

VOICE DIGITIZATION AND VOICE/DATA INTEGRATION

(Please read appropriate parts of Section 2.5.2 in book)

1. VOICE DIGITIZATION IN THE PSTN

The frequencies contained in telephone-quality analog speech waveforms extend up to a maximum of **less than 4KHz**. Instead of dealing with continuous-time analog waveforms, it is possible to represent (digitize) speech as a **sequence of binary digits**, making it appear to be “data”. This has several advantages. It allows integration of voice and data services. Furthermore, binary data can be switched and transmitted more reliably and conveniently within the network.

A key result enabling digitization of waveforms is the following:

The Nyquist Sampling Theorem

“Any waveform with a **maximum frequency** content of less than **W Hz** can be **reconstructed from its samples taken at a rate of a minimum of 2W samples/sec.**”

The reconstruction from samples, or *interpolation*, is effected by forming a sequence of narrow pulses, at the sampling rate, with the sampled amplitudes, and passing the pulse train through a low pass filter cutting off all frequencies beyond W Hz.

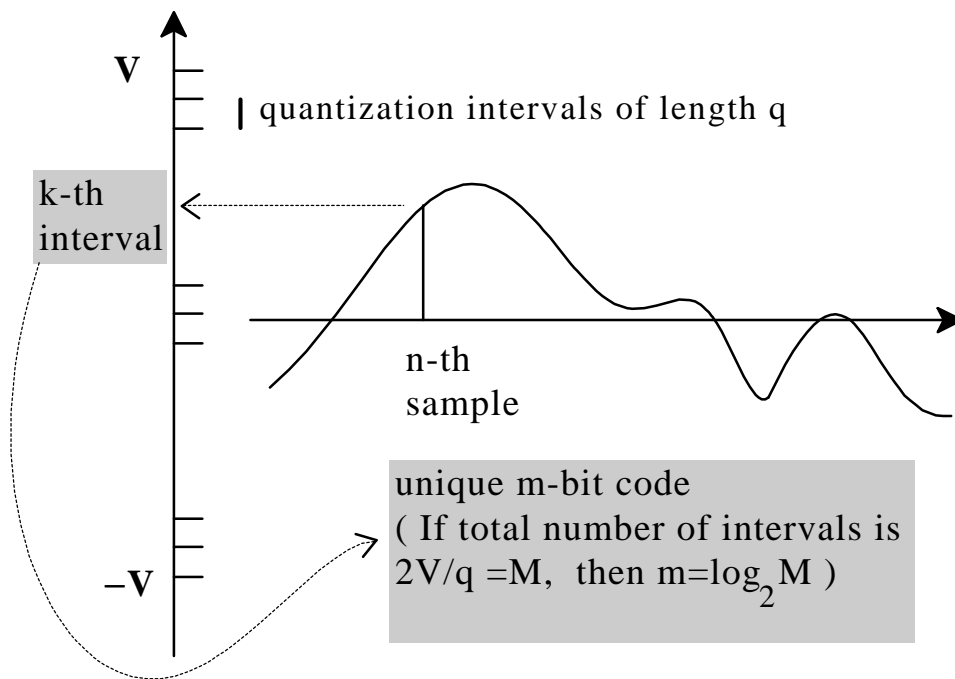
(Do not confuse this with Nyquist's signaling criterion which states that pulses can be sent over a bandwidth W channel at a maximum rate of 2W pulses per second if we are to recover exact amplitudes of individual sent pulses from received pulses. While this is a distinct result, the underlying mathematical basis is the same, and involves the “sinc” function. It is used as the pulse shape for the signaling criterion and as the interpolating function (provided by the low pass filter) in the sampling result.)

(visit <http://www2.ece.jhu.edu/wjr/index.html> for an interactive sampling exercise).

2. PULSE CODE MODULATION (PCM) IN THE PSTN – 64 KBPS DIGITIZED VOICE

- The digitization standard in the PSTN uses a **sampling rate of 8 KHz**. This is arrived at by application of the sampling theorem.

Let V (volts) be the max. amplitude the input speech is constrained to have. Each speech sample falls in one of M intervals spanning the amplitude range of $[-V, V]$. The sample is approximated as having a value at the mid-point of this interval. In the PSTN, $M=256$ and V is of the order of 1 volt. This allows sufficiently fine **quantization** of the input amplitude range.



- Each quantized sample is **encoded** as an 8-bit word ($8=\log_2 256$). This results in a bit rate of $8 \times 8 = 64$ Kbps. This is called **pulse code modulation**. (Do not confuse this with modulation for frequency shifting, as done earlier.)
- For best results, the quantization of the amplitude range $[-V, V]$ should **not** be done uniformly (equal-length intervals). *Low* amplitudes are *more* likely in ordinary speech signals, and a uniform approximation error over the full amplitude range from uniform quantization also results **in larger relative error for lower amplitudes**.

- In practice the quantization is **non-uniform**, with smaller quantization intervals at lower amplitudes. This is implemented by compression of the amplitude before uniform quantization (and subsequent expansion after reconstruction at the receiver).
- For North American systems, the μ -law compressor characteristic is used; it is the function $F(x) = \frac{\ln(1 + \mu|x|)}{\ln(1 + \mu)}$ $\text{sign}(x)$ with $\mu=255$. Actually, a 15-segment linear approximation to this characteristic is implemented. For European systems a slightly different (A-law) compressor is used.

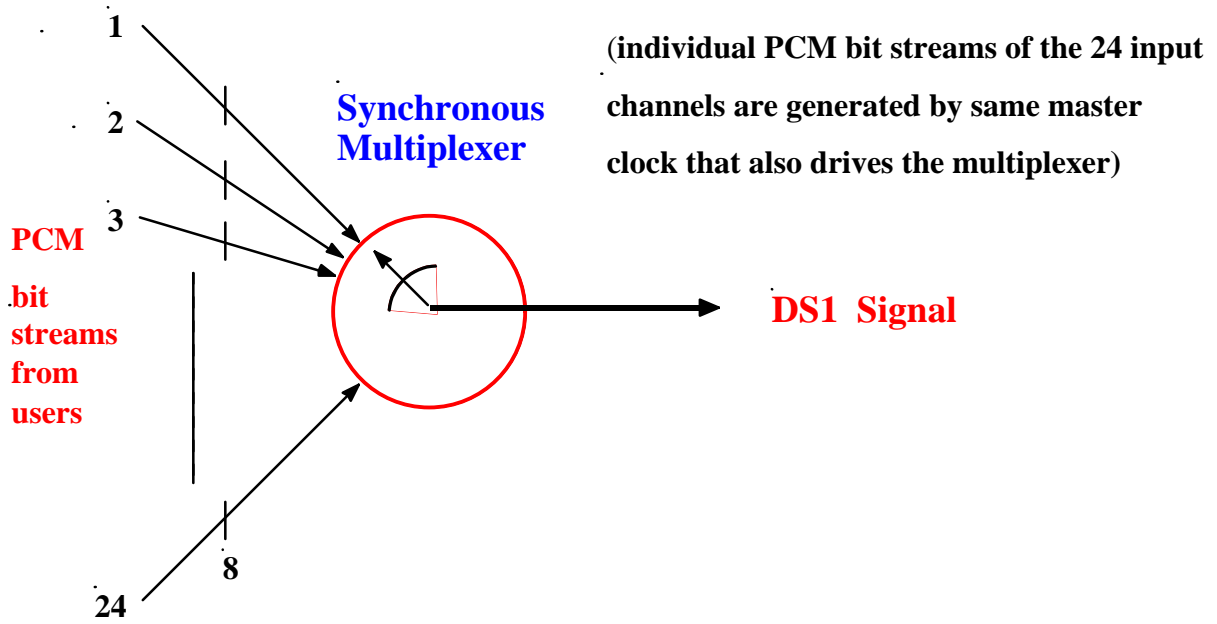
It is possible to use more sophisticated processing of the speech waveform and obtain lower bit-rate digital representations of the speech. For example, using **differential PCM (DPCM)**, we can get almost as good speech quality at 32 Kbps or less. A further modification of this is **adaptive DPCM** for lower bit rates.

3. TIME DIVISION MULTIPLEXING OF DIGITIZED SPEECH CHANNELS IN THE PSTN

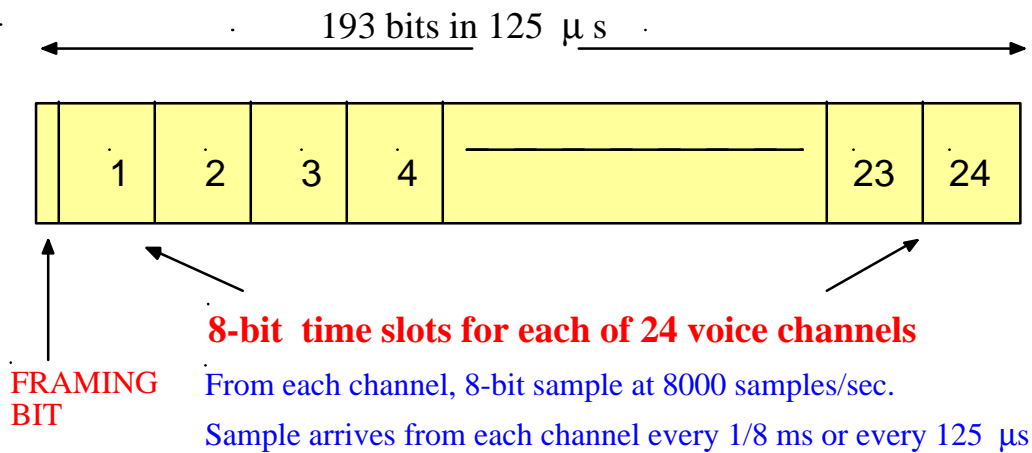
For transmission over a common medium (single high bandwidth line, coaxial cable, microwave line-of-sight link, etc.) we need to be able to multiplex a number of individual bit streams into a single higher rate stream. This is done in a hierarchy of stages, using **time division multiplexing (TDM)** to rates in the hundreds of Mbps over single links in the network.

(An older scheme for multiplexing a number of analog channels (each 4 KHz wide) used frequency translation (modulation) to separate individual channels on a high-bandwidth medium. The modulation employed was single sideband (SSB) for bandwidth conservation.)

In the U.S. and Japanese TDM standard hierarchy, the first stages of time-division multiplexing in the telephone system generate composite bit streams for groups of 24 individual digitized voice channels. This multiplexing results in a **frame** structure consisting of 193 bits in each 125 μs time interval, as explained below. The European standard is similar but different in detail.



24-Channel PCM Multiplexing and Frame Structure



Total bit rate of multiplexed channels is $8(24)+1=193$ bits per 125 μ s which is 1.544 Mbps. This is the DS1 signal carried on what are known as T1 transmission lines.

A **superframe** consists of a **sequence of 12 frames**. The first bit (S-bit) of each frame forms a 12 bit sub-stream which always follows the pattern 1-0-0-0-1-1-0-1-1-1-0-0. This allows the demultiplexer to obtain frame alignment (by looking for alternating 1's and 0's every 2 frames or $193 \times 2 = 386$ bits), and also superframe alignment once frame alignment is obtained, by looking for the pattern 0-0-1-1-1-0 in the S-bit of every other frame. Superframe alignment is necessary for the following reason: in each 6th and 12th frame of a superframe, the 8-th bit (least significant bit) of the sampled word is discarded and this bit position is used to carry signaling and control information for the system. This provides one bit every superframe in each of two signaling channels (6th and 12th frame last bits), for a rate of $8000/12 = 667$ bps per signaling channel.

Higher Order Multiplexing

The DS1 (T1) signal at 1.544 Mbps can then be multiplexed with other DS1 signals in a second stage. For example, the DS2 signal is derived by multiplexing 4 DS1 lines. The output rate is **not** exactly $4 \times 1.544 = 6.176$ Mbps; it is actually 136Kbps higher, or 6.312 Mbps. This is because of overhead (framing bits) and also to equalize timing variations. Since the 4 streams may have been generated by clocks with slightly different speeds (for DS1, the allowed variation is 50 parts per million or 77 bps), we have to form the output stream at a slightly higher rate to avoid dropping bits. The output stream may therefore contain extra "stuff" bits (dummy bits) in designated positions in the bit stream (an indication has to be provided as to the presence or otherwise of any stuffed bit), to make up the slightly faster rate for the DS2 signal.

This process is called plesiochronous (almost synchronous) multiplexing. The difficulty with this scheme is that it is not easy at any point in its transmission path to access any particular voice channel without going through the entire demultiplexing process. *The synchronous digital hierarchy (SDH)* is a newer scheme that uses synchronous multiplexing stages, using a master clock signal.

A note on the DS1 signal: The DS1 structure allows *data* to be transmitted instead of digitized voice, with one change. In view of the "robbed" bits in the 6th and 12th frames of each superframe, there would be a loss of data bits if data transmission was attempted at 64 Kbps on each channel. Instead,

we may drop the rate to 56 Kbps (use a dummy bit 8 in each byte). Notice that for voice, losing the least significant bit one out of every 6 samples is hardly perceptible. Finally, note that it is possible to lease a dedicated T1 line for data transmission at a higher rate.

INTEGRATED SERVICES DIGITAL NETWORK (ISDN)

The electrical waveform representing voice in the analog PSTN is quite different in character from a pulse stream representing digital data. This is why we have to use special modems to allow data transmission over a voice circuit.

There is a growing trend to integrate voice and data traffic in a single digital network. The *Integrated Services Digital Network (ISDN)* is an example of this integration. ISDN is a service offered by the telephone company which is fundamentally different from the older analog, 3 KHz wide, telephone service. ISDN in its basic form allows two 64 Kbps user channels as well as a 16 Kbps channel for control signaling (call set-up, etc.). Thus the subscriber is able to use a basic 128 Kbps data channel. Actually, the total rate for the multiplexed subscriber connection is

$$2 \times 64 + 16 + \text{overhead and framing bits} = 192 \text{ Kbps.}$$

A twisted-pair physical telephone line is used, it is leased by the subscriber for a monthly fee, and it terminates in the local telephone office plant in separate circuits for the ISDN network. Note that the telephone line itself is capable of carrying the high data rate.

(The physical interface standard between a DTE and the ISDN network termination equipment (DCE) specifies an 8-pin RJ45 connector. A balanced two-wire connection is provided for the data stream in each direction between the terminal equipment and the network equipment. The electrical signaling is in the form of baseband pulses.)

With an ISDN connection a user may send data and/or voice over the network to another ISDN terminal. More generally, ISDN allows circuit-switched service, packet-switched service, or dedicated connection. Each 64 Kbps channel may carry data, digitized voice, or a combination of lower rate data and digitized voice destined for the same endpoint.

ISDN is one example of the move towards digitization of voice to make it look more like data. This is the opposite of what a voiceband modem does; making a data stream look more like a waveform that the analog telephone link can accept. Even without ISDN, voice signals in the PSTN are converted to digital form (PCM) whenever the signal travels outside the local exchange area, and all the switching is done with digital switches. The 64 Kbps rate for voice digitization in the ISDN was actually influenced by this existing standard for digitized speech within the PSTN.