Selecting Mixed-Signal Components for Digital Communication Systems—An Introduction

by Dave Robertson*

Communications is about moving information from point A to point B, but the computer revolution is fundamentally changing the nature of communication. Information is increasingly created, manipulated, stored, and transmitted in digital form—even signals that are fundamentally analog. Audio recording/playback, wired telephony, wireless telephony, audio and video broadcast—all of these nominally analog communications media have adopted, or are adopting, digital standards. Entities responsible for providing communications networks, both wired and wireless, are faced with the staggering challenge of keeping up with the exponentially growing demand for digital communications traffic. More and more, communications is about moving *bits* from point A to point B.

Digital communications embraces an enormous variety of applications, with radically different constraints. The transmission medium can be a twisted pair of copper wire, coaxial cable, fiber-optic cable, or wireless—via any number of different frequency bands. The transmission rate can range from a few bits per second for an industrial control signal communicating across a factory floor to 32 kbits/second for compressed voice, 2 Mb/s for MPEG compressed video, 155 Mbps for a SONET data trunk, and beyond. Some transmission schemes are constrained by formal standards, others are free-lance or developmental. The richness of design and architectural alternatives produced by such variety boggles the mind. The digital communications topic is so vast as to defy a comprehensive treatment in anything less than a shelf of books.

A communications jargon and a bewildering array of acronyms have developed, making it sometimes difficult for the communications system engineer and the circuit hardware designer to communicate with one another. Components have often been selected based on voltage-oriented specifications in the time domain for systems whose specifications are expressed in frequency and power. Our purpose here, and in future articles, will be to take a fairly informal overview of some of the fundamentals, with an emphasis on tracing the sometimes complex relationship between component performance and system performance.

The "communications perspective" and analytic tool set have also contributed substantially in solving problems not commonly thought of as "communications" problems. For example, the approach has provided great insight into some of the speed/bandwidth limits inherent in disk-drive data-recovery problems, where the channel from A to B includes the writing and reading of data in a magnetic medium—and in moving data across a high speed bus on a processing board.

*His photo and a brief biography appear in Analog Dialogue 30-3, page 2.

Shannon's law—the fundamental constraint: In general, the objective of a digital communications system is:

- to move as much data as possible per second
- across the designated channel
- with as narrow a bandwidth as possible
- using the cheapest, lowest-power, smallest-space (etc.) equipment available.

System designers are concerned with each of these dimensions to different degrees. Claude Shannon, in 1948, established the theoretical limit on how rapidly data can be communicated:

This means that the maximum information that can be transmitted through a given channel in a given time increases linearly with the channel's bandwidth, and noise reduces the amount of information that can be effectively transmitted in a given bandwidth, but with a logarithmic sensitivity (a thousandfold increase in noise may result in a tenfold reduction in maximum channel capacity). Essentially, the "bucket" of information has two dimensions: bandwidth and signal-to-noise ratio (SNR). For a given capacity requirement, one could use a wide-bandwidth channel with relatively poor SNR, or a narrowband channel with relatively good SNR (Figure 1). In situations where bandwidth is plentiful, it is common to use cheap, bandwidth-hungry communications schemes because they tend to be insensitive to noise and implementation imperfections. However, as demand for data communication capacity increases (e.g., more cellular phones) bandwidth is becoming increasingly scarce. The trend in most systems is towards greater spectral efficiency, or bits capacity per unit of bandwidth used. By Shannon's law, this suggests moving to systems with better SNR and greater demands on the transmit and receive hardware and software.

Let's examine the dimensions of bandwidth (time/frequency domain) and SNR (voltage/power domain) a little more closely by considering some examples.

PCM: A simple (but common) case: Consider the simple case

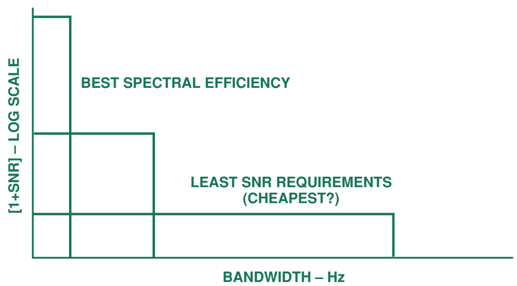


Figure 1. Shannon's capacity limit: equal theoretical capacity.

of transmitting the bit stream illustrated in Figure 2a, from a transmitter at location A to a receiver at location B (one may assume, that the transmission is via a pair of wires, though it could be any medium.) We will also assume that the transmitter and receiver have agreed upon both the voltage levels to be transmitted and the timing of the transmitted signals. The transmitter sends "high" and "low" voltages at the agreed-upon times, corresponding to 1s and 0s in its bit stream. The receiver applies a decision element (comparator) at the agreed-upon time to discriminate between a transmitted "high" and "low", thereby recovering the transmitted bit stream. This scheme is called *pulse code modulation* (or PCM). Application of the decision element is often referred to as "slicing" the input signal stream, since a determination of what bit is being sent is based on the value of the received signal at one instant in (slice of) time. To transmit more information down this wire, the transmitter increases the rate at which it updates its output signal, with the receiver increasing its "slicing" rate correspondingly.

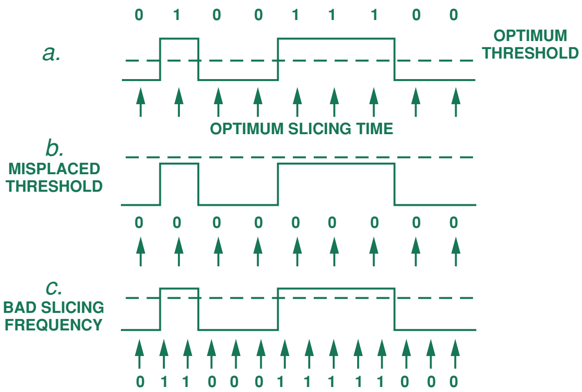


Figure 2. Simplified bit voltage transmission (PCM).

This simple case, familiar to anyone who has had an introductory course in digital circuit design, reveals several of the important elements in establishing a digital communications system. First, the transmitter and receiver must agree upon the "levels" that are to be transmitted: in this case, what voltage constitutes a transmitted "1", and what voltage level constitutes a transmitted "0". This allows the receiver to select the right threshold for its decision element; incorrect setting of this threshold means that the transmitted data will not be recovered (Figure 2b). Second, the transmitter and receiver must agree on the transmission frequency; if the receiver "slices" at a different rate than the bits are being transmitted, the correct bit sequence will not be recovered (2c). In fact, as we'll see in a moment, there must be agreement on both frequency and phase of the transmitted signal.

How difficult are these needs to implement? In a simplified world, one could assume that the transmitted signal is fairly "busy", without long strings of consecutive ones or zeros. The decision threshold could then be set at the "average" value of the incoming bit stream, which should be some value between the transmitted "1" and transmitted "0" (half-way between, if the density of ones and zeros are equal.) For timing, a phase-locked loop could be used—with a center frequency somewhere near the agreed-upon transmit frequency; it would "lock on" to the transmitted signal, thereby giving us an exact frequency to slice at. This process is usually called *clock recovery*; the format requirements on the transmit signal are related to the performance characteristics of the phase-locked-loop. Figure 3 illustrates the elements of this simplified pulse receiver.

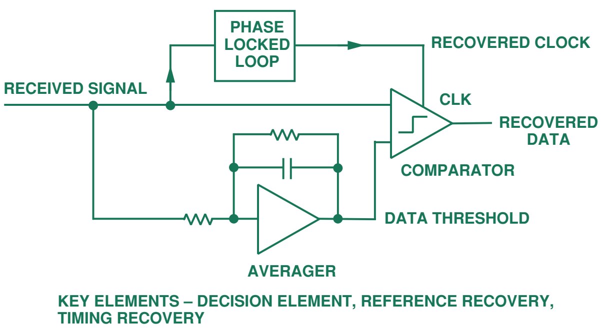


Figure 3. Idealized PCM.

Bandwidth Limitations: The real world is not quite so simple. One of the first important physical limitations to consider is that the transmission channel has finite bandwidth. Sharp-edged square wave pulses sent from the transmitter will be "rounded off" by a low bandwidth channel. The severity of this effect is a function of the channel bandwidth. (Figure 4). In the extreme case, the transmitted signal never gets to a logical "1" or "0", and the transmitted information is essentially lost. Another way of viewing this problem is to consider the impulse response of the channel. An infinite bandwidth channel passes an impulse undistorted (perhaps with just a pure time delay). As the bandwidth starts to decrease, the impulse response "spreads out". If we consider the

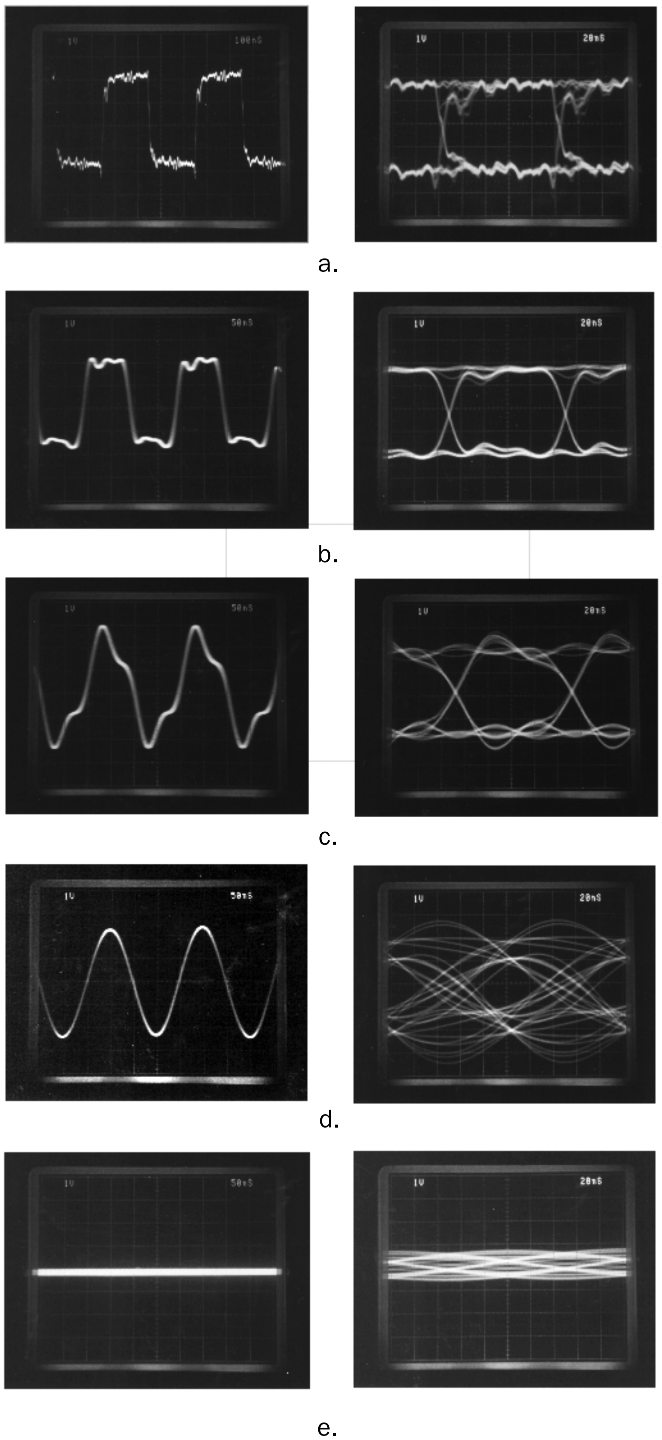


Figure 4. Scope waveforms vs. time (L) and eye diagrams (R).

bit signal to be a stream of impulses, inter-symbol interference (ISI) starts to appear; the impulses start to interfere with one-another as the response from one pulse extends into the next pulse. The voltage seen at the Receive end of the wire is no longer a simple function of the bit sent by the transmitter at time t_1, but is also dependent on the previous bit (sent at time t_0), and the following bit (sent at time t_2).

Figure 4 illustrates what might be seen with an oscilloscope connected to the Receive end of the line in the simple noisy communications system described above for the case where the bandwidth restriction is a first-order lag (single R-C). Two kinds of response are shown, a portion of the actual received pulse train and a plot triggered on each cycle so that the responses are all overlaid. This latter, known as an “eye” diagram, combines information about both bandwidth and noise; if the “eye” is open sufficiently for all traces, 1s can be easily distinguished from 0s. In the adequate bandwidth case of Figure 4a, one can see unambiguous 1s, 0s, and sharp transitions from 1 to 0. As the bandwidth is progressively reduced, (4b, 4c, 4d, 4e), the 1s and 0s start to collapse towards one another, increasing both timing- and voltage uncertainty. In reduced-bandwidth and/or excessive-noise cases, the bits bleed into one another, making it difficult to distinguish 1s from 0s; the “eye” is said to be *closed* (4e).

As one would expect, it is much easier to design a circuit to recover the bits from a signal like 4a than from 4d or 4e. Any misplacement of the decision element, either in threshold level or timing, will be disastrous in the bandlimited cases (d, e), while the wideband case would be fairly tolerant of such errors. As a rule of thumb, to send a pulse stream at rate F_S, a bandwidth of at least $F_S/2$ will be needed to maintain an open eye, and typically wider bandwidths will be used. This *excess bandwidth* is defined by the ratio of actual bandwidth to $F_S/2$. The bandwidth available is typically limited by the communication medium being used (whether 2000 ft. of twisted-pair wire, 10 mi of coaxial cable etc.), but it is also necessary to ensure that the signal processing circuitry in the transmitter and receiver do not limit the bandwidth.

Signal processing circuitry can often be used to help mitigate the effects of the intersymbol interference introduced by the bandlimited channel. Figure 5 shows a simplified block diagram of a bandlimited channel followed by an equalizer, followed by the bit “slicer”. The goal of the equalizer is to implement a transfer function that is effectively the inverse of the transmission channel over a portion of the band to extend the bandwidth. For example, if the transmission channel is acting as a low pass filter, the equalizer might implement a high-pass characteristic, such that a signal passing through the two elements will come out of the equalizer undistorted over a wider bandwidth.

Though straightforward in principle, this can be very difficult to implement in practice. To begin with, the transfer function of the transmission channel is not generally known with any great precision, nor is it constant from one situation to the next. (You and your neighbor down the street have different length phone wires running back to the phone company central office, and will therefore have slightly different bandwidths.) This means that these equalizers usually must be tunable or adaptive in some way. Furthermore, considering Figure 5 further, we see that a passive equalizer may flatten out the frequency response, but will also attenuate the signal. The signal can be re-amplified, but with a probable deterioration in signal-to-noise ratio. The ramifications of that approach will be considered in the next section. While they are not an easy cure-all, equalizers are an important part of many communications systems, particularly those seeking the maximum possible bit rate over a bandwidth-constrained channel. There are extremely sophisticated equalization schemes in use today, including decision feedback equalizers which, as their name suggests, use feedback from the output of the decision element to the equalization block in an attempt to eliminate trailing-edge intersymbol interference.[1]

[1]The field of disk-drive *read*-channel design is a hotbed of equalizer development in the ongoing struggle to improve access specs.

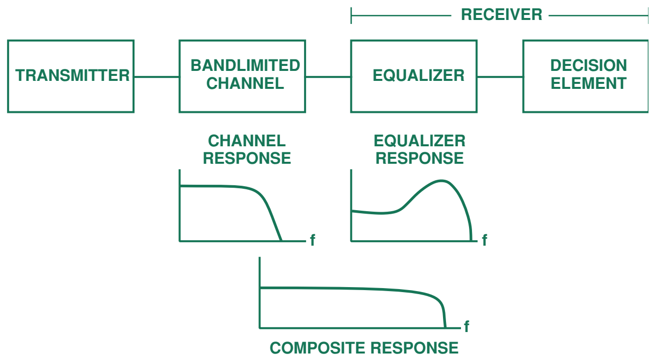


Figure 5. Channel equalization.

Multi-level symbols—sending more than one bit at a time: Since the bandwidth limit sets an upper bound on the number of pulses per second that can be effectively transmitted down the line, one could decide to get more data down the channel by transmitting two bits at a time. Instead of transmitting a “0” or “1” in a binary system, one might transmit and receive 4 distinct states, corresponding to a “0” (00), “1” (01), “2” (10), or “3” (11). The transmitter could be a simple 2-bit DAC, and the receiver could be a 2-bit ADC. (Figure 6). In this kind of modulation, called pulse-amplitude modulation (PAM), additional information has been encoded in the amplitude of the bit stream.

Communication is no longer one bit at a time; multiple-bit words, or *symbols*, are being sent with each transmission event. It is then necessary to distinguish between the system’s bit rate, or number of bits transmitted per second, and its symbol rate, or baud rate, which is the number of *symbols* transmitted per second. These two rates are simply related:

$$\textit{bit rate} = \textit{symbol rate} \text{ (baud)} \times \textit{bits/symbol}$$

The bandwidth limitations and intersymbol interference discussed in the last section put a limit on the realizable symbol rate, since they limit how closely spaced the “transmission events” can be in time. However, by sending multiple bits per symbol, one can increase the effective bit rate, employing a *higher-order modulation* scheme. The transmitter and receiver become significantly more complicated. The simple switch at the transmitter has now been replaced with a DAC, and the single comparator in the receiver is now an A/D converter. Furthermore, it is necessary to use more care to properly scale the amplitude of the received signal; more information is needed than just the sign. Making the simplifying assumption that the A/D converter, representing the receiver, is implemented as a straight flash converter, it is manifest that the receiver hardware complexity grows exponentially with the number of bits per symbol: one bit, one comparator; two bits, three comparators; three bits, seven comparators, etc. Depending on the particular application, circuit cost should not quite increase exponentially with bits per symbol, but it generally will be a steeper-than-linear increase. However, hardware complexity is not the only limiting factor on the number of bits per symbol that can be transmitted.

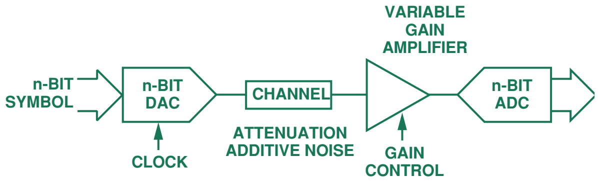


Figure 6. Simplified PAM transmitter/receiver.

NOISE LIMITATIONS

Consider again the simple case of one-bit-per-symbol PCM modulation. Assuming that 1 V is used to send a "1", and –1 V to send a "0", the simple receiver (Figure 3) is a comparator with its decision threshold at 0 V. In the case where the bit being received is a "0", and the channel bandwidth is wide enough so that there is virtually no intersymbol interference, in a noiseless environment, the voltage at the receiver is expected to be –1 V. Now introduce additive noise to the received signal (this could come from any number of sources, but for simplicity and generality, assume it to be gaussian white noise that could correspond to thermal noise). At the moment the decision element is applied, the voltage at the comparator will differ from –1 V by the additive noise. The noise will not be of real concern unless it contains values that will push the voltage level above 0 V. If the noise is large enough (and in the right sign) to do this, the decision element will respond that it has received a "1", producing a bit error. *In the eye diagram of Figure 4d, the noise would produce occasional closures of the "eye".*

If the system is modified to send a 4-bit (16-level) symbol, with the same peak-to-peak voltage, –1 V corresponds to "0" (0000), and +1 V corresponds to "15" (1111). Now the incremental threshold between "0" and the next higher level, "1", is much smaller: 16 distinct states must fit into the 2-V span, so the states will be roughly 125 mV apart, center-to-center. If the decision thresholds are placed optimally, the "center" of a state will be 62.5 mV away from adjacent thresholds. In this case, >62.5 mV of noise will cause a "bit error". If the initial assumption holds and the additive noise is gaussian in nature, one can predict from the rms noise value how often the noise will exceed this critical value. Figure 7 shows the error threshold of 62.5 mV for the probability density functions of two different rms noise values. From this, one can predict the bit error rate, or how often the received data will be interpreted incorrectly for a given transmitted bit rate.

Special care must be taken as to how the data is encoded: if the code 1000 is one threshold away from the code 0111, a small noise excursion would actually cause all 4 bits to be misinterpreted. For this reason, Gray code (which changes only one bit at a time between adjacent states—e.g., 00, 01, 11, 10) is often used to minimize the bit error impact from a misinterpretation between two adjacent states.

So, despite the increase in bit rate, there are limitations to using higher-order modulation schemes employing more bits per symbol: not only will the hardware become more complex, but, for a given noise level, bit errors will be more frequent. Whether the bit error rate is tolerable depends very much on the application; a digitized voice signal may sound reasonable with a bit error rate of 10^{-5}, while a critical image transmission might require 10^{-15}.

Bit errors can be detected and corrected by various coding and parity schemes, but the overhead introduced by these schemes eventually consumes the additional bit capacity gained from increasing the symbol size. One way to try to increase the signal-to-noise ratio (SNR) is to increase transmitted power; for example, increase signal amplitude from 2 V peak-to-peak to 20 V peak-to-peak, thereby increasing the "error threshold" to 625 mV. Unfortunately, increasing the transmitted power generally adds to the cost of the system. In many cases, the maximum power that can be transmitted in a given channel may be limited by regulatory authorities for safety reasons or to ensure that other services using the same or neighboring channels are not disturbed. Nevertheless, in systems that are straining to make use of all available capacity, the transmit power levels will generally be pushed to the maximum practical/legal levels.

Voltage noise is not the only kind of signal impairment that can degrade the receiver performance. If timing noise, or jitter, is introduced into the receiver "clock," the decision "slicer" will be applied at sub-optimal times, narrowing the "eye" (Figs. 4a-4d) horizontally. Depending on how close the channel is to being band-limited, this could significantly decrease the "error threshold," with increased sensitivity to voltage noise. Hence, SNR must be determined from the combination of voltage-domain and time-domain error sources.

This is the first in a series of articles offering an introduction to topics in communications. In the next issue, we'll discuss various modulation schemes and ways of multiplexing multiple users in the same channel.

For Further Reading: This article scratches the surface of a very complex field. If your appetite for information has been whetted, here are a few suggested texts (bibliographies within these books will fan out to a wider list):

Electronic Communication Systems—a complete course, 2nd edition, by William Schweber. Englewood Cliffs, NJ: Prentice Hall ©1994. A good basic introduction to communications fundamentals, with an emphasis on intuitive understanding and real-world examples. No more than one equation per page.

Digital Communication (2nd edition), by Edward Lee and David Messerschmitt. Norwell, MA: Kluwer Publishing, ©1994. A more comprehensive and analytical treatment of digital communications.

Wireless Digital Communications: Modulation and Spread-Spectrum Applications, by Dr. Kamilo Feher. Englewood Cliffs, NJ: Prentice Hall, ©1995. A fairly rigorous analysis of different wireless modulation schemes, with insights into particular strengths and weaknesses of each, and discussion of why particular schemes were chosen for certain standards.

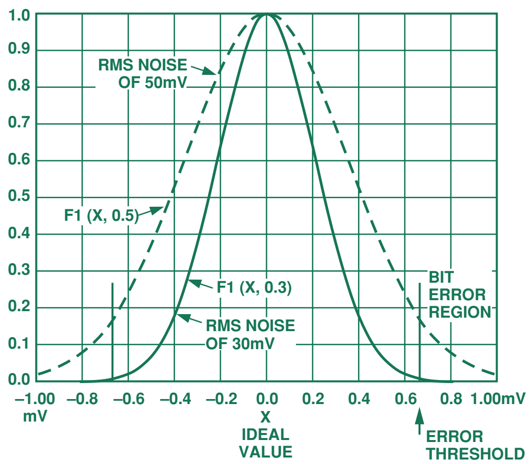


Figure 7. Ideal signal plus noise vs. error threshold: Threshold at 2 σ, and threshold at 1 σ.

Selecting Mixed-Signal Components for Digital Communication Systems-II

by Dave Robertson

Part I, in Analog Dialogue 30-3, provided an introduction to channel capacity and its dependence on bandwidth and SNR. This installment discusses a variety of modulation schemes, and the demands each places on signal processing components.

Digital Modulation Schemes: The first installment in this series showed how limitations of SNR and bandwidth constrain the bit capacity of a communication system that uses pulse amplitude to convey bit information. As a way to encode digital bits, pulse amplitude is one of many modulation schemes used in digital communications systems today; each has advantages and disadvantages. We define below some of the more common modulation types, highlighting their basic principles, and noting the typical component specifications that impact performance. The textbooks listed on page 12 can provide more complete descriptions of these modulation schemes.

PAM—pulse amplitude modulation: (discussed earlier) encodes the bit values in the amplitude of a stream of pulses sent down the channel. The theoretical bandwidth (in Hz) required is at least 1/2 the symbol rate; practical implementations use more bandwidth than this. PAM is typically a baseband modulation scheme: it produces a signal whose spectral content is centered on dc. The simplest case, where each symbol represents the presence or absence of a single bit, is called pulse-code modulation.

Since the bit value is encoded in the amplitude of the signal, gain and offset of the components in the signal path affect system performance. Higher-order modulation schemes using more than two levels will need correspondingly better amplitude accuracy in the system components. Offset, which can shift the signal from the proper level threshold, creating a biased tendency to misinterpret bits high (or low) in the presence of noise, should be controlled. Bandwidth of the components is also an important consideration. As shown earlier, limited bandwidth produces undesirable intersymbol interference. Filtering may be used to carefully control the bandwidth of a transmitted signal, but signal processing components should not unintentionally limit the bandwidth. Generally, components should have enough bandwidth so that the channel itself is the band-limiting factor, not the signal processing circuitry.

AM—amplitude modulation: closely related to PAM, straight AM represents transmitted data by varying the amplitude of a fixed-frequency carrier, usually a sine wave, of designated frequency, f_C. Conceptually, this can be produced by taking the basic PAM signal, band-limiting it to reduce harmonic content, and multiplying it by a carrier at a fixed frequency, f_C. The result is a double-sideband signal, centered on the carrier frequency, with bandwidth twice that of the bandlimited PAM signal.

As with the PAM case, components in the signal chain must be selected to maintain amplitude integrity within the band centered around the carrier frequency, f_C. In this case, analog components may be evaluated based on their linearity, THD (total harmonic distortion) or SFDR (spurious free dynamic range) performance at f_C. For multi-bit symbols with numerous distinct amplitude levels, noise may be an important consideration in component specification.

FM/FSK—frequency modulation/frequency shift keying: We've shown that amplitude modulation schemes (including PAM) can be very sensitive to voltage noise and distortion. Alternatively, information can be encoded in the *frequency* of the sine wave being sent, so that signal attenuation or other amplitude-based disturbance would not tend to corrupt the recovered data (FM radio's resistance to static and signal degradation compared to AM are well-known analog examples; similar principles apply for digital transmission). In a simple binary case of one-bit-per-symbol, the transmitted signal would shift between frequencies f_0 ("0") and f_1, ("1"), on either side of an average, or carrier, frequency—*frequency shift keying* (FSK). It is important to note that the transmitted signal bandwidth actually spreads over a larger bandwidth than just the span between f_0 and f_1, because the speed of transitioning between the two frequencies generates additional spectral content. To simplify receiver design, it is desirable that the symbol rate be substantially less than the difference between f_0 and f_1; this makes changes in frequency easier to detect.

Frequency modulation significantly reduces the sensitivity to amplitude errors in the signal path. Since all the useful information is held in the frequency domain, many FSK receivers feature a *limiter*, a high-gain circuit designed to convert a variable-amplitude sinusoidal signal to a more nearly constant-amplitude square wave, desensitizing the circuit to component non-linearities and making it easier for subsequent processing circuitry to detect the frequency of the signal (even by counting crossings within a given time interval). Signal bandwidth is at least as important as with AM: intersymbol interference still results from insufficient processing bandwidth. Because a carrier frequency must be processed, the required bandwidth is probably significantly larger than PAM modulation of the same data. These systems are typically more sensitive to timing errors, such as jitter, than to voltage noise.

PM/QPSK—phase modulation/quadrature phase shift keying: phase and frequency are closely related mathematically; in fact, phase is the integral of frequency (e.g., doubling frequency causes phase to accumulate at twice the original rate). In PM, the signal is encoded in the phase of a fixed-frequency carrier signal, f_C. This can be accomplished with a direct digital synthesizer (DDS) that generates a digital sine wave, whose phase is modulated by a control word. A D/A converter restores the sine wave to analog for transmission.

Another example of how a 2-bit phase-modulated symbol may be derived can be seen with two equal sinusoidal components at the same frequency: in-phase (I) and quadrature (Q), 90° apart, each representing digital "1" if non-inverted, "0" if inverted (shifted 180°). When they are added, their sum is a single wave at the same frequency with 4 unique phases, 90° apart (i.e., 45°, 135°, 225°, and 315°), corresponding to the phases of the I and Q waves. Figure 1 is a "unit-circle" or "satellite" plot, graphically representing these combinations. Systems embodying this principle of phase modulation are often referred to as quadrature phase-shift keying (QPSK). As with FM, the relationship between the bandwidth of the transmitted spectrum and the symbol rate is fairly complicated. There are several variations of phase modulation, including DQPSK (differential QPSK). These types of modulation schemes are popular in difficult environments such as cellular telephony, because the phase information can be maintained in the presence of noise and the distortion introduced by power amplifiers.

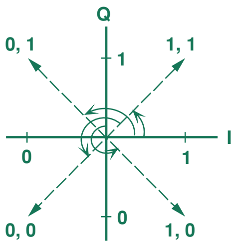


Figure 1. 2-bit QPSK phases.

As with FSK, components for PSK systems are typically selected based on bandwidth and other frequency domain specifications. Limiters may be used to eliminate amplitude noise. Timing errors, including jitter, effectively become "phase noise," making it more difficult to properly interpret the received signal. Modulator/demodulator units may be implemented in a quadrature arrangement, where the I and Q components are separated and processed separately through part of the signal chain. Here amplitude- and phase match between the I and Q paths are important specifications, since any mismatches map to an effective phase error.

QAM—Quadrature Amplitude Modulation: Returning to Figure 1, the representation of the four different phases of the carrier in a QPSK system, note that each of the phases also has an amplitude that is the vector sum of the I and Q amplitudes; since the amplitudes are equal, the amplitudes of the vector sums are equal. More bits per symbol could be transmitted if, instead of just two levels for I and Q, they were further quantized; then, by adding the differing amounts of sine (I axis) and cosine (Q axis) together, the combination in vector sums would modulate both amplitude and phase. Figure 2a shows the use of 2-bit quantization of both I and Q to realize 16 unique states of the carrier in each symbol, allowing transmission of 4 bits per symbol. This modulation could be produced by varying the phase and amplitude of the generated carrier directly using, for example, direct digital synthesis. More commonly, amplitude-modulated I and Q (sine and cosine) versions of the carrier are combined.

Hence the term *quadrature amplitude modulation* (QAM): the two quadrature versions of the carrier are separately amplitude modulated, then combined to form the amplitude- and phase-modulated resultant. The plot in Figure 2a, showing the various possible combinations of I and Q, is referred to as a "constellation." Note that very large constellations can, in concept, be used to represent many bits per symbol, with a required bandwidth similar to simple QPSK of the same symbol rate. The points of the constellation represent the transmitted signal and the expected value of the received signal; but noise or distortion will displace the received signal from its ideal position; it can be misinterpreted as a different constellation point if the error is large.

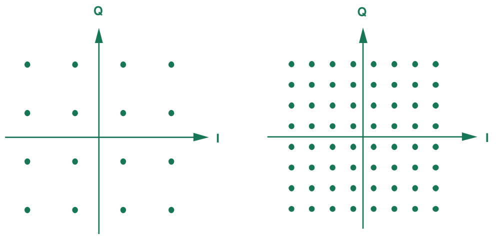


Figure 2. QAM constellations. a) 4 bits: 2-bit I and 2-bit Q. b) 6 bits: 3-bit I and 3-bit Q.

Figures 2a and 2b compare the 16-point constellation (2 bits I and Q) to a 64-point constellation (3 bits I and Q). At similar transmitted power levels the constellation points for the 6-bit case are twice as close together, therefore the "error threshold" is 1/2 as large and, for a given bit error rate, a 6-dB (approximately) better signal-to-noise ratio is required. The table shows typical SNR requirements for various sizes of QAM constellations to realize a 10^{-7} bit error rate. Note that binary I & Q information can be encoded [e.g., Gray code] so that points representing adjacent or nearby transmitted signal levels have similar bit patterns. In this way, misinterpreting a constellation point for one of its neighbors would corrupt only 1 or 2 bits of a multi-bit symbol.

Bits/Symbol (I, Q)	QAM Constellation Size	Required SNR
2 (1, 1)	4 (QPSK)	14.5 dB
3 (1, 2)	8	19.3 dB
4 (2, 2)	16	21.5 dB
5 (2, 3)	32	24.5 dB
6 (3, 3)	64	27.7 dB
7 (3, 4)	128	30.6 dB
8 (4, 4)	256	33.8 dB
10 (5, 5)	1024	39.8 dB
12 (6, 6)	4096	45.8 dB
15 (7, 8)	32768	54.8 dB

Here are some of the important specifications for components selected for QAM signal processing. *Bandwidth* should be sufficient to handle the carrier frequency, plus enough frequencies within the band to avoid introducing intersymbol interference. *Total harmonic distortion* (THD) at the carrier frequency is an important consideration, since distortion will tend to corrupt the amplitude information in the carrier. *Jitter* should be minimized to ensure that the phase information can be properly recovered. *Matching* of amplitude and phase between the I and Q processing blocks is important. Finally, *noise* (quantization and thermal) can be an important consideration, particularly for high-order constellations. Wherever practical, components should be selected to ensure that the channel itself is the noise-limiting part of the system, not the components of the signal processing system. QAM can be used to transmit many bits per symbol, but the trade-off is increased sensitivity to non-idealities in the communications channel and the signal processing components.

This provides a quick review of the basic modulation schemes. The many variations, combinations and enhancements of these approaches seek to deal with the characteristics of particular applications and the shortcomings of the various transmission techniques. They offer trade-offs between spectral efficiency, robustness, and implementation cost.

The next part of this series will explore multiplexing schemes and the variety of dynamic range requirements encountered in digital communications systems.

REFERENCES

Electronic Communication Systems—a complete course, 2nd edition, by William Schweber. Englewood Cliffs, NJ: Prentice Hall ©1996. A good basic introduction to communications fundamentals, with an emphasis on intuitive understanding and real-world examples. No more than one equation per page.

Digital Communication (2nd edition), by Edward Lee and David Messerschmitt. Norwell, MA: Kluwer Publishing, ©1994. A more comprehensive and analytical treatment of digital communications.

Wireless Digital Communications: Modulation and Spread-Spectrum Applications, by Dr. Kamilo Feher. Englewood Cliffs, NJ: Prentice Hall, ©1995. A fairly rigorous analysis of different wireless modulation schemes, with insights into particular strengths and weaknesses of each, and discussion of why particular schemes were chosen for certain standards.

Selecting Mixed Signal Components for Digital Communication Systems–III: Sharing the Channel

by Dave Robertson

Part I provided an introduction to the concept of channel capacity—and its dependence on bandwidth and SNR; Part II gave a brief summary of different types of modulation schemes. This segment discusses communications signal-processing issues that arise when multiple users share the same transmission medium .

SHARING THE CHANNEL

Selection of an appropriate modulation scheme is only part of the problem of defining a communications network. In most cases, the transmission medium must accommodate signals from more than a single transmitter. The most obvious case of such multiple use is the airwaves; they must carry a variety of wireless traffic, from broadcast radio and television, to cellular telephony, to CB and short-wave radio. Even a simple twisted-pair telephone wire, which represents a dedicated line between the phone company central office and a user, must carry both incoming and outgoing voice and data during a call.

In most cases, the key to effective multiplexing of independent transmissions is proper observance of "live and let live" protocols, enabling the effective transmission of the desired message without undue interference to other transmissions. There are a variety of approaches towards sharing a communications medium among multiple users; each has its own requirements affecting component selection. Most of these schemes are usable for both analog and digital communications; but the flexibility of time compression, and other features available in digital communications, opens up more options.

TDMA—time-division multiple access: perhaps the most obvious way of sharing the communications channel is to "take turns": only one transmitter at a time is allocated the channel. There must of course be some sort of protocol to establish who has the transmission privilege, when, how often, and for how long. A simple example is the walkie-talkie user's employment of the word "over" to indicate the termination of a transmission stream and freeing up the communications channel for other users to transmit.

A more formal arrangement is usually desirable, especially when each user is to be allotted a very brief—but repetitive—participation. An overall time period can be divided into designated "slots", with each of the transmitters assigned a different time slot for transmission (Figure 1). This kind of scheme requires synchronization of all the transmitters, plus a "supervisor" to assign time slots as new transmitters want to enter the channel—and to keep track of slots vacated. Some "overhead" space must be provided to allow for transitions between transmitter time slots; the better the synchronization, the less time lost to these transition periods. Time multiplexing also means that the stream of data from a given transmitter is not continuous, but in bursts. To represent a continuous conversation (say in a cellular phone call), the digitized information acquired during the period between transmissions must be time compressed, transmitted in a short burst, then expanded in the receiver to form a transparently continuous message.

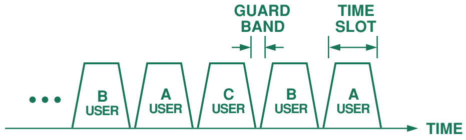


Figure 1. Illustration of time multiplexing, showing guard bands.

The analogy of a panel discussion is sometimes used to illustrate the nature of TDMA. A participant who interrupts out of turn or rattles on and on endlessly commits "violations of the TDMA protocol." The European GSM digital cellular telephony standard makes use of TDMA; each channel carries eight phone calls simultaneously in a repeated transmitting sequence of eight time slots.

Component selection for TDMA systems must involve careful consideration of bandwidths and settling times; long time constants of components with insufficient bandwidth will tend to cause signals to "bleed into" an adjacent user's time slot.

FDMA—frequency-division multiple access: anyone who receives TV or radio broadcasts at home is familiar with an example of frequency-division multiple access. In this case, multiple transmitters can simultaneously transmit without interference (at a given power level in a given geographical area) by keeping each frequency in their transmissions within a designated frequency slot. The receiver determines which channel is to be recovered by tuning to the desired frequency slot. It is important that each transmitter's frequency limits be strictly observed; any transgressions would create interference in the neighboring channels. (Figure 2)

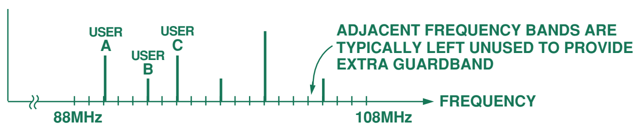


Figure 2. Illustration of frequency multiplexing, showing unused frequency bands to provide separation.

Using the conversational analogy, this might be like providing a set of booths, one for each speaker; if they speak quietly enough, all "transmitters" can broadcast simultaneously, and a listener may "tune in" by listening at the desired booth.

Almost all wireless applications are subject to frequency band constraints; national and international regulatory bodies, e.g., the FCC in the United States, license the transmitter to specific frequencies or restrict its class to specific bands. Wired applications like cable TV also use frequency separation to allow simultaneous transmission of hundreds of channels (both analog and digital).

Keeping within the specified frequency constraints has numerous ramifications for component selection. For example, some component in the system will be used as a precise frequency reference. It could be an absolute frequency reference, like a crystal, or it might contain a circuit that receives and "locks on" to an external reference frequency. Components in the transmission path must have carefully limited spectral content; this can be done through filtering—but it is also necessary to control component linearity, so as not to generate incidental "out of band" harmonics and other spurious frequency components.

CDMA—Carrier Division Multiple Access—Continuing the conversation analogies, suppose that 10 people are trying to carry on 5 simultaneous one-on-one conversations in a small room. Suppose further that one pair agrees to converse in English, another in French, the others in Chinese, Finnish, and Arabic—and all are monolingual. If you were a member of the English speaking pair, you would hear a din of background "babble", but the only intelligible information would be in English. So it's easy to see that all 5 conversations could take place simultaneously in the same room (though in practice, everyone would probably get a headache).

This is essentially a description of the underlying idea of carrier division multiple access. All users transmit and receive over the same frequency band, but each pair is assigned a unique code sequence. The digital bit stream you wish to send is modulated with this unique code sequence and transmitted. A receiver will receive the combined modulated bit streams of all the transmitters. If the receiver demodulates this composite signal with the same unique code, it essentially performs a cross-correlation operation: the bit stream that was modulated with the same code sequence will be recovered; all the other transmitted signals that were modulated with different codes will be rejected as "noise".

Modulation with the special code tends to spread the spectrum of the initial digital bit stream over a much wider bandwidth, which helps improve its immunity from interference. Despite this spectral spreading, spectral efficiency can be maintained, because multiple users can share the same bandwidth. Adding more users simply leads to the appearance of increased noise in the channel

Examples of CDMA systems include the IS95 Digital Cellular standard in the US and numerous military "spread spectrum" communications applications (an additional advantage of modulating the transmitted signal with a unique signal is that it is essentially encrypted; a receiver cannot recover the transmitted message without the unique modulation sequence). Though CDMA systems involve greater digital complexity, the performance requirements for their analog components are reduced. However, because multiple transmitters will be broadcasting in the same channel at the same time, it is usually desirable to minimize the contributions to background and spurious noise by transmitter components.

SDMA—Space Division Multiple Access: Returning to the conversation analogy, another way to carry on simultaneous one-on-one conversations in the same room is to move to opposite corners of the room and speak in relatively hushed tones. This captures the spirit of SDMA. In wireless applications, signal strength falls off rapidly with increasing distance from the transmitting antenna. At a great enough distance, the signal can be considered to have faded completely, from which point a new transmitter could reuse the same frequency or time slot for a different signal (Figure 3). In broadcast radio, the same frequency can be reused in different cities, provided that they are far enough apart.*

*The attenuation of signal with distance is a strong function of frequency: The higher the transmitter frequency, the faster the rolloff.

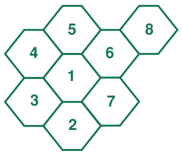


Figure 3. Illustration of geographical multiplexing, showing honeycomb of cells with base-station antennas at centers.

The concept of channel re-use with distance underlies the term "cellular telephony." Cell size is determined by the area of effective coverage by a given transmitter, and the same frequencies can be reused in other cells. In practice, however, patterns are designed so that adjacent cells will not re-use the same frequencies. Conventional antennas radiate in all directions, producing a circular coverage area and the "honeycomb" cellular pattern in Figure 3. Modern technology has added new dimensions to the concept of SDMA with the development of focused, or beam steering antennas. Phased-array technology can create a focused, directional signal transmission pattern aimed at either an individual target receiver or a particular target area (e.g., a specific highway at rush hour). This can allow more rapid re-use of frequency spectrum, thereby effectively increasing total capacity for wireless applications.

Advanced digital communications systems use combinations of these multiplexing schemes to effectively pack as much capacity as possible into the available transmission channel. For example, GSM cellular phones use TDMA, FDMA and SDMA to allocate traffic. Even many wired applications make use of TDMA and FDMA protocols. Although these multiplexing arrangements typically add to the system complexity, the effective increase in channel capacity more than offsets increases in component cost.

THE NEAR/FAR PROBLEM

In previous installments, we have discussed the impact of error rate and modulation scheme on the required dynamic range in a digital communications system. However, in many applications, the *multiplexing* arrangements create the ultimate demands on dynamic range in the communications receiver.

In any application, the strength of the received signal is a function of the strength of the transmitted signal, the distance from the transmitter, and numerous environmental factors relating to the transmission medium (be it wireless or wired). Most communications systems are designed to work over a variety of distances, and so have to be designed to accommodate a large variation in power of the received signal.

Consider, for example, a cellular telephony application. The receiver circuitry must be designed to recover the weak signal resulting from a transmission while at the very edge of the "cell". This capability to recover weak signals is often referred to as a receiver's *sensitivity*. To recover such weak signals, it seems appropriate to include gain stages in the receive circuitry. Consistent with good, low-noise design practice, one might expect to put the gain as early in the signal path as possible to quickly boost the signal above the noise floor of subsequent stages.

Unfortunately, this same receiver must also be capable of receiving the signal transmitted by a user standing directly under the base station's antenna. In the case of GSM, for example, this signal can be up to 90 dB stronger than the weakest signal. If the receiver has too much gain in the signal path, the strong signal can saturate the gain stages. For modulation schemes that include amplitude information (including AM and QAM), this will essentially destroy the signal. Phase and frequency modulation approaches may be more tolerant of this clipping, depending on the circumstances. (The clipping will still create distortion products which are sufficient to cause problems, even in phase-modulation schemes.)

A basic approach to addressing the near/far dynamic range problem is to use variable/programmable gain stages in the receive signal path. Automatic gain control (AGC) allows the gain to be adjusted in response to the strength of the received signal. An important design consideration, though, is how rapidly the gain needs to be adjusted. For example, in ADSL (asymmetric digital subscriber line—see sidebar) modems, the received signal strength changes as outdoor temperature changes affect the line impedance, so time constants of minutes would be tolerable. On the other hand cellular phone receivers must be designed to track the signals from fast moving vehicles that may be moving into or emerging from the shadows of buildings or other signal obstacles, so very rapid gain changes are required. TDMA systems put an additional demand on gain-ranging circuitry, because the near/far signals could be located in adjacent TDMA time slots; in this case, the circuitry would have to change gains and settle in the transition period between time slots.

FDMA systems offer a different kind of near/far challenge. Here, the worst case to consider is recovery of a weak signal in a frequency slot next to strong signal (Figure 4). Since both signals are present simultaneously as a composite at the input of a gain stage, the gain is set according to the level of the stronger signal; the weak signal could be lost in the noise floor (in this case, the noise floor could be thermal noise or quantization noise of an A/D converter.)

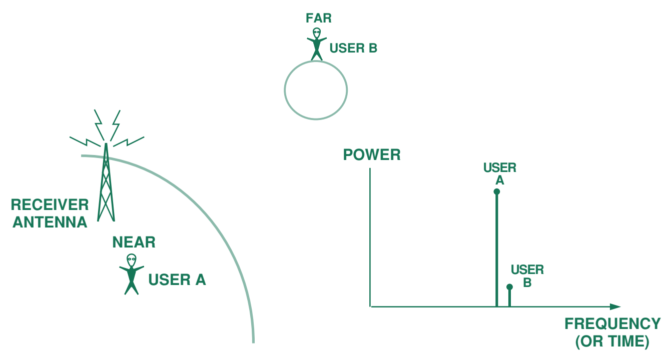


Figure 4. Near-far effect calls for the ability to handle wide dynamic range between adjacent channels.

Even if subsequent stages have a low enough noise floor to provide dynamic range to recover the weak signal, there must also be a very stringent constraint on the dynamic linearity of the gain stage; harmonics or other spurious responses of the strong signal that wind up in the wrong frequency bin could easily obliterate the weaker desired signal. To reduce this interference problem, most FDMA systems attempt to filter out unwanted signals early in the receive circuitry. The ability to discriminate against unwanted signals in adjacent frequency bands is usually referred to as a receiver's *selectivity*.

Most radio designs feature a cascaded series of filters and gain stages (some of which may be variable) to remove/attenuate strong interferers, then amplify the desired signal to a level that can be readily demodulated. Wideband radios, however, attempt to simultaneously recover all the signals in one receiver; they cannot use analog discrimination filters; accordingly, wideband receivers typically have the most stringent requirements on dynamic range in their analog circuitry and converters. Interestingly enough, even applications where you think you have the communications channel to yourself can suffer from simultaneous near/far signals. For example, in ADSL modems, the system must be designed for the scenario where the near-end echo (leakage from the local transmitter) appears as an interfering signal that is actually up to 60 dB stronger than the desired receive signal.

In CDMA systems the near/far problem is a little more difficult to describe. Since all signals are simultaneously transmitted in the same frequency space, filtering cannot be used to discriminate against unwanted signals (though it is still used to eliminate signals in adjacent bands). CDMA employs demodulation using a carrier unique to the desired signal to extract the desired from the unwanted signals; signals modulated with a different carrier appear as background noise. The ability to successfully recover the signal is set by the total noise energy—including that of the other carriers—in the band. Since filtering can't be used to discriminate, the best situation to strive for is to have all signals arrive at the base-station antenna at equal power. To achieve this, many CDMA systems communicate the received power levels back to the transmitters so that power of the individual signal components may be adjusted to equalize power levels at the base-station receiver. To help reduce their near/far problem, TDMA systems could also use this kind of power control, though it tends to require a more-sophisticated (i.e., costly) handset.

ASYMMETRIC DIGITAL SUBSCRIBER LINE

ADSL is one of the many technologies competing to bring broadband digital services into the home. The concept underlying ADSL is to take advantage of the twisted-pair wires that already provide almost universal telephone service to homes in the United States. Other services providing a two-way flow of information, such as ISDN (integrated services digital network), require an additional, dedicated wire to provide service.

ADSL uses frequency-division multiplexing (FDM) to convey modulated digital information in the frequency space between 20 kHz and 1.2 MHz, above the frequency space occupied by conventional voice traffic. This frequency separation allows an ADSL modem to operate without disturbing a phone call occurring at the same time—an extremely important feature.

The ANSI standard for ADSL provides for simultaneous upstream (outgoing from the home) and downstream (incoming to the home) transmission using either FDM (separating the upstream and downstream signals in frequency) or echo cancelling. Echo cancelling uses sophisticated signal processing (analog, digital, or both) to separate the strong transmitted signal from the weaker received signal, passing only the received signal to the demodulator. Using the conversational model, this is analogous to a person who can effectively talk and listen at the same time.

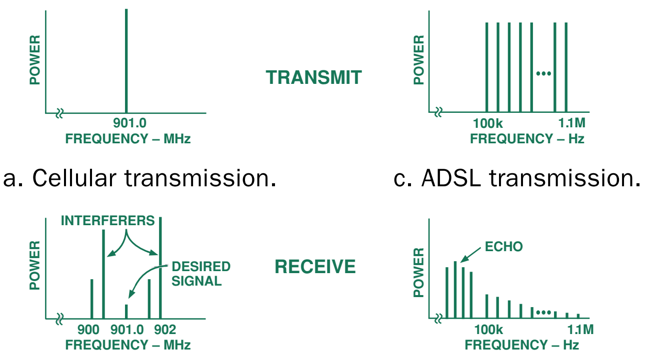
Selecting Mixed Signal Component for Digital Communications Systems

IV. Receiver Architecture Considerations

by Dave Robertson

Part I introduced the concept of channel capacity and its dependence on bandwidth and SNR; part II summarized briefly different types of modulation schemes; and part III discussed approaches to sharing the communications channel, including some of the problems associated with signal-strength variability. This installment considers some of the architectural trade-offs used in digital communications receiver design for dealing with dynamic range management and frequency translation problems.

System Constraints: In a digital communications system, the function of the receiver circuitry is to recover the transmitted signal and process it for introduction to the demodulator, which then recovers the digital bits that constitute the transmitted message. As the last installment illustrates, obstacles to signal recovery show up as the signal travels through the transmission medium. These "impairments" can include signal attenuation, reflections, distortion, and the introduction of "interferers" (other signals sharing the transmission medium). The nature of the transmission impairments is a strong function of the medium (wireless, coaxial cable, or twisted pair wire), the communications scheme being used (TDMA, FDMA, CDMA, etc.) and the particular circumstances of the transmitter/receiver pair (distance, geography, weather, etc.). In any event, the important receiver design considerations are present to some extent in all receivers, simply to differing degrees. For this discussion, two examples will be used to illustrate the various receiver design issues. Figure 1 illustrates the relevant portions of the signal spectrum at the transmitter outputs and receiver inputs for two very different systems: a GSM cellular telephony application (Figure 1a and 1b) and an ADSL twisted-pair modem application (Figure 1c and 1d).



a. Cellular transmission. c. ADSL transmission.

b. Cellular received signal. d. ADSL received signal.

Figure 1. Transmitted and received spectra.

GSM uses a combination of FDMA (frequency division multiple access) and TDMA (time division multiple access) for multiplexing and a variation of quadrature phase shift keying for modulation. In 1b, the amplitude is significantly reduced—a result of distance from the transmitter. In addition, several strong interfering signals are present—signals from other cellular transmitters in nearby bands that are physically closer to the receiver than the desired transmitter.

The ADSL modem in this example (Figure 1c) uses FDMA to separate upstream and downstream signals, and transmits its signal in a number of separate frequency bins, each having its own QAM (quadrature amplitude modulation) constellation (discrete multi-tone, or DMT modulation). The ADSL signal is attenuated by the twisted pair wire; attenuation is a strong function of frequency. In addition, an "interferer" is present. This might seem anomalous in a dedicated wire system, but in fact the interferer is the duplex (travelling in the opposite direction) signal of the modem leaking back into the receiver. This is generally referred to as *near-end echo*, and for long lines it may be much stronger than the received signal (Figure 1d).

These two examples illustrate critical functions of the receiver processing circuitry:

Sensitivity represents the receiver's ability to capture a weak signal and amplify it to a level that permits the demodulator to recover the transmitted bits. This involves a gain function. As was discussed in Part 3 of this series, signal strength may vary significantly, so some degree of variable or programmable gain is generally desired. The way gain is implemented in a receiver usually requires a tradeoff between noise, distortion, and cost. Low-noise design dictates that gain be implemented as early in the signal chain as possible; this is a fundamental principle of circuit design. When calculating the noise contribution from various noise sources in a system, the equivalent noise of each component is referred to one point in the system, typically the input—referred-to-input (RTI) noise. The RTI noise contribution of any given component is the component's noise divided by the total signal gain between the input and the component. Thus, the earlier the gain occurs in the signal path, the fewer stages there are to contribute significant amounts of noise.

Unfortunately, there are obstacles to taking large amounts of gain immediately. The first is distortion. If the signal is in the presence of large interferers (Figures 1b, 1d), the gain can't be increased beyond the point at which the large signal starts to produce distortion. The onset of distortion is described by a variety of component specifications, including THD (total harmonic distortion), IP3 (third-order intercept point: a virtual measurement of the signal strength at which the power of the 3rd-order distortion energy of the gain stage is as strong as the fundamental signal energy), IM3 (a measure of the power in the 3rd order intermodulation products), and others. For an A/D converter or digital processing, "clipping" at full-scale produces severe distortion. So these strong signals must usually be attenuated before all the desired gain can be realized (discussed below).

Cost is another limiting factor affecting where gain can occur in the signal chain. As a general rule of thumb, high-frequency signal processing is more expensive (in dollars and power) than low frequency or baseband signal processing. Hence, systems that include frequency translation are generally designed to try to implement as much of the required gain as possible at the IF or baseband frequencies (see below). Thus, to optimize the location of gain in the signal path, one must simultaneously trade off the constraints of noise, distortion, power dissipation, and cost.

Specifications used to evaluate gain stages include the gain available (linear ratio or dB) and some description of the noise of the component, either in RTI noise spectral density (in nV/√Hz) or as *noise figure* (basically, the ratio of the noise at the output divided by the noise at the input, for a given impedance level).

Selectivity indicates a receiver's ability to extract or select the desired signal in the presence of unwanted interferers, many of which may be stronger than the desired signals. For FDMA signals, selectivity is achieved through filtering with discrimination filters that block unwanted signals and pass the desired signal. Like gain, filtering is generally easier at lower frequencies. This makes intuitive sense; for example, a 200-kHz bandpass filter implemented at a 1-MHz center frequency would require a much lower Q than the same 200-kHz filter centered on 1 GHz. But filtering is sometimes easier in certain high-frequency ranges, using specialized filter technologies, such as ceramic or surface acoustic wave (SAW) filters.

As noted above, filtering will be required early in the signal path to attenuate the strong interferers. Such filters will need to combine the required frequency response and low noise. Figures of merit for a filter include bandwidth, stop-band rejection, pass-band flatness, and narrowness of the transition band (the region between pass-band and stop-band). Filter response shape will largely be determined by the channel spacing and signal strength variations of the communications channel. Most FDMA cellular standards seek to ease filter requirements by avoiding the use of adjacent frequency channels in the same or adjacent cells, to permit wider transition bands and lower-Q (cheaper) filters.

Part of the selectivity problem is *tuning*—the ability to change the desired channel, since in most applications the signal of interest could be in any one of a number of available frequency bands. Tuning may be accomplished by changing the filter bandpass frequencies, but it is more commonly realized as part of the mixing operation (see below).

Frequency planning (*mixing*): Radio frequencies are selected based on radio transmission characteristics and availability of bandwidth for use for a given service, such as FM radio or cellular telephony. As was noted earlier, signal processing at high radio frequencies tends to be expensive and difficult. Besides, this added trouble seems unnecessary, since in most cases the actual signal bandwidth is at most a few hundred kHz. So most radio receivers use frequency translation to bring the signal carriers down to lower, more manageable frequencies for most of the signal processing. The most common means of frequency translation is a *mixer* (Figure 2).

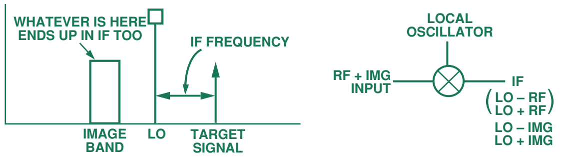


Figure 2. Mixing—the image problem.

Mixing means using a nonlinear operation, usually multiplying the input signal and a reference oscillator signal, to produce spectral images at the sum and difference frequencies. For example: if we "mix" an RF signal at 900 MHz with an oscillator at 890 MHz, the output of the mixer will have energy at 1790 MHz (sum of frequencies) and 10 MHz (their difference). The 10-MHz signal becomes the signal of interest at the 10-MHz *intermediate frequency* (IF), while the sum frequency is easily filtered out. If the oscillator frequency is increased to 891 MHz, it will translate an RF signal at 901 MHz to the IF; hence, channel selection, or tuning, can be realized by varying the oscillator frequency and tuning the output to the IF, using a fixed-frequency bandpass filter.

However, when mixing the 900-MHz RF with an 890-MHz local oscillator (LO), any 880-MHz interference present on the RF signal will also be translated to a difference frequency of 10 MHz. Clearly, any RF signal at the "image" frequency of 880 MHz must be suppressed well below the level of the desired signal before it enters the mixer. This suggests the need for a filter that passes 900 MHz and stops 880 MHz, with a transition band of twice the intermediate frequency. This illustrates one of the trade-offs for IF selection: lower IFs are easier to process, but the RF image-reject filter design becomes more difficult. Figures of merit for mixers include gain, noise, and distortion specifications like those used for gain stages, as well as the requirements on the oscillator signal input.

Other mechanisms of dealing with the image rejection problem are beyond the scope of this short treatment. One worth mentioning, though, because of its widespread use is *quadrature downconversion*. In-phase and quadrature representations of the input signal are mixed separately and combined in a way to produce constructive interference on the signal of interest and destructive interference on the unwanted image frequency. Quadrature mixing requires two (or more) signal processing channels well-matched in both amplitude and frequency response, because mismatches allow the unwanted image signal to leak into the output.

Equalization: Real-world transmission channels often have a more severe impact on signals than simple attenuation. Other channel artifacts include frequency-dependent amplitude and phase distortion, multi-path signal interference (prevalent in mobile/cellular applications), and bandlimiting/intersymbol interference from the receiver processing circuits. Many receiver systems feature "equalization" circuits, which provide signal processing that attempts to reverse channel impairments to make the signal more like the ideal transmitted signal. They can be as simple as a high frequency boost filter in a PAM system or as complicated as adaptive time- and-frequency-domain equalizers used in DMT ADSL systems. As capacity constraints push system architectures towards more complicated modulation schemes, equalization techniques, both in the analog and digital domains, are increasing in sophistication.

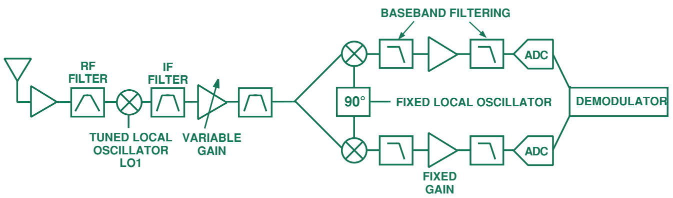
Diversity: In mobile applications, the interference patterns from a mobile transmitter can vary the strength of the signal at the basestation receiver, making the signal difficult or impossible to recover under certain conditions. To help reduce the odds of this occurring, many basestations are implemented with two or more receiving antennas separated by a fraction of the RF wavelength, such that destructive interference at one antenna should represent constructive interference at the other. This diversity improves reception at the cost of duplicating circuitry. Diversity channels need not be closely matched (matching is required for quadrature channels), but the system must have signal processing circuitry to determine which of the diversity paths to select. *Phased-array* receivers take the diversity concept to the ultimate, combining the signal from an array of receivers with the proper phase delays to intentionally create constructive interference between the multiple signal paths, thereby improving the receiver's sensitivity.

Conventional Receiver Design: Figure 3a illustrates a possible architecture for a GSM receiver path, and Figure 3b illustrates that of an ADSL modem. As noted earlier, the task of the receive circuitry is to provide signal conditioning to prepare the input signal for introduction to the demodulator. Various aspects of this signal conditioning can be accomplished with either digital or analog processing. These two examples illustrate fairly traditional approaches, where the bulk of signal processing is done in the analog domain to reduce the performance requirements on the A/D converter. In both examples, the demodulation itself is done digitally. This is not always necessary; many of the simpler modulation standards can be demodulated with analog blocks. However, digital demodulation architectures are becoming more common, and are all but required for complicated modulation schemes (like ADSL).

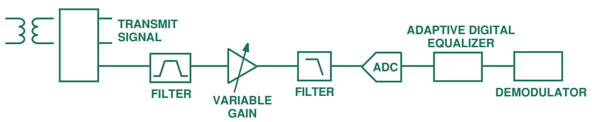
The GSM receiver signal path shown in Figure 3a illustrates the use of alternating gain and filter stages to provide the required selectivity and sensitivity. Channel selection, or tuning, is accomplished by varying the frequency of the first local oscillator, LO1. Variable gain and more filtering is applied at the IF frequency. This is a narrowband IF system, designed to have only a single carrier present in the IF processing. The IF signal is mixed down to baseband, where it is filtered once more and fed to a sigma-delta A/D converter. More filtering is applied in the digital domain, and the GMSK signal is digitally demodulated to recover the transmitted bit stream.

The ADSL receiver has different requirements. Frequency translation is not required, since the signal uses relatively low frequencies (dc to 1.1 MHz). The first block is the "hybrid", a special topology designed to extract the weak received signal from the strong transmitted signal (which becomes an interferer—see Figure 1d). After a gain stage, a filter attempts to attenuate the echo (which is in a different frequency band than the desired signal.) After the filter, a variable-gain stage is used to boost the signal to as large a level as possible before it is applied to the A/D converter for digitization. In this system, equalization is done in both the time and frequency domains before the signal is demodulated. This example shows the equalization taking place digitally (after the A/D converter), where it is easier to implement the required adaptive filters.

New twists—receivers "go digital": Advances in VLSI technology are making more-sophisticated receiver architectures practical; they enable greater traffic density and more flexibility—even receivers that are capable of handling multiple modulation standards. An important trend in this development is to do more and more of the signal processing in the digital domain. This means that the A/D "moves forward" in the signal chain, closer to the



a. GSM receiver.



b. ADSL modem receiver.

Figure 3. Typical receiver architectures.

antenna. Since less gain, filtering and frequency translation is done prior to the A/D, its requirements for resolution, sampling frequency, bandwidth, and distortion increase significantly.

An example of this sophistication in modems is the use of *echo cancellation*. The spectrum of Figure 1d shows the strong interferer that dominates the dynamic range of the received signal. In the case of a modem, this interference is not a random signal, but the duplex signal that the modem is transmitting back upstream. Since this signal is known, signal processing could be used to synthesize the expected echo on the receive line, and subtract it from the received signal, thereby cancelling its interference. Unfortunately, the echo has a strong dependence on the line impedance, which varies from user to user—and even varies with the weather. To get reasonable cancellation of the echo, some sort of adaptive loop must be implemented. This adaptivity is easier to do in the digital domain, but it requires an ADC with sufficient dynamic range to simultaneously digitize the weak received signal and the echo; in the case of ADSL, this suggests a 16 bit A/D converter with 1.1 MHz of bandwidth. (e.g., the AD9260). As a significant reward for this higher level of performance with a sufficiently accurate echo canceller, upstream and downstream data can simultaneously occupy the same frequencies, dramatically increasing the modem's capacity, particularly on long lines.

In the case of GSM, there are various approaches to advanced receivers. As the ADC moves forward in the signal chain, instead of capturing a baseband signal around dc, it has to digitize the IF signal, which would typically be in the range of 70 MHz to 250 MHz. Since the bandwidth of interest is only a few hundred kHz, it is unnecessary (and undesirable) to run the ADC at 500 MHz; instead, undersampling is used. If the ADC is clocked at 20 MHz with the signal of interest at 75 MHz, the signal will alias down to 5 MHz (= 4 × 20 – 75) MHz; essentially, the undersampling operation of the ADC acts like a mixer. As with a mixer, there is an image problem, so signal content at 65 MHz (= 3 × 20 + 5 MHz) and 85 MHz (= 4 × 20 + 5 MHz) would need to be filtered out ahead of the ADC. (An AD6600 dual-channel gain-ranging ADC—available by winter—would be useful here).

An even greater advancement on cellular receivers is to implement a wideband receiver. In the example shown in Figure 3b, the single carrier of interest is selected by varying the LO frequency and using very selective filters in the IF signal processing. A wideband radio (available soon) seeks to digitize *all* the carriers, allowing the tuning and signal-extraction functions to be implemented digitally. This imposes severe requirements on the ADC's performance. If a 15-MHz-wide cellular band is to be digitized, an ADC sampling rate of 30-40 MSPS is required. Furthermore, to deal with the near/far problem, the converter dynamic range must be large enough to simultaneously digitize both strong and weak signals without either clipping the strong signals or losing the weak signals in the converter quantization noise. The converter requirements for a wideband radio vary with the cellular standard—anywhere from 12 bits, 40 MSPS for the U.S. AMPS standard (AD9042) to 18 bits, 70 MHz for GSM. The great advantages to this kind of implementation make the tradeoff worthwhile; one receiver can be used to simultaneously capture multiple transmissions, and—since the selection filtering is done digitally—programmable filters and demodulators can be used to support a multi-standard receiver. In radio industry jargon, this is a move towards the "software radio", where most of the radio processing is digital.

Selecting Mixed-Signal Components for Digital Communications Systems—Part V

Aliases, images, and spurs

by Dave Robertson

Part I (Analog Dialogue 30-3) provided an introduction to the concept of channel capacity, and its dependence on bandwidth and SNR; part II (30-4) briefly summarized different types of modulation schemes; part III (31-1) discussed different approaches to sharing the communications channel, including some of the problems associated with signal strength variability. Part IV (31-2) examined some of the architectural trade-offs used in digital communications receivers, including the problems with frequency translation and the factors contributing to dynamic range requirements. This final installment considers issues relating to the interface between continuous-time and sampled data, and discusses sources of spurious signals, particularly in the transmit path.

Digital communications systems must usually meet specifications and constraints in both the time domain (e.g., settling time) and the frequency domain (e.g., signal-to-noise ratio). As an added complication, designers of systems that operate across the boundary of continuous time and discrete time (sampled) signals must contend with aliasing and imaging problems. Virtually all digital communications systems fall into this class, and sampled-data constraints can have a significant impact on system performance. In most digital communications systems, the continuous-time-to-discrete-time interface occurs in the digital-to-analog (DAC) and analog-to-digital (ADC) conversion process, which is the interface between the digital and analog domains. The nature of this interface requires clear understanding, since the level-sensitive artifacts associated with conversion between digital and analog domains (e.g., quantization) are often confused with the time-sensitive problems of conversion between discrete time and continuous time (e.g., aliasing). The two phenomena are different, and the subtle distinctions can be important in designing and debugging systems. (Note: all digital signals must inherently be discrete-time, but analog signal processing, though generally continuous-time, may also be in discrete time—for example, with switched-capacitor circuits.)

The Nyquist theorem expresses the fundamental limitation in trying to represent a continuous-time signal with discrete samples. Basically, data with a sample rate of F_s samples per second can effectively represent a signal of bandwidth up to $F_s/2$ Hz. Sampling signals with greater bandwidth produces *aliasing*: signal content at frequencies greater than $F_s/2$ is folded, or aliased, back into the $F_s/2$ band. This can create serious problems: once the data has been sampled, there is no way to determine which signal components are from the desired band and which are aliased.

Most digital communications systems deal with band-limited signals, either because of fundamental channel bandwidths (as in an ADSL twisted-pair modem) or regulatory constraints (as with radio broadcasting and cellular telephony). In many cases, the signal bandwidth is very carefully defined as part of the standard for the application; for example, the GSM standard for cellular telephony defines a signal bandwidth of about 200 kHz, IS-95 cellular telephony uses a bandwidth of 1.25 MHz, and a DMT-ADSL twisted-pair modem utilizes a bandwidth of 1.1 MHz . In each case, the Nyquist criterion can be used to establish the *minimum* acceptable data rate to unambiguously represent these signals: 400 kHz, 2.5 MHz, and 2.2 MHz, respectively. Filtering must be used carefully to eliminate signal content outside of this desired bandwidth. The analog filter preceding an ADC is usually referred to as an *anti-alias filter*, since its function is to attenuate signals beyond the Nyquist bandwidth prior to the sampling action of the A/D converter. An equivalent filtering function follows a D/A converter, often referred to as a *smoothing filter*, or *reconstruction filter*. This continuous-time analog filter attenuates the unwanted frequency images that occur at the output of the D/A converter.

At first glance, the requirements of an anti-alias filter are fairly straightforward: the passband must of course accurately pass the desired input signals. The stopband must attenuate any interferer outside the passband sufficiently that its residue (remnant *after* the filter) will not hurt the system performance when aliased into the passband after sampling by the A/D converter. Actual design of anti-alias filters can be very challenging. If out-of-band interferers are both very strong and very near the pass frequency of the desired signal, the requirements for filter stopband and narrowness of the transition band can be quite severe. Severe filter requirements call for high-order filters using topologies that feature aggressive filter roll-off. Unfortunately, topologies of filters having such characteristics (e.g., Chebychev) typically place costly requirements on component match and tend to introduce phase distortion at the edge of the passband, jeopardizing signal recovery.

Designers must also be aware of distortion requirements for anti-alias filters: in general, the pass-band distortion of the analog anti-alias filters should be at least as good as the A/D converter (since any out-of-band harmonics introduced will be aliased). Even if strong interferers are not present, *noise* must be considered in anti-alias filter design. Out-of-band noise is aliased back into the baseband, just like out-of-band interferers. For example, if the filter preceding the converter has a bandwidth of twice the Nyquist band, signal-to-noise (SNR) will be degraded by 3 dB (assuming white noise), while a bandwidth of 4× Nyquist would introduce a degradation of 6 dB. Of course, if SNR is more than adequate, wide-band noise may not be a dominant constraint.

Aliasing has a frequency translation aspect, which can be exploited to advantage through the technique of *undersampling*. To understand undersampling, one must consider the definition of the Nyquist constraint carefully. Note that sampling a signal of *bandwidth*, $F_s/2$, requires a minimum sample rate $\geq F_s$. This $F_s/2$ bandwidth can theoretically be located anywhere in the frequency spectrum [e.g., NF_s to $(N+1/2)F_s$], not simply from dc to $F_s/2$. The aliasing action, like a mixer, can be used to translate an RF or IF frequency down to the baseband. Essentially, signals in the bands $NF_s<\text{signal}<(N+1/2)F_s$ will be translated down intact, signals in the bands $(N-1/2)F_s<\text{signal}<NF_s$ will be translated "flipped" in frequency (see Figure 1) This "flipping" action is identical to the effect seen in high-side injection mixing, and needs to be considered carefully if aliasing is to be used as part of the signal processing. The anti-alias filter in a conventional baseband system is a low-pass filter. In undersampling systems, the anti-alias filter must be a bandpass function.

Undersampling offers several more challenges for the A/D converter designer: the higher speed input signals not only require wider input bandwidth on the A/D converter's sample-and-hold (SHA) circuit; they also impose tighter requirements on the jitter performance of the A/D converter and its sampling clock. To illustrate, compare a baseband system sampling a 100-kHz sine-wave signal and an IF undersampling system sampling a 100-MHz sine-wave signal. In the baseband system, a jitter error of 100 ps produces a maximum signal error of 0.003% of full scale (peak-to-peak)—probably of no concern. In the IF undersampling case, the same 100-ps error produces a maximum signal error of 3% of full scale.

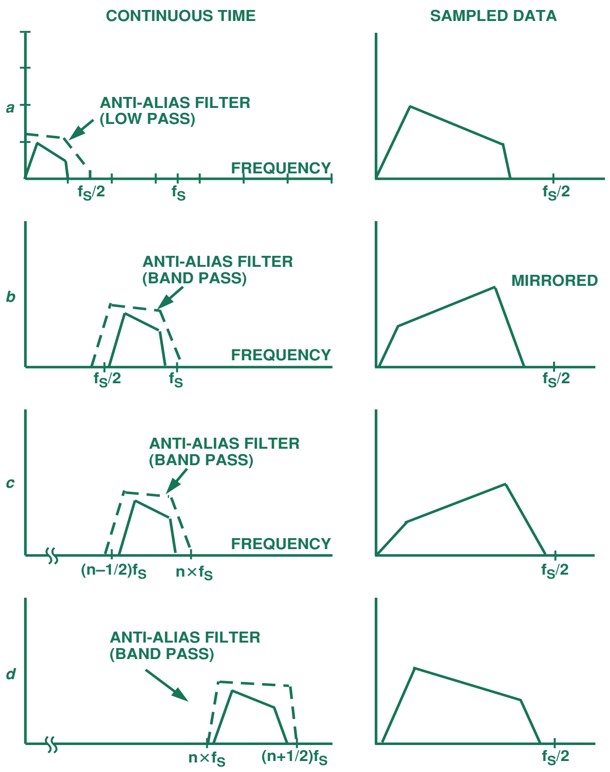


Figure 1. Aliasing, and frequency translation through undersampling.

Oversampling is not quite the opposite of undersampling (in fact, it is possible to have a system that is simultaneously oversampling and undersampling). Oversampling involves sampling the desired signal at a rate greater than that suggested by the Nyquist criterion: for example, sampling a 200-kHz signal at 1.6 MHz, rather than the minimum 400 kHz required. The oversampling ratio is defined:

$$OSR = sample\ rate/(2 \times input\ bandwidth)$$

Oversampling offers several attractive advantages (Figure 2). The higher sampling rate may significantly ease the transition band requirements of the anti-alias filter. In the example above, sampling a 200-kHz bandwidth signal at 400 kHz requires a "perfect" brick-wall anti-alias filter, since interferers at 201 kHz will alias in-band to 199 kHz. (Since "perfect" filters are impossible, most systems employ some degree of oversampling, or rely on system specifications to provide frequency guard-bands, which rule out interferers at immediately adjacent frequencies). On the other hand, sampling at 1.6 MHz moves the first critical alias frequency out to 1.4 MHz, allowing up to 1.2 MHz of transition band for the anti-alias filter.

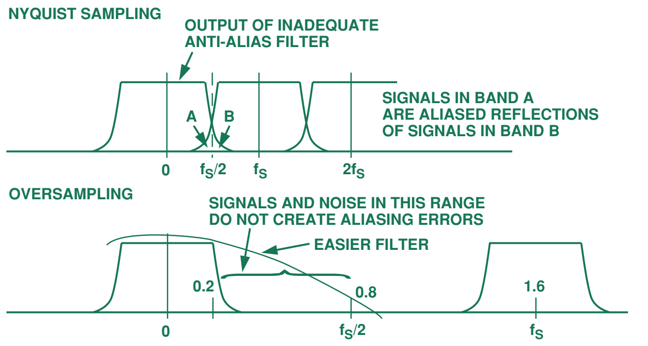


Figure 2. Oversampling makes filtering easier.

Of course, if interferers at frequencies close to 200 kHz are very strong compared to the desired signal, additional dynamic range will be required in the converter to allow it to capture both signals without clipping (see part IV, *Analog Dialogue* 31-2, for a discussion of dynamic range issues.) After conversion, oversampled data may be passed directly to a digital demodulator, or *decimated* to a data rate closer to Nyquist. Decimation involves reducing the digital sampling rate through a digital filtering operation analogous to the analog anti-aliasing filter. A well-designed digital decimation filter provides the additional advantage of reducing the quantization noise from the A/D conversion. For a conventional A/D converter, a *conversion gain* correspnding to a 3-dB reduction in quantization noise is realized for every octave (factor-of-two) decimation. Using the 1.6-MHz sample rate for oversampling as above, and decimating down to the Nyquist rate of 400 kHz, we can realize up to 6 dB in SNR gain (two octaves).

Noise-shaping converters, such as sigma-delta modulators, are a special case of oversampling converters. The sampling rate of the modulator is its high-speed clock rate, and the antialiasing filter can be quite simple. Sigma delta modulators use feedback circuitry to shape the frequency content of quantization noise, pushing it to frequencies away from the signal band of interest, where it can be filtered away. This is possible only in an oversampled system, since by definition oversampled systems provide frequency space beyond the signal band of interest. Where conventional converters allow for a 3-dB/octave conversion gain through decimation, sigma-delta converters can provide 9-, 15-, 21- or more dB/octave gain, depending on the nature of the modulator design (high-order loops, or cascade architectures, provide more-aggressive performance gains).

In a conventional converter, quantization noise is often approximated as "white"—spread evenly across the frequency spectrum. For an N-bit converter, the full-scale signal-to-quantization noise ratio (SQNR) will be $(6.02\,N + 1.76)$ dB over the bandwidth from 0 to $F_s/2$. The "white" noise approximation works reasonably well for most cases, but trouble can arise when the clock and single-tone analog frequency are related through simple integer ratios—for example, when the analog input is exactly 1/4 the clock rate. In such cases, the quantization noise tends to "clump" into spurs, a considerable departure from white noise.

While much has been written in recent years about anti-aliasing and undersampling operations for A/D converters, corresponding filter problems at the output of D/A converters have enjoyed far less visibility. In the case of a D/A converter, it is not unpredictable interferers that are a concern, but the very predictable frequency images of the DAC output signal. For a better understanding of the DAC image phenomenon, Figure 3(a,b) illustrates an ideal sine wave and DAC output in both the time and frequency domains. It is important to realize that these frequency images are *not* the result of amplitude quantization: they exist even with a "perfect" high-resolution DAC. The cause of the images is the fact that the D/A converter output exactly matches the desired signal only *once* during each clock cycle. During the rest of the clock cycle, the DAC output and ideal signal differ, creating error energy. The corresponding frequency plot for this time-domain error appears as a set of Fourier-series image frequencies (c). For an output signal at frequency F_{out} synthesized with a DAC updated at F_{clock}, images appear at $NF_{clock} \pm F_{out}$. The amplitude of these images rolls off with increasing frequency according to

$$\frac{\sin \pi(F_{out}/F_{clock})}{\pi(F_{out}/F_{clock})}$$

leaving "nulls" of very weak image energy around the integer multiples of the clock frequency. Most DAC outputs will feature some degree of clock feedthrough, which may exhibit itself as spectral energy at multiples of the clock. This produces a frequency spectrum like the one shown in Figure 4.

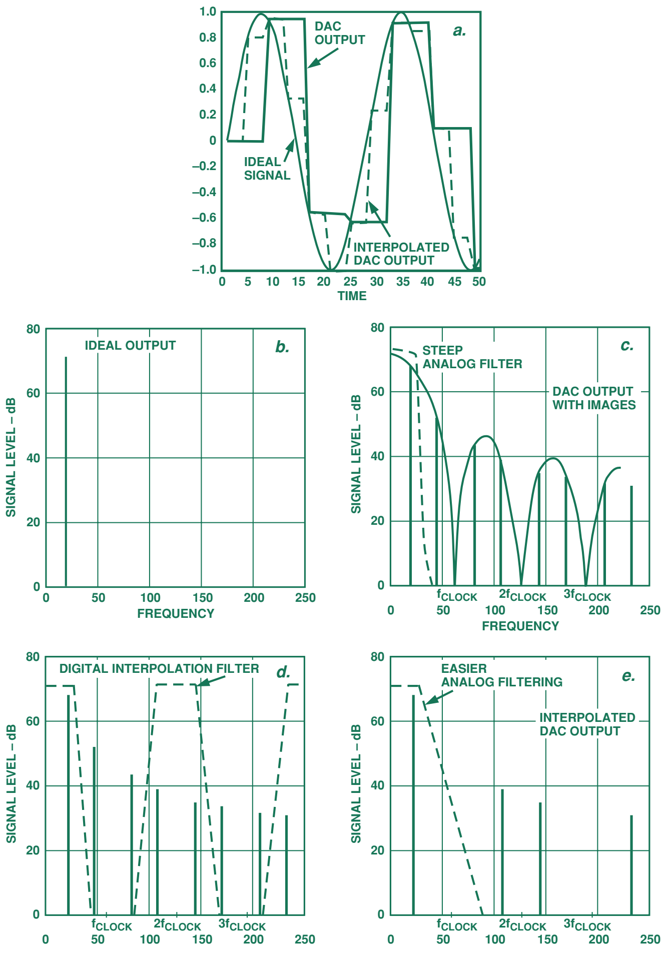


Figure 3. Time domain and frequency domain representation of continuous time and discrete sampled sine wave, and an interpolated discrete sampled sine wave.

The task of the DAC reconstruction filter is to pass the highest desired output frequency, F_{outmax}, and block the lowest image frequency, located at $F_{clock} - F_{outmax}$, implying a smoothing filter transition band of $F_{clock} - 2F_{outmax}$.

This suggests that as one tries to synthesize signals close to the Nyquist limit ($F_{outmax} = F_{clock}/2$), the filter transition gets impossibly steep. To keep the filter problem tractable, many designers use the rule of thumb that the DAC clock should be at least three times the maximum desired output frequency. In addition to the filter difficulties, higher frequency outputs may become noticeably attenuated by the sinx/x envelope: a signal at $F_{clock}/3$ is attenuated by 1.65 dB, a signal at $F_{clock}/2$ is attenuated by 3.92 dB.

Oversampling can ameliorate the D/A filter problem, just as it helps in the ADC case. (More so, in fact, since one need not worry about the strong-interferer problem.) The D/A requires an *interpolation filter*. A digital interpolation filter increases the effective data rate of the D/A by generating intermediate digital samples of the desired signal, as shown in Figure 3(a). The frequency-domain results are shown in (d,e): in this case 2× interpolation has suppressed the DAC output's first two images, increasing the available transition bandwidth for the reconstruction filter from $F_{clock} - 2F_{outmax}$ to $2F_{clock} - 2F_{outmax}$. This allows simplification of the filter and may allow more-conservative pole placement—to reduce the passband phase distortion problems that are the frequent side effects of analog filters. Digital interpolation filters may be implemented with programmable DSP, with ASICs, even by integration with the D/A converter (e.g., AD9761, AD9774). Just as with analog filters, critical performance considerations for the interpolation filters are passband flatness, stop-band rejection (how much are the images suppressed?) and narrowness of the transition band (how much of the theoretical Nyquist bandwidth ($F_{clock}/2$) is allowed in the passband?)

DACs can be used in undersampling applications, but with less efficacy than are ADCs. Instead of using a low-pass reconstruction filter to *reject* unwanted images, a bandpass reconstruction filter can be used to *select* one of the images (instead of the fundamental). This is analogous to the ADC undersampling, but with a few complications. As Figure 3 shows, the image amplitudes are actually points on a sinx/x envelope in the frequency domain. The decreasing amplitude of sinx/x with frequency suggests that the higher frequency images will be attenuated, and the amount of attenuation may vary greatly depending on where the output frequency is located with respect to multiples of the clock frequency. The sinx/x envelope is the result of the DAC's "zero-order-hold" effect (the DAC output remains fixed at target output for most of clock cycle). This is advantageous for baseband DACs, but for an undersampling application, a "return-to-zero" DAC that outputs ideal impulses would not suffer from attenuation at the higher frequencies. Since ideal impulses are physically impractical, actual return-to-zero DACs will have some rollof of their frequency-domain envelopes. This effect can be pre-compensated with digital filtering, but degradation of DAC dynamic performance at higher output frequencies generally limits the attractiveness of DAC undersampling approaches.

Frequency-domain images are but one of the many sources of spurious energy in a DAC output spectrum. While the images discussed above exist even when the D/A converter is itself "perfect", most of the other sources of spurious energy are the result of D/A converter non-idealities. In communications applications, the transmitter signal processing must ensure that these spurious outputs fall below specified levels to ensure that they do not create interference with other signals in the communications medium. Several specifications can be used to measure the dynamic performance of D/A converters in the frequency domain (see Figure 4):

- *Spurious-free dynamic range* (SFDR)—the difference in signal strength (dB) between the desired signal (could be single tone or multi-tone) and the highest spurious signal in the band being measured (Figure 4). Often, the strongest spurious response is one of the harmonics of the desired output signal. In some applications, the SFDR may be specified over a very narrow range that does not include any harmonics. For narrowband transmitters, where the DAC is processing a signal that looks similar to a single strong tone, SFDR is often the primary spec of interest.
- *Total harmonic distortion* (THD)—while SFDR indicates the strength of the highest single spur in a measured band, THD adds the energy of all the harmonic spurs (say, the first 8).
- *Two-tone intermodulation distortion* (IMD)—if the D/A converter has nonlinearities, it will produce a mixing action between synthesized signals. For example, if a nonlinear DAC tries to synthesize signals at 1.1 and 1.2 MHz, second-order intermodulation products will be generated at 100 kHz (difference frequency) and 2.3 MHz (sum frequency). Third-order intermodulation products will be generated at 1.3 MHz (2 × 1.2 – 1.1) and 1.0 MHz (2 × 1.1 – 1.2). The application determines which intermodulation products present the greatest problems, but the third-order products are generally more troublesome, because their frequencies tend to be very close to those of the original signals.
- *Signal-to-noise-plus-distortion* (SINAD)—THD measures just the unwanted harmonic energy. SINAD measures all the non-signal based energy in the specified portion of the spectrum, including thermal noise, quantization noise, harmonic spurs, and non-harmonically related spurious signals. CDMA (code-division, multiple-access) systems, for example, are concerned with the total noise energy in a specified bandwidth: SINAD is a more-accurate figure of merit for these applications. SINAD is probably the most difficult measurement to make, since many spectrum analyzers don't have low-enough input noise. The most straightforward way to measure a DAC's SINAD is with an ADC of significantly superior performance.

These specifications, or others derived from them, represent the primary measures of a DAC's performance in signal-synthesis applications. Besides these, there are a number of conventional DAC specifications, many associated with video DACs or other applications, that are still prevalent on DAC data sheets. These include integral nonlinearity (INL), differential nonlinearity (DNL), glitch energy (more accurately, glitch *impulse*), settling time, differential gain and differential phase. While there may be some correlation between these time-domain specifications and the true dynamic measures, the time-domain specs aren't as good at predicting dynamic performance.

Even when looking at dynamic characteristics, such as SFDR and SINAD, it is very important to keep in mind the specific nature of the signal to be synthesized. Simple modulation approaches like QPSK tend to produce strong narrowband signals. The DAC's SFDR performance recreating a single tone near full scale will probably be a good indicator of the part's suitability for the application. On the other hand, modern systems often feature signals with much different characteristics, such as simultaneously synthesized multiple tones (for wideband radios or discrete-multi-tone (DMT) modulation schemes) and direct sequence spread-spectrum modulations (such as CDMA). These more-complicated signals, which tend to spend much more time in the vicinity of the DAC's mid- and lower-scale transitions, are sensitive to different aspects of D/A converter performance than systems synthesizing strong single-tone sine waves. Since simulation models are not yet sophisticated enough to properly capture the subtleties of these differences, the safest approach is to characterize the DAC under conditions that closely mimic the end application. Such requirements for characterization over a large variety of conditions accounts for the growth in the size and richness of the datasheets for D/A converters.

For Further Reading:

For detailed discussion of discrete time artifacts and the Nyquist Theorem: Oppenheim, Alan V. and Schaeffer, Ronald W, *Discrete-Time Signal Processing*. Englewood Cliffs, NJ: Prentice Hall, 1989.

For more details on sigma-delta signal processing and noise shaping: Norsworthy, Steven R, Schreier; Richard; Temes, Gabor C., *Delta-Sigma Data Converters: Theory, Design, and Simulation*. New York: IEEE Press, 1997.

For more details on DAC spectral phenomena: Hendriks, Paul, "Specifying Communication DACs", *IEEE Spectrum* magazine, July, 1997, pages 58–69.

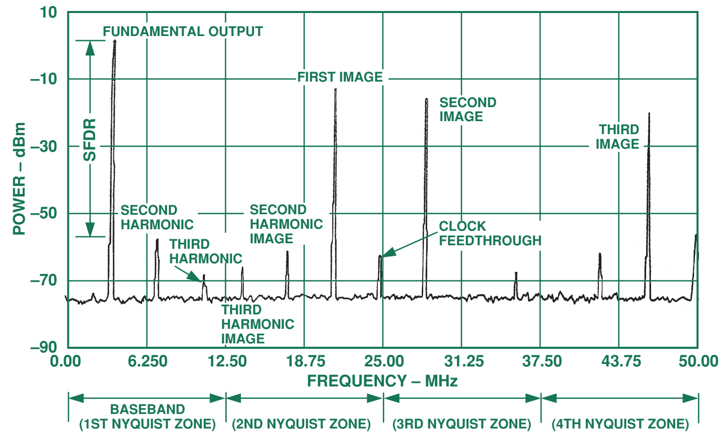


Figure 4. Different error effects in the output spectrum of a DAC.