
SECTION X

MIXED SIGNAL PROCESSING APPLICATIONS

MIXED SIGNAL PROCESSING APPLICATIONS

■ HIGH PERFORMANCE MODEMS

V.32 Modem Overview

V.32 Modem Transmitter

V.32 Modem Receiver

I/O Ports and Codecs for V.32 Modems

■ DIGITAL MOBILE RADIO

Overview

The GSM System

Speech Codec

Discontinuous Transmission (DTX)

GSM System Upconversion and Downconversion

■ DIGITAL AUDIO STUDIO RECORDING

■ COMPACT DISC (CD) PLAYER ELECTRONICS

SECTION X

MIXED SIGNAL PROCESSING APPLICATIONS

HIGH PERFORMANCE MODEMS

Modems (*Modulator/Demodulator*) are widely used to transmit and receive digital data using analog modulation over the General Switched Telephone Network (GSTN) as well as private lines. Although the data to be transmitted is digital, the telephone channel is designed to carry voice signals having a bandwidth of approximately 300 to 3300Hz. The telephone transmission channel suffers from delay distortion, noise, crosstalk, near-end and far-end echoes, and other imperfections listed in Figure 10.1. While certain levels of these signal degradations are perfectly acceptable for voice communication,

they can cause high error rates in digital data transmission. The fundamental purpose of the transmitter portion of the modem is to prepare the digital data for transmission over the analog voice line. The purpose of the receiver portion of the modem is to receive the signal which contains the analog representation of the data, and reconstruct the original digital data at an acceptable error rate. High performance modems make use of digital techniques to perform such functions as modulation, demodulation, error detection and correction, equalization, and echo cancellation.

IMPERFECTIONS IN THE TELEPHONE CHANNEL

- Attenuation
- Bandwidth Flatness
- Harmonic Distortion
- Echoes (Near-End and Far-End)
- Phase Jitter
- Phase Distortion, Group Delay Variation
- Noise
- Impedance Mismatches
- Frequency Offset
- Phase and Gain Hits

Figure 10.1

A block diagram of a telephone channel is shown in Figure 10.2. Most voiceband telephone connections involve several connections through the telephone network. The 2-

wire subscriber line available at most sites is generally converted to a 4-wire signal at the telephone central office. The signal is converted back to a 2-wire signal at the far-end

TELEPHONE CHANNEL BLOCK DIAGRAM

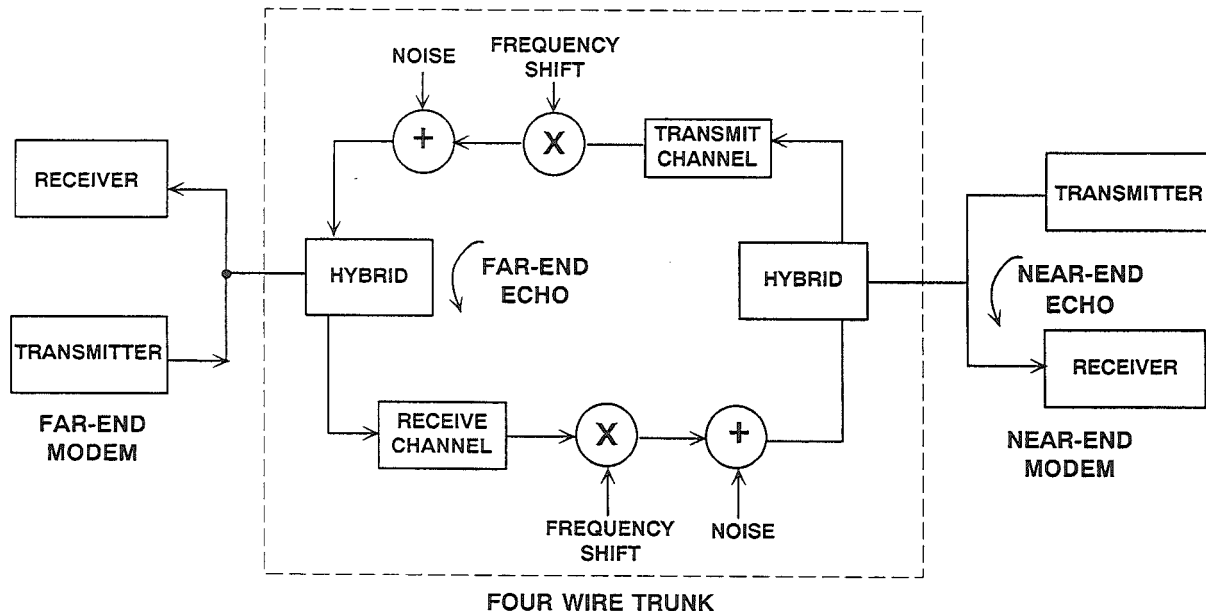


Figure 10.2

subscriber line. The 2- to 4-wire interface is implemented with a circuit called a *hybrid*. The hybrid intentionally inserts impedance mismatches to prevent oscillations on the 4-wire trunk line. The mismatch forces a portion of the transmitted signal to be reflected or echoed back to the transmitter. This echo can corrupt data the transmitter receives from the far-end modem.

Half-duplex modems are capable of passing signals in either direction on a 2-wire line, but not simultaneously. *Full-duplex* modems operate on a 2-wire line and can transmit and receive data simultaneously. Full-duplex operation requires the ability to separate a receive signal from the reflection (echo) of the transmitted signal. This is accomplished by assigning the signals in the two directions different frequency bands separated by filtering, or by echo cancelling in which a locally synthesized replica of the reflected transmitted signal is subtracted from the composite receive signal.

There are two types of echo in a typical voiceband telephone connection. The first echo is the reflection from the near-end hybrid, and the second echo is from the far-end hybrid.

In long distance telephone transmissions, the transmitted signal is heterodyned to and from a carrier frequency. Since local oscillators in the network are not exactly matched, the carrier frequency of the far-end echo may be offset from the frequency of the transmitted carrier signal. In modern applications this shift can affect the degree to which the echo signal can be cancelled. It is therefore desirable for the echo canceller to compensate for this frequency offset.

For transmission over the telephone voice network, the digital signal is modulated onto an audio sinewave carrier, producing a modulated tone signal. The frequency of the carrier is chosen to be well within the telephone band. The transmitting modem modulates the audio carrier with the trans-

mit data signal, and the receiving modem demodulates the tone to recover the receive data signal.

The baseband data signal may be used to modulate the amplitude, the frequency, or the phase of the audio carrier, depending on the data rate required. These three types of modulation are known as amplitude shift keying (ASK), frequency shift keying (FSK), and phase shift keying (PSK). In its simplest form the modulated carrier takes on one of two states - that is, one of two amplitudes, one of two frequencies, or one of two phases. The two states represent a logic 0 or a logic 1.

Low- to medium-speed data links usually use FSK up to 1200 bits/s. Multiphase PSK

are used for 2400 bits/s and 4800 bits/s links. PSK utilizes bandwidth more efficiently than FSK but is more costly to implement. ASK is least efficient and is used only for very low speed links (less than 100 bits/s) For 9600 bits/s, a combination of PSK and ASK is used, known as Quadrature Amplitude Modulation (QAM).

Assuming 7-bit ASCII and 4 bits/character overhead (start, parity, and two stop bits), a data transmission rate of 300 bits/s translates to approximately 27 characters/s. This is faster than a person can type but is too slow for transferring large files or for many applications requiring graphics.

MODULATION METHODS FOR MODEMS

- Amplitude Shift Keying (ASK): Up to 100 bits/s
- Frequency Shift Keying (FSK): Up to 1200 bits/s
- Phase Shift Keying (PSK): Up to 4800 bits/s
- Quadrature Amplitude Modulation (QAM): Up to 9600 bits/s

Figure 10.3

The International Telegraph and Telephone Consultative Committee (CCITT in French) has established standards and

recommendations for modems which are given in Figure 10.4.

CCITT RECOMMENDATIONS FOR TELEPHONE MODEMS

CCITT REC.	DATE	SPEED (BITS/s)	HALF DUPLEX/ FULL DUPLEX/ ECHO CANCEL	GSTN/ PRIVATE	MODULATION METHOD
V.21	1964	200	FDX	GSTN	FSK
V.22	1980	1200	FDX	GSTN	PSK
V.22bis	1984	2400	FDX	GSTN	16QAM
V.23	1964	1200	HDX	GSTN	FSK
V.26	1968	2400	HDX	Private	PSK
V.26 bis	1972	2400	HDX	GSTN	PSK
V.26 ter	1984	2400	FDX(EC)	GSTN	PSK
V.27 bis	1976	4800	HDX	Private	8PSK
V.27 ter	1976	4800	HDX	GSTN	8PSK
V.29	1976	9600	HDX	Private	16AM/PM
V.32	1984	9600	FDX(EC)	GSTN	32QAM
V.33	-	14400	HDX	Private	64QAM

Figure 10.4

V.32 MODEM OVERVIEW

The goal in designing high performance modems is to achieve the highest data transfer rate possible over the GSTN and avoid the expense of using dedicated conditioned private telephone lines. The V.32 recommendation describes a full-duplex (simultaneous transmission and reception) synchronous modem that operates on the Public Switched Telephone Network (GSTN). The V.32 modem communicates at a rate of 9600 bits/s utilizing quadrature amplitude modulation (QAM). Four-bit symbols (bauds) modulate a

carrier frequency of 1800Hz with a modulation rate of 2400 bauds/s. The modulation of 4-bit symbols at a rate of 2400 symbols/s yields the 9600 bits/s specification. These 4-bit symbols are transmitted using 32-state trellis-encoded QAM. The trellis encoding provides an extra bit per symbol for forward error correction. This additional bit dramatically increases the noise performance of the modem. Characteristics of the V. 32 modem are summarized in Figure 10.5

V.32 MODEM CHARACTERISTICS

- 9600 bits/second Bit Rate on GSTN
- 1800Hz Carrier Frequency (Transmit and Receive)
- 4 Bits/Symbol, 2400Hz Symbol Rate
- 32-QAM, Trellis Coded, 4 Bit Data + Redundancy Bit
- Transmit/Receive Isolation Using Echo Cancellation
- Extensive Use of DSP Techniques

Figure 10.5

A simplified block diagram for a V.32 modem is shown in Figure 10.6. The diagram shows that the bulk of the signal processing is done digitally. Both the trans-

mit and receive portions of the modem subject the digital signals to a number of DSP algorithms which can be efficiently run on modern processors.

V.32 MODEM BLOCK DIAGRAM

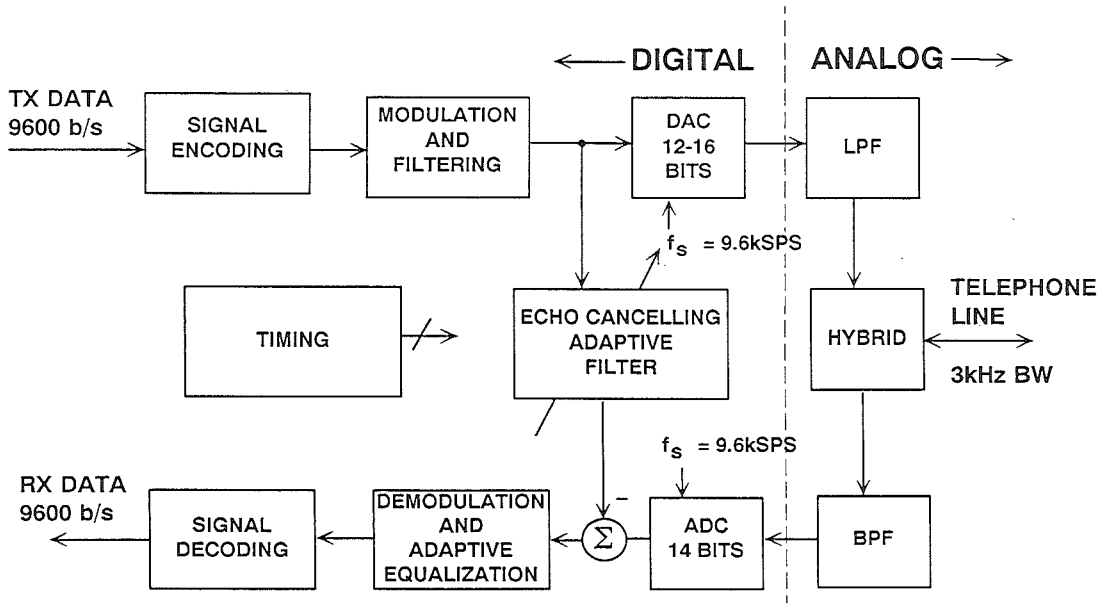


Figure 10.6

V.32 MODEM TRANSMITTER

A block diagram of the V.32 transmitter is shown in Figure 10.7. The input serial bit stream is first scrambled. Scrambling takes the input bit stream and produces a pseudo-random sequence. The purpose of the scrambler is to whiten the spectrum of the transmitted data. Without the scrambler, a long series of identical symbols could cause the receiver to lose carrier lock. Scrambling makes the transmitted spectrum resemble white noise, to utilize the bandwidth of the channel more efficiently, makes carrier recovery and timing synchronization easy, and makes adaptive equalization and echo cancellation possible.

The scrambled bit stream is divided into groups of four bits. The first two bits of each 4-bit group are first differentially encoded and then convolutionally encoded. This produces a 5-bit trellis-coded symbol in which the extra bit is a redundantly coded bit.

The 5-bit symbols are then mapped into the signal space using trellis-coding as defined in the V.32 recommendation. The signal space mapping produces two coordinates, one for the real part of the QAM modulator and one for the imaginary part. A diagram of the resulting V.32 signal constellation is shown in Figure 10.8.

Used prior to modulation, the digital pulse shaping filters attenuate frequencies above the Nyquist frequency that are generated in the signal mapping process. These filters are designed to have zero crossings at the appropriate frequencies to cancel intersymbol interference. The pulse shape filter is based on the impulse response of a raised cosine function as shown in Figure 10.9. The value T is equal to the reciprocal of the symbol rate (2400 symbols/second). For a sampling rate of 9600Hz and a symbol rate of 2400Hz, a 17-tap FIR filter can be used.

V.32 MODEM TRANSMITTER

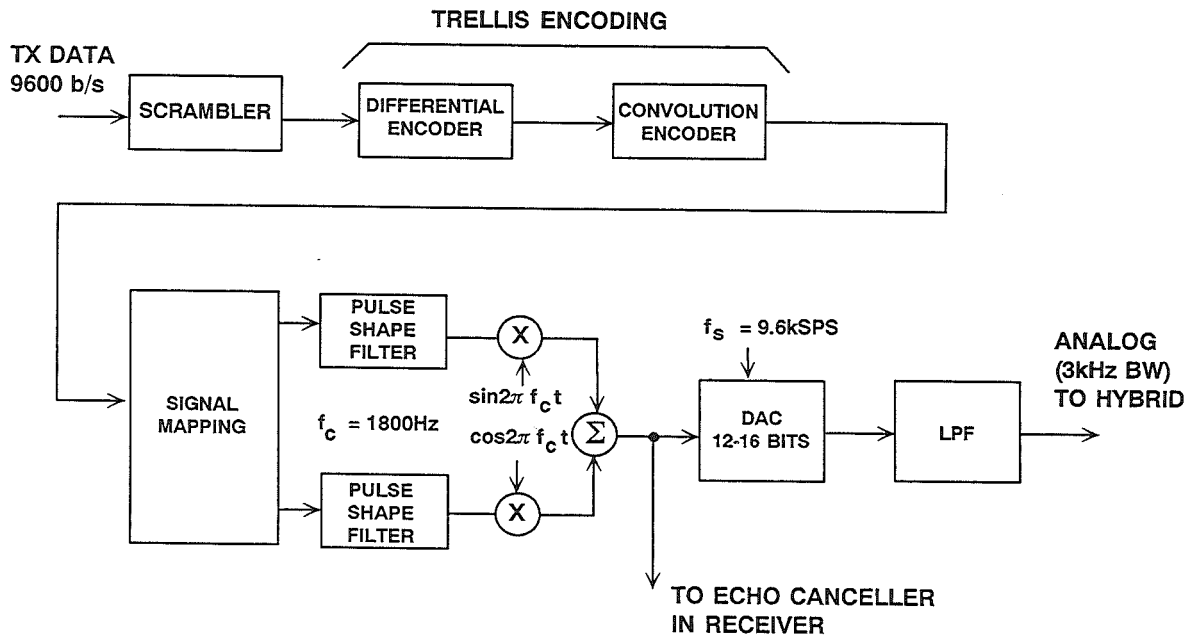


Figure 10.7

V.32 MODEM SIGNAL CONSTELLATION

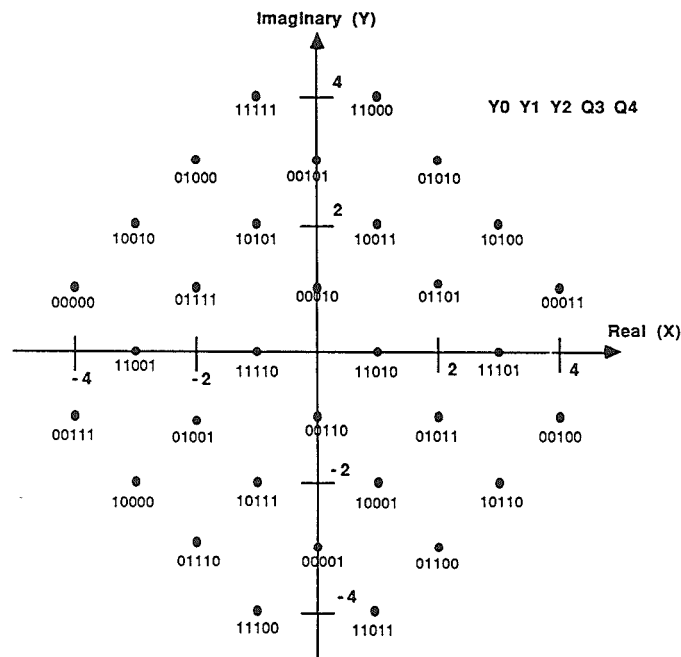


Figure 10.8

The modulation for the V.32 coding scheme is quadrature amplitude modulation (QAM). Modulation is easily implemented in modern DSP processors. The process of modulation requires the access of a sine or

cosine value, the access of an input symbol (x or y coordinate) and a multiplication. The parallel architecture of the ADSP-2101 permits all three operations to be performed in a single 80ns cycle.

The output of the digital QAM modulator drives a 12- to 16-bit DAC which is updated at 9.6kSPS . The output of the DAC is

passed through a 3.5kHz analog lowpass filter and to the 2-wire telephone line for transmission over the GSTN.

PULSE SHAPING FILTER IMPULSE RESPONSE

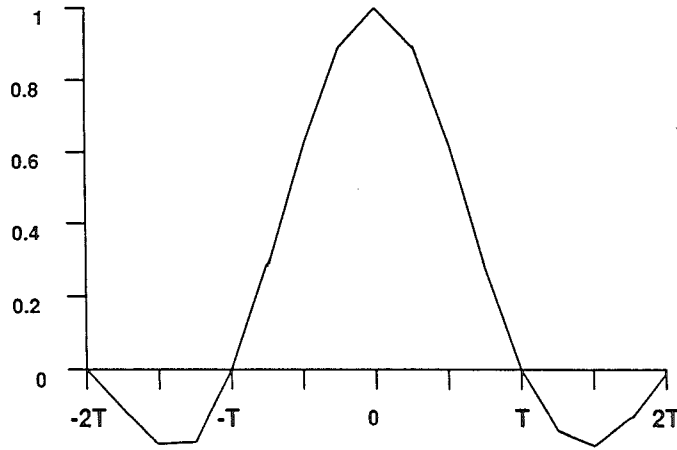


Figure 10.9

V.32 MODEM RECEIVER

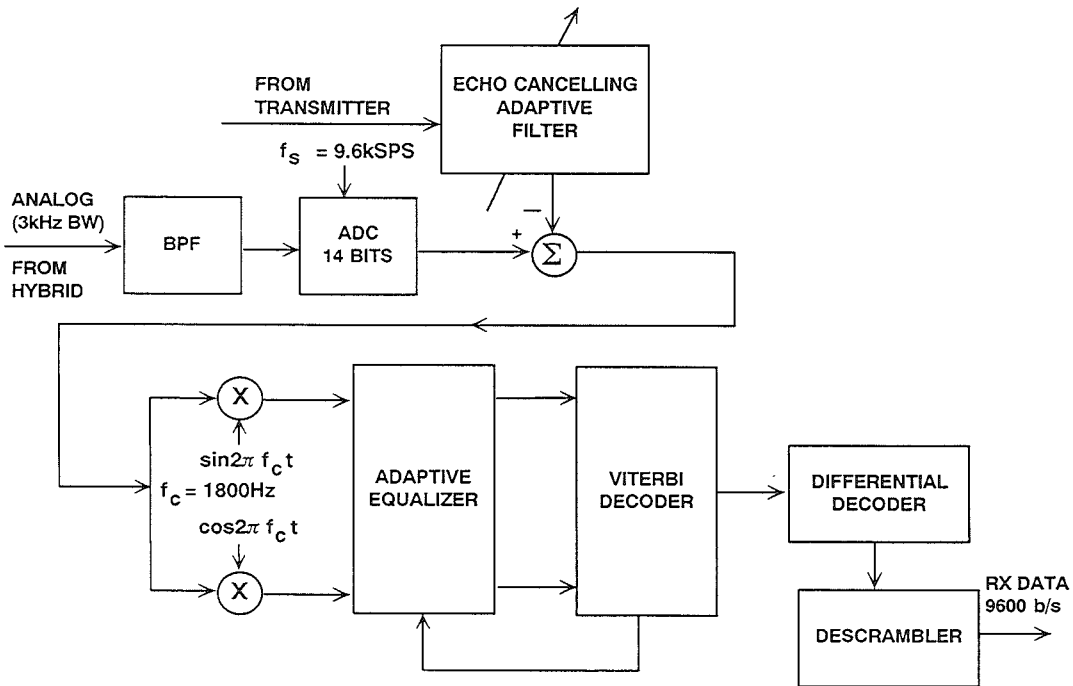


Figure 10.10

V.32 MODEM RECEIVER

A block diagram of the V.32 modem receiver is shown in Figure 10.10. The receiver is made up of several functional blocks: the input antialiasing filter and ADC, a demodulator, an adaptive equalizer, a Viterbi decoder, an echo canceller, a differential decoder, and a descrambler. The receiver DSP algorithms are both memory-intensive and computation-intensive. The ADSP-2101 addresses both needs, providing 2K of program memory RAM (for both code and data) on chip, 1K of data memory RAM on chip and an instruction execution rate of 12.5MIPS.

The antialiasing filter and ADC in the receiver need to have a dynamic range from the largest echo signal to the smallest. The received signal can be as low as -40dBm, while the near-end echo can be as high as -6dBm. In order to insure that the analog front end of the receiver does not contribute any significant impairment to the channel under these conditions, an instantaneous dynamic range of 84dB (14 bits) and an SNR of 72dB is required.

In order to compensate for amplitude and phase distortion in the telephone channel, equalization is required to recover the transmitted data at an acceptably low bit error rate. In order to respond to rapidly changing conditions on the telephone line, adaptive equalization is required for the V.32 modem receiver. An adaptive equalizer can be implemented digitally in an FIR filter whose coefficients are continuously updated based on current line conditions. A 64-tap fractionally spaced equalizer provides the performance necessary for V.32 applications.

Separation between the transmit and receive signal in the V.32 modem is accomplished using echo cancellation. Echo cancellation is mandatory since both the calling and the answering modem use the same carrier frequency of 1800Hz. Both near-end and far-end echo must be cancelled in order to yield reliable communication. Echo cancellation is achieved by subtracting an estimate of the echo return signal from the actual received signal. The predicted echo is

AD7869 14-BIT, 83kSPS I/O PORT BLOCK DIAGRAM

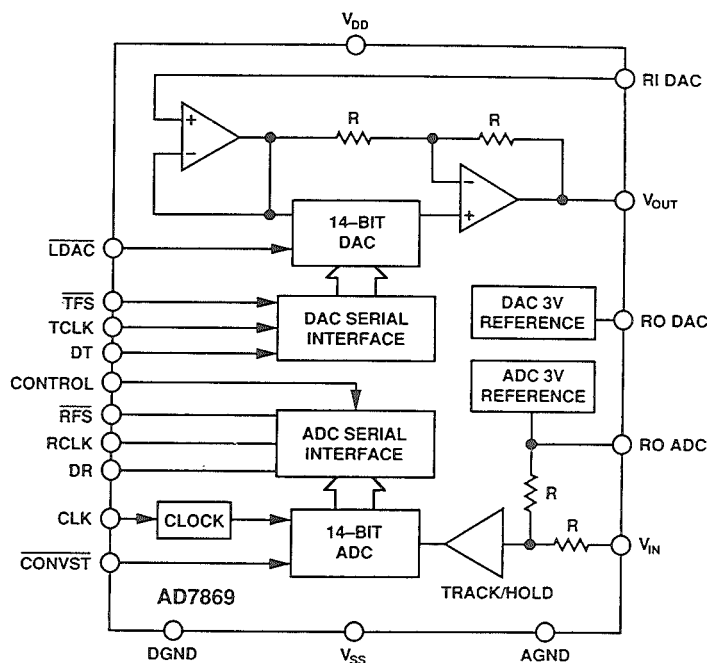


Figure 10.11

determined by feeding the transmitted signal into an adaptive filter with a transfer function that approximates the telephone channel. The adaptive filter commonly used in echo cancellers is the FIR filter (chosen for its stability and linear phase response), where the taps are determined using the least-mean-square (LMS) algorithm during a training sequence executed prior to full-duplex communications. The echo canceller must be able to cancel 16ms of echo. At 9600 samples/second, a 154-tap FIR filter is required to cancel the echo. Assuming that the canceller and frequency shifter have converged during the training period, about 200 cycles are required to cancel an echo in a V.32 modem.

The most common technique for decoding the received data is Viterbi decoding. Named after its inventor, the Viterbi algorithm is a general-purpose technique for making an

error-corrected decision. Viterbi decoding provides a certain degree of error correction by examining the received bit pattern over time to deduce the value that was the most likely to have been transmitted at a particular time. Viterbi decoding is computation-intensive. A history for each of the possible symbols sent at each symbol interval has to be maintained. For the V.32 modem, the symbol history spans 20 symbol intervals. At each symbol interval, the length of the path backward in time from each possible received symbol to a symbol sent some time ago is calculated. After 20 symbol intervals, the symbol that has the shortest path back to the original signal is chosen to be the current decoded symbol. A complete description of Viterbi decoding and its implementation on the ADSP-2100 family of DSP processors is given in Reference 2.

V.32 ANALOG FRONT END

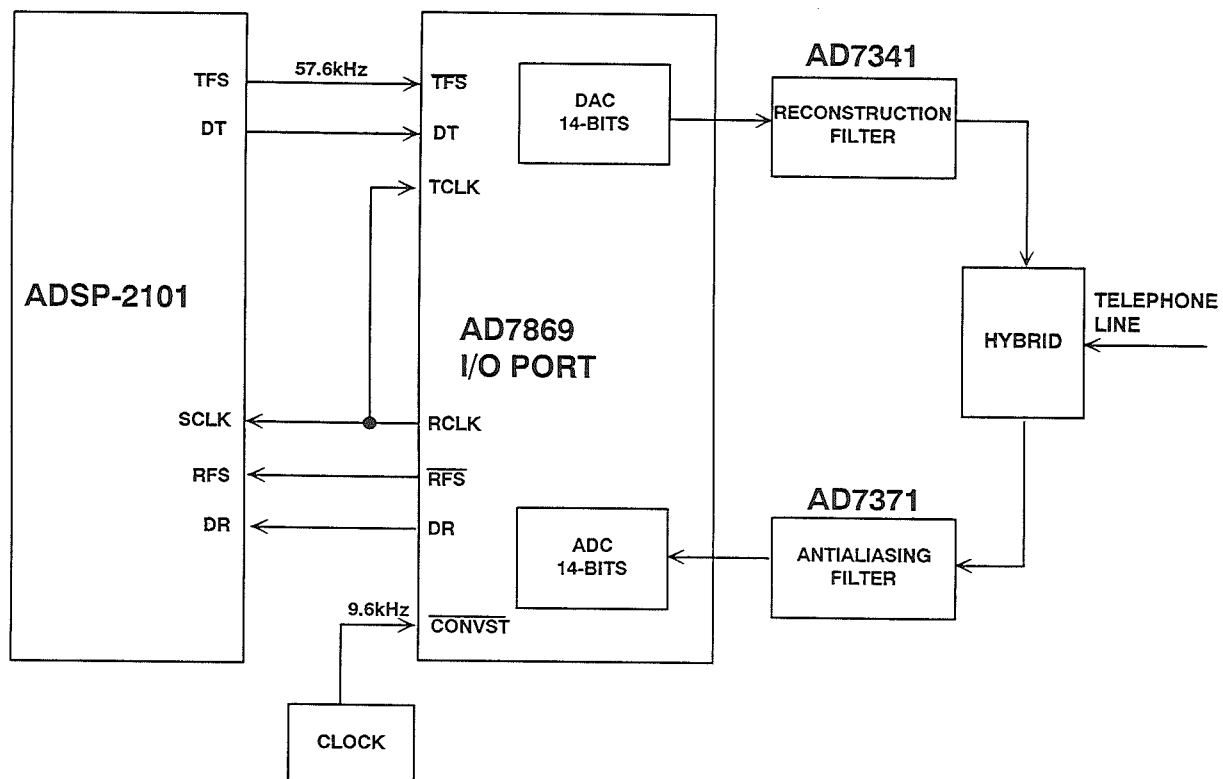


Figure 10.12

I/O PORTS AND CODECS FOR V.32 MODEMS

The AD7869 is a complete 14-bit, 83kSPS I/O Port with a zero-chip serial interface to most DSP processors such as the ADSP-2101, TMS3020/C25, and DSP56000. The SNR (including distortion) of the AD7869 is 80dB which meets the requirements for the V.32 transmit and receive channels. A block diagram of the device is shown in Figure 10.11

A block diagram of a complete analog front-end for a V.32 modem is shown in Figure 10.12. The serial interface is shown with the ADSP-2101 DSP processor.

The AD7341/AD7371 is a switched capacitor voiceband reconstruction/antialiasing filter chip set designed to be used in conjunction with the AD7869 to implement a complete V.32 modem analog front end. The

SCFs are clocked at a 57.6kHz rate, representing an oversampling ratio of 6X with respect to the ADC sampling rate of 9.6kSPS. The AD7341 is the transmit reconstruction filter. It implements the filter function using a seventh-order lowpass SCF and a second-order continuous time filter. The cutoff frequency is 3.5kHz. The AD7371 is the receive ADC antialiasing filter. It is a band-pass filter with a lower cutoff frequency of 180Hz and an upper cutoff frequency of 3.5kHz. The filter function is implemented using a second-order lowpass continuous time filter, a fourth-order highpass SCF and a seventh-order lowpass SCF. Key specifications for the AD7341/AD7371 SCF chip set are summarized in Figure 10.13.

KEY SPECIFICATIONS FOR THE AD7341/AD7371 MODEM FILTER CHIP SET

- **Stopband Attenuation:** 70dB, $f \geq 6.1\text{kHz}$
40dB, $f \leq 60\text{Hz}$
- **In-Band Signal to Noise Ratio:** 75dB
- **Total Harmonic Distortion:** -75dB
- **Differential Group Delay:** 350 μs
- **Programmable Attenuation (AD7341) :** 0 to 38dB
- **Programmable Gain (AD7371):** 0 to 24dB

Figure 10.13

The ADSP-28msp01 is a complete analog front end for high performance modems. The device has an architecture similar to the ADSP-28msp02 voiceband codec. The ADSP-28msp01 contains a 16 bit sigma-delta ADC and DAC and is capable of sampling rates of

7.2, 8.0, and 9.6kSPS with SNR and THD performance of 84dB. The extensive support of bit, baud, and convert clocks allow the ADSP-28msp01 to support many modem standards such as the V.32. Key specifications are summarized in Figure 10.14.

ADSP-28msp01 INTEGRATED MODEM ANALOG FRONT END KEY SPECIFICATIONS

- 16 Bit Sigma-Delta ADC and DAC
- On-Chip Antialiasing and Anti-Imaging Filters
- On-Chip Clock Generation Circuitry
- 84 dB THD and SNR
- Programmable Sampling Frequency of 7.2, 8.0, and 9.6kSPS
- DSP Compatible Serial Port
- 28 Pin DIP/SOIC

Figure 10.14

DIGITAL MOBILE RADIO

OVERVIEW

The rapidly growing number of cellular mobile phones in the United States has created significant system performance problems, especially in crowded metropolitan areas such as New York and Los Angeles. Call blocking during rush hour, flaws in call processing (disconnects and misconnects), and undesirable interchannel crosstalk are only a few. In addition, the current system lacks privacy and security, and data trans-

mission over a mobile link is almost impossible at rates above 1200 bits/s. These factors have led to the search for a more efficient and robust system based on digital techniques. Several digital approaches are being considered in the United States, while the Pan-European Digital Cellular Radio System (also known as *Groupe Speciale Mobile*, or GSM) has been defined and will be introduced in Europe in 1991.

PROBLEMS WITH CURRENT ANALOG CELLULAR RADIO

- Call Blocking During Busy Hours
- Misconnects and Disconnects due to Rapidly Fading Signals
- Lack of Privacy and Security
- Data Transmission Limited to 1200 bits/s

Figure 10.15

The current system in the United States is a cellular system based on Frequency Division Multiple Access (FDMA). A region is broken up into cells, with each cell having its own base station and its own group of assigned frequencies (see Figure 10.16). Be-

cause the radius of each cell is small (10 miles, for example) low power transmitters and receivers can be used. The cellular system lends itself to frequency reuse, since cells which are far enough apart can utilize the same band of frequencies without inter-

CELLULAR RADIO FREQUENCY REUSE

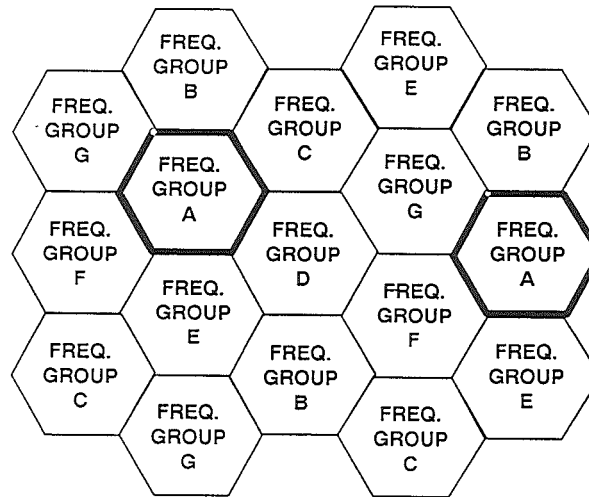


Figure 10.16

ference. The base stations must be linked together with an elaborate central control network so that a call may be handed-off to another cell when the signal strength from the mobile unit becomes too low for the current cell to handle.

The frequency spectrum allocation for cellular radio in the United States is approximately 825 to 850MHz and 870 to 895MHz. Conventional architectures (both analog and digital) are channelized. The total spectrum is divided up into a large number of relatively narrow channels, defined by a carrier

frequency. The carrier frequency is frequency-modulated with the voice signal using analog techniques. Each full-duplex channel requires a pair of frequencies, each with a bandwidth of approximately 30kHz. A user is assigned both frequencies for the duration of the call. The forward and reverse channel are widely separated, to help the radio keep the transmit and receive functions separated. The 40MHz allocated to cellular service can therefore be divided up into 666 frequency pairs, each serving one full-duplex circuit.

FREQUENCY DIVISION MULTIPLE ACCESS MOBILE RADIO SYSTEM

- Uses 825-850MHz and 870-895MHz Spectrum
- 30kHz Transmit, 30kHz Receive
- Analog Frequency Modulation (FM)
- Approximately 700 Users Capacity

Figure 10.17

Time Division Multiple Access (TDMA) allocates bandwidth on a time-slot basis. In the proposed United States TDMA system, the entire 30kHz channel is assigned to a particular transmission, but only for a short period of time. A 3:1 multiplexing scheme means that three conversations can take place with TDMA using the same amount of bandwidth as one analog cellular conversation does. Each transmit/receive sequence occurs on time slots lasting 6.7ms. The TDMA system relies on an extensive amount of DSP technology to reduce the coded speech bit-rate as well as to prepare the digital data for transmission over the analog medium. The TDMA approach has been chosen for the Pan-European GSM system and will be discussed later in more detail.

The second digital approach being considered in the United States is called Code Division Multiple Access (CDMA). This technique has been used in secure military

communications for a number of years under the name of *spread spectrum*. In spread spectrum, the transmitter transmits in a pseudo-random sequence of frequency hops over a relatively wide frequency range. The receiver has access to the same random sequence and can decode the transmission. The effect of adding additional users on the system is to decrease the overall signal to noise ratio for all the users. With this technique, the effect of allowing more calls than the normal capacity is to increase the bit-error rate for all users. New callers can keep coming in, interference levels will rise gradually, until at some point the process will become self-regulating: the quality of the voice link will become so bad that users will cut short or refrain from making additional calls. No one is ever blocked in the conventional sense, as they are in FDMA or TDMA systems when all channels or slots are full.

DIGITAL MOBILE RADIO APPROACHES

- Time Division Multiple Access (TDMA) - User Allocation Based on Time Slots: At Least 3X More Capacity than FDMA
- Code Division Multiple Access (CDMA) - Base on Spread Spectrum Technology: More Users Cause Graceful Degradation in Bit-Error Rate
- Both TDMA and CDMA Make Extensive Use of DSP in Speech Encoding and Channel Coding for Transmission

Figure 10.18

Both TDMA and CDMA systems make extensive use of DSP algorithms in both speech encoding and in preparing the signal for transmission. In the receiver, DSP techniques are used for demodulation and decoding the speech signal. The remainder of this

section will concentrate on speech processing and channel coding as they relate to the Pan-European GSM system. This will serve to illustrate the fundamental principles which are applicable to all digital mobile radio systems.

THE GSM SYSTEM

Figure 10.19 shows a simplified block diagram of the GSM Pan-European Digital Cellular Telephone System. The *speech encoder and decoder* and *discontinuous transmission* function will be described in detail. Up conversion and downconversion

portion of the system contain a digital modem similar to the V.32 recommendation previously discussed. Similar functions are performed digitally such as equalization, convolutional coding, Viterbi decoding, modulation and demodulation.

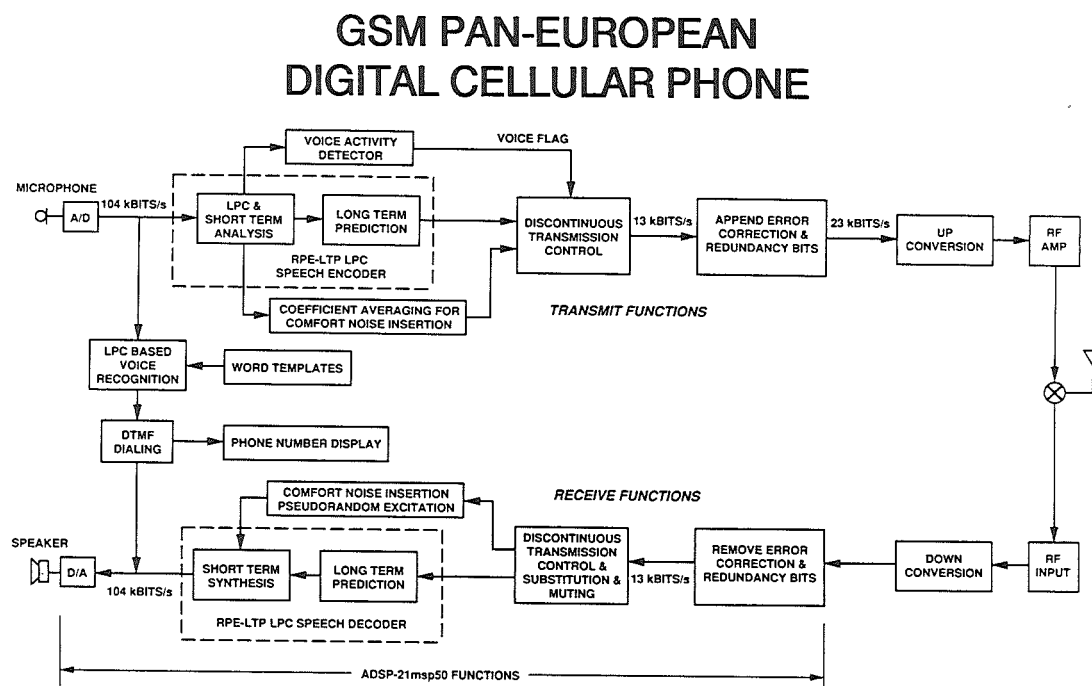


Figure 10.19

SPEECH CODEC

The standard for encoding voice has been set in the T-Carrier digital transmission system. In this system, speech is logarithmically encoded to 8 bits at a sampling rate of 8kSPS. The logarithmic encoding and decoding to 8 bits is equivalent to linear encoding and decoding to 13 bits of resolution. This produces a bit-rate of 104kb/s. The Speech Encoder portion of the GSM system compresses the speech signal to 13kb/s, and the

decoder expands the compressed signal at the receiver. The terms *codec* and *transcoder* are both often used to refer to the entire encoding and decoding speech compression function. The speech encoder is based on an enhanced version of linear predictive coding (LPC). The LPC algorithm uses a model of the human vocal tract that represents the throat as a series of concentric cylinders of various diameters. An excitation (breath) is

forced into the cylinders. This model can be mathematically represented by a series of simultaneous equations which describe the cylinders.

The excitation signal is passed through the cylinders, producing an output signal. In the human body, the excitation signal is air moving over the vocal cords or through a constriction in the vocal tract. In a digital system, the excitation signal is a series of pulses for vocal excitation, or noise for a constriction. The signal is input to a digital lattice filter. Each filter coefficient represents the size of a cylinder.

An LPC system is characterized by the number of cylinders it uses in the model. Eight cylinders are used in the GSM system, and eight reflection coefficients must be generated.

Early LPC systems worked well enough to understand the encoded speech, but often the quality was too poor to recognize the voice of the speaker. The GSM LPC system employs two advanced techniques that improve the quality of the encoded speech. These techniques are *regular pulse excitation* (RPE) and *long term prediction* (LTP). When these techniques are used, the resulting quality of encoded speech is nearly equal to that of logarithmic pulse code modulation (companded PCM as in the T-Carrier system).

The actual input to the speech encoder is a series of 13-bit samples of uniform PCM speech data. The sampling rate is 8kHz. The speech encoder operates on a 20ms window (160 samples) and reduces it to 76 coefficients (260 bits total), resulting in an encoded data rate of 13kb/s.

SPEECH COMPRESSION IN THE GSM SYSTEM

- Input Data: 13bit Samples at 8KSPS = 104kbits/s
- Output Data for Each 20ms Window: 76 Filter Coefficients, 260 bits Total = 13kbits/s

Figure 10.20

DISCONTINUOUS TRANSMISSION (DTX)

Discontinuous transmission (DTX) allows the system to shut off transmission during the pauses between words. This reduces transmitter power consumption and increases the overall GSM system's capacity.

Low power consumption prolongs battery life in the mobile station and is an important consideration for hand-held portable phones. Call capacity is increased by reducing the

interference between channels, leading to better spectral efficiency. In a typical conversation each speaker talks for less than 40% of the time, and it has been estimated that DTX can approximately double the call capacity of the radio system.

The required DTX functions are summarized in Figure 10.21.

DISCONTINUOUS TRANSMISSION (DTX) FUNCTIONS

- **Voice Activity Detection (VAD) to Detect Speech**
- **Comfort Noise Insertion (CNI) to Synthesize Artificial Car Noise During Pauses Between Words**
- **Output Muting When Lost Speech Frames Are Received**

Figure 10.21

The voice activity detector (VAD) is located at the transmitter; its job is to distinguish between speech superimposed on the background noise and noise with no speech present. The input to the voice activity detector is a set of parameters computed by the speech encoder. The VAD uses this information to decide whether or not each 20ms frame of the encoder contains speech.

Comfort noise insertion (CNI) is performed at the receiver. The comfort noise is generated when the DTX has switched off the transmitter; it is similar in amplitude and spectrum to the background noise at the transmitter. The purpose of the CNI is to eliminate the unpleasant effect of switching between speech with noise, and silence. If you were listening to a transmission without CNI, you would hear rapid alternating between speech in a high-noise background (i.e. in a car), and silence. This effect greatly reduces the intelligibility of the conversation.

When DTX is in operation, each burst of speech is transmitted followed by a *silence descriptor* (SID) frame before the transmission is switched off. The SID serves as an end of speech marker for the receive side. It contains characteristic parameters of the background noise at the transmitter, such as spectrum information derived through the use of linear predictive coding.

The SID frame is used by the receiver's comfort noise generator to obtain a digital

filter which, when excited by pseudo-random noise, will produce noise similar to the background noise at the transmitter. This comfort noise is inserted into the gaps between received speech bursts. The comfort noise characteristics are updated at regular intervals by the transmission of SID frames during speech pauses.

Redundant bits are then added by the processor for error detection and correction at the receiver, increasing the final encoded bit rate to 22.8kb/s. The bits within one window, and their redundant bits, are interleaved and spread across several windows for robustness.

The ADSP-21msp50 Mixed Signal Processor shown in Figure 10.22 can perform all of the above tasks within the 20ms sampling window because of its optimized DSP architecture and the special on-chip peripherals associated with it. The sigma-delta converters provide the necessary interface to the speaker and microphone. The parallel host interface port communicates with a host processor, which is responsible for loading the ADSP-21msp50 with the appropriate programs during power-up, dialing, and actual conversation phases of a complete call. The ADSP-21msp50 has 1K words of (16-bit) data memory static RAM and 2K words of (24-bit) program memory static RAM on chip. The device operates at a 13MHz clock rate and has a low power mode and a power down

ADSP-21msp50 BLOCK DIAGRAM

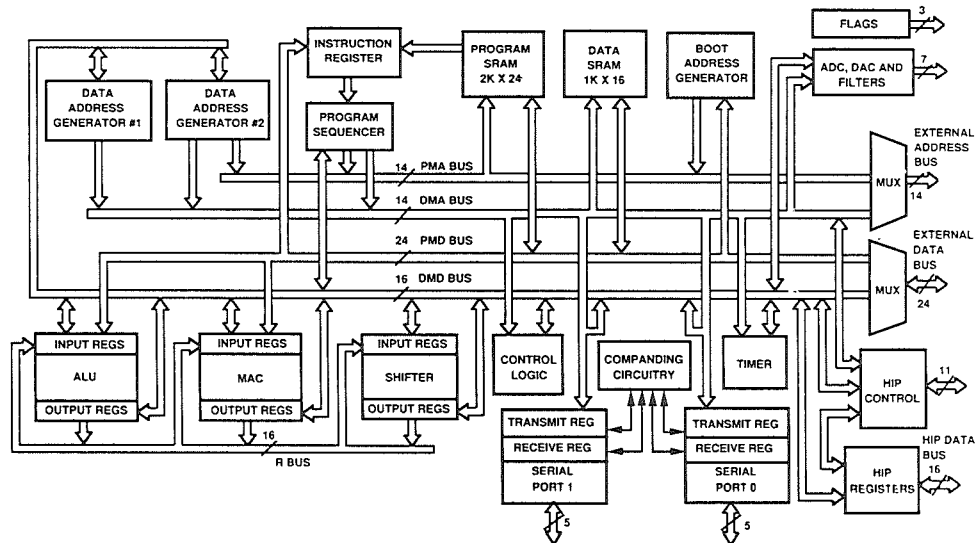


Figure 10.22

mode (less than 1mW in power down). The ADSP-21msp50 combines the core ADSP-2100 architecture (three computational units, data address generators, and a program sequencer) with two serial ports, a programmable timer, host interface port, an

analog interface, extensive interrupt capabilities. Key features of the ADSP-21msp50 are summarized in Figure 10.23, and the benchmark performance in the GSM system is shown in Figure 10.24.

ADSP-21msp50 MIXED SIGNAL PROCESSOR KEY SPECS

- On-Chip 16-Bit Sigma-Delta ADC and DAC
- 65dB SNR and THD
- 8kSPS Sampling Frequency, 1MHz Clock (125X Oversampling)
- 2K Words Program Memory Ram (24-bits)
- 1K Words Data Memory Ram (16-bits)
- 13MIPS Performance
- Host Interface Port
- ADSP-2100 Family Compatible Instruction Set
- Low Power and Power Down Mode

Figure 10.23

ADSP-21msp50 GSM BENCHMARKS

FUNCTION	CYCLE COUNT MAXIMUM WORST CASE	TIME REQUIRED OUT OF 20ms WINDOW	PROCESSOR LOADING
RPE-LTP LPC Encoder	49300	3.8ms	19.0%
RPE-LTP LPC Decoder	14400	1.1ms	5.5%
Voice Activity Detector	2141	0.17ms	0.9%
Total Functions	65841	5.07ms	25.4%
Free			74.6%

Internal Program Memory Required: 1988 words
Internal Data Memory Required: 964 words

Figure 10.24

GSM SYSTEM UPCONVERSION AND DOWNCONVERSION

A block diagram of the GSM system with particular emphasis on the upconversion and downconversion circuitry is shown in Figure 10.25. The transmit data coming from the speech processor contains error correction and redundancy bits. The bit rate at this point in the system is 23kb/s. The channel coder and filters prepare the data to fit the TDMA format of the GSM system. Figure

10.26 shows how each 200kHz of frequency spectrum contains data from 8 users. Each user is assigned a time slot of 0.577ms during which time a burst of 156 data bits are transmitted at a modulation frequency of approximately 270kHz. Modulation is accomplished using Gaussian Minimum Shift Keying (GMSK), a form of frequency shift keying which minimizes spectral leakage.

GSM BLOCK DIAGRAM: UPCONVERSION DETAILS

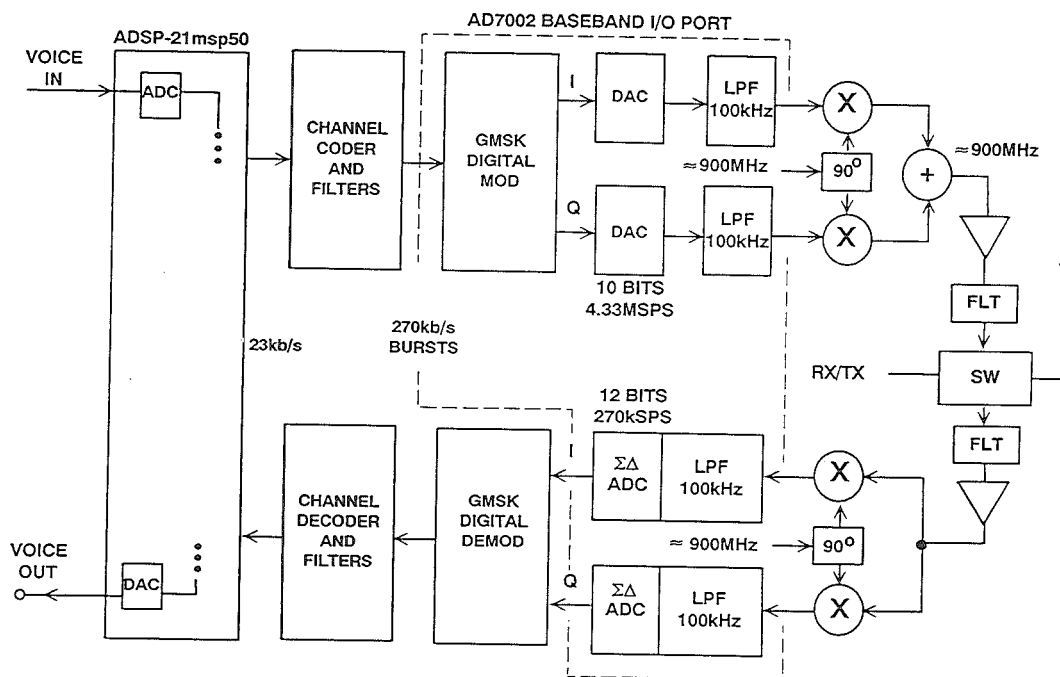


Figure 10.25

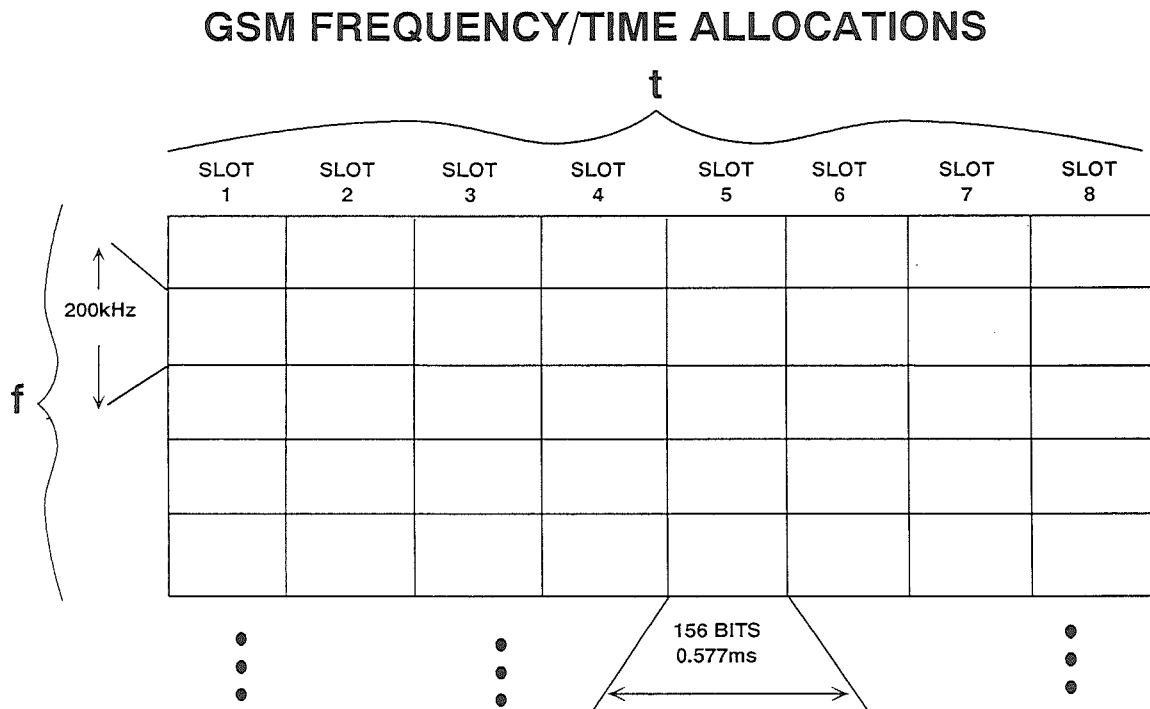


Figure 10.26

The modulation is done digitally and converted into an I and Q signal. The modulator outputs drive two 10-bit DACs whose filtered output drives the RF modulators. The DACs are oversampled by a factor of 16 in order to simplify the anti-imaging analog filter requirements. The combined I and Q signal drives the RF amplifier, filter, and the antenna.

The received signal is filtered, amplified, and fed to an I/Q RF demodulator which recovers the I and Q signals. The baseband I and Q signals are converted by two 12-bit DAC at an effective sampling rate of 270kSPS. The I and Q signals are then demodulated by the GMSK digital demodulator. The 270kb/s burst is sent to the channel decoder and filters and then to the speech processor.

The AD7002 is a complete GSM Baseband I/O Port which performs the functions shown in Figure 10.25. The transmit path contains two 10-bit oversampled (16X) DACs followed by fourth-order anti-imaging filters. The DACs are driven by a digital modulator containing a GMSK-coded ROM. The receive path contains two high-performance 12 bit sigma-delta ADCs having a throughput rate of 270kSPS. The sigma-delta ADCs contain a 288-tap FIR filter having linear phase response and a 3dB point of 122kHz. Three auxiliary DACs are included for such functions as AFC, AGC and carrier shaping. The device dissipates approximately 100mW and has flexible power-down or sleep modes. Key specifications are summarized in Figure 10.27.

AD7002 GSM BASEBAND I/O PORT KEY SPECIFICATIONS

- **Transmit Path:** **GMSK I/Q Digital Modulator**
 Dual 10 Bit, 4.33MSPS Oversampled DACs
 Dual Anti-Imaging Filters
- **Receive Path:** **Dual 12 Bit 270kSPS Sigma-Delta ADCs**
 288-Tap 100kHz Linear Phase FIR Filter
- **3 Auxiliary DACs for AFC, AGC**
- **Low Power: 100mW**
- **Sleep Mode**

Figure 10.27

DIGITAL AUDIO STUDIO RECORDING

The activities related to studio recording are complex and varied. Generally, multiple channels are used, with each track dedicated to one or more sources (instruments/voices). All channels need not be recorded at the same time. Each channel is subjected to extensive processing such as gain control, filtering, non-linear compression or expansion, reverberation, spectral equalization, and other special-effects enhancements. The contributing channels are then mixed together to obtain a final arrangement with the desired overall effect.

Traditionally, channel processing and mixing were implemented entirely in the analog domain—with numerous disadvan-

tages. Each channel's information—stored as an analog signal on magnetic tape—degrades as the cutting, splicing, and re-recording process progresses, undermining the benefits of the processing. The limited performance range available with analog processing sets a ceiling on the signal enhancement that can be obtained. Also, analog circuitry can only handle one channel at a time; multi-channel mixers are expensive and difficult to control. Finally, if analog processing hardware is used, overall mixing flexibility can be achieved only through hardware modifications. In practice, this means that the mixing process loses its ability to creatively explore special effects.

DIGITAL AUDIO STUDIO TECHNIQUES

- Digital Recording: 16, 18 or 20 Bits for ADC
- Digital Mixing
- Gain Control
- Reverberation and Special Effects
- Equalization using Digital Filters

Figure 10.28

Increasingly, audio processing is relying on digital techniques to improve audio quality. The first step in this transition was digital recording, which became prevalent in the early 1980s. Audio signals are first converted to digital form before being stored on magnetic tape. Digital recording eliminates several sources of degradation that hamper analog recordings, including the effects of non-linearities and additive noise in the magnetic materials used for recording, and wow and flutter in the tape playback mechanism.

In studio mixing applications, however, digital recording does not eliminate all complications. In the mixing and enhancement process, information is passed from one tape

to another—requiring both ADCs and DACs, a source of noise. These conversions are no longer necessary if all processing and mixing are handled with DSP techniques.

In the DSP-based studio recording system shown in Figure 10.29, signals are converted to digital as early as possible, usually to 16, 18, or 20-bit resolution. After conversion, the audio processing is handled digitally with high performance DSP processors. Gain factors are handled with digital multiplication. Filtering and equalization can be handled with linear-phase FIR filters. Dynamic-range control is easily included in the system by using a multiplier for non-linear compression/expansion computations.

DIGITAL AUDIO STUDIO SYSTEM

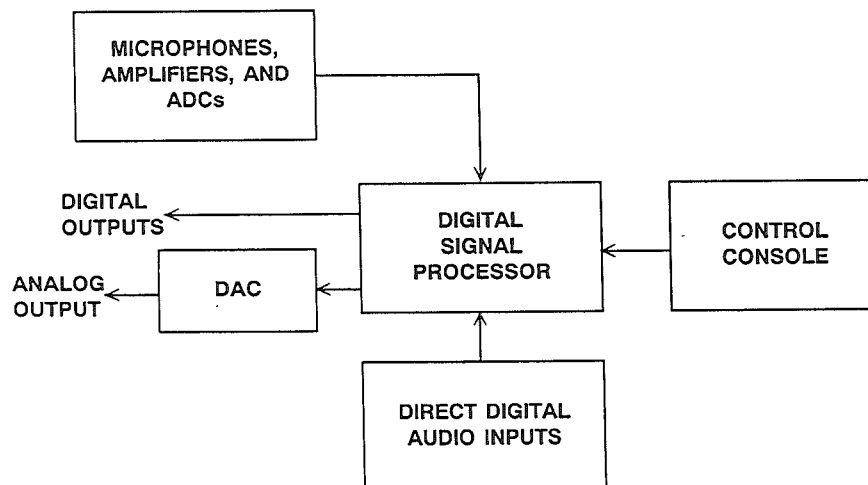


Figure 10.29

The traditional mixing process is also easily implemented in a DSP-based system. Digital channels to be mixed are simply added together. Relative time delay lags can be easily introduced into the channel flows, allowing phase delays to be equalized. Channel interconnections—which have to be hardwired in an analog processor—can be easily reconfigured in a DSP system.

In addition to improving on traditional operation, a DSP studio recording system opens up numerous new options. Unusual special effects are readily included in the system. Reverberation effects can be modeled, simulated, and integrated into the final recording. Digital reverberation can give concert hall or cathedral ambience to what might have been recorded in a dry studio. An FFT routine's spectral analysis of the signal forms the basis

for adaptive digital filters that provide optimal equalization.

With the advent of compact disc (CD) and digital audio tape (DAT) players, there is no requirement for digital-to-analog conversion anywhere in the studio recording process, except for monitoring and optimization purposes. The final digital recording can be transferred directly to the CD or DAT in digital form with no loss in fidelity.

Although 18 and 20-bit ADCs may be used in the recording process, the standard for CD and DAT has been set at 16 bits. Additional bits may be used in the DSP studio processing to allow for roundoff errors, overflows, etc., but the final recording is truncated to 16 bits per sample on the CD or DAT. The sampling-rate standard for CD recordings is 44.1kSPS, and 48kSPS for DAT.

DIGITAL AUDIO RECORDING STANDARDS

- **16-20 bits ADC Resolution, Truncated to 16 bits for Compact Disc**
- **44.1kSPS Sampling Rate for CD Players**
- **48kSPS Sampling Rate for Digital Audio Tape (DAT) Players**

Figure 10.30

Performance of audio systems is primarily measured in terms of three dynamic specifications: Total harmonic distortion plus noise

(THD+N), D-Range distortion, and signal-to-noise ratio. The definitions of these specifications are given in Figure 10.31.

KEY AUDIO PERFORMANCE SPECIFICATIONS

- **THD + N:** Ratio of the Square Root of the Sum of the Squares of the Values of the Harmonics and Noise to the Value of the Fundamental Input Frequency Expressed in % or dB
- **D-Range Distortion:** Ratio of the Distortion Plus Noise to the Signal at a Signal Amplitude of -60dB. Add 60dB to the Ratio to Obtain D-Range Distortion Value
- **Signal-to-Noise Ratio:** Ratio of the Amplitude of the Output with No Signal Present to the Amplitude of the Output When a Fullscale Output is Present

Figure 10.31

Specifications for the AD1876 16-bit 100kSPS ADC are given in Figure 10.32, and specifications for the AD1879 18-bit sigma-

delta ADC are given in Figure 10.33. These two ADCs are fully specified in terms of digital audio parameters.

AD1876 16-BIT 100kSPS SAMPLING ADC

- Autocalibrating
- 0.001% THD (100dB)
- 90dB S/(N + D)
- 2x Audio Oversampling Capability
- Power Dissipation: 250mW

Figure 10.32

AD1879 DUAL 18-BIT SIGMA-DELTA AUDIO ADC

- Signal-to-Noise Ratio: -106dB FS
- 0.0017% THD (95dB) at 1kHz
- Interchannel Crosstalk: -110dB at 1kHz
- Decimator Filter Passband Ripple: 0.001dB
- Decimator Filter Stopband Attenuation: 115dB
- Oversampling Ratio: 64x
- Output Word Rate: 48kSPS

Figure 10.33

COMPACT DISC (CD) PLAYER ELECTRONICS

A simplified block diagram of the read electronics for a typical CD player is shown in Figure 10.34. The read electronics takes the data from the CD read head and performs the necessary data qualification, error detection and error correction. Data from the read electronics is in serial format, 16 bits per sample, at an effective sampling rate of 44.1kHz per channel. Data for the two channels is usually multiplexed in a single

1.4112MHz bit stream. In theory, it is possible to reconstruct the audio signal using two 16-bit DACs preceded by a digital demultiplexer and parallel-to-serial converters operating at an update rate of 44.1kHz followed by analog anti-imaging filters. First-generation CD players used this approach as shown in Figure 10.35. An alternative is a single 16 bit DAC with output multiplexing for left and right channel.

COMPACT DISC PLAYER READ ELECTRONICS

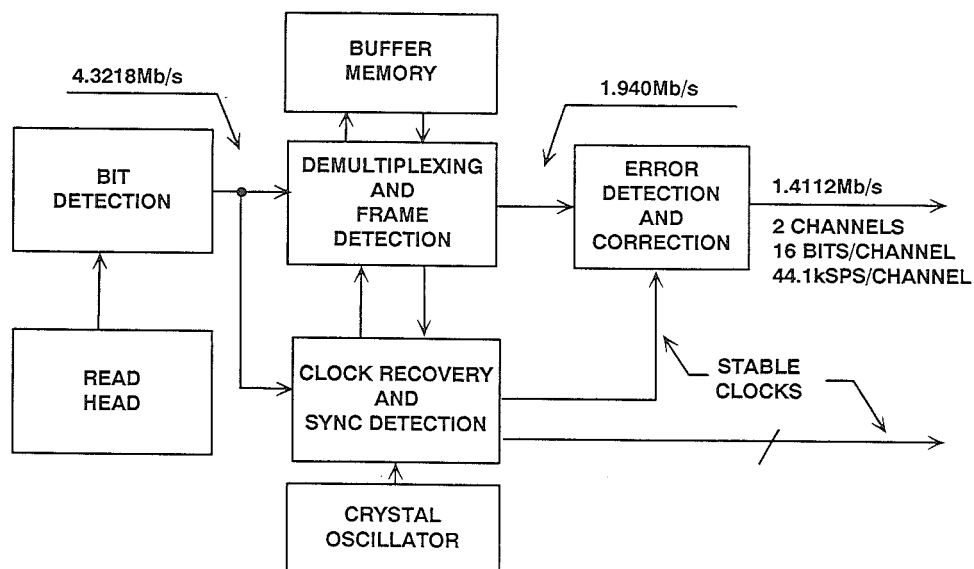


Figure 10.34

FIRST-GENERATION CD RECONSTRUCTION ELECTRONICS

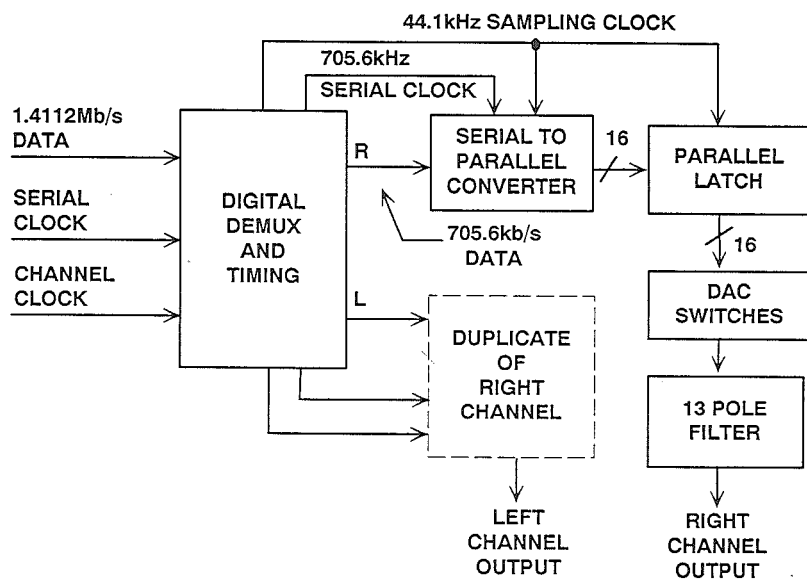


Figure 10.35

8X OVERSAMPLED 18-BIT CD RECONSTRUCTION ELECTRONICS

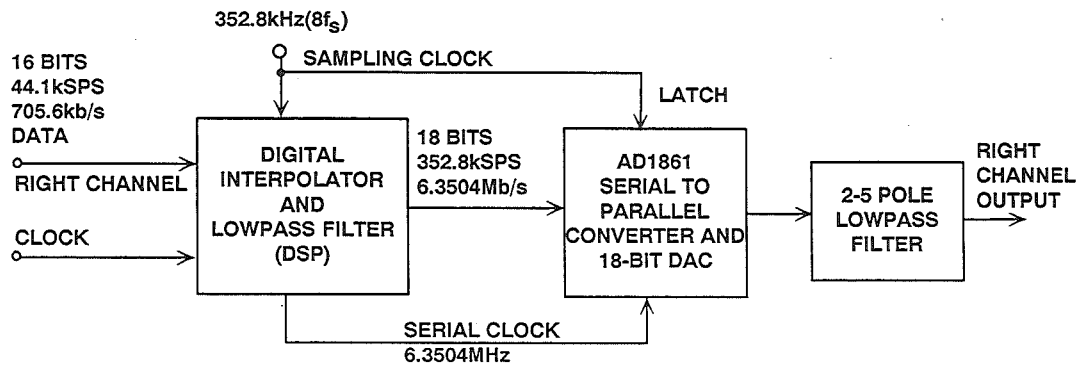


Figure 10.36

Sampling at 44.1kHz places severe requirements on the antialiasing filters. The audio bandwidth extends from 20Hz to 20,000Hz, and the filters must exhibit a flat frequency response over this frequency. In order to prevent aliasing, the filters must have at least 40dB to 50dB attenuation at 22.05kHz. This implies a complicated and costly 9- to 13-pole analog filter. In addition, these higher-order filters typically have non-linear phase response which is undesirable in audio applications. For this reason, the principles of oversampling and digital filtering are now in widespread use to simplify the design of the analog filter as well as increase the overall signal-to-noise ratio.

Second-generation CD players typically used oversampling ratios of 2x (88.2kSPS) or 4x(176.4kSPS) in conjunction with linear phase FIR digital interpolation filter chips. Third-generation players are using 8x oversampling (352.8kSPS) as shown in Figure 10.36, and the trend for future players will probably be 16x (705.6kSPS) or higher.

In addition to easing the requirements on the anti-imaging filter, oversampling followed by digital filtering spreads the quantization noise over a wider bandwidth, giving an improvement in SNR of $10 \log_{10}(K)$, where K is the oversampling ratio. This implies that for an oversampling ratio of 8x, there is a theoretical 9dB (or 1.5bits) improvement in SNR. It is possible to carry the arithmetic in the digital interpolation filter out to 18 bits, drive an 18-bit audio DAC with the result, and realize this improvement in practical CD players. If 16x oversampling were used, the theoretical improvement in SNR would be 12dB, or 2 bits (See Figure 10.37). A block diagram of a the complete reconstruction channel of an 8x oversampled 18 bit CD player is shown in Figure 10.38. The design is based on the dual 18 bit AD1865 DAC. Because of the 8x oversampling ratio, the output filter is the simple 3-pole filter which is shown in Figure 10.39.

EFFECTS OF OVERSAMPLING AND DIGITAL FILTERING ON CD PLAYER DESIGN

OVERSAMPLING RATIO K	THEORETICAL INCREASE IN SNR	USEFUL BITS OF DAC RESOLUTION	NUMBER OF POLES REQUIRED IN ANALOG FILTER
1	0dB	16	10
2	3dB	16	5
4	6dB	16/18	4
8	9dB	16/18/20	3
16	12dB	16/18/20	2

Figure 10.37

8X OVERSAMPLED CD RECONSTRUCTION ELECTRONICS USING DUAL 18-BIT DAC

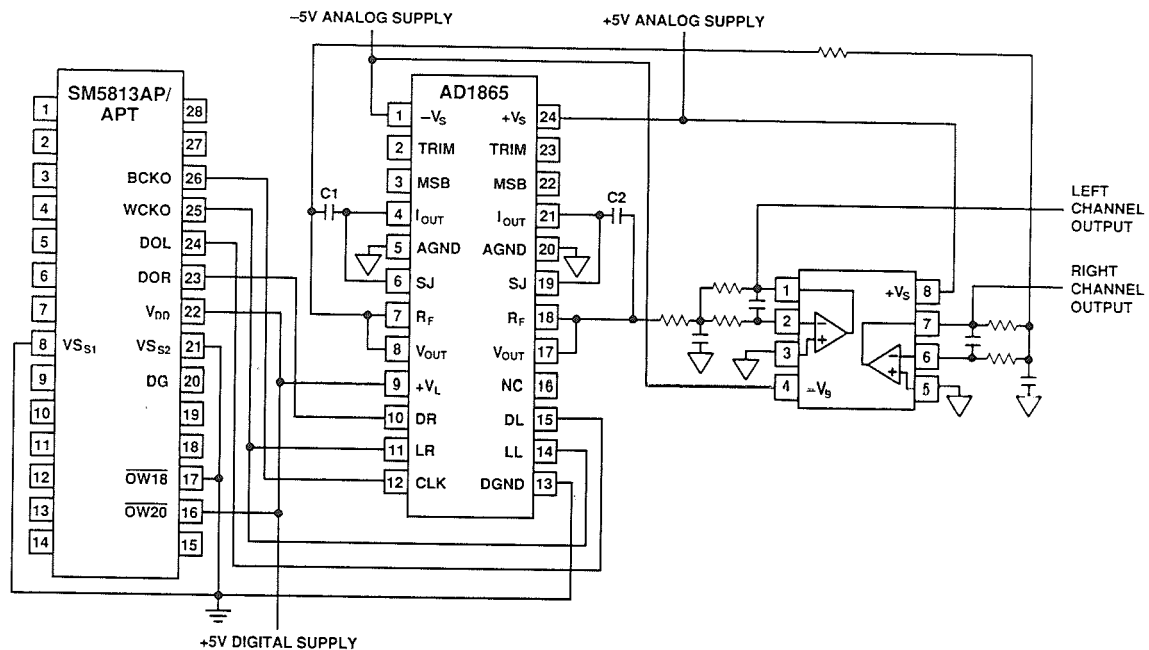


Figure 10.38

3-POLE ANTIALIASING FILTER FOR 18-BIT, 8X OVERSAMPLING

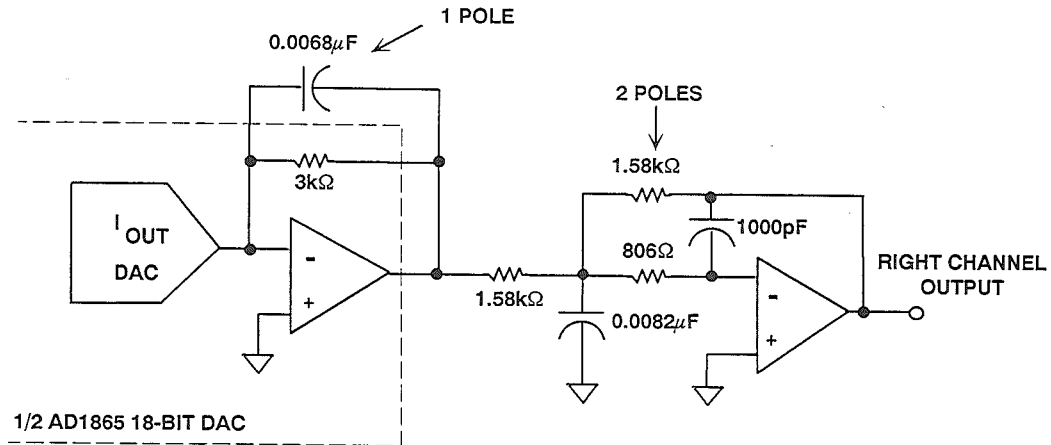


Figure 10.39

HIGH PERFORMANCE 20-BIT 8X OVERSAMPLING CD RECONSTRUCTION ELECTRONICS

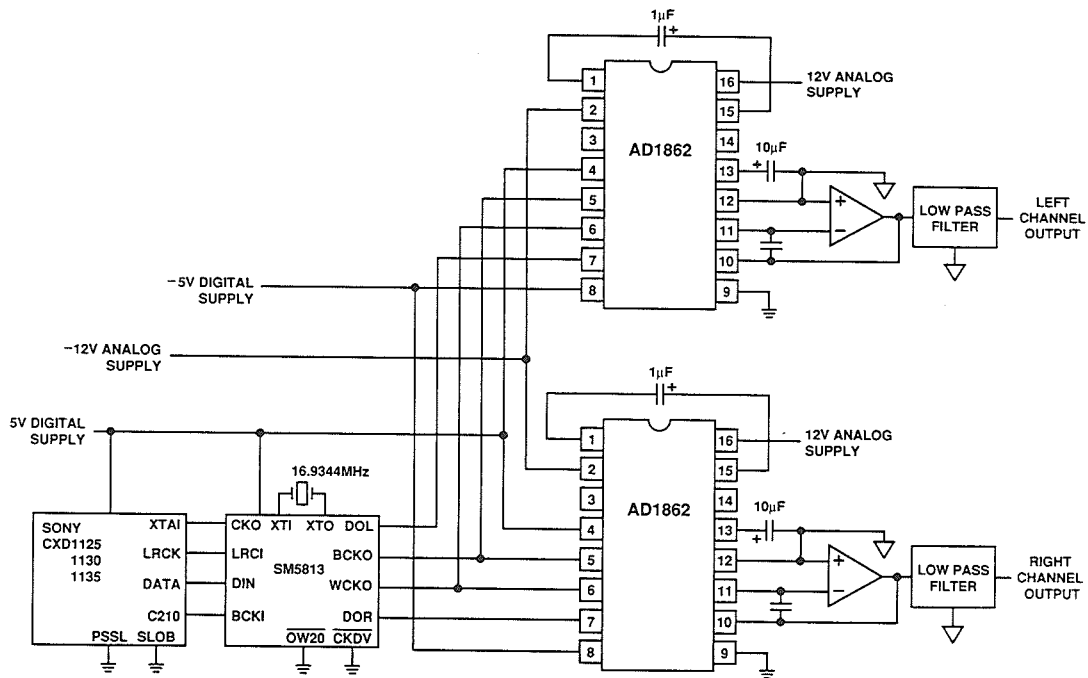


Figure 10.40

An additional reason for using DACs with greater than 16-bit resolution is that the process of digital interpolation and filtering adds truncation noise when the digital filter rounds off the interpolated values. This noise is reduced by using 18- and even 20-bit DACs to preserve accuracy in the interpolated values. A block diagram of a 20-bit, 8x oversampling CD filter and DAC configuration is shown in Figure 10.40. Because of the 8x oversampling ratio, a 5-pole lowpass filter is sufficient to maintain the required performance. Typical THD + N performance for the system shown in Figure 10.40 (including the output filter) is shown in Figure 10.41. A variety of CD digital interpolation filter chips are currently available from manufacturers such as Yamaha, NPC, and Sony.

There are a number of audio DACs currently available on the market ranging from

16 to 20-bit resolution. Newer devices are capable of sampling rates up to 768kSPS, allowing 16x oversampling. Unlike traditional DACs, audio DACs are specified in terms of ac parameters such as THD + N, SNR, and D-Range Distortion because traditional dc specifications are not as critical for audio applications. All audio DACs accept serial inputs and have internal serial-to-parallel converters followed by a parallel latch. Two clock inputs are therefore required to operate an audio DAC. A serial clock is needed to strobe the serial data into the serial-to-parallel converter, and a latch-enable clock is required to strobe the parallel latch. A simplified block diagram of a typical digital audio DAC is shown in Figure 10.42.

THD+N PERFORMANCE OF 20-BIT 8X CD ELECTRONICS USING AD1862

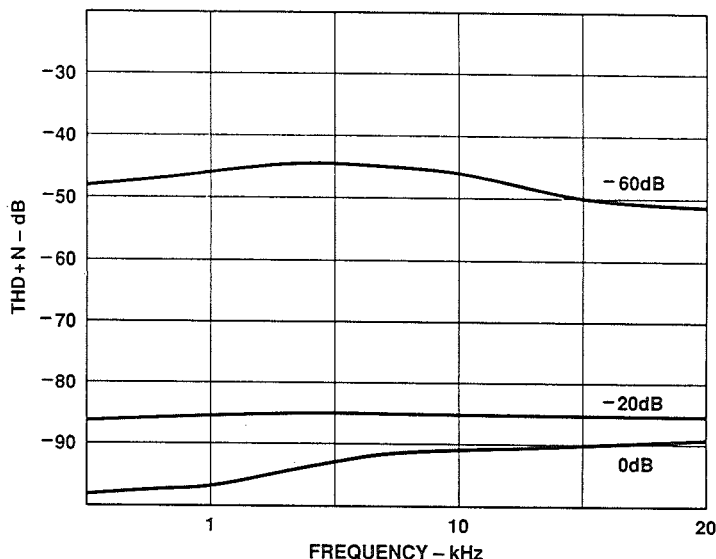


Figure 10.41

TYPICAL DC AUDIO DAC (SINGLE CHANNEL, 18 BITS, 8X)

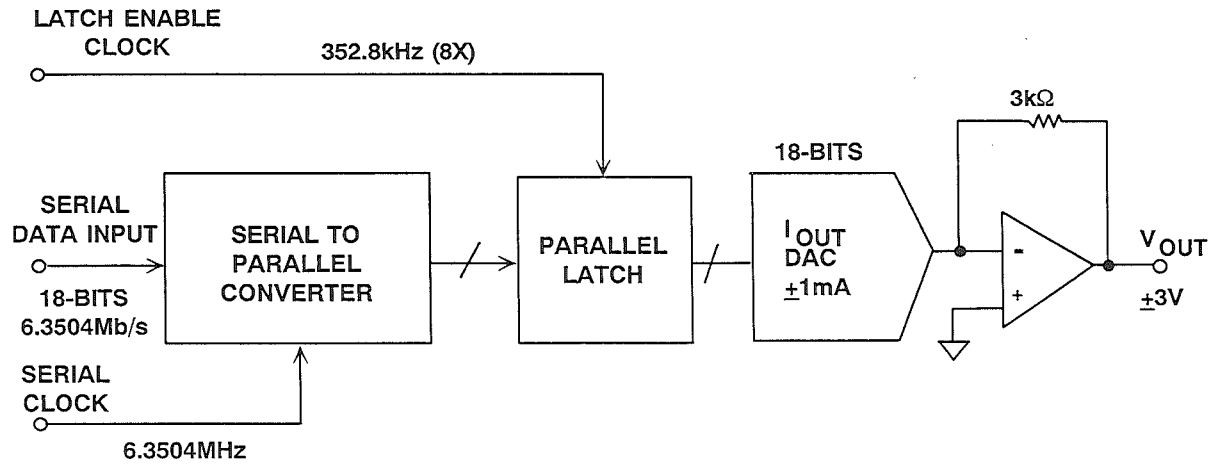


Figure 10.42

REFERENCES

1. John Bingham, **The Theory and Practice of Modem Design**, John-Wiley, 1988.
2. **ADSP-2100 Family Applications Handbook Vol. 4**, Analog Devices, 1990.
3. John Reidy and Mike Curtin, **Understanding and Applying the AD7341/AD7371 Switched Capacitor Filters**, Analog Devices Application Note E1373-15-5/90.
4. George Calhoun, **Digital Cellular Radio**, Artech House, Norwood MA, 1988
5. Nick Morley, *New Cellular Scheme Muscles In*, **EDN News**, November 15, 1990, p. 1.
6. P. Vary, K. Hellwig, R. Hofmann, R. J. Sluyter, C. Galand, M. Rosso, *Speech Codec for the European Mobile Radio System*, **ICASSP 1988 Proceedings**, p. 227.
7. D. K. Freeman, G. Cosier, C. B. Southcott, I. Boyd, *The Voice Activity Detector for the Pan-European Digital Cellular System*, **ICASSP 1989 Proceedings**, p. 369.
8. J. Haspeslagh, D. Sallaerts, et. al, *A 270-kb/s 35mW Modulator IC for GSM Cellular Radio Hand-Held Terminals*, **IEEE Journal of Solid State Circuits**, Vol. 25, No. 6, December, 1990, pp. 1450-1457.
9. Ken Pohlmann, *The Compact Disc Formats: Technology and Applications*, **J. Audio Eng. Soc.**, Vol. 36, No.4, April 1988, pp 250-280.
10. Gary Davis and Ralph Jones, **Sound Reinforcement Handbook**, Second Edition, Written for Yamaha, Hal Leonard Publishing Company, Milwaukee, 1989.