COMP ENG 4TL4:

Digital Signal Processing

Notes for Lecture #3

Wednesday, September 10, 2003

1.4 Quantization

Digital systems can only represent sample amplitudes with a finite set of prescribed values, and thus it is necessary for A/D converters to *quantize* the values of the samples $x[n]$.

A typical form of quantization uses *uniform* quantization steps, where the input voltage is either *rounded* or *truncated*.

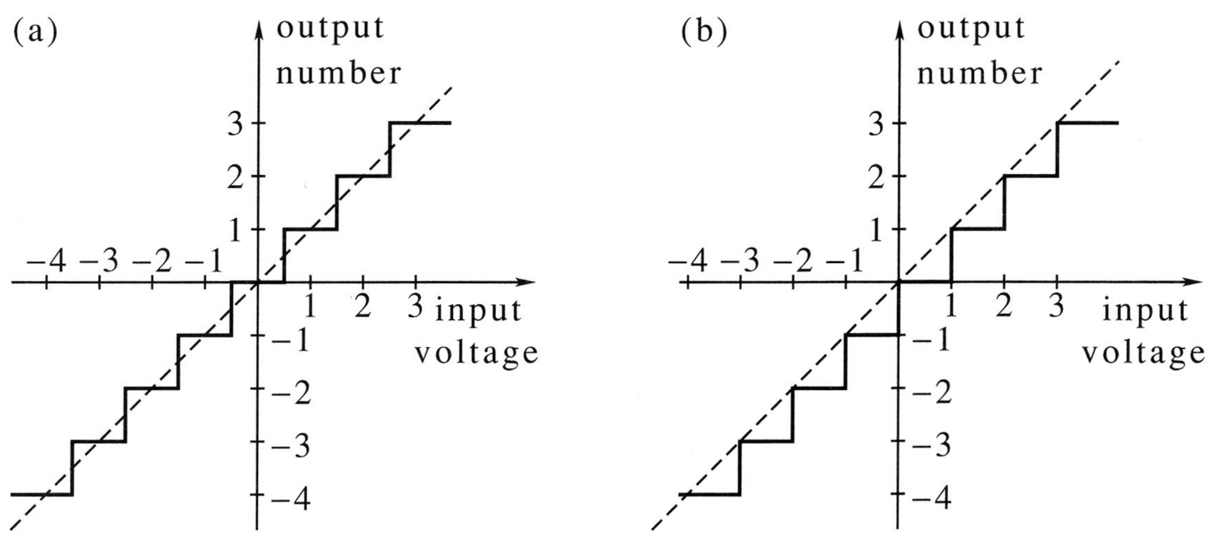

Figure 3.20 Quantization in analog-to-digital converter: (a) rounding; (b) truncation. Staircase lines show the actual responses; dashed lines show the ideal responses. *(Porat)*

A 3-bit rounding uniform quantizer with quantization steps of Δ is shown below.

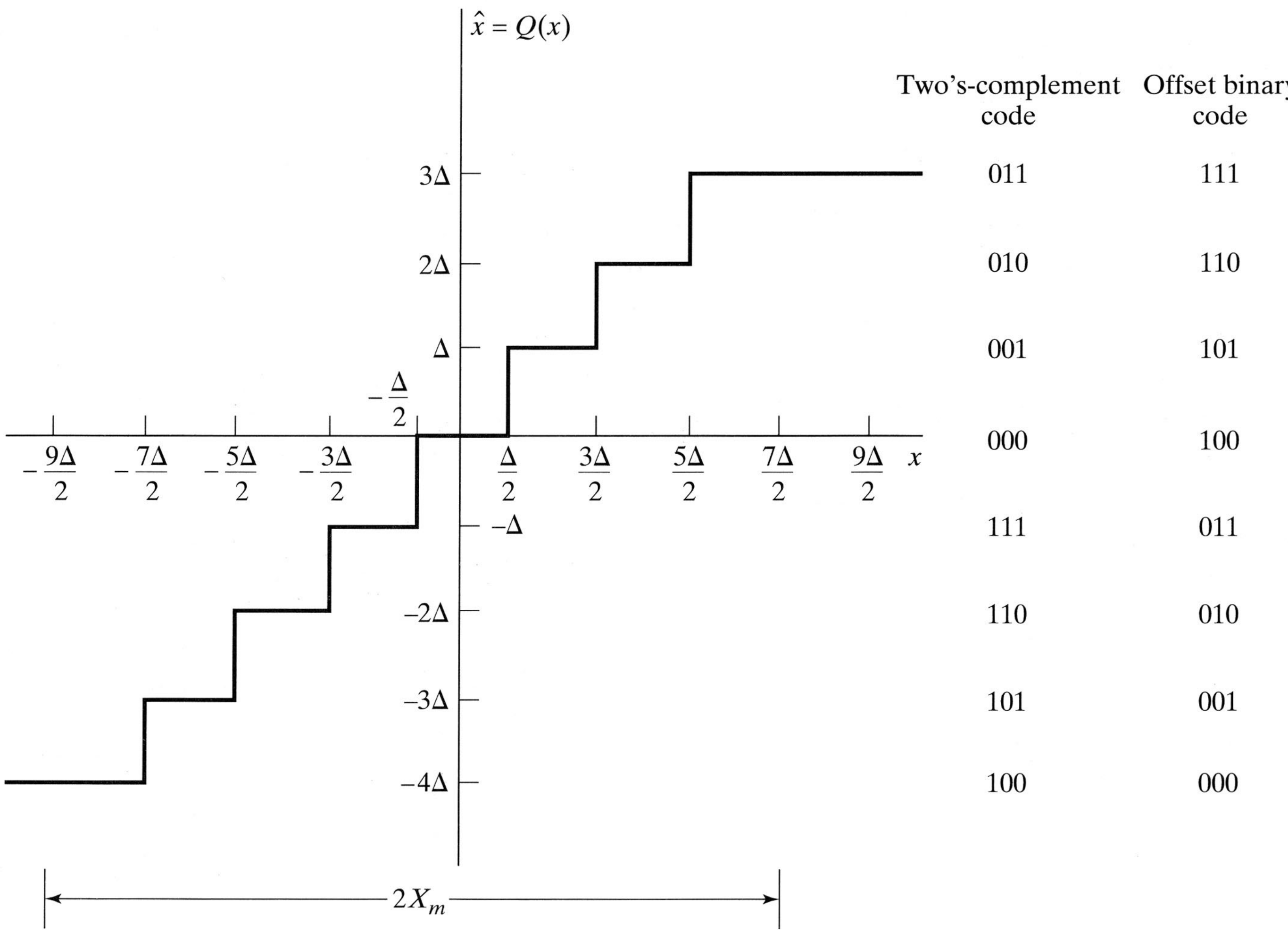

Figure 4.48 Typical quantizer for A/D conversion.

(Oppenheim and Schafer)

Two forms of *quantization error* $e[n]$ exist:

1. Quantization noise: due to rounding or truncation over the range of quantizer outputs; and
2. Saturation ("peak clipping"): due to the input exceeding the maximum or minimum quantizer output.

Both types of error are illustrated for the case of a sinusoidal signal in the figure on the next slide.

Quantization noise can be minimized by choosing a sufficiently small quantization step. We will derive a method for quantifying how small is "sufficient".

Saturation can be avoided by carefully matching the full-scale range of an A/D converter to anticipated input signal amplitude ranges.

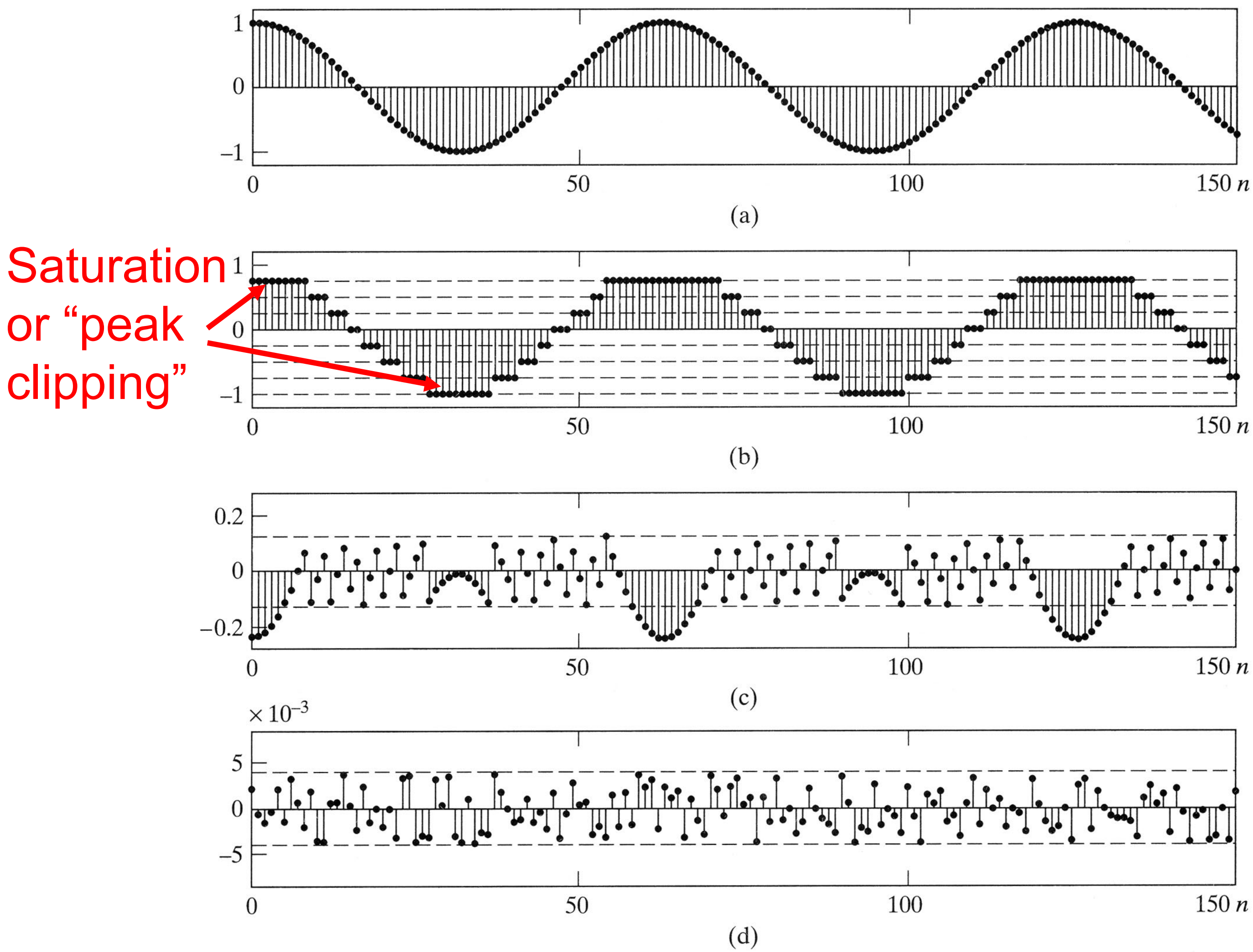

Figure 4.51 Example of quantization noise. (a) Unquantized samples of the signal $x[n] = 0.99\cos(n/10)$. (b) Quantized samples of the cosine waveform in part (a) with a 3-bit quantizer. (c) Quantization error sequence for 3-bit quantization of the signal in (a). (d) Quantization error sequence for 8-bit quantization of the signal in (a).

(Oppenheim and Schafer)

For rounding uniform quantizers, the amplitude of the quantization noise is in the range $-\Delta/2 < e[n] \leq \Delta/2$.

For small Δ it is reasonable to assume that $e[n]$ is a random variable uniformly distributed over $(-\Delta/2, \Delta/2]$.

For fairly complicated signals, it is reasonable to assume that successive quantization noise values are uncorrelated and that $e[n]$ is uncorrelated with $x[n]$.

Thus, $e[n]$ is assumed to be a uniformly distributed white-noise sequence with a mean of zero and variance:

$$\sigma_e^2 = \int_{-\Delta/2}^{\Delta/2} e^2 \frac{1}{\Delta} \mathrm{d}e = \frac{\Delta^2}{12}.$$

For a $(B+1)$-bit quantizer with full-scale X_m, the noise variance (or power) is:

$$\sigma_e^2 = \frac{2^{-2B} X_m^2}{12}.$$

The signal-to-noise ratio (SNR) of a $(B+1)$-bit quantizer is:

$$\begin{aligned}\mathsf{SNR} &= 10\log_{10}\left(\frac{\sigma_x^2}{\sigma_e^2}\right) = 10\log_{10}\left(\frac{12\cdot 2^{2B}\sigma_x^2}{X_m^2}\right)\\ &= 6.02B + 10.8 - 20\log_{10}\left(\frac{X_m}{\sigma_x}\right),\end{aligned}$$

where ${\sigma_x}^2$ is the signal variance (or power).

Analog signals such as speech and music typically have a Gaussian (or super-Gaussian) amplitude distribution, and consequently samples rarely exceed 3 or 4 times the standard deviation.

To avoid saturation (peak clipping) we might set $X_m = 4\sigma_x$, in which case:

$$\mathsf{SNR} \approx 6B - 1.25 \ \ \mathsf{dB}.$$

This uniform quantization scheme is called *pulse code modulation* (PCM).

Advantages:

- no coding delay
- not signal specific

Disadvantages:

- high bit rates

e.g. Wireless telephony requires 11 bits for "toll quality" $\equiv$ analog telephone quality.
If $f_s = 10{,}000 \text{ Hz} \Rightarrow \text{bit rate} = 110{,}000 \text{ bps}$, which may be impractical for wireless systems.

However, consider a CD player that uses 16-bit PCM $\Rightarrow \text{SNR} \approx 88.75 \text{ dB} \;\&\; \text{bit rate} \approx 320{,}000 \text{ bps}$, which is acceptable for wired applications.

Uniform quantization is suboptimal for many applications. Consider the probability density function (p.d.f.) of speech:

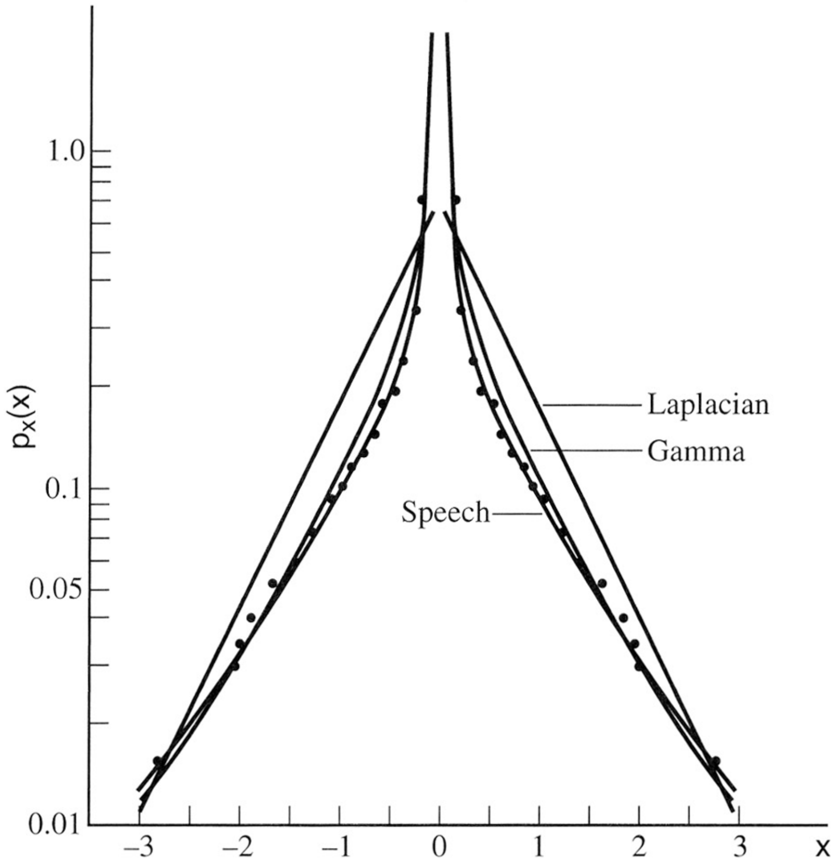

Figure 12.2 Comparison of histograms from real speech and gamma and Laplacian probability density fits to real speech. The densities are normalized to have mean $m_x = 0$ and variance $\sigma_x^2 = 1$. Dots (and the corresponding fitted curve) denote the histogram of the speech.

SOURCE: M.D. Paez and T.H. Glisson, "Minimum Mean-Squared Error Quantization in Speech" [66]. ©1972, IEEE. Used by permission.

(Quatieri)

If the p.d.f. of signals to be quantized is known, then an optimal *nonuniform* quantization scheme can be derived:

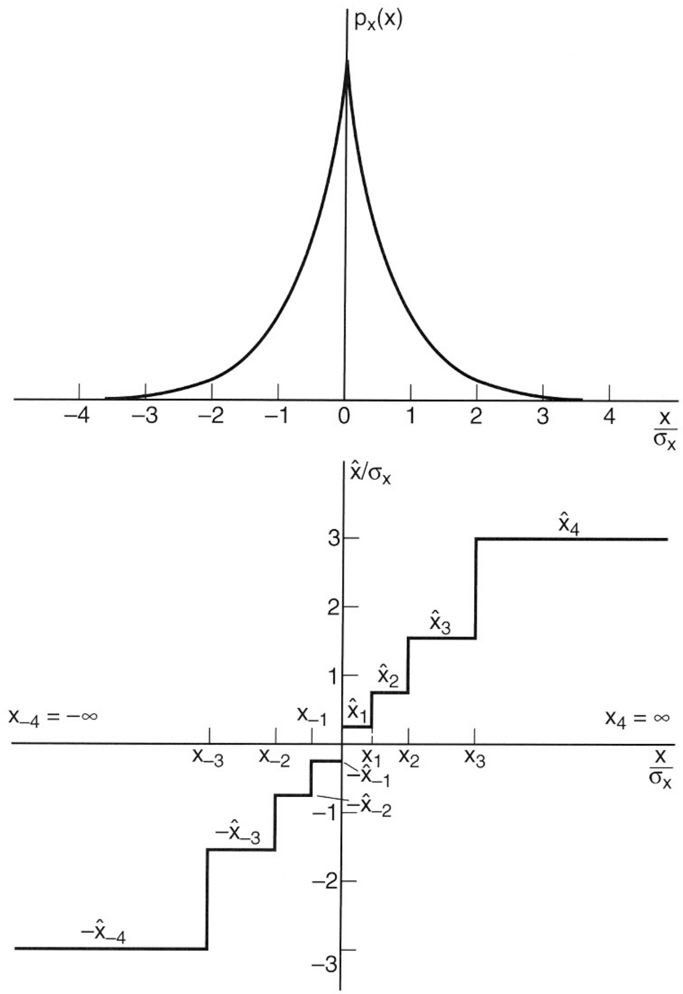

Figure 12.6 3-bit nonuniform quantizer: (a) Laplacian pdf; (b) decision and reconstruction levels.

SOURCE: L.R. Rabiner and R.W. Schafer, *Digital Processing of Speech Signals* [71]. ©1978, Pearson Education, Inc. Used by permission.

(Quatieri)

Sometime the p.d.f. of signals to be quantized is known (or can be assumed), but the signal variance (power) may vary over time. In this case, an adaptive *nonuniform* quantization scheme can be employed:

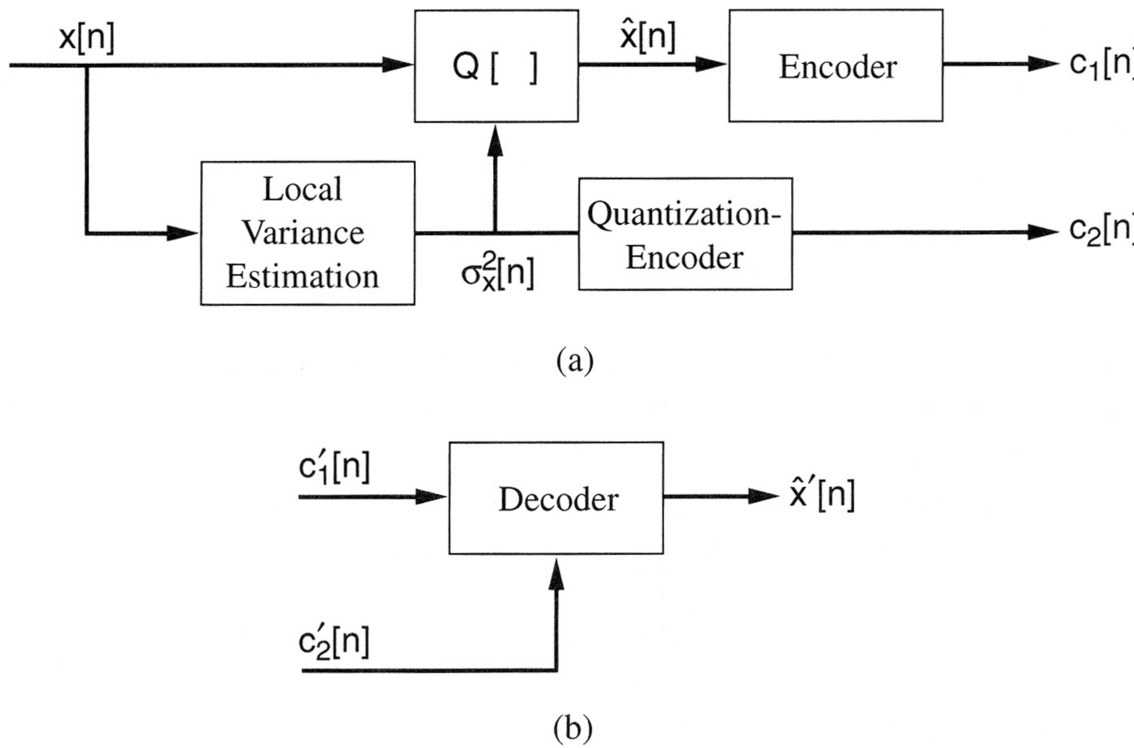

Figure 12.11 Adapting a nonuniform quantizer to a local pdf. By measuring the local variance $\sigma_x^2[n]$, we characterize the assumed Gaussian pdf. $c_1[n]$ and $c_2[n]$ are codewords for the quantized signal $\hat{x}[n]$ and time-varying variance $\hat{\sigma}^2[n]$, respectively. This feed-forward structure is one of a number of adaptive quantizers that exploit local variance.

(Quatieri)

An alternative to nonuniform quantization is *companding*, in which a nonlinearity is used to produce a new discrete-time signal that has a uniform distribution, and a uniform quantizer can consequently be applied:

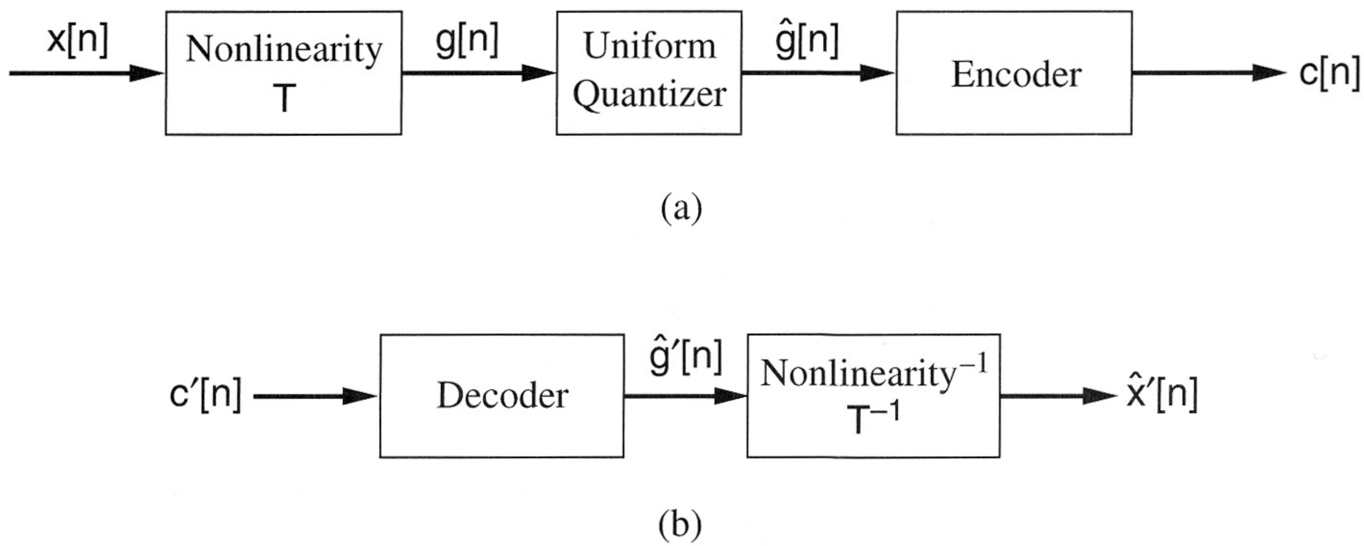

Figure 12.10 The method of companding in coding and decoding: (a) coding stage consisting of a nonlinearity followed by uniform quantization and encoding; (b) an inverse nonlinearity occurring after decoding.

(Quatieri)

Other nonlinearities approximate the companding operation, but are easier to implement and do not require a p.d.f. measurement.

One example is μ-law companding, ubiquitous in waveform coding, which uses the nonlinearity:

$$T\left(x\left[n\right]\right) = X_m \frac{\log\left(1+\mu\frac{|x[n]|}{X_m}\right)}{\log\left(1+\mu\right)} \operatorname{sign}\left(x\left[n\right]\right).$$

Together with uniform quantization, this nonlinearity (for large values of μ, e.g., 255) yields an SNR approximately independent of X_m and σ_x over a large range of signal input. For example, toll quality speech is obtainable with "μ-law PCM" utilizing μ-law companding followed by a 7-bit uniform quantizer, which would require 11-bit quantization without the companding operation.

Table 12.1 Comparison of 3-bit adaptive and nonadaptive quantization schemes [60]. Adaptive schemes use feed-forward adaptation.

SOURCE: Table from L.R. Rabiner and R.W. Schafer, *Digital Processing of Speech Signals* [71]. Data by Noll [60].

Nonuniform Quantizers	*Nonadaptive SNR (dB)*	*Adaptive* ($M = 128$) *SNR (dB)*	*Adaptive* ($M = 1024$) *SNR (dB)*
μ-law ($\mu = 100$, $x_{\max} = 8\sigma_x$)	9.5	–	–
Gaussian	7.3	15.0	12.1
Laplacian	9.9	13.3	12.8
Uniform Quantizers			
Gaussian	6.7	14.7	11.3
Laplacian	7.4	13.4	11.5

(Quatieri)