Second Edition

Principles of Digital and Analog Communications

Jerry D. Gibson

Department of Electrical Engineering
Texas A&M University

Macmillan Publishing Company
NEW YORK

Maxwell Macmillan Canada
TORONTO

Maxwell Macmillan International
NEW YORK OXFORD SINGAPORE SYDNEY

TABLE 9.2.1 Optimum Step Sizes for Uniform Quantization of a Gaussian PDF

Number of Levels (L)	Step Size (Δ_{opt})	Minimum Mean-Squared Error (D)	SQNR (dB)
4	0.9957	0.1188	9.25
8	0.5860	0.03744	14.27
16	0.3352	0.01154	19.38

Source: J. Max, "Quantizing for Minimum Distortion," *IRE Trans. Inf. Theory*, © 1960 IEEE.

9.3 Nonuniform Quantization

The most important nonuniform quantization method to date has been the logarithmic quantization used in the telephone network for speech digitization for over 20 years. The general idea behind this type of quantization is that for a fixed, uniform quantizer, an input signal with an amplitude less than full load will have a lower SQNR than a signal whose amplitude occupies the full dynamic range of the quantizer (but without overload). This fact is illustrated by Example 9.2.2. Such a variation in performance (SQNR) as a function of quantizer input signal amplitude is particularly detrimental for speech, since low-amplitude signals can be very important perceptually. There is the additional consideration for speech signals that low amplitudes are more probable than larger amplitudes, since speech is generally stated to have a gamma or Laplacian probability density, which is highly peaked about zero.

Therefore, for speech signals a type of nonlinear quantization was invented called *logarithmic companding*. Initially, this scheme was implemented by passing the analog speech signal through a characteristic of the form

$$F_\mu(s) = \frac{\ln[1 + \mu|s|]}{\ln[1 + \mu]} \operatorname{sgn}(s), \qquad -1 \le s \le 1, \tag{9.3.1}$$

where s is the normalized speech signal and μ is a parameter, usually selected to be $\mu = 100$, or more recently, $\mu = 255$. The function $F_\mu(s)$ is shown in Fig. 9.3.1. Notice that $F_\mu(s)$ tends to amplify small amplitudes more than larger amplitudes whenever $\mu > 0$. The output of $F_\mu(s)$ then served as input to a uniform, n-bit quantizer. To resynthesize the speech signal, the quantizer output $\hat{s}$ was passed through the inverse function of Eq. (9.3.1) given by

$$F_\mu^{-1}(\hat{s}) = \frac{1}{\mu}[(1 + \mu)^{|\hat{s}|} - 1] \operatorname{sgn}(\hat{s}), \tag{9.3.2}$$

where, of course, $-1 \le \hat{s} \le 1$.

The performance in SQNR of this system for $\mu = 100$ and $n = 7$ bits is shown in Fig. 9.3.2. It is evident from this figure that the SQNR is relatively flat over

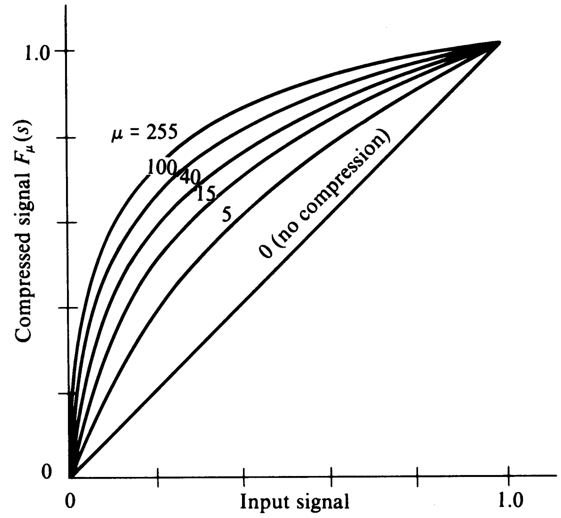

FIGURE 9.3.1 Logarithmic compression characteristics.

a wide dynamic range of input signal power (amplitudes), and hence low-amplitude signals are reproduced almost as well as higher-amplitude signals. In particular, we see from Eqs. (9.2.8) and (9.2.9) that for linear quantization, if we decrease the input signal power by 12 dB (from a peak value of $V/2$ to $V/8$), the output SQNR decreases by 12 dB. However, as we move along the "sine wave" curve in Fig. 9.3.2 from 0 to -12 dB on the input power axis, the SQNR decreases only by about 2 dB. The companding clearly improves the SQNR for low-amplitude signals.

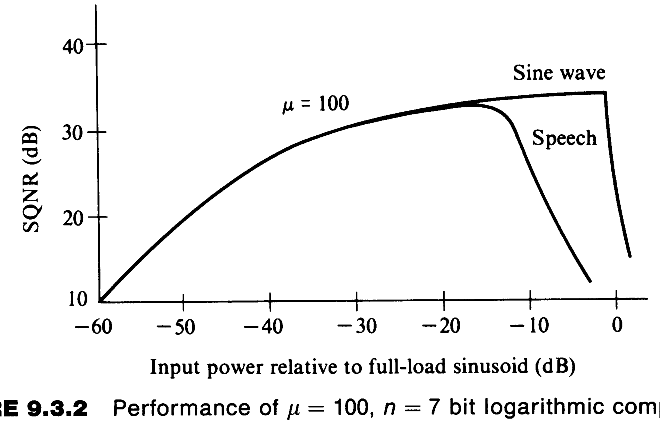

FIGURE 9.3.2 Performance of $\mu = 100$, $n = 7$ bit logarithmic companding.

9.6 Codecs and Channel Banks

At the present time, speech digitization is most often accomplished in the telephone network by devices called *codecs* (*co*ders/*dec*oders), which are contained in terminals designated as D-type *channel banks*. The codecs transform the analog speech into 8-bit PCM form, and the channel banks combine the PCM version of numerous voice channels into a single data stream using *time-division multiplexing* (TDM). In this section we briefly describe a few important details of the (U.S.) standard PCM codec and the various channel banks.

We begin by considering the codec used in the D2, D3, and D4 channel banks. This codec bandpass filters the analog voice signal to 200 to 3400 Hz and samples the filtered signal at a rate of 8000 samples/sec. Each of these samples is then quantized to 8-bit accuracy using a nonlinear quantizer based on the μ-law logarithmic characteristic discussed in Section 9.3. In particular, the quantizer is based on a 16-segment piecewise linear approximation to the $\mu = 255$ logarithmic companding characteristic in Eq. (9.3.1). There are eight positive and eight negative segments, but since the two segments around zero are collinear, it is often referred to as a 15-segment approximation. There are 16 equal quantization steps for each segment.

Table 9.6.1 specifies the nonlinear quantizer characteristic assuming that the maximum magnitude is scaled to 8159. The code for each quantization level is also shown in the table. For the 8-bit representation, the first bit is a polarity bit (1 denotes positive values, 0 represents negative values), the next 3 bits indicate the segment number, and the final 4 bits designate the particular step within a segment. Note from Table 9.6.1 that the segment codes and quantization steps for each segment are binary numbers that proceed from largest to smallest as the magnitude increases. Since lower amplitudes occur more often for speech than larger amplitudes, this tends to increase the density (number) of 1's. This is useful because the bipolar transmission line code described in Section 9.5 is employed with these codecs, and thus the number of pulses is increased. The higher density of pulses provides a strong timing component and aids synchronization.

In fact, another constraint is placed on the code assigned to quantizer output levels in Table 9.6.1. If an input sample falls within the most negative range of input amplitudes (quantization bin), the table indicates that the all 0's codeword would be transmitted. However, to guarantee a certain density of 1's, and hence bipolar pulses, the all 0's code is replaced by the codeword 00000010. Although it would cause less of an error to replace the all 0's codeword by 00000001, this is not done, for reasons that will be explained shortly.

Note now that we sampled the analog input voice signal (after filtering) at a rate of 8000 samples/sec, and thus if we pass each sample through the quantizer represented by Table 9.6.1, we get a bit rate of (8000 samples/sec) $\times$ (8 bits/sample) $=$ 64,000 bits/sec or 64 kbits/sec for each voice channel. Referring to Fig. 9.5.4, we see that for the bipolar code and $1/T = 64{,}000$, the required bandwidth is 64 kHz. Thus, by using PCM (the codec), we have expanded

TABLE 9.6.1 Quantizer Characteristic and Code Assignment for D2, D3, and D4 Channel Bank Codecs[a]

Input Amplitude Range:	Step Size:	Polarity Bit:	Quantization Segment Code:	Quantizer Step Code:	Output Value:
0–1	1	1	111	1111	0
1–3				1110	2
3–5	2	1	111	1101	4
⋮				⋮	⋮
29–31				0000	30
31–35				1111	33
⋮	4	1	110	⋮	⋮
91–95				0000	93
95–103				1111	99
⋮	8	1	101	⋮	⋮
215–223				0000	219
223–239				1111	231
⋮	16	1	100	⋮	⋮
463–479				0000	471
479–511				1111	495
⋮	32	1	011	⋮	⋮
959–991				0000	975
991–1055				1111	1023
⋮	64	1	010	⋮	⋮
1951–2015				0000	1983
2015–2143				1111	2079
⋮	128	1	001	⋮	⋮
3935–4063				0000	3999
4063–4319				1111	4191
⋮	256	1	000	⋮	⋮
7903–8159				0000	8031

[a] Positive inputs only; assumed symmetric about zero.

the bandwidth required by a single voice channel from approximately 4 kHz (3.4 kHz at the 3-dB point) to 64 kHz. We must be getting something in return for this extra bandwidth, and in the development in the remainder of this section and in Section 9.7, the primary advantages of PCM transmission are pointed out.

Before proceeding further, it is necessary to define clearly what is meant by time-division multiplexing and how it is used in the digital transmission of speech. In frequency-division multiplexing (FDM), we are given a specified band of frequencies, and we allocate nonoverlapping portions of this band to several different messages or channels. In TDM, we are given a specified time interval, and we allot nonoverlapping subintervals of this larger time slot to binary codewords generated by different codecs.

As a specific example of TDM, we mention what is called the T1 carrier system, which combines 24 PCM voice channels into a single data stream. Since each channel requires 8 bits, we thus have 192 bits for each set of 24 voice channels, usually called a *frame*. One bit is added to this total for synchronization purposes, so that 193 bits are transmitted per frame. Since each voice channel produces a new (8-bit) binary word 8000 times/sec, our frame rate is 8000 frames/sec and the transmitted bit rate for T1 carrier systems is 1.544 Mbits/sec.

Other than synchronization in the usual sense, the added "framing" bit has another purpose. When a telephone call is placed, there is a requirement for what is called *control signaling* information, such as "on-hook" or "off-hook" signals. For T1 carrier systems (with D2, D3, or D4 channel banks) this information is carried in the least significant bit position of each PCM word in every sixth frame. Therefore, in every sixth frame, the samples representing the 24 voice channels have only 7 bits of accuracy, while the intervening five frames use the full 8 bits. As a result, the PCM systems are sometimes said to employ $7\frac{5}{6}$-bit encoding. Since the least significant bit is sometimes "stolen" for signaling, this is why the most negative quantizer output level is encoded as 00000010 rather than 00000001. We would have gained nothing (in terms of a higher density of 1's) if the latter code appeared in the sixth frame.

The T1 systems constitute the first level (or lowest rate) in what is called the *digital hierarchy* in the telephone network. In recent years the various rates (or levels) in the digital hierarchy have been given DS designations, as shown in Table 9.6.2 and Fig. 9.6.1. From Fig. 9.6.1 we see that two DS1 links can be combined to form one DS1C link, four DS1 lines can be combined to form one DS2 link, or 28 DS1 links can be combined to form one DS3 link. Finally, six DS3 links can be combined to form one DS4 signal. Starting with the DS1 rate, which TDMs 24 voice channels in a frame, the method used to combine

TABLE 9.6.2 Data Rates and Line Codes in the Digital Hierarchy

Signal	Bit Rate (Mbits/sec)	Line Code
DS1	1.544	Bipolar
DS1C	3.152	Bipolar
DS2	6.312	B6ZS
DS3	44.736	B3ZS
DS4	274.176	Polar

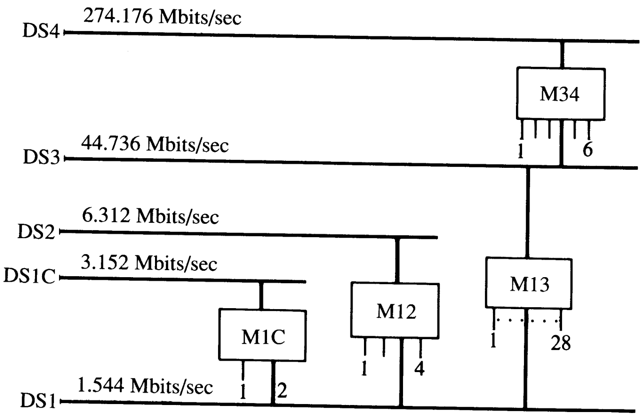

FIGURE 9.6.1 Digital hierarchy multiplexing plan.

all these signals is time-division multiplexing. Note also that in forming the next-highest level of the hierarchy, there seems to be some bits left over; that is, combining four DS1 signals should require 1.544 Mbits/sec $\times$ 4 = 6.176 Mbits/sec, but a DS2 line uses 6.312 Mbits/sec. The "extra" 0.136 Mbit/sec is used for synchronization and framing information, so that the multiplexing can be "undone."

It is also pointed out that as shown in Table 9.6.2, the line codes change as the data rate is increased, except in going from DS1 to DS1C signals. This is because the specifications on timing extraction, dc wander, and hardware change as the bit rate is increased. The line codes designated in Table 9.6.2 should be familiar to the reader from Section 9.5.

9.7 Repeaters

A principal advantage of employing time-division multiplexing of PCM signals to transmit analog waveforms such as speech is that PCM signals can be transmitted, in theory, over any distance without *any* degradation by noise simply by employing devices called *regenerative repeaters.* The concept that forms thebasis for repeaters can be explained as follows. Given a particular line code, only a finite number of fixed levels are allowable. For example, the bipolar line code has allowable levels of $+V$, $-V$, or 0 in each pulse interval, and no other voltage values are used to designate a 1 or 0. As these pulses proceed through a transmission medium (say, a pair of wires), they are smeared by deterministic distortion (amplitude and time delay) and they are subjected to additive noise as well as other random impairments. If the pair of wires is long enough, the pulses will eventually overlap so much and become so distorted that their original identity is irretrievably lost. If, however, we detect