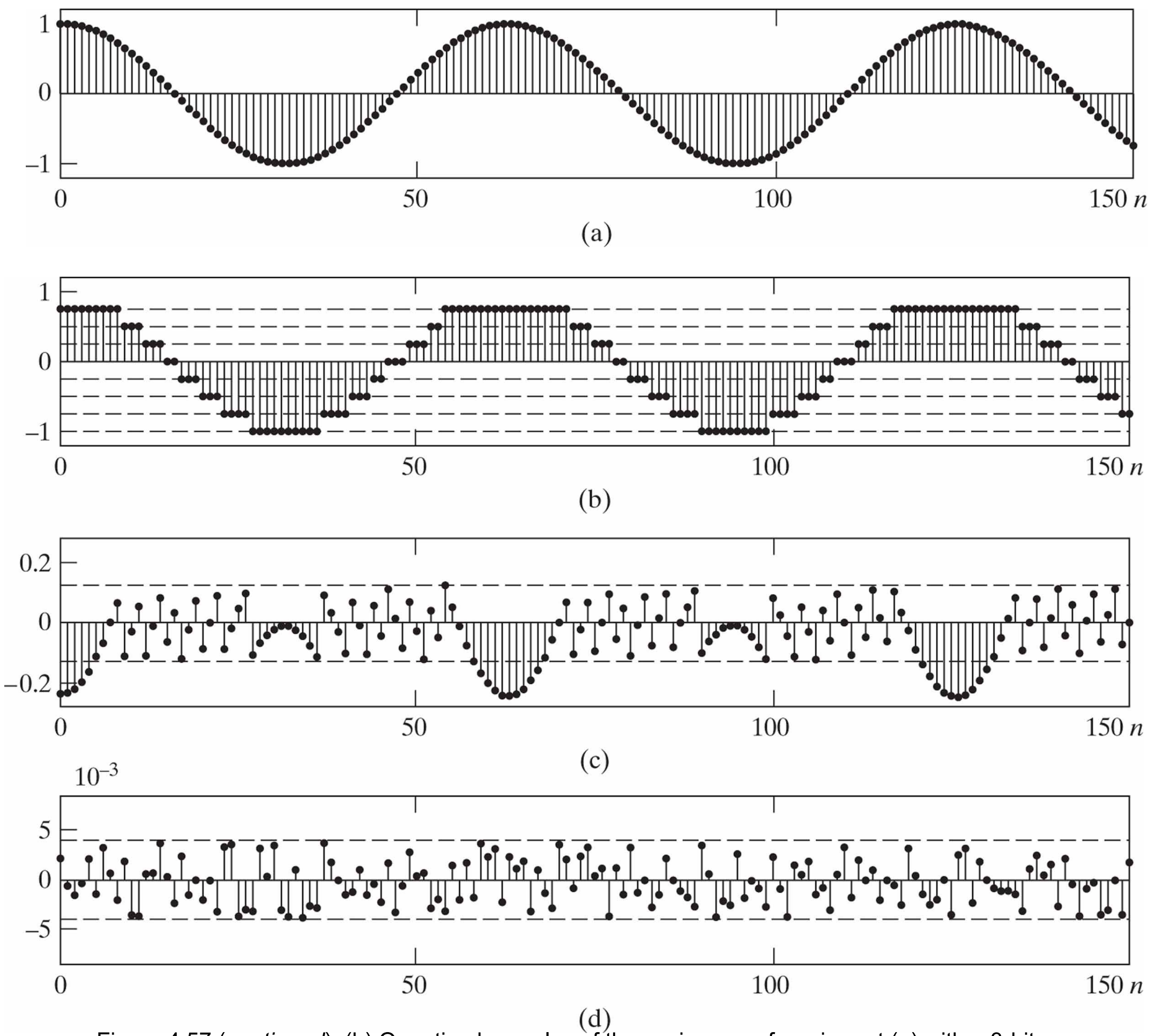


Figure 4.57 (*continued*) (b) Quantized samples of the cosine waveform in part (a) with a 3-bit quantizer.
(c) Quantization error sequence for 3-bit quantization of the signal in (a). (d) Quantization error sequence for 8-bit quantization of the signal in (a).

Error analysis

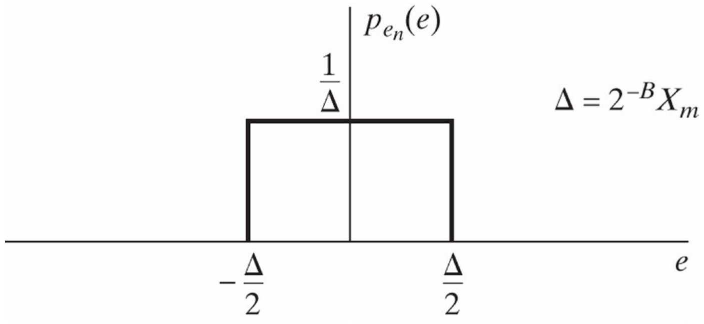


Figure 4.58 Probability density function of quantization error for a rounding quantizer such as that of Figure 4.54.

- Rounding uniform quantiser
  - $^{-\Delta}/_2 < e[n] < {^{\Delta}/_2}$
- For small Δ, treat $e[n]$ as random variable distributed from $^{-\Delta}/_2$ to $^{\Delta}/_2$
- Assume successive noise samples uncorrelated with each other and $e[n]$ uncorrelated with $x[n]$
  - Valid for complicated signals like audio etc.
- Variance of $e[n]$ is:
  - $\sigma_e^2 = \int_{-\Delta/2}^{\Delta/2} e^2 \frac{1}{\Delta} de = \frac{\Delta^2}{12}$
- For $(B+1)$-bit quantiser with full scale value X_M, the noise variance or power is therefore:
  - $\sigma_e^2 = \frac{2^{-2B} X_M^2}{12}$

Required Reading & other material

- Oppenheim & Schafer, Chapter 4, section 8.2, 8.3
- Noise from quantisation - audio samples & explorations
  - https://www.youtube.com/watch?v=UaKho805vCE
  - https://www.youtube.com/watch?v=1KBLguIXL30