Analog Dialogue

A forum for the exchange of circuits, systems, and software for real-world signal processing

Volume 35, 2001

ANALOG DEVICES

About *Analog Dialogue*

Analog Dialogue is the free technical magazine of Analog Devices, Inc., published continuously for thirty-five years, starting in 1967. It discusses products, applications, technology, and techniques for analog, digital, and mixed-signal processing.

Volume 35 incorporates all articles published during 2001 in the Worldwide Web editions **www.analog.com/analogdialogue**—and more. All recent issues, starting with Volume 29, Number 2 (1995) have been archived on that website and can be accessed freely. Two special anniversary issues in the archives contain useful articles extracted from earlier editions, going all the way back to Volume 1, Number 1. This issue revisits two series that have proved popular in the recent quinquennium.

Analog Dialogue's objectives have always been to inform engineers, scientists, and electronic technicians about new ADI products and technologies and to help them understand and competently apply these products.

The frequent Web editions have at least three further objectives:

- To provide digests that alert readers to prereleased and newly available products.
- To provide a set of links to important and rapidly proliferating sources of information and activity fermenting within the ADI website (**www.analog.com**).
- To listen to reader suggestions and provide answers to their questions.

Thus, *Analog Dialogue* is more than a magazine; its links and tendrils to all parts of our external website make its bookmark a favorite "high-pass-filtered" point of entry to the **analog.com** site—the virtual world of Analog Devices.

Our hope is that readers will think of ADI publications as "Great Stuff" and will consider the *Analog Dialogue* bookmark a favorite alternative path to answer the question, "What's new in technology at ADI?"

Welcome! Read and enjoy!
We encourage your feedback!

Dan Sheingold
dan.sheingold@analog.com
Editor, *Analog Dialogue*

IN THIS ISSUE

Analog Dialogue Volume 35, 2001

This annual issue of *Analog Dialogue* contains reprints of all the articles published in the on-line editions during 2001. As a special bonus, in celebration of our 35th year in print, we are adding reprints of two collections that have proven popular during the past five years: Dave Robertson's five-part 1996–1997 series, "Selecting Mixed-Signal Components for Digital Communication Systems," and all installments of the "Ask The Applications Engineer" series, published on line (and in print) since the 30th anniversary collection in 1996. (You can find that publication in our *Analog Dialogue* on-line Archives.)

35TH ANNIVERSARY BONUS MATERIAL

Cover: *The cover illustration was designed and executed by Kristine Chmiel-Lafleur, of Communications Services, Analog Devices, Inc.*

Editor's Notes

At the close of 2001, we celebrate the completion of *Analog Dialogue*'s 35th year in print—and the third year of its electronic version. This season also marks the beginning of the undersigned's 34th year at Analog Devices as Editor of this journal—and his 53rd year of advocating analog technology. In the early days, at Philbrick, this featured K3 Analog Computor (*sic*) building blocks. But they would soon be overshadowed in fame and cannibalized in fortune by the yet smaller, still undifferentiated, analog building block that lurked within, dreaming of an identity.

It materialized in 1952 as Philbrick's versatile and ubiquitous *K2-W plug-in Differential Operational Amplifier*, which celebrates the jubilee of its birth in 2002. The K2-W, with its two 12AX7 tubes plugged in at the top, was the harbinger of a new era of compact precision at low cost in modular measurement, control, and much else. By plugging into an octal tube socket in the same way as a tube (and in the same way that the later ICs plugged into transistor-type sockets), this "smart tube" could also be considered a forerunner of the *integrated circuit*. At the heart of circuits and modules that perform analog computing, it could even be said that op amps resulted in the transformation of *analog computers* into components—in the same way that microprocessors would later turn *digital* computers into components.

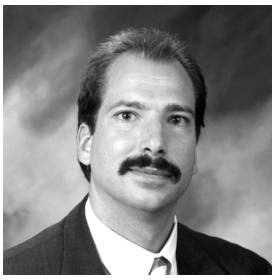
OK. Enough of analogy! Let's get on with *Analog*.

Dan Sheingold

AUTHORS

John Cowles (page 5) works at ADI Northwest Labs in Beaverton, Oregon, designing RF-IF products for the communications market. He received a PhD in EE from the University of Michigan in 1994. Before joining ADI in 1998, he worked at TRW, in Redondo Beach, CA, on high-speed GaAs and InP technologies. At ADI, he designed the AD8302, which was nominated for Product of the Year by both *Electronic Products* and *EDN* magazines. He has over 50 technical publications on high-speed devices and circuits.

Mike Curtin (page 17) is the applications manager for RF products in Ireland. Mike graduated from the University of Limerick with a BSc in Electronic Engineering. He worked in various local companies before joining Analog Devices in 1983 as an Applications Engineer. He has provided technical support for a wide variety of Limerick-developed products, including high-speed ADCs, sigma-delta ADCs, and DACs, an activity interrupted by a two-year assignment as a field applications engineer in the U.S. In his spare time, he listens to music and enjoys a game of snooker or indoor football.



Joe DiPilato (pages 9, 13), is the product-line manager for communication-specific integrated products within ADI's High-Speed Converter (HSC) group. Since joining ADI in 1987, he has been involved with industrial, instrumentation, and communication-focused products—and for the last eight years, HSC standard- and communication-specific products, including TxDAC® and MxFE™ brands. He has authored numerous articles. Joe has a BSEE from Worcester Polytechnic Institute (1982) and an MBA from Anna Maria College (1984). He is active in his church, and enjoys swimming and quality beach time with his wife Lisa and son Nicholas.

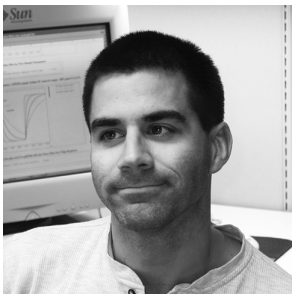
Zoltan Frasch (page 37) is principal applications engineer for ADI's Display-Drive Electronics group. He earned his BSEE from the University of Toronto, Canada, in 1976 and joined ADI in 1999.

Barrie Gilbert (page 5), the first ADI Fellow, has "spent a lifetime in the pursuit of analog elegance." He joined Analog Devices in 1972, was appointed ADI Fellow in 1979, and manages the Northwest Labs in Beaverton, Oregon. Barrie was born in Bournemouth, England, in 1937. Before joining ADI, he worked with first-generation transistors at SRDE in 1954, and at Mullard, Ltd.; later at Tektronix and Plessey Research Labs. Barrie is an IEEE Fellow (1984) and has received numerous awards. He has some 50 issued patents, has authored about 40 papers, is co-author of several books, and is a reviewer for several journals. He was awarded an Honorary Doctorate of Engineering from Oregon State University in 1997.

Paul Hendriks (page 9) is a senior applications engineer for ADI's Standard Linear Products Division (Wilmington, MA). He has been working in the High-Speed Converter product group for the past 8 years, focusing on the use of high-speed converters in communication applications. Since joining ADI in 1989, he has authored numerous articles, design ideas, and product data sheets. Paul received a BSEE in 1986 from McGill University.

Paul Kearney (page 30) is a design engineer in ADI's Precision-Converter product line, in Limerick, Ireland. He was graduated with a BE in Electrical Engineering from University College Cork, Ireland, and joined ADI immediately thereafter, in 1990. Before his present assignment, in 1998, he worked as a product engineer.

Martin Kessler (page 13) first encountered ADI in 1990, as a coop student in Munich, Germany. In 1994, he joined the company full-time, providing application support for all product lines. Subsequently, Martin moved to ADI Wilmington, where he focuses his efforts on digital communication-specific ICs and mixed-signal front ends. Hobbies include skiing, hiking, backpacking, and playing the guitar.



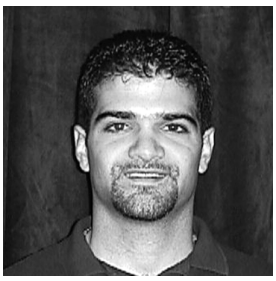
Charles ("Chuck") Kitchin (page 24) is a hardware applications engineer at ADI in Wilmington, MA. His main responsibilities include writing technical publications and developing new applications circuits. He has published over 70 technical articles, three books, and a large number of application notes. Chuck graduated with an ASET from Wentworth Institute, in Boston, and then continued studying electrical engineering at the University of Lowell. His avocations include astronomy, amateur radio, and oil painting.

Alan Li (page 35) is an applications engineer in the General-Purpose Converter and Reference product line of the Precision Converter Division. He has a BSEE from the Florida International University and is currently pursuing a graduate degree at San Jose State University. He has also worked at National Semiconductor and Fairchild Semiconductor, focusing on power-electronics applications. His leisure, activities include playing soccer and listening to music.

Iuri Mehr (page 13) is a design engineer for communications-specific products in ADI's High-Speed Converter Group, Wilmington, MA. He received the MSEE degree from Washington State University in 1994. After a two-year stint as Design Engineer for Crystal Semiconductor, he joined ADI in 1996, in the Standard Products group. In 1997, he joined the High-Speed Converter group. He serves as an Associate Editor for *IEEE Transactions on Circuits and Systems II*.

Albert O'Grady (page 27) is a staff applications engineer at ADI's facility in Limerick, Ireland, providing applications support for general-purpose digital-to-analog converters (DACs) and high-resolution, low-bandwidth, sigma-delta analog-to-digital converters (ADCs). Albert holds a BEng from the University of Limerick. In his spare time, Albert enjoys reading and plays badminton and tennis.

Giuseppe Olivadoti (page 41) is the technical marketing specialist for the DSP Development Tools product line. He holds a BS in Electrical Engineering with a concentration in Computer Engineering from Northeastern University.



Catherine Redmond (page 21) is an applications engineer at ADI's facility in Limerick, Ireland, providing applications support for general-purpose D/A converters and Switch/Multiplexer products. Catherine holds a BEng from Cork Institute of Technology and joined ADI in 1997. Her interests include listening to music, reading, and travelling.

Richard Schreier (page 9) is a design engineer in ADI's High-Speed Converters group, in Wilmington, MA. He received a PhD degree from the University of Toronto in 1991. Before joining ADI, he was an Assistant Professor at Oregon State University in Corvallis, Oregon, from 1991 to 1997. Richard is a coedi-tor of a 1997 IEEE Press book on delta-sigma converters; he also authored a freeware Matlab toolbox for the high-level design and simulation of delta-sigma modulators.

Peter Shih (page 44) is an applications engineer with the Central Applications group in Wilmington, MA, working closely with product-line applications engineers to provide customer support for linear products. He joined ADI in May 2000, after graduating from Tufts University with a BSEE. His interests include playing volleyball, traveling, and eating!

Ed Spence (page 37) is product manager of ADI's display-drive electronics strategy, based in Wilmington, MA. He received a BSEE from the University of Lowell, MA. Ed joined ADI in 1989.

Harvey Weinberg (page 44) is an applications engineer for ADI's Micromachined Products division in Cambridge, MA. He has a BEng in Electrical Engineering from Concordia University, Montreal, Canada. For nine years, prior to working as an applications engineer, he designed analog- and micro-controller-based instrumentation for the process-control industry.

NEW FELLOWS

We are pleased to note the introduction of *four* new Fellows at our 2001 General Technical Conference. Fellow, at Analog Devices (ADI), represents the highest level of achievement that a technical contributor can attain, on a par with Vice President. The criteria for promotion to Fellow are very demanding. Fellows will have earned universal respect and recognition from the technical community for unusual talent and identifiable innovation at the state of the art. Their creative technical contributions in product or process technology, or software, will have led to commercial success with a major impact on the company's net revenues and earnings.

Attributes include roles as mentor, consultant, entrepreneur, organizational bridge, teacher, and ambassador. Fellows must also be effective as leaders and members of teams—and in perceiving customer needs. In the year 2001, the unprecedented number of *four* outstanding individuals—**Denis Doyle**, **Paul Ferguson**, **Josh Kablotsky**, and **Larry Singer**—were identified as having the rare combination of technical abilities, accomplishments, and personal qualities to qualify them to enhance our existing roster of Fellows: *Bob Adams* (1999), *Woody Beckford* (1997), *Derek Bowers* (1991), *Paul Brokaw* (1979), *Lew Counts* (1983), *Barrie Gilbert* (1979), *Roy Gosser* (1998), *Bill Hunt* (1998), *Jody Lapham* (1988), *Chris Mangelsdorf* (1998), *Jack Memishian* (1980), *Doug Mercer* (1995), *Frank Murden* (1999), *Mohammad Nasser* (1993), *Wyn Palmer* (1991), *Carl Roberts* (1992), *Paul Ruggerio* (1994), *Brad Scharf* (1993), *David Smart* (2000), *Jake Steigerwald* (1999), *Mike Timko* (1982), *Bob Tsang* (1988), *Mike Tuthill* (1988), *Jim Wilson* (1993), and *Scott Wurcer* (1996).

DENIS DOYLE

Denis is a major contributor to process development and operations at ADI's Limerick (Ireland) wafer fab. His principal contributions have been in the evolution of BiCMOS processes from 1-μm to 0.35-μm feature size. He has also worked on other projects, ranging from EPROMs and ESD to process transfers between ADI and our foundries.

Denis Doyle received his BE (Elect) from University College, Cork, Ireland, in 1985, followed by the MEngSc and PhD—for his work in small-geometry bipolar devices at the National Microelectronics Research Centre (NMRC), in Cork, Ireland. In 1991, he joined the Process Development group at ADI, Limerick, where he has since worked on and led process developments. Denis's current project is on the development of a 0.6-μm, 30-V CMOS/BiCMOS process.

PAUL FERGUSON

Paul is an inventive circuit designer who has contributed heavily in the field of switched-capacitor circuits and higher-order sigma-delta ADCs and DACs. He was instrumental in the development of new inventions and technologies that led to the emergence of ADI as the leader in the PC business audio segment. Among these are new ideas for offset calibration, which led to a reduction in "speaker pop" phenomena, once common throughout audio applications. He was also a pioneer in standard-cell layout and novel higher-order sigma-delta loop architectures at ADI.

Besides his contributions as an inventor and designer, he also excels as a teacher, mentor, and team builder, has developed strong relationships with universities, has delivered many technical papers and lectures, and has served a key role in recruiting promising graduates. He has an AB from Dartmouth College and an SM in EE from MIT—and is in pursuit of his PhD from Oregon State University. He holds 16 patents and is currently a key player in mixed-signal development for ADI's RF and wireless systems, including the SoftFone™ and Othello™ chipsets. Paul lives in North Andover, MA, with his wife, Amy, and their three children.

JOSH KABLOTSKY

Joining the Development Tools group in 1990, Josh has contributed heavily to DSP software development and created numerous software development methodologies and tools that are now in daily use throughout the company. He has developed DSP software for successful ADI products in diverse communications and computer segments—and is now focused on DSP-based video and image processing. He has been responsible for remote-access and voice-over networks businesses, consulted on processor architectures, and has been ADI's primary contact to organizations such as ITU-T and TIA.

He has contributed to development in many product areas, including DSP answering-machine software, voice compression and tone detection, acoustic echo cancellation, modem development, state-machine design, and remote data gathering. Through application by others of his development methodologies, he has also contributed indirectly to a much wider universe, including digital still cameras, color laser printers, 3G wireless infrastructure, wireless terminals, and IADs. He has a BSEE from Cornell University, and currently lives in Sharon, MA, with his wife, Deborah.

LARRY SINGER

Larry is most recently the lead A/D converter designer for the High-Speed Converter Standard Products group in Wilmington, MA. He is a graduate of MIT, with SB (1985) and SM (1987) degrees in EE. He joined ADI upon graduation, after being recruited by Fellow-to-be Paul Ferguson. Since then he has developed many of ADI's converters that join high speed and high resolution, including both successive-approximation and high-speed pipelined types.

Early in his career, he solved challenging problems in the design of practical 12- and 14-bit ADCs in bipolar-CMOS (BiCMOS), that used laser-trimmed thin-film-on-chip resistors. He then moved on to moderate resolution (to 12 bits) video (20-MSPS) converters, designing pipelined converters in high-speed FLAMOS, an advanced bipolar-CMOS process—with laser-trimmed resistors. Since then his major contributions have been in innovative CMOS designs, pushing the boundaries of capacitor matching and segmentation, and showing that performance up to 14 bits could be realized without calibration. With his efforts, CMOS devices are closing the speed and dynamic-performance gap with bipolar and BiCMOS devices. Larry is married, has two children, and enjoys drumming and playing volleyball.

Accurate Gain/Phase Measurement at Radio Frequencies up to 2.5 GHz

by John Cowles (john.cowles@analog.com) and
Barrie Gilbert (barrie.gilbert@analog.com)

INTRODUCTION

Measuring Signals

Electronic circuits fall into two broad categories—those that *process* and transform signals and those that *measure* signals. Their functions are often combined, as in the IF section of a receiver—which processes the signal (by amplification and demodulation), and also delivers an indication of received signal-strength (the RSSI function), a slowly-varying voltage that may be displayed and/or used for automatic control of variables such as gain and frequency (AGC and AFC).

Circuits that measure RF signal strength, whose fundamental metric is *power*, are generally called detectors, but only a thermopile (bolometer) measures this quantity *directly*. Integrated-circuit detectors invariably operate on a voltage sample of the signal to be measured. Circuits of this class are classified by the type of signal transformation they provide. In 1976, Analog Devices supplied the first monolithic "true-rms" detectors for use at moderate frequencies. Now this product line includes devices, such as the AD8361, that have extended this capability to the *microwave* domain. The accurate determination of signal power, independent of its waveshape (stated otherwise, its *probability density function*) is important in modern communications systems such as CDMA. Unlike thermal detectors, these true-rms detectors use *analog computation* to directly implement the relevant equations—at gigahertz frequencies.

Another valuable type of RF detector (also using computation) is the *demodulating logarithmic amplifier*. As the name suggests, it *amplifies* the signal, which allows devices of this class to measure small signals, and it *demodulates* the alternating RF waveform to a slowly-varying "quasi-dc" output. However, unlike the rms detectors, whose output is proportional to the root-mean-squared value of the input voltage, logarithmic detectors deliver an output proportional to the *decibel value* of the signal level, referenced to a fixed voltage, V_{INT} (defined below). The output, usually a voltage, may be interpreted in terms of either voltage or power, simply by using a different value for a scaling parameter called the "slope."

For RF log amps, it is necessary to use voltage metrics for all the signal and scaling parameters. To define the input level, we will use dBV (here meaning decibels relative to 1 V rms) rather than refer to "power," in dBm (decibels relative to 1 mW). This is unambiguous, independent of the choice of impedance at the input interface, and appropriate for an IC detector. For example, 0 dBV corresponds to a sine wave of 2.83-V p-p amplitude; similarly, –60 dBV refers to a 2.83-mV p-p sine wave.

The operation of these RF logarithmic detectors conforms to a function like this:

$$V_{OUT} = V_{SLP} \log(V_{SIG}/V_{INT}) \quad (1)$$

If base-10 logarithms are chosen [$\log_{10}(10) = 1$ decade], with decibels in mind, the slope voltage, V_{SLP}, can be viewed in terms of "volts-per-decade" in the scaling of the log of the voltage ratio. Since there are 20 decibels in a decade, the corresponding "volts/dB" is just one-twentieth of this voltage. Thus, for a V_{SLP} of 400 mV/decade, the slope can also be expressed as 20 mV/dB. The second scaling parameter, called the "intercept," V_{INT}, is the input voltage at which the log argument is unity. At this voltage, independent of choice of base, the output would be zero, since $\log(1) = 0$. In practice, the finite available gain in an RF log amp, the presence of noise, and other practical limitations result in a value for V_{INT} that is an *extrapolated* value, typically only a few microvolts, and fixed by the design.

A question then arises as to the precise interpretation of what V_{INT} represents. Is this quantity "volts dc," or perhaps "volts rms"? Or is it some other metric, such as a simple average value, or the peak value? For measurements of ratios from one level to another, the value of V_{INT} is unimportant. However, where it is required to determine the *absolute* level of V_{SIG}, the measurement accuracy depends directly on the value of V_{INT} in just the same way as a reference voltage in, say, a DVM.

A close study of RF logarithmic amplifiers, which use the technique[1] known as "progressive compression," shows another effect not encountered in classical log amp practice, namely, that the effective value of V_{INT} strongly depends on the *waveform* of the input signal. For that reason, we choose to define V_{INT} for a sinusoidal input, and then provide conversion factors for various other waveforms.

In practice, the control of V_{INT} in an untrimmed production IC cannot be as accurate as is often needed in precision RF metrication. Laser trimming, first used for an RF log amp in the AD640/641, and more recently in products such as the AD8306, can provide very accurate calibration, using a sine-wave input during calibration. However, while the use of the appropriate conversion factor for a *known* waveform can maintain good accuracy, there remains the basic problem of waveform-dependence. This poses a problem in contemporary systems where the waveform is both unknown and can vary rapidly.

Measurement of Signal Ratios to 2.5 GHz

This problem has been addressed, in the AD8302, by the use of two identical log amps integrated in monolithic form, as shown in Figure 1. Each channel is capable of measuring signals over a 60-dB range, from very low frequencies up to 2.5 GHz. The defining function for the *amplitude* ("gain") output is

$$V_{MAG} = V_{SLP} \log (V_A/V_B) + V_{CP} \quad (2)$$

where V_A and V_B are two independent signals, applied to the two identical input ports of the AD8302, and V_{CP} is the center-point, defined as the value of the output, V_{MAG}, for a level difference of 0 dB. (V_{SLP} and V_{CP} are design choices, made with ease-of-use in mind; both are traceable to a band-gap reference).

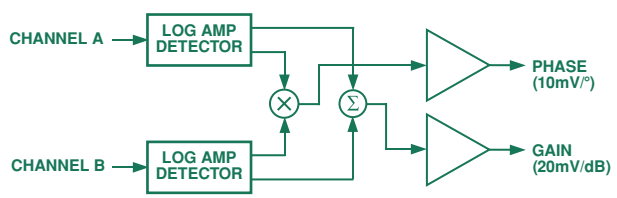
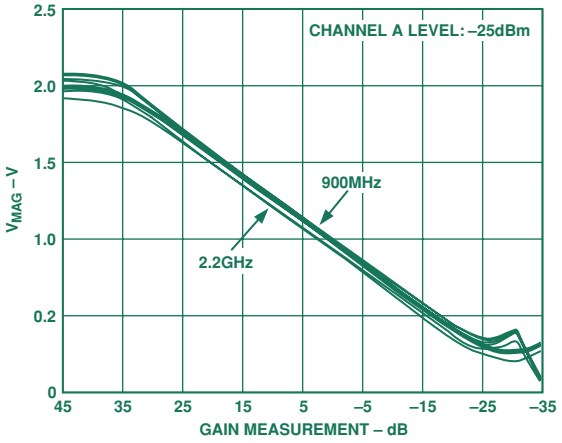


Figure 1. The AD8302 comprises a pair of accurately matched log amps and a high-frequency phase detector.

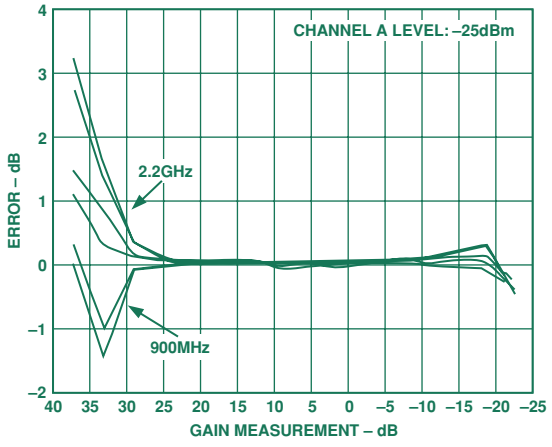
The customary fixed intercept of Equation 1, is eliminated in the AD8302 by taking the difference of the outputs of the two separate log amps. This key step computes the log of the ratio $(V_A/V_B)\times(V_{INTB}/V_{INTA})$; and, since the log amps are identical, the second term is very accurately unity, independently of temperature, supply voltage, and numerous production variances.

This elegant elimination of a fixed intercept results in a highly accurate measurement of signal level, in many applications. The primary limitation to the accuracy of the log argument is now the matching of the two co-integrated channels. This novel "C" structure[2] opens up many measurement possibilities that would otherwise require the use of two distinct log amps, with their inherent differences in slope and intercept calibration. The AD8302 is the first IC to permit the *direct measurement of ac signal ratios*. This unique capability for measuring gain/loss and the relative phase (see below) between two signal ports, over a very wide range of frequencies, will be of value in many other applications.

Figure 2 illustrates the output voltage variation as a function of signal ratio (which, for example, may correspond to the *gain or loss* of a channel being monitored) at frequencies ranging from 900 MHz to 2.2 GHz. The signal level presented to Channel B is fixed while that at Channel A is varied from –30 dB to +30 dB relative to Channel B. The output, V_{MAG}, demonstrates the precise slope, V_{SLP}, of 20 mV/dB and a center-point, V_{CP}, of 900 mV. The very small deviation from an ideal logarithmic law (Figure 2b) demonstrates the value of using co-integrated log amps.



(a)



(b)

Figure 2. Measurements of signal-level ratios (a) exhibit errors (b) of less than 0.2 dB up to high frequencies.

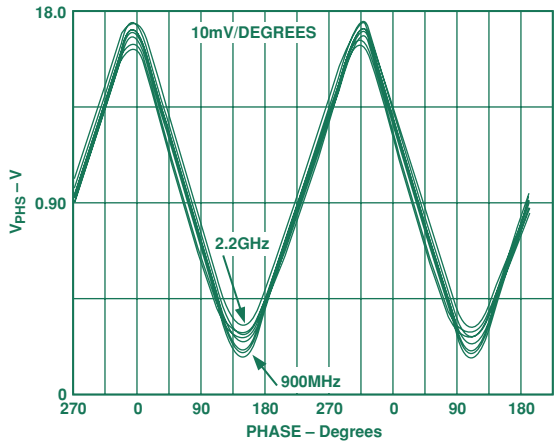
Measurement of Relative Phase to 2.5 GHz

The AD8302 can also measure the *phase difference* between two signals, from low frequencies up to 2.5 GHz. Each of the individual log amps generates a "hard-limited" output at its final stage. These signals are applied to the two inputs of a novel multiplier-style phase detector having exact symmetry with regard to its two inputs and a 180° range. The phase output, V_{PHS}, is given by,

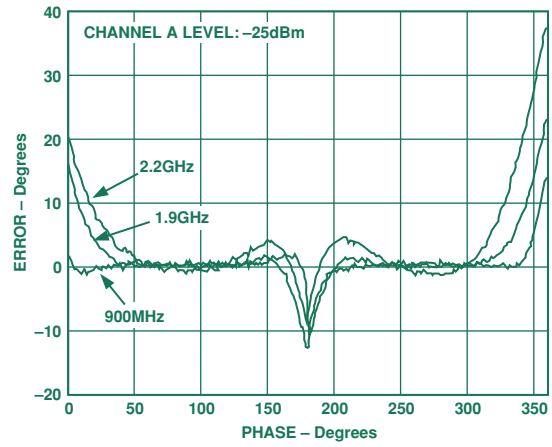
$$V_{PHS} = \pm V_{\Phi}(\Phi - 90°) + V_{CP} \quad (3)$$

where V_{Φ} is the scaling voltage for the phase output and Φ is the phase difference between the two inputs. The choice of sign depends on which two quadrants constitute the 180° phase interval. With the inclusion of this feature, the AD8302 becomes a "network analyzer on a chip."

Figure 3 illustrates phase measurement at 900 MHz, 1.9 GHz, and 2.2 GHz. Here, the phase difference was generated as a "slip," by slightly offsetting the two input frequencies and allowing the angle to accumulate. The slope of the V_{PHS} output is 10 mV/°C, centered at a V_{CP} of 900 mV. The alternating sign of the slope is apparent as the phase slips through 180° intervals. Figure 3b shows the measurement error. The rapid increase in error near 0° and 180° is due mainly to dead-zones caused by the finite rise and fall times of the hard-limited signals. The unique ability of the AD8302 to accurately measure phase at these frequencies is a result of the excellent balance of its two tightly-integrated log amps.



(a)



(b)

Figure 3. The phase measurement (a) exhibits low errors (b) over wide angular ranges and up to high frequencies.

Using the AD8302

These new capabilities for measuring gain/loss and the relative phase between two signal ports will be of value in many RF and IF applications. The functionality, versatility, and compact form-factor of this "*network analyzer on a chip*" are ideally suited for *in-situ* diagnostics and monitoring of system parameters and for feedback and feed-forward linearization and control of subsystems. These are a few applications of the AD8302.

The measurement of absolute signal level is now possible using a *known ac reference*. As shown in Figure 4, the reference signal applied to Channel B creates an effective intercept of value V_B. When the two signals have similar waveforms, the measurement can be very accurate. Even the error due to uncertainty in the slope voltage can be minimized (eliminated, in principle) if one can ensure that the two inputs are close-to-equal in amplitude. This will often be a simple matter to arrange, using an attenuator pad on the larger signal to position the ratio V_A/V_B close to unity. Centering techniques are valuable when the highest accuracy is needed or where very large dynamic ranges must be handled.

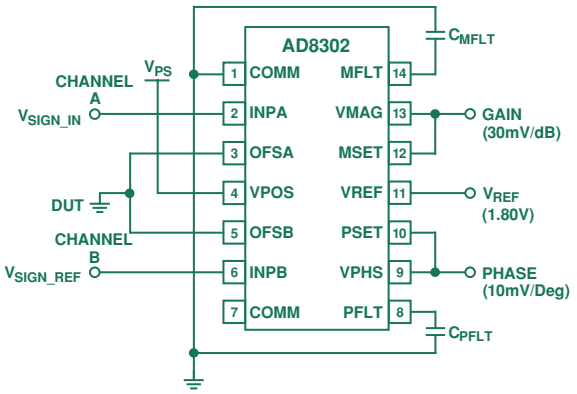


Figure 4. Absolute measurements of signal level using an AC reference on Channel B as the intercept for Channel A.

The most useful application of the AD8302 is in monitoring and reporting the gain or loss of a functional block or subsystem. In the example shown in Figure 5, samples of the input and output signals of a 500-MHz IF amplifier with a nominal gain of 20 dB are monitored. By using attenuators and couplers, the two signals are conditioned to be of the same general magnitude. The gain response shows the mid-scale low frequency value, which corresponds to a 20-dB level difference at the amplifier and a 3-dB bandwidth of approximately 500 MHz. The functional block in this example could have been a frequency-translation device, such as a mixer. In that case, the two inputs would be at different frequencies, and the measured quantity would be the conversion gain. Since the waveforms remain similar, that source of error is again eliminated. However, when the input frequencies differ greatly, a systematic offset may occur due to inequalities in the impedance match and the frequency dependence of the scaling of the two log amps in the gigahertz region.

In many communication systems, there is an unpredictable load presented to an external interface port. Variations in this load can

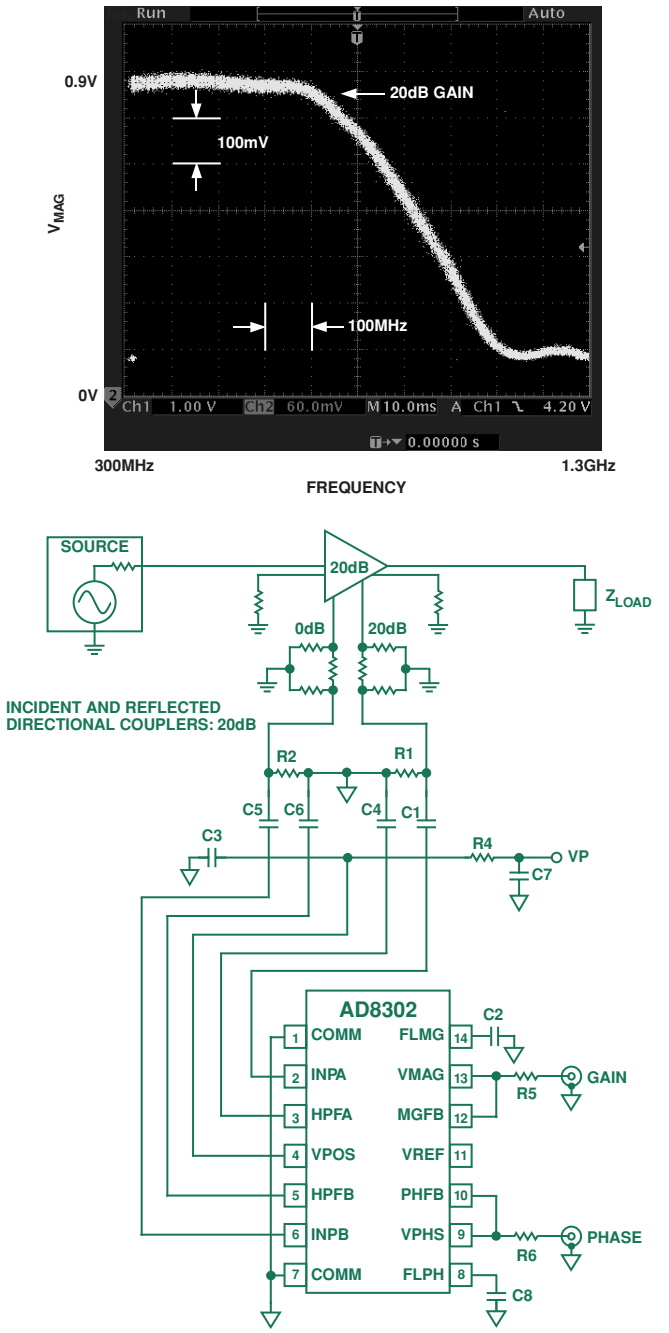


Figure 5. The AD8302 monitoring the frequency response of an amplifier under test and reporting the gain.

lead to changes in system performance, or even to catastrophic failure in extreme cases. It is of great value to provide the means of monitoring the load impedance—or reflection coefficient in RF terms—without perturbing it. In Figure 6, the AD8302 is configured to measure the reflection coefficient of an arbitrary load which, in this case, is a PIN-diode whose bias is swept to change its impedance. The notch in the response curve represents a near-match to the 50-Ω characteristic line impedance, where the reflected signal is almost zero.

[1]A detailed description of the theory of operation of log amps can be found in the AD640 data sheet.
[2]Patent pending.

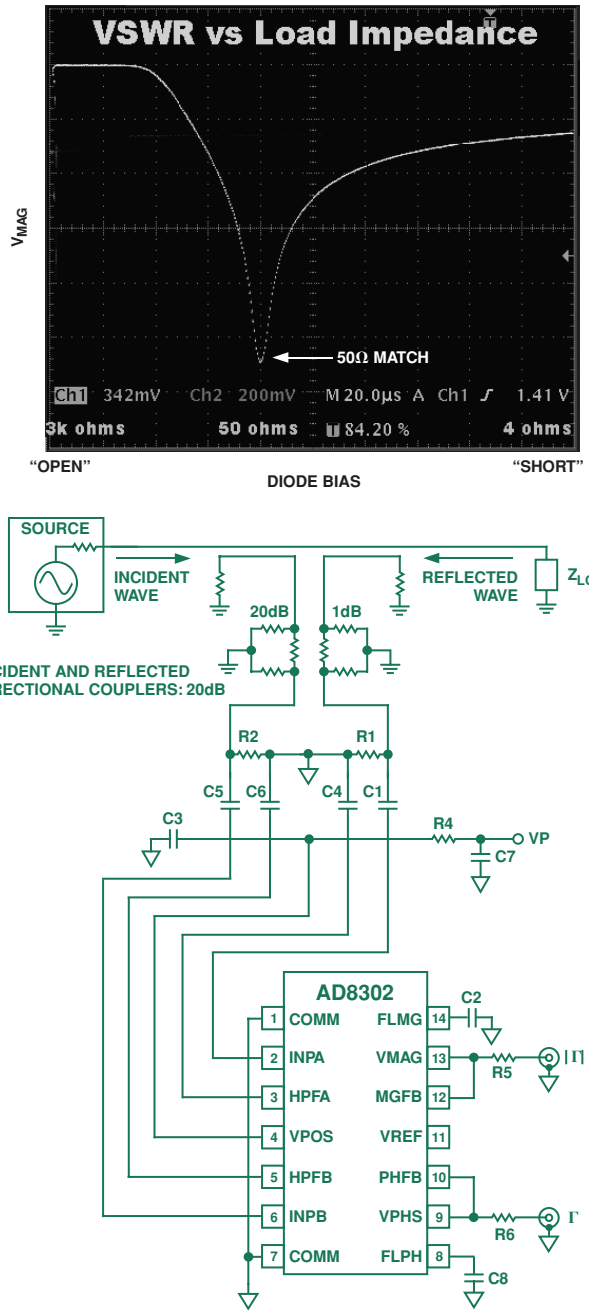


Figure 6. The AD8302 monitors the reflection coefficient of a load—a PIN diode whose impedance is manipulated by its bias.

Versatility and Ease of Use

The AD8302 offers several other modes of operation, the result of careful planning and the fundamentally versatile nature of this unusual structure. The previous examples have demonstrated the AD8302 in its typical *measurement* mode, where the V_{MAG} and V_{PHS} outputs report the signal level and phase difference between its inputs. However, the built-in scaling and center-points of the transfer functions can be adjusted using external resistors and the 1.80-V internal reference provided at the VREF pin.

By disconnecting the output pins from the feedback pins, MSET and PSET, a gain- and phase-*comparator* is realized, as shown in Figure 7. Here, the V_{MAG} and V_{PHS} outputs toggle between 0 V and the maximum output voltage of 1.8 V, depending on whether the signal level and phase difference are greater than or less than the thresholds presented to the MSET and PSET pins.

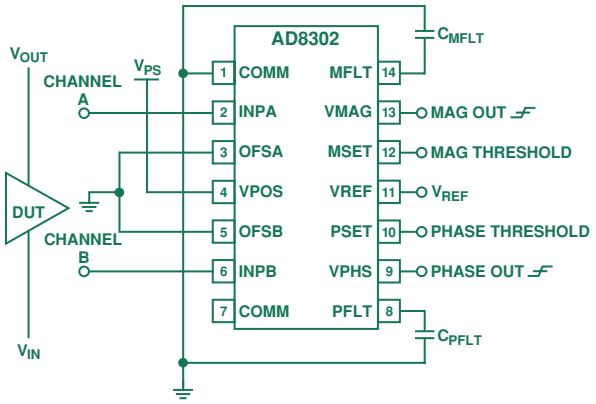


Figure 7. The AD8302, configured as a gain- and phase-comparator with controllable thresholds.

In the *controller* mode, shown in Figure 8, the VMAG and VPHS pins drive gain/phase-adjusters that are included in the signal chain being monitored so as to servo the overall gain and phase of the system toward the desired set-points presented to the MSET and PSET pins.

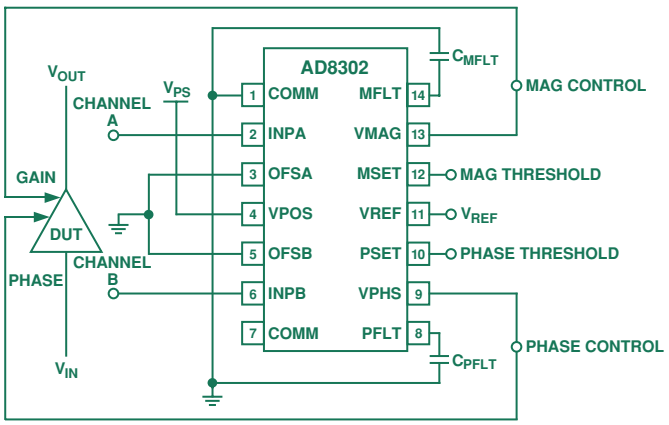


Figure 8. The AD8302 in a control loop that drives the gain and phase of a DUT towards prescribed set-points.

The AD8302 has a broad frequency range at its inputs, ranging from arbitrarily low frequencies (even audio!) up to 2.5 GHz. The wide dynamic range of the log amps accommodates not only large changes in relative signal level but also variations in the absolute levels. The output quantities representing the amplitude and phase difference have a maximum small-signal envelope bandwidth of 30 MHz; this can optionally be reduced by adding external filter capacitors.

The AD8302 provides this powerful computational function for the first time in monolithic form using an advanced bipolar process. The excellent log-amp matching, high-frequency capability, and precise scaling of gain and phase measurement, all in a small footprint, open up new opportunities for *in-situ* monitoring and controlling of RF and IF systems in a noninvasive fashion. Operation from 2.7-V to 5.5-V supply voltages is provided at a current of only 20 mA. The product is available in a 14-pin TSSOP package.

ACKNOWLEDGMENTS

The authors would like to recognize the tireless efforts of Tom Kelly and Shirine Eslamdoust in product engineering, and Ron Simonson and Rick Cory for applications development.

High Performance Narrowband Receiver Design Simplified by IF Digitizing Subsystem in LQFP

By Paul Hendriks, Richard Schreier, and Joe DiPilato
(paul.hendricks@analog.com; richard.schreier@analog.com; joe.dipilato@analog.com)

INTRODUCTION

Mobile radios are used for public safety and emergency services—police, fire and ambulances—as well as for private services such as fleet management. Increasingly, in order to provide enhanced services, along with improved spectral efficiency and coverage, the design of these radios has moved from traditional analog-based modulation schemes, such as FM and PM, to digital modulation approaches.

Receivers for these radios must be capable of accurately digitizing a low-level, high-frequency signal in the presence of large interfering signals. In radios using some narrowband land mobile standards, interfering signals can be 70 dB greater than the desired channel, with frequency offsets as little as 25 kHz. Since these systems usually are not cellular, the geographical coverage range of mobile radios is also an important feature—they must possess excellent sensitivity to recover low-level signals originating from subscribers at the fringe of the coverage range. As a further complication, these radios are often portable with high rates of usage; they demand low power consumption using smaller, longer-lived batteries.

As an aid to equipment designers, Analog Devices has made available the AD9870 IF Digitizing Subsystem, an IC designed to meet the demanding requirements of land mobile radio, and similar narrowband radio applications, with superheterodyne architectures employing analog and/or digital modulation schemes. The AD9870 integrates the entire IF strip with minimal external components. It can accept an IF signal at frequencies as high as 300 MHz, with bandwidths up to 150 kHz, and provides a serial data output containing 16-bit I and Q data, which can then be demodulated with a host processor. The AD9870 is intended for both base stations and subscriber units, combining the dynamic range required by base stations with the low power consumption needed by portable radios.

The big problem in all receivers is dynamic range

The dynamic range of a receiver determines its ability to recover low-level signals in the presence of larger signals, known as *blockers* and *interferers*. Figure 1 shows the various sources that can reduce the effective dynamic range of any radio receiver.

Assume for the moment that the only signal present in the spectrum is the "small target signal." The minimum detectable signal or sensitivity will be determined by the signal bandwidth (B), the receiver's detection threshold (SNR_{MIN}), the receiver's noise figure (NF), and inherent thermal noise limitations (kTB). At a temperature of 290 K, the sensitivity can be estimated with the following equation:

$$Sensitivity = SNR_{MIN} + 10\log(B) + NF + (-174\ dBm/Hz)$$

Following are some of the potential noise sources:

Low-frequency 1/f noise becomes an issue if insufficient gain is applied to the target signal prior to down-conversion to frequencies below the 1/f corner of the process technology. DC components caused by offsets and second order distortion can also be problematic.

Large interferers can have their energy spread over a broad range of frequencies by the phase noise of the receiver's LO through a process known as "reciprocal mixing." The larger the interferer and the closer it is to the target signal, the more likely the target signal will be corrupted by the noise transfer mechanism. Also, if this interferer is large enough to induce nonlinearities in the receiver's front-end circuitry, it is possible for a spurious component to mix back into the target signal's passband. The "half-IF" problem is a specific case afflicting receivers with poor second-order linearity—in which an interferer falling halfway between the LO and the target signal generates a second order component that mixes with the LO's second harmonic to generate a spur falling on the target signal. The IIP2 specification of a receiver allows a receiver designer to quantify the "half-IF" spur. The difference, or Δ, between the interferer level, P_{IN}, and the resulting second-order spur is IIP2 – P_{IN}. With an IIP2 of 45 dBm, the AD9870 is mostly immune to this "half-IF" problem.

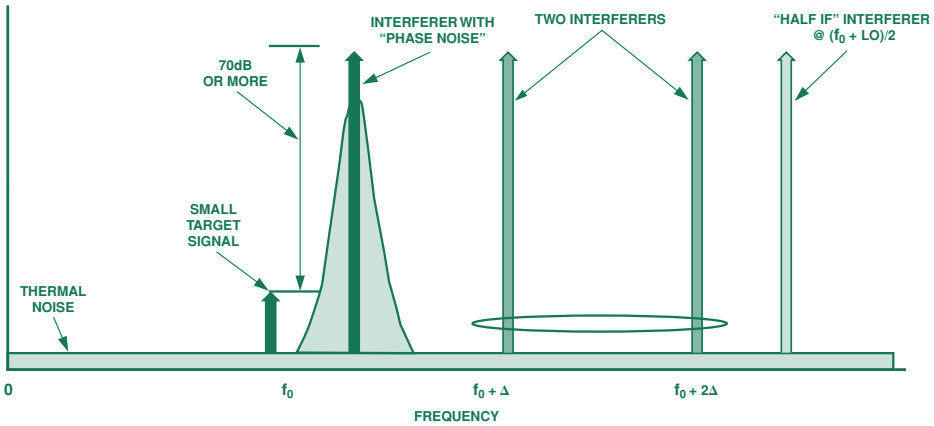


Figure 1. The "Big Problem" in all receivers is dynamic range.

Two large interferers at equally spaced frequency offsets (i.e., $f_0 + \Delta$ and $f_0 + 2\Delta$) from the target signal will result in a spurious component falling on top of the target signal through a process of intemodulation. The linearity of a receiver in this scenario is captured in its IIP3 specification with higher numbers representing a higher tolerance to third-order intermodulation. The difference, or Δ, between the two equal interferers, P_{IN}, and the resulting third-order intermodulation component is $2 \times (IIP3 - P_{IN})$. The AD9870 has a respectable IIP3 performance of –1 dBm, thus tolerating interferers as high as –45 dBm before degrading the receivers sensitivity.

Superheterodyne Architecture

To cope with large interferers that would otherwise degrade the receiver's ability to recover a target low-level signal, a superheterodyne architecture is used to translate an RF signal down to one or more intermediate (IF) frequencies where filtering of the adjacent interferer signals as well as amplification and gain control of the target signal is more practical. The superheterodyne scheme has been employed since World War I and is to this day the most popular of radio receiver architectures. A generic version employing this architecture, common among narrowband digital receivers, is shown by the signal-chain in Figure 2.

Prior to RF-to-IF down-conversion, a band-select filter (duplexer) and/or image reject filter selects the entire RF band within which the target signal operates. The low-noise amplifier (LNA), which provides amplification of the intended RF band prior to down-conversion, is critical in determining the receiver's sensitivity. The down-converted IF spectrum following the RF mixer often contains an array of signals of varying strengths in addition to the target signal. Channel selection and amplification occurs at IF: the target signal is selected from among the other signals via one or more crystal or SAW-type passive filters. After filtering, the target signal undergoes further amplification, with its signal strength stabilized at a preset level by an AGC loop to optimize the quadrature demodulation process. In many digital receivers, an IF analog quadrature modulator separates the IF signal into its quadrature baseband I and Q components, which are then digitized by a dual ADC. In such cases, the modulation accuracy of the demodulated signal is quite sensitive to analog offsets, quadrature LO mismatch, and I/Q gain mismatch in the quadrature modulator and dual ADC.

AD9870 Architecture

The AD9870 IF digitizing subsystem reduces the complexity of a typical superheterodyne receiver by integrating most of the IF, baseband, and some digital post processing functional blocks as shown in Figure 3.

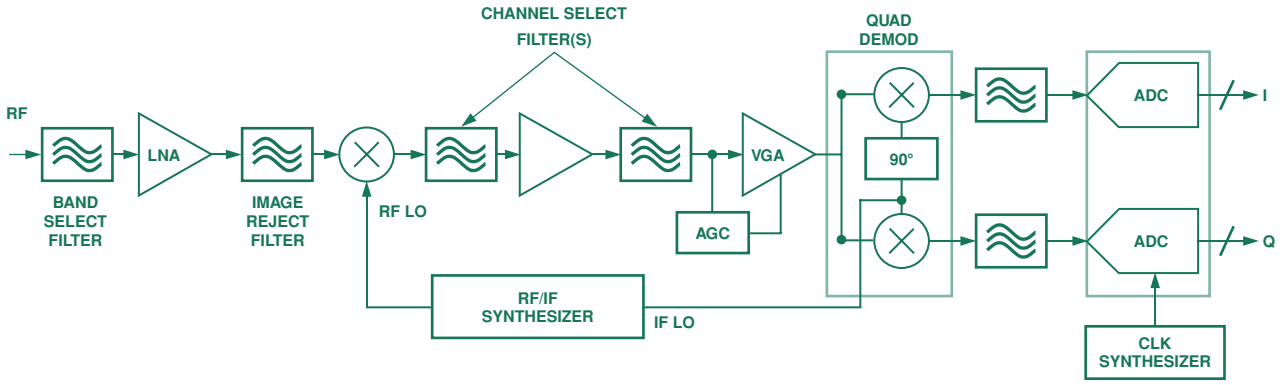


Figure 2. Typical superheterodyne architecture for a digital receiver.

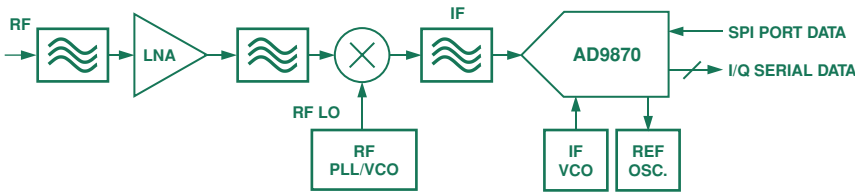


Figure 3. The AD9870 simplifies the digital receiver while enhancing performance.

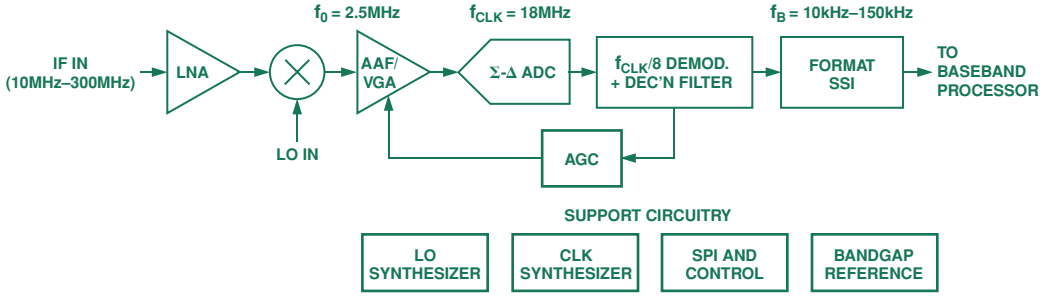


Figure 4. Functional block diagram of the AD9870 shows the level of integration.

The AD9870 differs from the typical superheterodyne architecture by employing a wide-dynamic-range *bandpass sigma-delta ADC* to sample a second-IF signal, along with any neighboring interferers. The demodulation of the target IF signal is performed with digital accuracy and stability, while the intrusive nearby interferers can be suppressed via digital filtering.

Figure 4 shows a functional block diagram of the AD9870. Functioning similarly to the RF portion of the superheterodyne architecture, an *LNA and mixer* are used to amplify and down-convert the target signal centered at the first-IF frequency to a lower second-IF frequency suitable for digitization by the bandpass ADC.

The LNA and mixer provide approximately 10.5 dB of gain, while preserving system dynamic range with an input noise figure of 9 dB and third-order intercept of 0 dBm. The high input impedance (360 Ω) simplifies interfacing to crystal or SAW filters. An *on-chip LO PLL synthesizer* can be used in conjunction with an external loop filter and VCO to generate a tunable LO frequency.

The second-IF signal is centered at exactly $1/8^{th}$ the bandpass ADC sample rate (i.e., $IF2 = f_{CLK}/8$) to allow for a simple $f_S/8$ digital quadrature demodulation scheme. Upon down-conversion to the second-IF, the signal is processed by a tunable (and programmable) *active third-order anti-alias filter (AAF)* to suppress signals which could appear within alias bands of the sampling ADC (i.e., $N \times f_{CLK}/8 \pm f_{CLK}/8$. The AAF tuning circuitry can support ADC sample rates between 13 and 18 MHz, with the 3 dB cut-off frequency typically set and tuned to slightly beyond the second-IF (i.e., $f_{-3dB} = f_{CLK}/3.2$).

Embedded in the AAF is a *variable-gain amplifier (VGA)* that provides up to 26 dB of gain range (Figure 5). The VGA gain, which extends the dynamic range of the AD9870, can either be programmed directly or controlled by an automatic gain-control (AGC) loop. The AGC loop is typically invoked under strong signal conditions to prevent "overloading" or clipping of the A/D converter by maintaining a programmable fixed-signal level at the ADC input. The AD9870 implements the AGC function with a highly effective hybrid approach, as shown in Figure 5: the analog and digital domains work together in signal estimation and control.

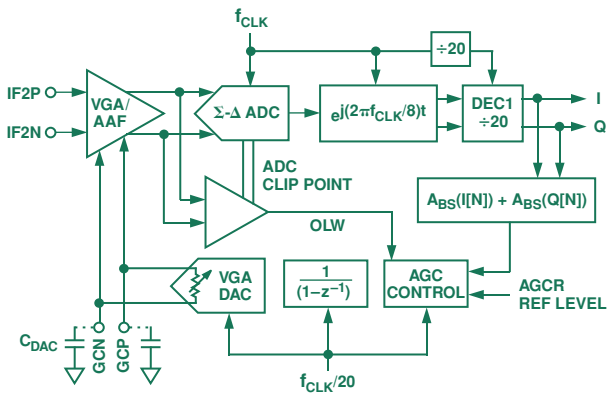


Figure 5. A "hybrid" AGC control loop extends the dynamic range of the AD9870.

In situations where a strong target signal or interferer falls within the bandwidth of the first-stage decimate-by-20 digital filter, the signal is estimated digitally and compared to a programmed reference level (AGCR). The difference between the two levels is fed to a digital integrator, which updates a control DAC to adjust the analog voltage of the VGA. Since a strong interferer falling outside of the passband of the first-stage digital filter can not be accurately estimated, an analog loop based on a simple differential comparator monitors the input to the ADC and assumes control of the loop during any overrange condition, to reduce the VGA gain.

An external capacitor is used to "smooth" the transitions of the DAC, with a time constant established by its capacitance and the internal source resistance of the DAC. The R-C cutoff frequency is typically set well outside the control system's loop bandwidth to ensure continual digital control of the loop dynamics. The control loop bandwidth is digitally programmable with attack and decay times variable over a wide range and ability to react to any overload condition.

The instantaneous dynamic range of any narrowband receiver signal chain containing a VGA is dependent on the particular gain setting of the VGA, since the noise contributed by each stage in the signal path to the "overall" input-referred noise decreases as the gain of the preceding stage increases. This implies that input noise described by its noise figure, NF, is typically dominated by the first few stages (i.e., LNA and mixer); noise sources at the end of the signal chain (i.e., the ADC) have minimal effect upon the system's NF, *provided* that there is sufficient gain between these blocks.

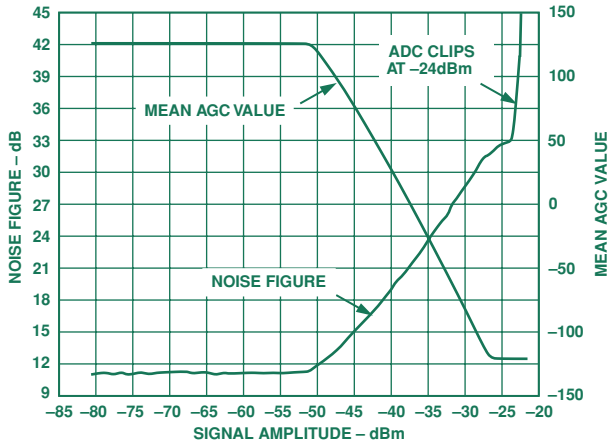


Figure 6. Dynamic range of AD9870 depends on VGA setting.

In the case of the AD9870, the VGA's gain is nominally adjustable over a 25 dB range. Figure 6 shows how the AD9870's noise figure is impacted by the VGA gain setting as a target signal's (or interferer's) input power is increased from –85 dBm to –23 dBm. Under small-signal conditions, the VGA is set to max gain; the AD9870's noise figure is set by the LNA/Mixer as well as the VGA's input noise. However, as the signal power is increased, a point is reached (depending on the AGC reference level) at which the VGA's gain begins to decrease to prevent ADC clipping. At this

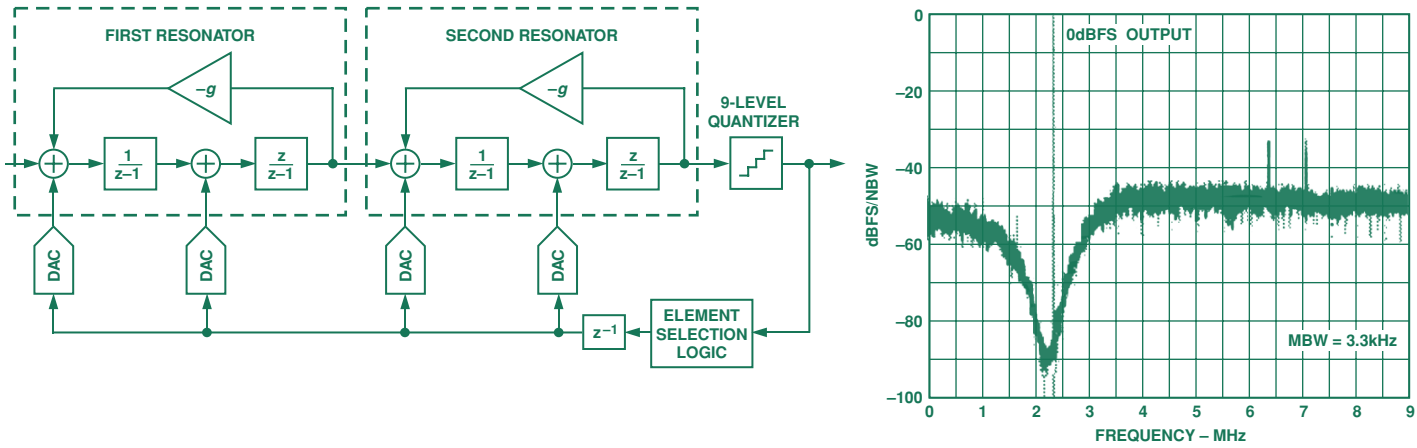


Figure 7. Multibit fourth-order bandpass Σ-Δ ADC results in deep notch at $f_{CLK}/8$.

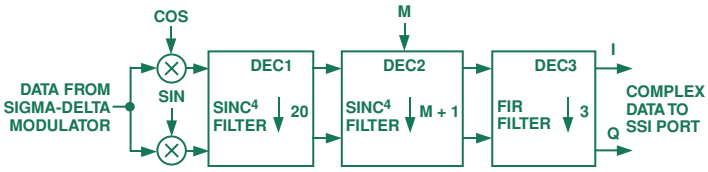


Figure 8. Digital quadrature demodulation, followed by programmable decimation filters, provides baseband I/Q data.

point, the VGA gain is reduced dB for dB as the signal power is further increased. Also, in this region, the input signal level to the ADC remains constant and the noise of the ADC begins to dominate such that the system's NF degrades also at a 1 dB per dB rate. As the signal power continues to increase, a point is reached (i.e., –26 dBm) at which the gain of the VGA is set to its absolute minimum and further increases in signal level are seen at the ADC input until clipping occurs (i.e., –24 dBm).

The bandpass sigma-delta ADC (Figure 7) is the "heart" of the AD9870 that makes a low second-IF digitization approach feasible and practical in an IC intended for radio systems requiring high dynamic range with minimal power consumption. This ADC, together with the back-end digital decimation filters, achieves nearly 14.5-ENOB performance within a 10-kHz bandwidth, while sampling a signal centered at frequencies as high as 2.25 MHz. It achieves these specifications while drawing a mere 13 mA from a 3.0 V power supply.

The sigma-delta ADC is based on a fourth-order switched-capacitor, multi-bit modulator consisting of two cascaded resonators that provide two complex pairs of zeros in the noise transfer function (NTF) falling near $f_{CLK}/8$. The location of these complex zeros at the second-IF frequency, along with the multibit feedback path, help ensure a low noise floor in a narrow region (±73.3% of $f_{CLK}/8$) around the second-IF frequency.

The digital output data from the ADC is fed into the digital signal processing section of the AD9870 (Figure 8). This section consists of an $f_{CLK}/8$ complex (or quadrature) demodulator followed by three linear-phase FIR filters. The complex demodulator separates the target second-IF signal centered at $f_{CLK}/8$ into its I/Q components prior to filtering.

The output spectrum of the complex demodulator consists of the target signal, now centered at "dc," along with any undesirable interferers and/or noise not sufficiently filtered in the analog domain. A series of decimation filters are used to remove these undesirable components, while simultaneously reducing the data rate in accordance to the target channel's bandwidth. Depending on the modulation scheme, the complex data rate (hence decimation factor) is set to be at least a factor of two greater than the channel bandwidth to allow for further post-processing. Both DEC1 and DEC2 use a cascaded fourth-order comb filter topology; DEC2's decimation factor is user programmable to accommodate different channel bandwidths. DEC3 is a decimate-by-3 FIR filter; it sets the close-in transition-band characteristics of the composite filter. The 16-bit I-and-Q output of DEC3 is fed into the synchronous serial-interface (SSI) function, which formats the data into a serial bit stream and embeds other optional information — AGC, signal strength, and synchronization—into the bit stream.

AVAILABILITY

The AD9870 was released to production in winter 2001. It is available in a 48-lead LQFP package for $16.96 in 1K volume.* The AD9870 datasheet is available on Analog Device's website (www.analog.com). An evaluation board and the associated software are also available.

*Recommended resale price USD. Prices are subject to change without notice. For specific price quotations, get in touch with our sales offices or distributors.

AD9873 Mixed-Signal Front End (MxFE) for Broadband Digital Set-Top Boxes

By Iuri Mehr, Joe DiPilato, and Martin Kessler

Widespread deployment of TV cable has led to extensive research in providing better quality and increased variety in TV programming and cable-modem functionality. This effort resulted in the development of digital set-top boxes from several established vendors, including Scientific Atlanta and Motorola (General Instrument). Instead of analog *vestigial-sideband* modulated (VSB) channels, digital set-top boxes receive TV programming and exchange information with the head-end station using *quadrature amplitude modulation* (QAM). Transmitting analog information in digital bits is not only more robust but also makes use of available bandwidth more efficiently. Figure 1 shows a digital set-top box connected to the head-end and to various devices inside the residence (or office).

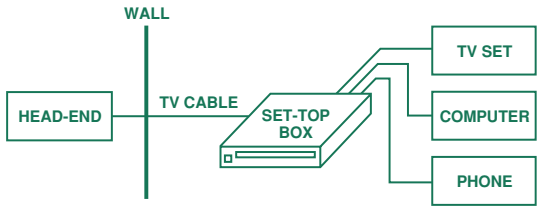


Figure 1. Cable set-top box gateway configuration.

Several services can be unified in this fashion, including Internet access, cable TV and even phone services. High data rates allow streaming-in MPEG movies as well as high-quality telephony (voice-packet) service.

A digital set-top box, like that shown in Figure 2, comprises several major subsystems to implement such functions as a TV tuner, base-band transceivers, a channel 3/4 modulator (for compatibility with analog TV sets), MPEG and NTSC decoders and encoders, physical layer (PHY), and media access control (MAC) for cable modems. Since Internet access implies an upstream channel, a cable driver is included; it can be implemented using a member of the AD832x family. The box can also include an *out-of-band* (OOB) control channel and a phone line interface.

The multiplicity and complexity of all these blocks impose significant challenges on designers at both the component and board levels. The large amount of digital processing required, combined with the high-quality reception requirements of high-definition digital TV, pose numerous challenges to digital set-top box architects. In addition, compatibility with analog TV calls for clean analog signal processing from the wall TV cable outlet to the TV set. Therefore, selecting a proper partition for integrated functions becomes a key requirement for combining high-quality TV reception and high data rates in a cable modem at low cost.

The mixed-signal front end, which can be implemented using the AD9873, is central to the set-top box (Figure 2).

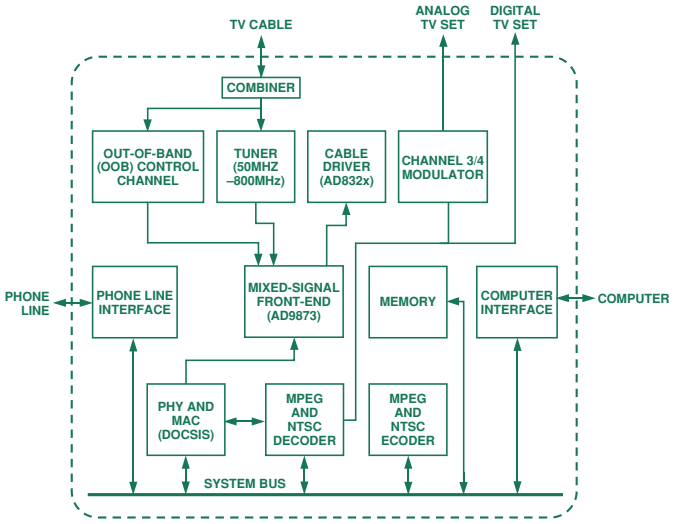


Figure 2. Inside a typical digital set-top box.

Mixed-Signal Front End

The definition of a mixed-signal front end for a set-top box must take into account the amount of functionality required from the *transmit* and *receive* data paths. Low cost is of vital importance, so selecting a proper technology is key to a successful design. In addition, time to market is equally important for both the IC vendor and the OEM. ASICs that include significant digital and analog content are often difficult to schedule due to the time needed to handle the inherent design challenges and the frequent need for customer feedback. The designers of the Analog Devices AD9873 took advantage of both their experience in set-top-box technology and their inventory of high-performance topological block core designs of the kind that would be needed to integrate the essential high-performance analog and mixed-signal functions on a single chip.

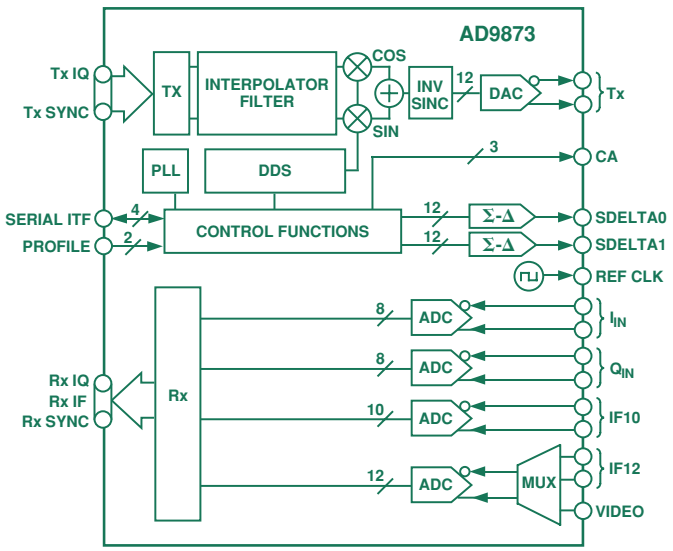


Figure 3. AD9873 functional block diagram.

Figure 3 is a block diagram of the AD9873 Analog Front End Converter for Set-Top Boxes and Cable Modems. The *receive* data path contains several analog-to-digital converters (ADCs) to accommodate the various set-top box functions described earlier. A pair of 8-bit ADCs is used to convert quadrature inputs from the demodulated OOB channel. They are designed for modest

performance—better than 7 effective number of bits (ENOB) when sampling at less than 16 MHz—since the OOB data utilizes low-complexity modulation (QPSK) in a narrow-band channel (<1 MHz). A tighter specification is required from the 10-bit ADC, due to its major role in digitizing cable modem data. This type of data is broadcast using higher-order QAM modulation, which requires a higher signal-to-noise ratio. Hence, the converter needs to exhibit better than 9 ENOB when sampling an input signal of up to 10 MHz at 33 MSPS. The fourth ADC, a 12-bit converter, sampling at 33 MHz and providing better than 10.5 ENOB for inputs up to the Nyquist rate, can digitize high-definition TV signals. For single-ended video signals multiplexed to the same input, a programmable black-level clamp is provided. The outputs of all these converters are multiplexed to reduce the number of package pins.

The *transmit* data path contains a demultiplexed interface, which receives I/Q baseband data, typically sampled at about 13 MHz (up to 16 MHz). Since interpolation is a powerful tool for reducing DAC output filter requirements (used successfully in the AD9772 and AD9856), three interpolation filters are used. The interpolation factor can be programmed for 12 or 16, bringing the data rate up to 230 MHz. The overall interpolator frequency response is determined by two half-band filters and a cascaded integrator comb filter (CIC). Following the interpolator, a quadrature modulator is implemented using *direct digital synthesis* (DDS) to generate the sine and cosine waveforms. Before being fed into the DAC, the signal can be compensated for the $\sin(x)/x$ roll-off, which results from the D/A conversion process. This operation is optional, since the roll-off becomes noticeable only toward the end of the synthesized carrier frequency range. The DDS can produce a low-spurious-content complex carrier at frequencies up to approximately one-third of the sampling rate, that is, up to 70 MHz.

The ADCs are clocked directly from a low-frequency crystal; its frequency is stepped up by an on-board programmable phase-locked-loop (PLL) to provide the high-speed clock required for the DAC. This approach reduces undesirable clock jitter when sampling the ADCs and eliminates the problems and expense of an off-chip high-frequency oscillator. The programmable PLL also provides the system clock to other blocks within the set-top box. Auxiliary digital sigma-delta outputs facilitate automatic gain control or timing recovery functions. Many of the device parameters are programmable through a 3- or 4-pin serial interface.

In order to seamlessly interface with a member of the AD832x cable driver family, a separate 3-wire interface is included, and several profile registers (which can be loaded through the serial interface) are designed to speed up changes in transmit gain data and carrier frequency. This can be achieved by using dedicated external pins that address a particular profile register bank. Figure 4 shows how the AD9873 would be used in a complete digital set-top box application.

Designers of broadband modems require the combination of small form factors, high performance levels, and low cost. Because of the cost and area impact of heat management within the box, they cannot afford to dissipate watts of power in the transmit or receive path. To build the large-scale digital integrated circuits that meet these requirements in broadband modem designs, state of the art, low-voltage lithographies are needed. However, they are not suitable for high-performance analog and mixed-signal circuitry. Products like the AD9873 provide a solution to this problem by offering the possibility of using two small highly integrated chips—a digital ASIC and a mixed-signal "everything else"—that appropriately partition the large scale digital IC from the high-performance mixed-signal component.

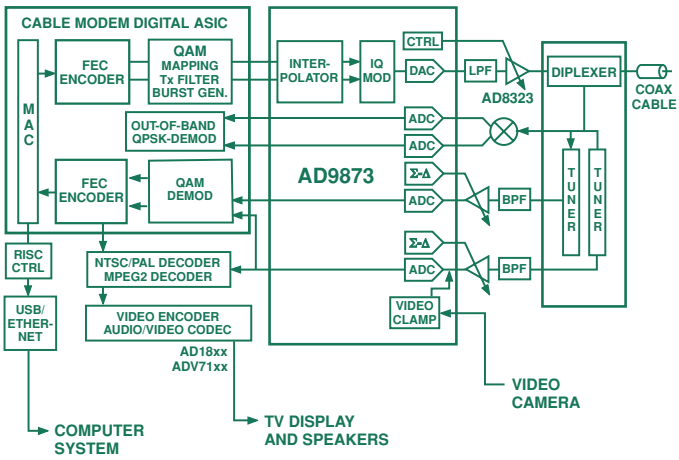


Figure 4. Intelligent system partitioning helps to solve the challenge of optimizing price, performance, size, and power in broadband modem designs.

Figure 5 demonstrates how this approach leads the trend in broadband communication applications away from single-chip solutions that unsuccessfully attempt to integrate the large scale digital processing with the high-performance mixed-signal devices. Emerging broadband modems require both more powerful digital processing (>MIPS) and higher performance mixed-signal (> dynamic range and bandwidth) devices. The large scale integrated digital devices used in these applications need the utmost state-of-the-art (fine geometry), low voltage, CMOS processes, while the mixed-signal devices depend on higher voltage CMOS processes that are optimized for handling mixed signals with high performance. As the first device of a new family of broadband MxFEs, the AD9873 will allow designers to take advantage of "smart partitioning."

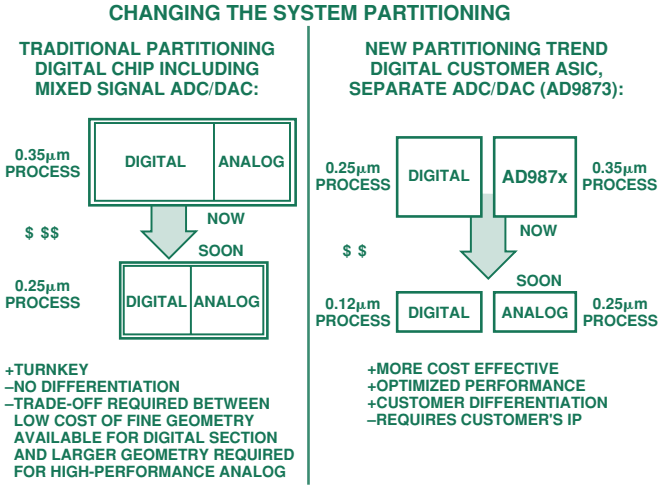


Figure 5. Smart partitioning model.

Here's why it works: Deep submicron geometry processes do not readily support the voltage levels required by high-performance D/A and A/D converters, and coupling of digital noise into the analog signal chain will corrupt signal fidelity. There are times when trying to put everything on a single chip results in a higher-price and/or lower-performance solution. Trying to mix high-speed and wide dynamic range mixed-signal devices with very large scale digital processing is a perfect case in point. It will always require compromises either in digital area (cost), power consumption, or mixed-signal performance. The AD9873 broadband Mixed-Signal Front End, and the other MxFE products that will follow in its wake, gives designers the benefit of high integration, low cost, and low power consumption, without compromising performance.

The AD9873 applies this optimized mixed-signal technology and "smart" partitioning to provide excellent dynamic performance for a variety of modulation formats—FSK, QPSK, 16/32/64/256 QAM, OFDM, spread spectrum, etc. The digital ASIC, which includes the modulation encoding, can be implemented on the most cost effective and finest geometry possible. With this cost-effective approach, system designers can keep more "value added" in their own digital ASIC, making best use of their system expertise, proprietary algorithms, and intellectual property. The AD9873's mixed-signal partitioning resolves the cost and performance trade-off issues related to integrating mixed-signal circuits within VLSI digital ASICs by getting them off the chip.

Other Applications for the AD9873 Mixed-Signal Front End

Besides cable set-top boxes, the AD9873 is well suited for a variety of other standard and proprietary broadband communications applications, as depicted in Figure 6. Here is a list of other applications where the AD9873 can be used:

- Cable Modem
- Digital Communications
- Data and Video Modems
- Power Line Modem
- Satellite Systems
- PC Multimedia
- Broadband Wireless Communication
- Home Networking

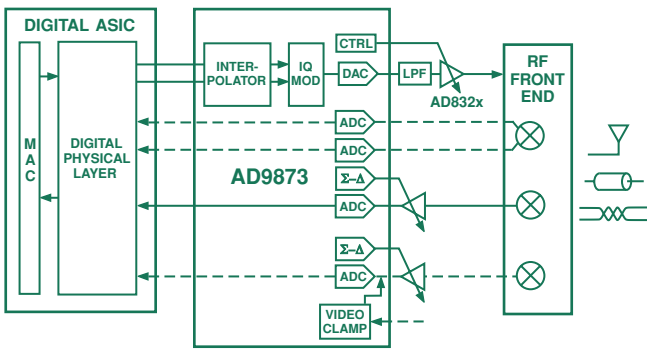


Figure 6. Broadband modems over cable, power line, or wireless, using the AD9873.

AD9873 Key Features and Performance

- 232 MHz Quadrature Digital Upconverter:
  - DC to 70 MHz Output Bandwidth
  - 12-Bit Direct IF D/A Converter
  - Direct Digital Synthesis
  - Interpolation and Sin(X)/X Filters
- 12-Bit 33 MSPS Direct IF ADC
- 10-Bit 33 MSPS Direct IF ADC
- Dual 8-Bit 16.5 MSPS I&Q ADCs
- Dual 12-Bit Sigma-Delta Control DACs
- Video Input with Clamp Circuitry
- Direct Interface to AD8321/AD8313 PGA Cable Driver
- Programmable PLL Clock Multiplier
- Single 3.3 V Supply Operation
- Power-Down Modes
- 100-Lead MQFP

Performance of the AD9873 was characterized with respect to the commercial temperature range; however, it can be safely used from –40°C to +75°C. Figure 7 shows a spectral plot of the 12-bit ADC performance with a 10-MHz input.

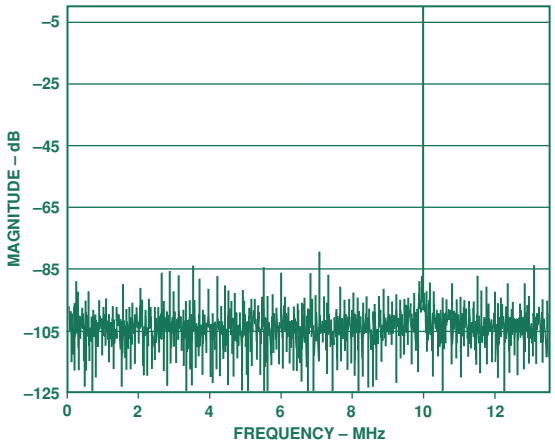


Figure 7. AD9873's 12-bit ADC performance plot with 10-MHz input.

Figure 8 shows the spectral plot of the DAC producing a 42-MHz 16-QAM signal. Figure 9 shows the constellation and eye diagram of a 64-QAM signal generated by the AD9873.

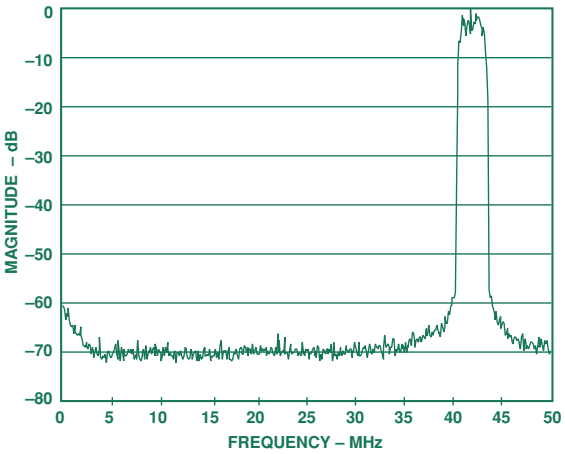


Figure 8. AD9873 DAC performance plot.

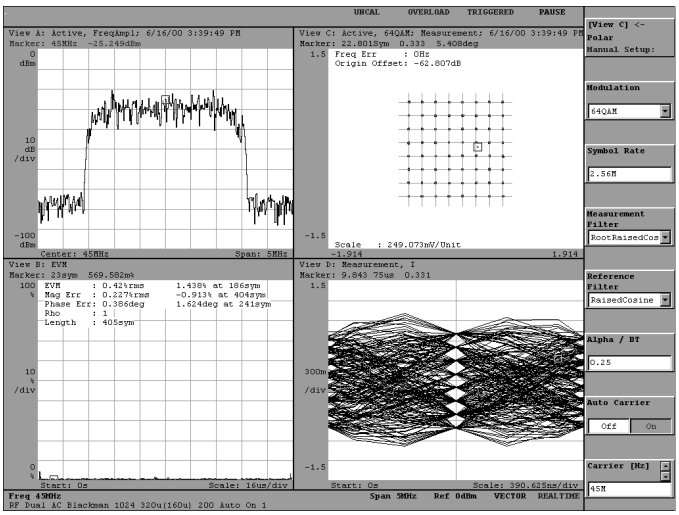


Figure 9. AD9873 64-QAM constellation plot.

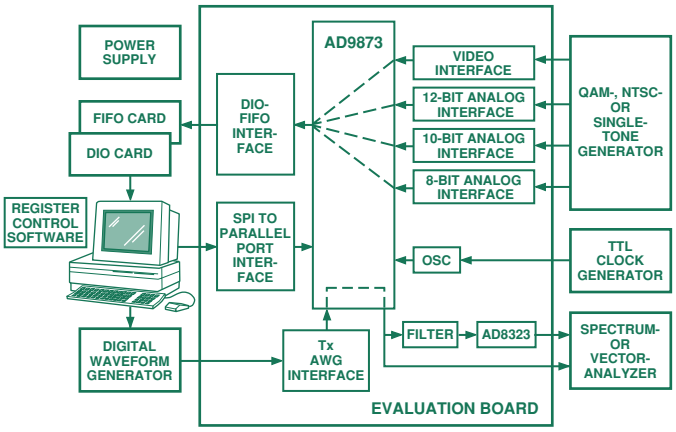


Figure 10. AD9873 evaluation setup.

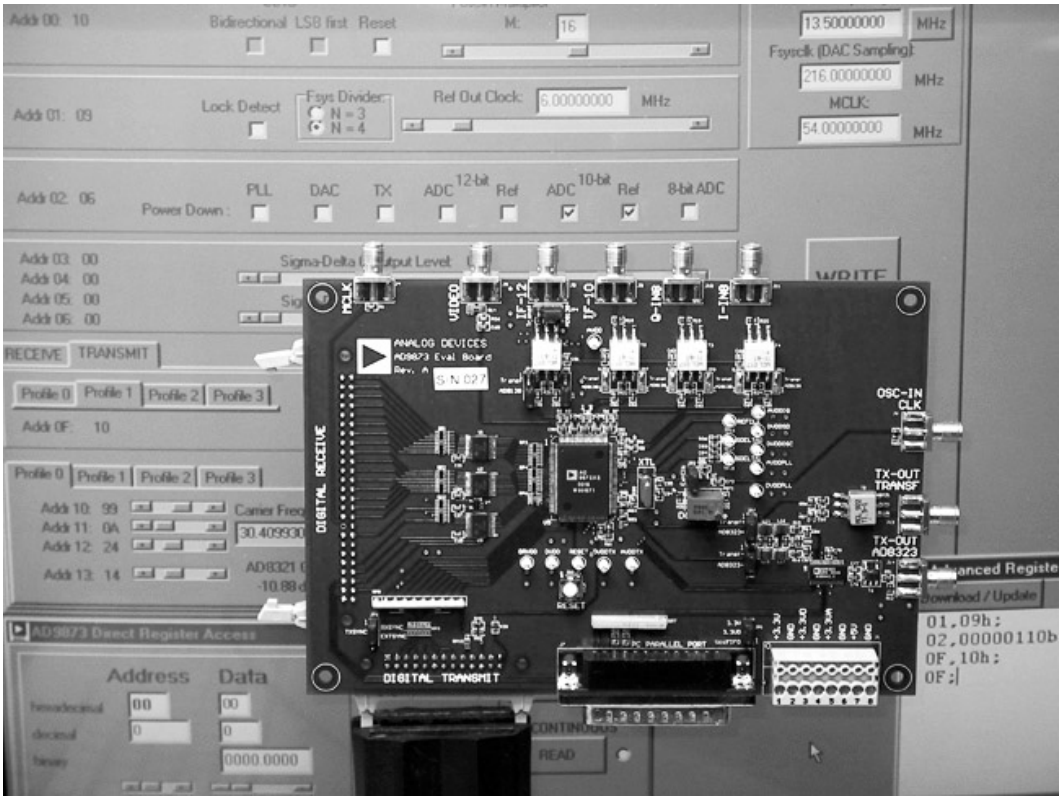


Figure 11. AD9873 evaluation board and software interface.

Evaluation Board and Software

The AD9873 Evaluation board and software allow users to easily program and quickly evaluate the AD9873 for a specific modem application.

AVAILABILITY

The AD9873 was released to production in summer 2000. It is available in a 100-lead MQFP package, is priced at $16.58 (1000s), and sells for less then $10 in high volume.

Design a Direct 6-GHz Local Oscillator with a Wideband Integer-*N* PLL Synthesizer

By Mike Curtin (mike.curtin@analog.com)

INTRODUCTION

Establishing a new benchmark for speed and RF phase-noise performance, the ADF4106 Phase-Locked-Loop Synthesizer is fully specified to operate at frequencies up to 6.0 GHz. This allows designs for the 5.4-GHz to 5.8-GHz upper ISM band to be greatly simplified. Fabricated on an advanced 0.35-µm BiCMOS process, it displaces the pin- and software-compatible 4-GHz ADF4113 as the fastest available integer-N synthesizer—and can achieve 3-dB lower phase noise to boot! It requires only a 3.3-V supply, yet its V_P pin is specified at up to 5.5 V for compatibility with tuning voltage levels often required by modular VCOs used in base stations.

The ADF4106 frequency synthesizer (Figure 1) can be used to implement local oscillators (LOs) in the up- and down-conversion sections of wireless receivers and transmitters. It consists of a low-noise digital phase frequency detector (PFD), a precision charge pump, a programmable reference divider, programmable A and B counters, and a dual-modulus prescaler (P/P + 1). The A (6 bit) and B (13 bit) counters, in conjunction with the dual-modulus prescaler (P/P + 1), implement an N-divider (N = BP + A). In addition, the 14-bit reference (R) counter allows selectable REF_{IN} frequencies at the PFD input. A complete phase-locked loop (PLL) can be implemented if the synthesizer is used with an external loop filter and voltage-controlled oscillator (VCO). Its very high bandwidth means that frequency doublers can be eliminated in many high-frequency systems, simplifying system architecture and lowering cost.

Wide bandwidth allows it to function as a 6-GHz local oscillator

The standard PLL system architecture used by the ADF4106 and its predecessor, the ADF4113, is shown in Figure 2. Since the maximum operating frequency of the ADF4113 is about 4 GHz, higher frequencies require the use of a frequency doubler—which usually calls for an extra RF amplifier to produce an adequate level for the doubler. Use of the ADF4106 eliminates the frequency doubler and its associated circuitry, achieving a much simpler and more power-efficient LO. For example, the design shown in Figure 3 generates RF output frequencies with 1-MHz channel separation from 5.4 GHz up to 6.0 GHz. The phase noise measured at the upper end is –83 dBc/Hz.

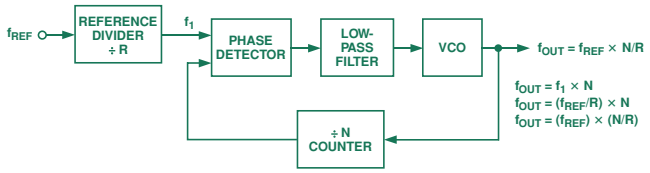


Figure 2. Standard PLL architecture.

Because the input impedance of the ADF4106 at this high operating frequency is very close to 50 Ω, a 50-Ω terminating resistor at the RF input is not needed for maximum power transfer efficiency. When operating at lower frequencies, the *s*-parameters in the data sheet give the impedance values needed for matching.

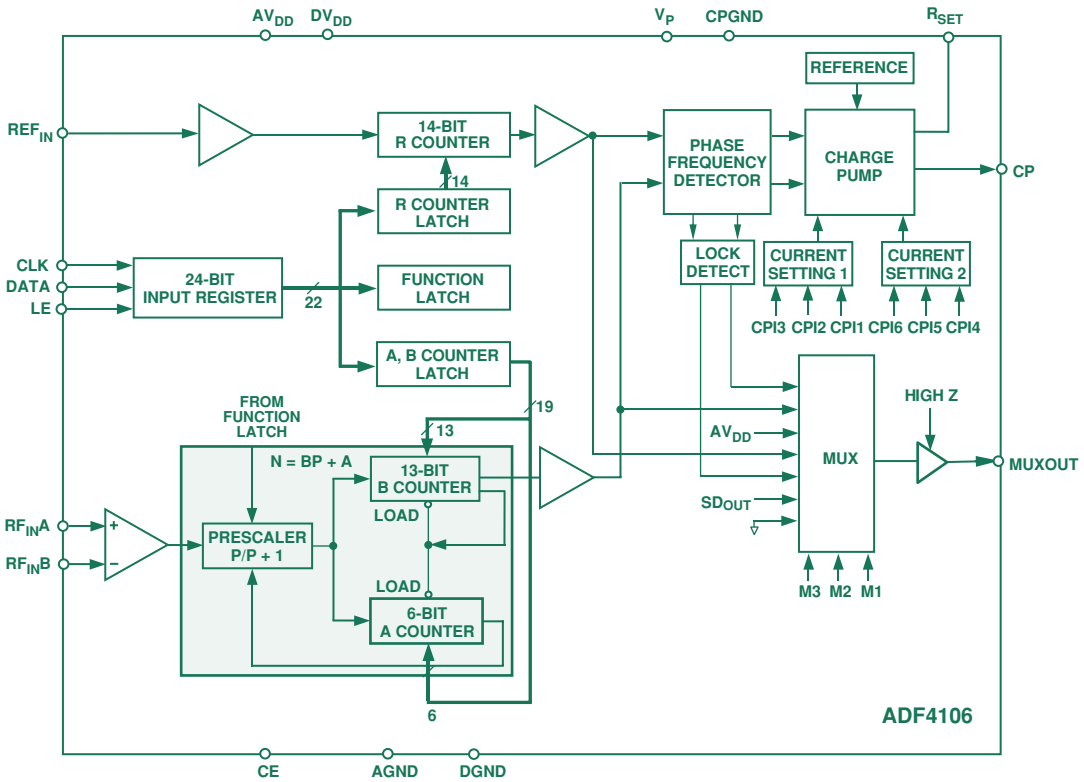


Figure 1. Functional block diagram of the ADF4106.

Low phase noise allows it to work as a low-noise, fast-settling 1.5-GHz local oscillator

The ADF4106, in conjunction with a wide bandwidth divider, can improve the phase noise and lock time of a standard local oscillator circuit at frequencies below 2.0 GHz.

A typical wireless system might be generating frequencies in 200-kHz increments from 1450 MHz to 1500 MHz. Using an integer-N architecture to do this, a phase/frequency detector reference frequency of 200 kHz is needed, and the N value would vary from 7250 (1450 MHz) to 7500 (1500 MHz).

Using the ADF4106 for best performance would give a phase noise figure of –88 dBc/Hz. Typical reference spurs in such a system would be –88 dBc at 200 kHz and –90 dBc at 400 kHz. Implementing a loop bandwidth of 20 kHz, typical lock time to 10 degrees of phase error would be 250 μs.

However, the wideband operation possible with the ADF4106 allows an alternative architecture to be considered, shown in Figure 4a. In this configuration, the core PLL is operated at a multiple of the final desired output frequency. In the example given above, the final desired frequency range is 1450 MHz to 1500 MHz. A multiple within the device's frequency range is 5800 MHz to 6000 MHz (four times the desired output band). In the proposed scheme, shown in Figure 4a, f_{PFD} operates at 800 kHz, the f_{VCO} band is 5800 MHz to 6000 MHz, and the final system LO output is obtained by dividing f_{VCO} by four.

$$f_{OUT} = (f_{PFD} \times N)/X \quad (1)$$

Some consequences of using this architecture follow.

Phase-noise reduction (see page 4)

The synthesizer phase noise has a 10 logf_{PFD} relationship. This means that for every doubling of the PFD frequency, there will be 3-dB degradation in the synthesizer phase noise. However, the output from the VCO will be divided down, and its phase noise obeys a 20 log$X$ rule. So, for every doubling of X, there will be a gain of 6 dB in phase noise performance. If the PFD frequency is quadrupled as above, f_{VCO} is divided by four to end up with the correct f_{OUT}. Thus 6 dB will be lost due to the quadrupling of f_{PFD} and 12 dB is gained due to the division by four. This results in an overall gain of 6 dB in phase-noise performance, using Figure 4a, compared to the use of the standard architecture. In this example, the resulting phase noise would be –94 dBc/Hz.

Reference spur reduction

In an integer-N PLL, spurious frequencies occur at integer multiples of the PFD frequency at the VCO output. In Figure 4a, using f_{VCO}, these spurs will be at f_{PFD}, $2f_{PFD}$, $3f_{PFD}$, etc. However, at f_{OUT}, the fundamental frequency is divided by X (X = 4) but the spurious frequencies still exist at the integer multiples of the PFD frequency. Note, however, that they are reduced in amplitude by 20 log$X$ (20 log4 = 12 dB). See Figure 4b.

So, using the architecture of Figure 4a with X = 4 and generating an f_{OUT} of 1450 MHz to 1500 MHz with 200-kHz spacing, the frequency spurs will exist at integer multiples of 800 kHz, the PFD frequency at levels below –90 dBc. Note that although the step frequency is 200 kHz, the lowest frequency spur is at 800 kHz.

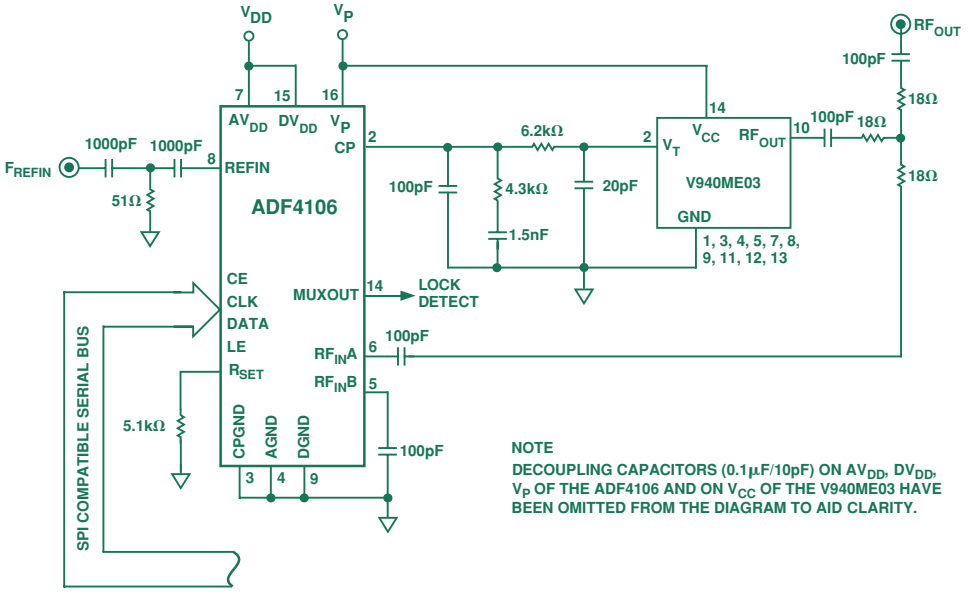


Figure 3. The ADF4106 used to implement a 6.0-GHz local oscillator.

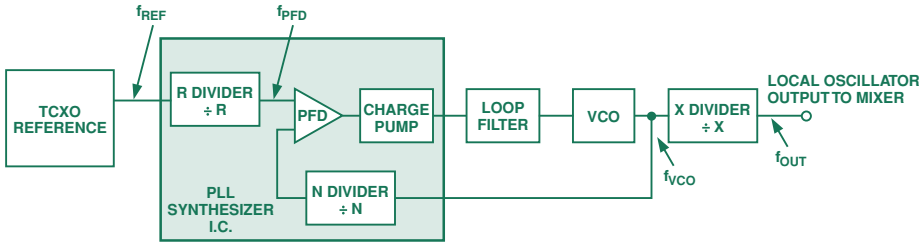


Figure 4a. Architecture for improved lock time, phase noise, and reference spurs.

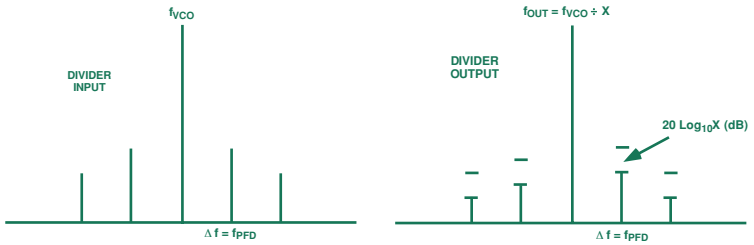


Figure 4b. Comparing the output spectrum at f_{VCO} and f_{OUT} of Figure 4a.

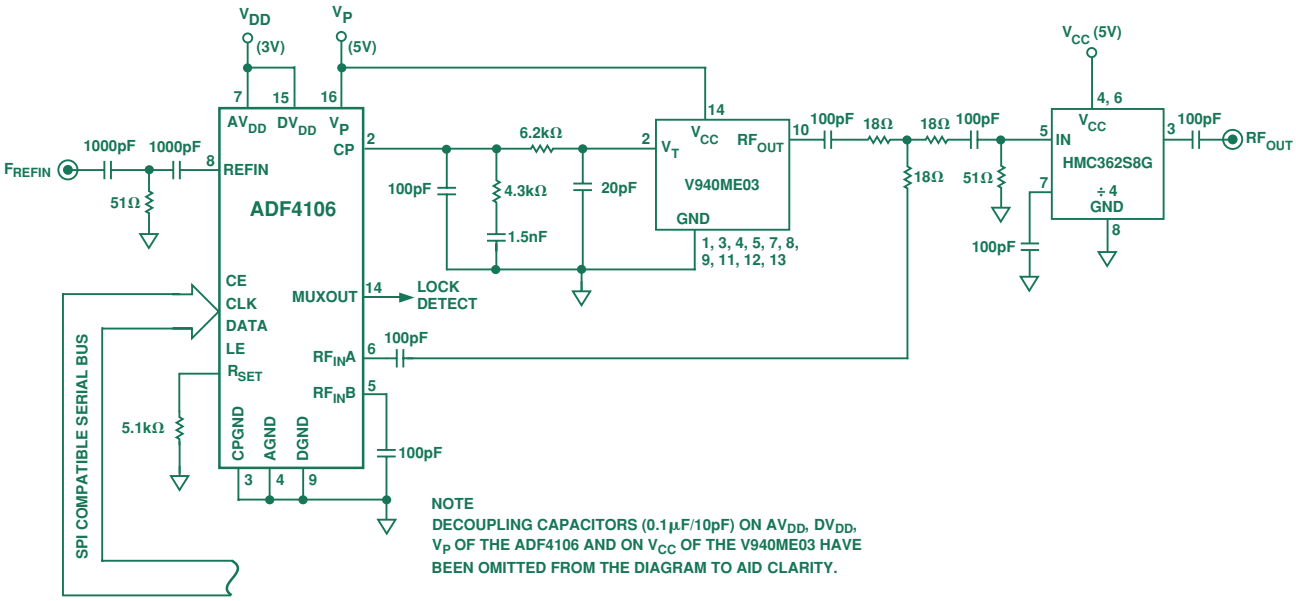


Figure 5. Using the ADF4106 with an output divider to generate a 1.5-GHz local oscillator.

Shorter lock time

Since the PFD in Figure 4a is operating at a higher frequency, phase comparisons are occurring at a higher rate; this will cause the loop to lock faster. In addition, because of the higher PFD frequency, a wider loop bandwidth is possible, and this too helps in improving the lock time. In this example, the lock time is about 70 µs to within 10 degrees of phase error for a PLL loop bandwidth of 80 kHz.

The actual implementation of Figure 4a is shown in Figure 5.

To summarize, the circuit of Figure 5 provides the following performance:

Phase Noise	–94 dBc/Hz @ 1-kHz offset
Reference Spurs	<–100 dBc (system noise floor) @ 200-kHz, 400-kHz, 600-kHz offsets
	–90 dBc @ 800-kHz offset
Lock Time	70 µs to within 10 degrees phase error

The price of this improved performance is the extra cost of the output divider and the extra power consumption of the system as a whole (the HMC typically adds 68 mA to the ADF4106's 13-mA current requirement). Thus improved performance must be a critical requirement for selecting this architecture. The extra board space needed for implementation is minimal since the HMC comes in an 8-lead SOIC package.

BANDWIDTH

The 0.35-µm BiCMOS fabrication process and careful application of RF design techniques permit the prescaler section of the ADF4106 to operate at up to 6.0 GHz with an input level of –10 dBm (referred to 50 Ω), guaranteed over the industrial temperature range (–40 to +85°C). Figure 6 shows a typical sensitivity plot for the ADF4106 in a TSSOP package at –40°C, +25°C, and +85°C. It can clearly be seen that performance to 6 GHz is well within the limits of the device with signals below –15 dBm.

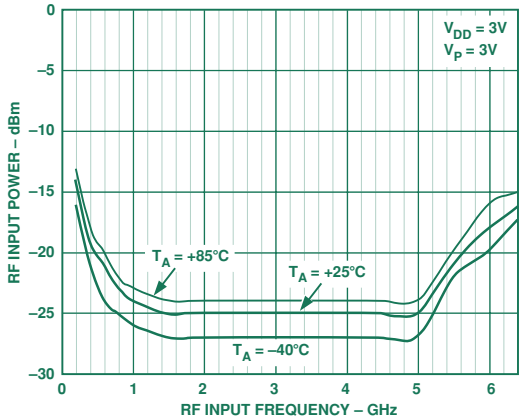


Figure 6. ADF4106 sensitivity vs. frequency.

PHASE NOISE

Phase noise, a measure of the purity of the local oscillator signal, is the single most critical specification in the local oscillator section of radios—with a direct bearing on receiver sensitivity. It is the ratio to output carrier power of the noise power in a 1-Hz bandwidth at a given offset from the carrier. Expressed as a log ratio, the units of phase noise are dBc/Hz. Phase noise is typically measured with a spectrum analyzer.

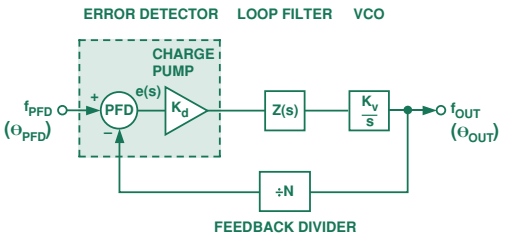


Figure 7. Basic phase-locked loop model.

The circuit of Figure 7 is used as the circuit model for the discussion of phase noise.

Total phase noise in a phase-locked loop (dB) can be expressed as follows:

$$PN_{TOTAL} = PN_{SYNTH} + 20\log N + 10\log f_{PFD} \quad (2)$$

where

PN_{TOTAL} is the total phase noise of the PLL

PN_{SYNTH} is the phase noise due to the PLL synthesizer circuit itself

20 log$N$ is the increase of phase noise due to the frequency magnification associated with the feedback ratio, 1/N.

10 logf_{PFD} is the increase of noise associated with the incoming PFD frequency.

The graph in Figure 8 shows the ADF4106's phase noise characteristics as a function of PFD frequency, f_{PFD}.

With a given measured total noise, synthesizer noise can be inferred as:

$$PN_{SYNTH} = PN_{TOTAL} + 20\log N + 10\log f_{PFD} \quad (3)$$

This provides a figure of merit for the PLL Synthesizer circuit itself, irrespective of the noise contributed by PLL N value and PFD frequency, since these would be the same for any similar circuit being compared. For the ADF4106, this figure is –219 dBc/Hz, a 3-dB improvement on the ADF4113, which had been the best available integer-N synthesizer in terms of phase noise.

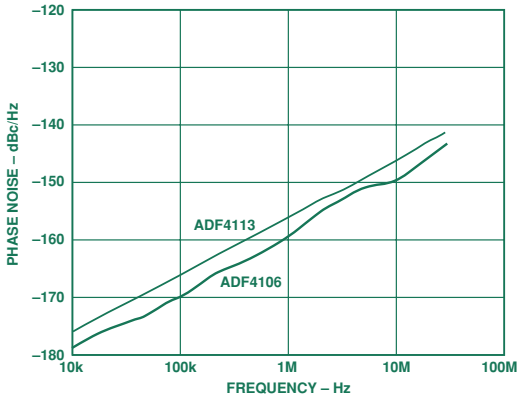


Figure 8. ADF4106 phase noise vs. PFD frequency.

With this phase-noise figure of merit, an engineer can work out the total PLL phase noise for any given PFD frequency and RF output frequency. For example, consider generation of a local oscillator signal with frequencies from 1700 MHz to 1800 MHz and channel spacing of 200 kHz. Using equation (2), the close-in phase noise using the ADF4106 as the PLL synthesizer is

$$PN_{TOTAL} = -219 + 20\log(9000) + 10\log(200 \times 10^3)$$

$$= (-219 + 79 + 53)\ \text{dBc/Hz}$$

$$= -87\ \text{dBc/Hz}$$

Figure 8 shows that the ADF4106 obeys the 10 logf_{PFD} "rule" (PFD phase noise substantially linear with log frequency) fairly consistently all the way to 30 MHz. Some integer-N devices begin to degrade rapidly once the PFD frequency goes above 1 MHz.

Note that the –219 dBc/Hz figure of merit is obtained by extrapolating Figure 8 back to 1 Hz. The graph can be used to quickly identify the performance possible in a given PLL setup once the N value is known. For example, 200 kHz on the graph corresponds to phase noise of –166 dBc/Hz. Adding 20 log$N$ (79 dBc) gives PLL phase noise of –87 dBc/Hz.

ACKNOWLEDGEMENTS

A note of thanks to Bill Hunt for valuable editorial comments, and to Brendan Daly, who verified the circuits used. The author would also like to acknowledge all the ADI customers who have provided valuable feedback on this device.

References

[1]. Mini-Circuits Corporation, *VCO Designer's Handbook*, 1996.

[2]. L.W. Couch, *Digital and Analog Communications Systems*. Macmillan Publishing Company, New York, 1990.

[3]. P. Vizmuller, *RF Design Guide*, Artech House, 1995.

[4]. R.L. Best, *Phase Locked Loops: Design, Simulation and Applications*, 3rd Edition, McGraw Hill, 1997.

[5]. Bannerjee, Dean, "PLL Performance, Simulation and Design," National Semiconductor website.

[6]. Analog Devices, Inc., ADF4106 Data Sheet. Online at www.analog.com.

[7]. Hittite Microwave Corporation, HMC362S8G Data Sheet.

[8] Curtin, Mike, "Phase-Locked Loops," *Analog Dialogue*, Volume 33, 1999. Three-part series in *Analog Dialogue* 33-3, 33-5, and 33-7 (1999) online at www.analog.com/publications/magazines/Dialogue/archives.html.

[9] Goldberg, Bar-Giora, *Digital Frequency Synthesis Demystified*, (LLH, 1999).

[10] Egan, William F., *Frequency Synthesis by Phase Lock: 2nd Edition*, John Wiley and Sons (1999).

Winning the Battle Against Latch-up in CMOS Analog Switches

by Catherine Redmond
(catherine.redmond@analog.com)

INTRODUCTION

This article will briefly describe the causes, mechanisms, and consequences of latch-up and discuss available prevention methods. Although our aim is to give an understanding of latch-up as it occurs in CMOS switches, similar principles apply to many other CMOS devices. *Latch-up* may be defined as the creation of a low-impedance path between power supply rails as a result of triggering a parasitic device. In this condition, excessive current flow is possible, and a potentially destructive situation exists. After even a very short period of time in this condition, the device in which it occurs can be destroyed or weakened; and potential damage can occur to other components in the system. Latch-up may be caused by a number of triggering factors, to be discussed below—including overvoltage spikes or transients, exceeding maximum ratings, and incorrect power sequencing.

Cause

For an understanding of latch-up, it is desirable to briefly review the basics and understand the participating components. As already stated, latch-up occurs as a result of triggering a parasitic device—in effect an SCR (silicon controlled rectifier), a four-layer pnpn device formed by at least one pnp and at least one npn transistor connected as shown in Figure 1.

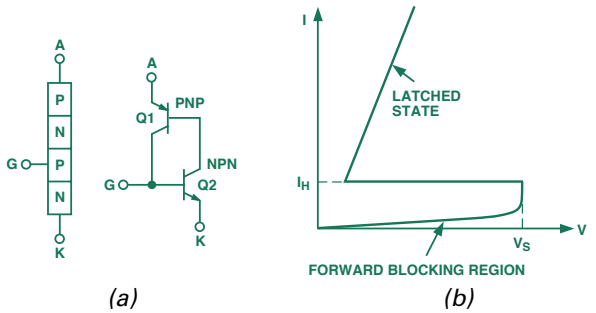


Figure 1. a) Transistor equivalent of an SCR.
b) Current voltage characteristic of an SCR.

An SCR is a normally *off* device in a "blocking state," in which negligible current flows. Its behavior is similar to that of a forward-biased diode, but conducts from anode, A, to cathode, K, only if a control signal is applied to the gate, G. In its normally off state, the SCR presents a high impedance path between supplies. When triggered into its conducting state as a result of excitation applied to the gate, the SCR is said to be "latched." It enters this state as a result of current from the gate injected into the base of Q2, which causes current flow in the base-emitter junction of Q1. Q1 turns on causing further current to be injected into base of Q2. This positive feedback condition ensures that both transistors saturate; and the current flowing through each transistor ensures that the other remains in saturation.

When thus latched, and no longer dependent on the trigger source applied to the gate (G), a continual low-impedance path exists between anode and cathode. Since the triggering source need not be constant, it could simply be a spike or a glitch; removing it will not turn off the SCR. As long as the current through the SCR is sufficiently large, it will remain in its latched state. If, however, the current can be reduced to a point where it falls below a holding-current value, I_H, the SCR switches off. Figure 1b shows the current-to-voltage transfer function for an SCR. In order to bring the device out of its conductive state, either the voltage applied across the SCR must be reduced to a value where each transistor turns off, or the current through the SCR must be reduced below its holding current.

A CMOS switch channel effectively consists of PMOS and NMOS devices connected in parallel; control signals to turn it off and on are applied via drivers. Since all these MOS devices are located close together on the die, it is possible that with appropriate excitation, parasitic SCR devices may conduct—a form of behavior possible with any CMOS circuit. Figure 2 illustrates a simplified cross section showing two CMOS structures, one PMOS and one NMOS; these could be connected together as an inverter or as the switch channel. The parasitic transistors responsible for latch-up behavior, Q1 (vertical PNP) and Q2 (lateral NPN) are also shown.

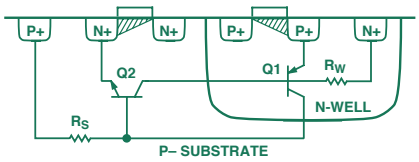


Figure 2. Cross-section of PMOS and NMOS devices, showing parasitic transistors Q1 and Q2.

P– substrate is used in devices from the ADG7xx family of switches and multiplexers, while devices from ADG4xx and ADG5xx families use N+ substrate. From Figure 2, it can be seen that a reinterpretation of the silicon configuration shows that the inherent parasitic bipolar transistors, Q1 & Q2, produce the parasitic SCR structure discussed above (Figure 3).

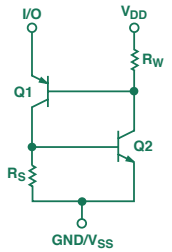


Figure 3. Rearrangement of the way we view the parasitic bipolars of Figure 2 shows an SCR structure.

Triggering mechanisms

Having described the architecture that makes latch-up possible, we now discuss the events that can trigger such behavior. SCR latch-up can occur through one of the following mechanisms.

- *Supply voltages exceeding the absolute maximum ratings.* These ratings in the data sheet are an indication of the maximum voltage that can safely be applied to the switch. Anything in excess may result in breakdown of an internal junction and hence damage to the device. In addition, operation of the switch under conditions close to the maximum ratings may degrade long term reliability. It is important to note that these ratings apply at all

times, including when the switch is being powered on and off. The triggering mode could result from transients on supply rails.

- *Input/output pin voltage exceeding either supply rail by more than a diode drop.*
  This could occur as a result of a fault on a channel or input—if a part of the system is powered on prior to the supplies being present at the switch (or similar CMOS components in the system). The powered part of the circuit would be sending signals to other devices in the design which may not be able to handle the voltage levels presented. The resulting voltage levels could exceed the maximum rating of the device and possibly result in latch-up. Again, this could occur as a result of spikes or glitches on input or output channels.
- *Poorly managed multiple power supplies.*
  Switches that have multiple power supplies tend to be more susceptible to latch-up resulting from improper power supply sequencing. Such switches usually have two analog supplies, V_{DD} and V_{SS}, and a digital supply, V_L. In some cases, when the digital supply is applied prior to the other supplies, it may be possible for maximum ratings to be exceeded and the device to enter a latch-up state. In general, for those devices that require an external digital supply, V_L, we recommend that when power is being applied to and removed from the device, care should be taken to ensure the maximum ratings are not exceeded.

When any of the triggering mechanisms described above occur, the parasitic SCR structure of Figure 1a may begin to conduct, producing a low impedance state between power supply rails. If there is no current limit mechanism on the supplies, excessive current will flow through this SCR structure and through the switch. This could destroy the switch and other components if allowed to persist. With high current levels, a device would not have to remain in a latch-up state for very long; even very brief latch-up can result in permanent damage if current is not limited.

Protection and prevention

But such a fate is not inevitable in CMOS circuitry. The simplest way of preventing latch-up occurring is to adhere to the absolute maximum ratings. If this is not always possible, there are other methods of designing a latch-up-proof system.

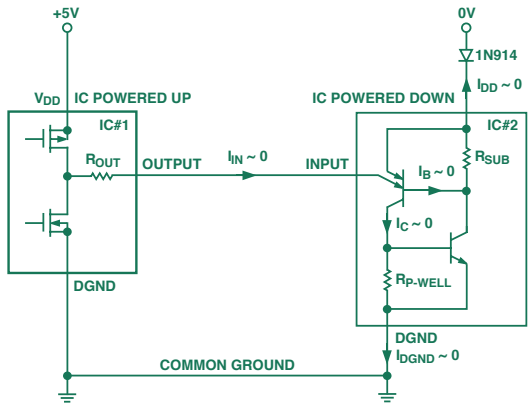


Figure 4. Addition of a diode in series with V_{DD} prevents SCR triggering.

Here are some options for protecting against and preventing latch-up: Where it is possible for digital or analog inputs to exceed the V_{DD} supply—either while power is being applied or during operation—the addition of a diode connected in series with V_{DD} prevents base current from flowing, thus avoiding SCR triggering and hence latch-up. While Figure 4 shows the case where the *digital* input is exceeding the supply of the switch, IC #2, the diode also protects against overvoltages applied to the switch's *analog* signal path.

Now, consider a switch with multiple supplies, where for example, the digital supply, V_L, may be applied to the device prior to other supplies, exceeding the maximum ratings and exposing the circuit to the potential for latch-up. Internal ESD (electrostatic-discharge-limiting) diodes may get turned on, so the simple addition of a Schottky diode, connected between V_L and V_{DD} (Figure 5) will adequately prevent SCR conduction and subsequent latch-up. This works very well; it ensures that when V_L and V_{DD} are applied to the switch, V_{DD} is always within a diode drop (0.3 V for Schottky) of V_L, so the maximum ratings are not exceeded.

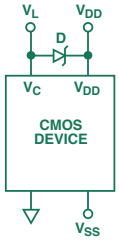


Figure 5. Addition of a Schottky diode from V_L to V_{DD} ensures maximum ratings are not exceeded.

Where the addition of an extra component is not a viable option, due to cost or limited board space, switches are available that have been manufactured with a process that ensures that they are latch-up proof. The process uses an insulating oxide layer (trench) between the NMOS and PMOS devices of each switch. This oxide layer is both horizontal and vertical and produces complete isolation between MOS devices, as shown in Figure 6.

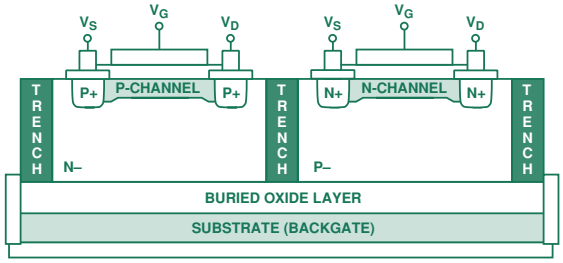


Figure 6. Cross section of switch manufactured with trench processing.

This eliminates the parasitic bipolar devices between transistors, resulting in a latch-up-proof switch. "Latch-up-proof" means that no matter what way the power is sequenced to the device, latch-up cannot occur.

Table 1 lists Analog Devices switches, multiplexers, and channel protectors that have such processing. Although all the devices listed are latch-up-proof, not all are designed to handle *overvoltages* outside the supply rails, as the table indicates. In addition to these latch-up proof switches, there are other devices that can tolerate under- and overvoltages, with power applied, of +40 V/–25 V in excess of supplies and +55 V/–40 V with power not applied to the device. These devices are specifically designed to ensure that they can handle faults in the event of power-on or -off conditions. They also employ the insulating oxide layer to protect against latch-up. They are available for use either as multiplexers or as *channel protectors*.

The multiplexers use a structure having n-channel, p-channel, and n-channel MOSFETs in series (Figure 7) to provide both device- and signal-source protection in the event of an overvoltage or power loss. The multiplexer can withstand continuous overvoltage inputs from –40 V to +55 V. When one of the analog inputs or outputs exceeds the power supplies, one of its MOSFETs will switch off, the multiplexer input (or output) appears as an open circuit, and the output is clamped to within the supply rail, thereby preventing the overvoltage from damaging any circuitry following the multiplexer. This protects the multiplexer, the circuitry it drives, and the sensors or signal sources which drive the multiplexer. Figure 7 shows what happens on one channel of the ADG438F in the event of a positive overvoltage. Because the fault protection works regardless of the presence of supplies, the muxes are also ideal for use in applications where power sequencing cannot always be guaranteed to protect analog inputs (e.g., hot-insertion rack systems).

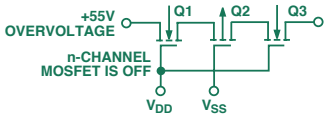


Figure 7. +55 V overvoltage applied to the input channel of ADG438F/ADG439F multiplexer in ON state.

Similarly, *channel protectors* are used to protect sensitive components from voltage transients in the signal path whether or not the power supplies are present. They are built like the fault-protected muxes described above. When powered, the channel is always in the ON condition, but in the event of a fault, it clamps the output to within the supply rails, as shown in Figure 8.

Channel protectors are generally placed in series with the signal path ahead of standard CMOS processed devices to ensure that potential faults can be tolerated without damage to components in the system. A common way of protecting a channel from potential faults, in either a powered or nonpowered condition, is to connect diodes and current limiting resistors between the channel and the supplies. While it is an effective solution, it requires three extra components per channel, plus the board space to accommodate them. A channel protector would be an equally effective but simpler solution in a single small package.

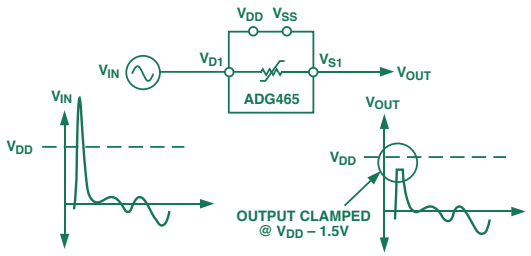


Figure 8. Channel protector clamps overvoltages to within power supply rail voltage and protects sensitive components.

For example, a channel protector could be used in conjunction with an ADC, switch, multiplexer, or other device to ensure that all the channels are protected, both in the event of an over- or undervoltage, and a fault when the system is unpowered. These devices can withstand continuous voltage inputs from –40 V to +40 V. Because the channel protection works regardless of the presence of supplies, channel protectors are also ideal for use in applications where power sequencing cannot always be guaranteed to protect analog inputs. (A familiar example is hot-insertion rack systems.)

CONCLUSION

Inasmuch as no application can tolerate latch-up, it is necessary to be aware of its possibility, understand it, protect against it, and take measures to prevent it from happening. Given some thought and the use of available methods and components, it is indeed possible to assemble a latch-up-proof system. While discrete solutions—such as diodes—could be used, devices like latch-up-proof switches, fault protected multiplexers, and channel protectors may provide a simpler, more compact, and more generally suitable solution, resulting in a robust system likely to give fewer problems in the field.

Table 1. Latch-up proof Analog Devices switches, multiplexers, and channel protectors.

Part Number	Function	Latch-up Proof	Over/Under-Voltage Capability	Package[1]
ADG431A	Quad SPST (NC)	Yes	No	R-16
ADG432A	Quad SPST (NO)	Yes	No	R-16
ADG433A	Quad SPST (2 NC, 2 NO)	Yes	No	R-16
ADG441	Quad SPST (NC)	Yes	No	R-16, N-16
ADG442	Quad SPST (NO)	Yes	No	R-16, N-16
ADG444	Quad SPST (2 NC, 2 NO)	Yes	No	R-16, N-16
ADG511A	Quad SPST (±5 V, +5 V, +3 V)	Yes	No	R-16
ADG512A	Quad SPST (±5 V, +5 V, +3 V)	Yes	No	R-16
ADG513A	Quad SPST (±5 V, +5 V, +3 V)	Yes	No	R-16
ADG438F	Octal 8-1 Channel Multiplexer	Yes	Yes	R-16, N-16
ADG508F	Octal 8-1 Channel Multiplexer	Yes	Yes	RN-16, RW-16, N-16
ADG439F	Differential 4-1 Channel Mux	Yes	Yes	R-16, N-16
ADG509F	Differential 4-1 Channel Mux	Yes	Yes	RN-16, RW-16, N-16
ADG465	Single Channel Protector	Yes	Yes	RT-6, RM-8
ADG466	Triple Channel Protector	Yes	Yes	RM-8, R-8, N-8
ADG467	Octal Channel Protector	Yes	Yes	RS-20, R-18

[1]N = DIP, R/RN = 0.15" SOIC, RW = 0.3" SOIC, RS = SSOP, RM = μSOIC, RT = SOT-23 NC = Normally Closed; NO = Normally Open

Avoiding Op Amp Instability Problems In Single-Supply Applications

by Charles Kitchin

SINGLE OR DUAL SUPPLY?

Although it is advantageous to implement op amp circuits with balanced dual supplies, there are many practical applications where, for energy conservation or other reasons, single-supply operation is necessary or desirable. For example, battery power, in automotive and marine equipment, provides only a single polarity. Even line-powered equipment, such as computers, may have only a single-polarity built-in supply, furnishing 5 V dc or 12 V dc for the system. In processing analog signals, a common feature of single-supply operation is the need for additional components in each stage for appropriate signal biasing. If this is not carefully thought through and executed, instability and other problems may be encountered.

Common Problems with Resistor Biasing

Single-supply op amp applications have inherent problems that are not usually encountered in dual-supply circuits. The fundamental issue is, if the signal is to swing both positive and negative with respect to "common," this zero-signal reference voltage must be at a fixed level between the supply rails. The principal advantage of dual supplies is that their common connection provides a stable, low-impedance zero reference. The two supply voltages are usually equal and opposite (and often tracking), but that is not an absolute necessity. With a single supply, such a node must be created artificially, by introducing additional circuitry to provide some form of biasing, to maintain signal common at an appropriate midsupply voltage.

Since it is usually desirable for large output values to limit symmetrically, the bias is usually established at the midpoint of the rated amplifier output range or, for convenience, at one-half the supply voltage. The most effective way to achieve this is with a regulator, as in Figure 6; however, a popular method involves tapping the supply voltage with a pair of resistors. Though apparently simple, there are problems with this.

Illustrating the problem, the circuit of Figure 1, which has several design weaknesses, is an ac-coupled noninverting amplifier. The signal is capacitively coupled in and out. The average level of the ac-coupled input is biased to $V_S/2$ by the R_A–R_B divider pair, and the in-band gain is G = 1 + R2/R1. The dc "noise gain" is reduced to unity by capacitively coupling the feedback with a zero established by R1 and C1, so that the dc level of the output is equal to the bias voltage. This avoids distortion due to excessive amplification of the amplifier's input offset voltage. The amplifier's closed-loop gain rolls off from (1 + R2/R1) at high frequency to unity at dc, with break frequencies at f = 1/[2π R1 C1] and f = 1/[2π (R1 + R2) C1], introducing phase shifts that add to those associated with the input- and output-coupling circuits.

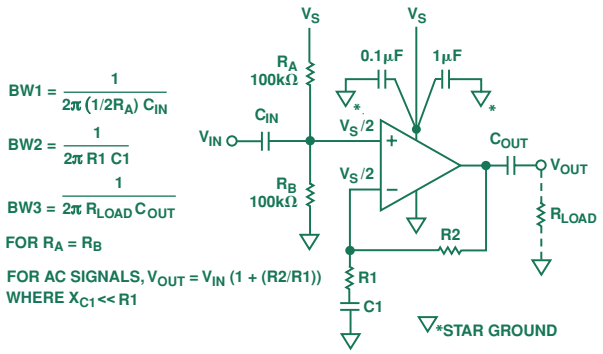


Figure 1. A potentially unstable single-supply op amp circuit.

This simple circuit has additional potentially serious limitations. First, the op amp's inherent ability to reject supply-voltage variations is of no avail, as any change in supply voltage will directly change the $V_S/2$ biasing voltage set by the resistor divider. While this does not present a problem at dc, any common-mode noise appearing at the power-supply terminals will be amplified, along with the input signal (except at the lowest frequencies). With a gain of 100, 20 *millivolts* of 60 Hz ripple and hum will be amplified up to a 1 V level at the output.

Even worse, instability can occur in circuits where the op amp must supply large output currents into a load. Unless the power supply is well regulated (and well bypassed), significant signal voltages will appear on the supply line. With the op amp's noninverting input referenced directly off the supply line, these signals will be fed directly back into the op amp, often in a phase relationship that will produce "motor boating" or other forms of oscillation.

While the use of extremely careful layout, multicapacitor power supply bypassing, star grounds, and a printed circuit board "power plane," all help to reduce noise and maintain circuit stability, it is better to employ circuit design changes that will improve power supply rejection. A few are suggested here.

Decoupling the Biasing Network from the Supply

One step toward a solution is to bypass the bias-voltage divider, and provide a separate input return resistor, modifying the circuit as shown in Figure 2. The tap point on the voltage divider is now bypassed for ac signals by capacitor C2, to restore the ac power-supply rejection. Resistor R_{IN}, which replaces $R_A/2$ as the circuit's input impedance for ac signals, also provides a dc return path for the + input.

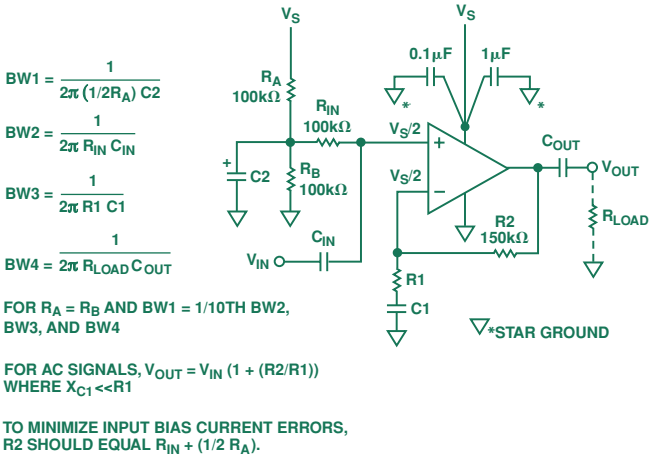


Figure 2. A decoupled single-supply op amp biasing circuit.

The values of R_A and R_B should, of course, be as low as feasible; the 100 kΩ values chosen here are intended to conserve supply current, as one might wish to do in a battery-powered application. The bypass capacitor value should also be carefully chosen. With a 100 kΩ/100 kΩ voltage divider for R_A and R_B and a 0.1 μF or similar capacitance value for C2, the –3 dB bandwidth of this network's impedance, set by the parallel combination of R_A, R_B, and C2, is equal to $1/[2\pi (R_A/2)\ C2] = 32$ Hz. Although this is an improvement on Figure 1, the common-mode rejection drops below 32 Hz, allowing substantial feedback through the power supply at low signal frequencies. This requires a larger capacitor to avoid "motorboating" and other manifestations of instability.

A practical approach is to increase the value of capacitor C2 so it is large enough to effectively bypass the voltage divider at all frequencies within the circuit's passband. A good rule of thumb is to set this pole at one-tenth the –3 dB input bandwidth, set by $R_{IN}C_{IN}$ and R_1C_1.

The amplifier's gain at dc is still unity. Even so, the op amp's input bias currents need to be considered. R_{IN}, in series with the R_A/R_B voltage divider, adds considerable resistance in series with the op amp's positive input terminal. Maintaining the op amp's output close to midsupply, using common voltage-feedback op amps that have symmetrical balanced inputs, can be achieved by balancing this resistance by the choice of R2.

Depending on the supply voltage, typical values that provide a reasonable compromise between increased supply current and increased sensitivity to amplifier bias current, range from 100 kΩ for 15 V or 12 V single supplies, down to 42 kΩ for a 5 V supply and 27 kΩ for 3.3 V.

Amplifiers designed for high-frequency applications (especially current-feedback types) need to use low input and feedback resistances in order to maintain bandwidth in the presence of stray capacitance. An op amp such as the AD811, which was designed for video speed applications, will typically have optimum performance using a 1 kΩ resistor for R2. Therefore, these types of applications need to use much smaller resistor values in the R_A/R_B voltage divider (and higher bypass capacitances) to minimize input bias current and avoid low-frequency instability.

Because of their low bias current, the need for balancing input resistors is not as great in applications with modern FET-input op amps, unless the circuit is required to operate over a very wide temperature range. In that case, balancing the resistance in the op amp's input terminals is still a wise precaution.

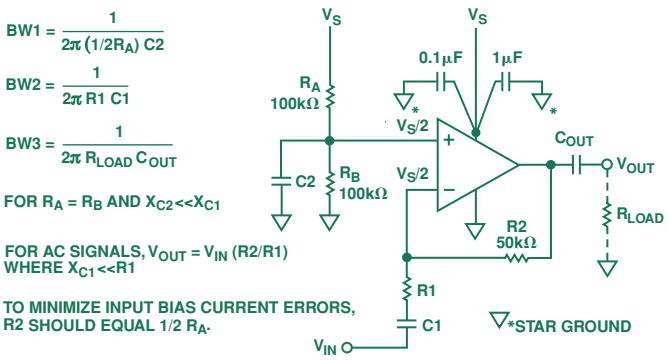


Figure 3. A decoupled single-supply inverting amplifier circuit.

Figure 3 shows how biasing and bypassing might be applied in the case of an inverting amplifier.

The resistor divider biasing technique is low in cost and keeps the op amp's dc output voltage at $V_S/2$, but the op amp's common-mode rejection still depends on the RC time constant formed by $R_A \| R_B$ and capacitor C2. Using a C2 value that provides at least 10 times the RC time constant of the input RC coupling network (R1/C1 and R_{IN}/C_{IN}) will help ensure a reasonable common-mode rejection ratio. With 100 kΩ resistors for R_A and R_B, practical values of C2 can be kept fairly small as long as the circuit bandwidth is not too low.

Zener Diode Biasing

A more effective way to provide the necessary $V_S/2$ biasing for single-supply operation is to use a Zener-diode regulator, such as that shown in Figure 4. Here current is supplied to the Zener diode through resistor R. Capacitor C_N helps reduce Zener-generated noise from appearing at the op amp input.

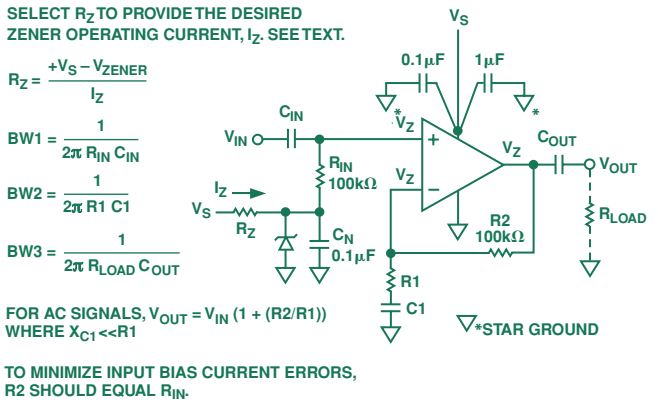


Figure 4. A noninverting single-supply amplifier using Zener diode biasing.

A Zener with an operating voltage close to $V_S/2$ should be chosen. Resistor R_Z needs to be selected to provide a high enough current to operate the Zener at its stable rated voltage and to keep the Zener output noise low. Yet, it is also important to minimize power consumption (and heating) and to avoid damage to the Zener. As the op amp input draws little current from the reference, it is a good idea to choose a low-power diode. A 250 mW-rated device is best, but the more common 500 mW types are also acceptable. The ideal Zener current varies with each manufacturer, but practical I_Z levels between 500 μA (250 mW device) and 5 mA (500 mW device) are usually a good compromise for this application.

Within the operating limits of the Zener, the circuit of Figure 4 basically provides low reference-level impedance, which restores the op amp's power supply rejection. The benefits are substantial, but there is a price: more power is consumed, and the op amp's dc output is fixed by the Zener voltage, rather than at $V_S/2$. If the power supply voltage drops substantially, asymmetrical clipping can occur on large signals. Also, input bias currents still need to be considered. Resistors R_{IN} and R2 should be close to the same value to prevent input bias currents from creating substantial offset voltage error.

Figure 5 is an inverting amplifier circuit using the same Zener biasing method.

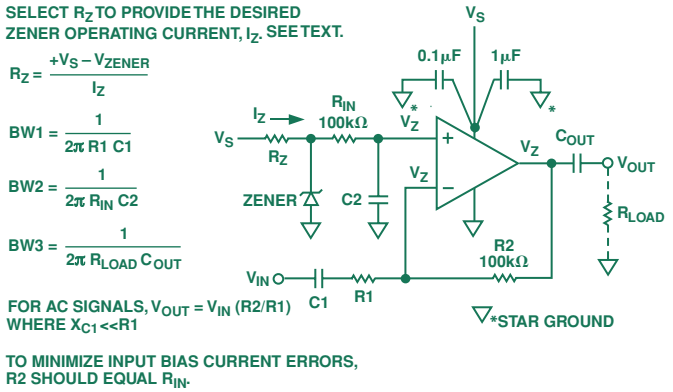


Figure 5. An inverting single-supply amplifier using Zener-diode biasing.

Table 1 shows some common Zener diode types that can be chosen to provide halfway supply bias for various supply voltage levels. For convenience, practical R_Z values are provided to furnish 5 mA and 0.5 mA device currents in Circuits 4 and 5. For lower circuit noise, the optimum Zener current can be selected by referring to the manufacturer's data sheet.

Table 1. Suggested Zener-diode part numbers (Motorola types) and R_Z values for use in Figures 4 and 5.

Supply Voltage	Reference Voltage	Diode Type	Zener Current	R_Z Value Ω
15 V	7.5 V	1N4100	0.5 mA	15k
15 V	7.5 V	1N4693	5 mA	1.5k
12 V	6.2 V	1N4627	0.5 mA	11.5k
12 V	6.2 V	1N4691	5 mA	1.15k
9 V	4.3 V	1N4623	0.5 mA	9.31k
9 V	4.3 V	1N4687	5 mA	931
5 V	2.4 V	1N4617	0.5 mA	5.23k
5 V	2.7 V	1N4682	5 mA	464

Op Amp Biasing Using a Linear Voltage Regulator

For op amp circuits operating from the 3.3 V standard, a 1.65 V biasing voltage is needed. Zener diodes are commonly available only down to 2.4 V, although the 1.225 V AD589 and AD1580 bandgap shunt regulators can be used like Zener diodes to provide a fixed—though not centered—voltage at low impedance. The easiest way to provide arbitrary values of biasing voltage at low impedance (for example, $V_S/2$) is to use a linear voltage regulator, such as the ADM663A or ADM666A, as shown in Figure 6. Its output can be adjusted from 1.3 V to 16 V, and it will provide low-impedance biasing for single-supply voltages from 2 V to 16.5 V.

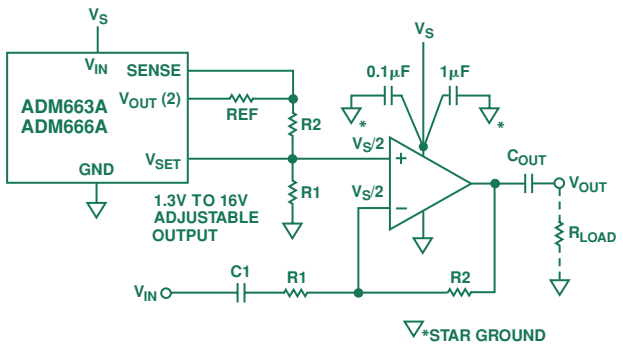


Figure 6. An op amp single-supply biasing circuit using a linear voltage regulator.

DC-Coupled Single-Supply Circuits

So far, only ac-coupled op amp circuits have been discussed. Although with the use of suitable large input and output coupling capacitors, an ac-coupled circuit can operate at frequencies well below 1 Hz, some applications require true dc input and output coupling. Circuits that provide constant dc voltage at low impedance, such as the Zener diodes and regulators discussed above, can be used to provide "ground-level" voltages.

Alternatively, the $V_S/2$ biasing resistors of Figures 1 through 3 can be buffered by an op amp to provide a low-impedance "phantom ground" circuit as shown in Figure 7. If a low-voltage battery, say 3.3 V, is the supply source, the op amp should be a "rail to rail" device that is able to operate effectively over the full supply-voltage range. The op amp also needs to be able to supply a positive or negative output current large enough to satisfy the main circuit's load requirement. Capacitor C2 bypasses the voltage divider to attenuate resistor noise. This circuit does not need to provide power supply rejection, because it will always drive the common terminal ("ground") at one-half the supply voltage

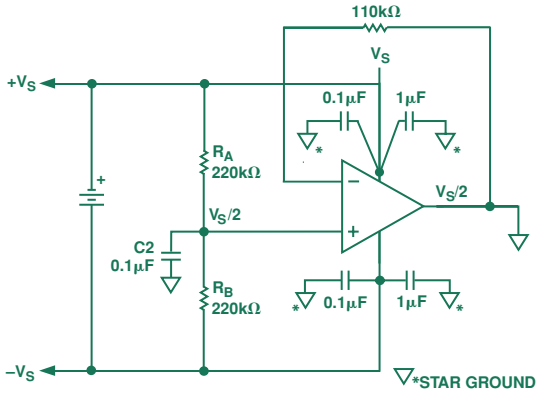


Figure 7. Using an op amp to provide a "phantom ground" for battery-powered direct-coupled applications.

Circuit Turn-On Time Issues

One final issue that needs to be considered is circuit turn-on time. The approximate turn-on time will depend on the RC time constant of the lowest-bandwidth filter being used.

The circuits with passive biasing shown here all should require the $R_A \| R_B$–C2 voltage-divider network to have a 10× longer time constant than that of the input or output circuit. This is to simplify the circuit design (since up to three different RC poles set the input bandwidth). This long-time constant also helps keep the biasing network from "turning on" before the op amp's input and output networks, thus allowing the op amp's output to gradually climb from zero volts to $V_S/2$ without being driven to the positive supply rail. The required 3 dB corner frequency is 1/10th that of R1/C1 and R_{LOAD}/C_{OUT}. For example: in Figure 2, for a circuit BW of 10 Hz and a gain of 10, a C2 value of 3 µF provides a 3 dB BW of 1 Hz.

With $R_A \| R_B$ = 50,000 Ω, a 3 µF capacitor provides an RC time constant of 0.15 seconds. So the op amp's output will take about 0.2 seconds to 0.3 seconds to settle reasonably close to $V_S/2$. Meanwhile, the input and output RC networks, will charge up ten times faster.

In applications where the circuit's turn-on time may become excessively long, a Zener or active biasing method may be a better choice.

Improved Data Acquisition Using a New Monolithic Dual-ADC Front End

by Albert O'Grady

Designers of measurement systems such as temperature-compensated weigh scales and RTD instrumentation, can perform their task more effectively with a new single-chip A/D converter/analog front end. The AD7719 features the high resolution inherent in sigma-delta converters; and for augmented throughput, contains *two* ADCs (24- and 16-bit resolution), permitting simultaneous parallel conversions of two input variables without the latency inherent in analog multiplexing schemes. It employs a signal-chopping approach that provides signal conditioning with stable gain and minimal offset, coupled with built-in calibration, to eliminate the need for calibration in the field. It conveniently includes a matched pair of current sources, to simplify transducer excitation and improve measurement accuracy when using resistive transducers. Other useful features include circuitry that can be switched off and low-side switches that can be used to conserve power when converters or transducer power are not being used, plus a digital I/O port for monitoring and control of external devices.

The AD7719, a complete analog front end for low-frequency measurement applications, is the newest addition to the Analog Devices family of high-resolution, low-bandwidth, sigma-delta converters. It builds on the experience gained from employing the previous generations of sigma-delta converters in applications ranging from weigh scales to portable instrumentation, pressure, temperature and transducer measurement, smart transmitters, liquid/gas chromatography, and industrial process control. The enhanced features mentioned above (see Figure 1) address many issues that commonly recur in the design of such high-performance data acquisition systems.

The AD7719 contains two independent, high-resolution, sigma-delta ADCs. Each analog-to-digital conversion is accomplished by a second-order sigma-delta modulator with a programmable sinc^3 filter. In addition, it makes available switchable matched 200 μA excitation current sources, low-side power switches, digital I/O port, and a temperature sensor. The 24-bit main channel, with a programmable-gain amplifier (PGA) that has gains from 1 to 128, accepts fully differential, unipolar, and bipolar input signals with ranges up to 1.024 × REFIN1 volts. The reference input is differential and can provide ratiometric conversion. The main analog input channel can be internally buffered to provide a very high input impedance; this allows input signals to be applied directly from a transducer without the need for external signal conditioning. The 16-bit auxiliary channel is unbuffered and offers an input signal range of REFIN2 or one-half REFIN2.

The device operates from a 32,768 Hz (32K) crystal, with an on-board PLL generating all of the required internal operating frequencies. The output data rate from the AD7719 is software

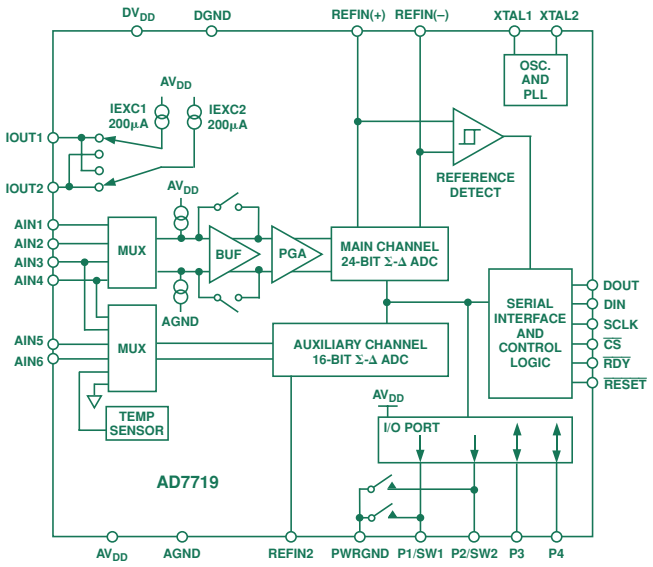


Figure 1. AD7719 functional block diagram.

programmable. This allows the digital filter notches to be placed at user-defined frequencies. For example, with a programmed update rate of 19.8 Hz, rejection notches at 50 Hz and 60 Hz can be achieved simultaneously.

The peak-to-peak resolution depends on the programmed gain and the output data rate. The device operates from a single 3 V or 5 V supply. When operating from a 3 V supply, the power dissipation is 4.5 mW with both ADCs continuously in use. Dissipation can be reduced by disabling one or both ADCs when appropriate. The AD7719 is available housed in space-saving 28-lead SOIC and TSSOP packages.

Signal-Processing Chain

The ADCs employ sigma-delta conversion to realize up to 24 bits of no-missing-codes performance. The sigma-delta modulator converts the sampled input signal into a digital pulse train whose duty cycle contains the digital information. A sinc^3 programmable low-pass filter then decimates the modulator output data stream to give a valid data conversion result at programmable output rates from 5.35 Hz (186.77 ms period) to 105.03 Hz (9.52 ms). A chopping scheme is employed to minimize ADC channel offset, gain, and drift errors. A block diagram of the main ADC input channel is shown in Figure 2. The signal chain for the auxiliary ADC is similar to that of Figure 2 but omits the buffer and PGA blocks.

The sampling frequency of the modulator loop is many times higher than the bandwidth of the input signal (oversampled). The integrator in the modulator shapes the quantization noise (which results from the analog-to-digital conversion) so that the noise is concentrated near one-half of the modulator frequency. The output of the sigma-delta modulator feeds directly into the digital filter, which band-limits the response to a frequency significantly lower than one-half of the modulator frequency. In this manner, the 1-bit output of the comparator is translated into a band-limited, low noise output from the ADC. The filter's cutoff frequency and decimated output data rate are programmable via the sinc-filter (SF) control word loaded to the filter register.

The alternating digital output values that result from the input chopping are summed in a final summing stage to average out dc offsets and low-frequency noise. Each output word from the filter is summed and averaged with the previous filter output to produce

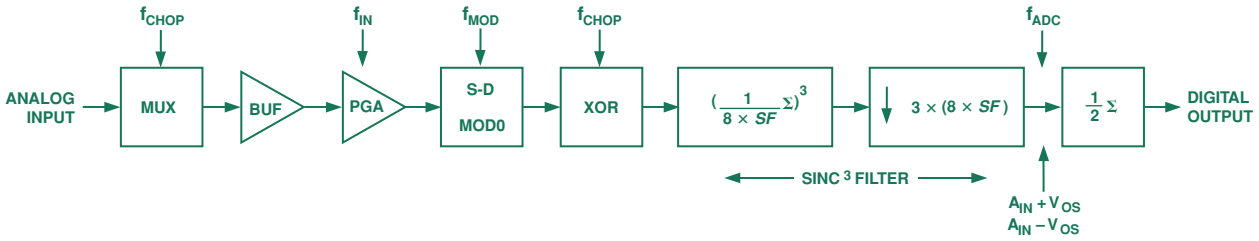


Figure 2. Main ADC signal chain.

a new valid output result to be written to the ADC data register. The resulting very low dc offset and offset- and gain-drift specifications are quite beneficial in applications where drift, noise rejection, and optimum EMI rejection are important.

Besides reduction of quantization noise, the digital filter also provides normal-mode rejection of 100 dB at 50 Hz and 60 Hz (±1 Hz) for respective filter control-word settings of 82 and 68. For applications requiring substantial rejection at both 50 Hz and 60 Hz, the filter's response at the default programmed setting of 69 (data-update rate of 19.8 Hz) has notches close to both frequencies, with >100 dB rejection at 60 Hz and >60 dB at 50 Hz, as shown in Figure 3.

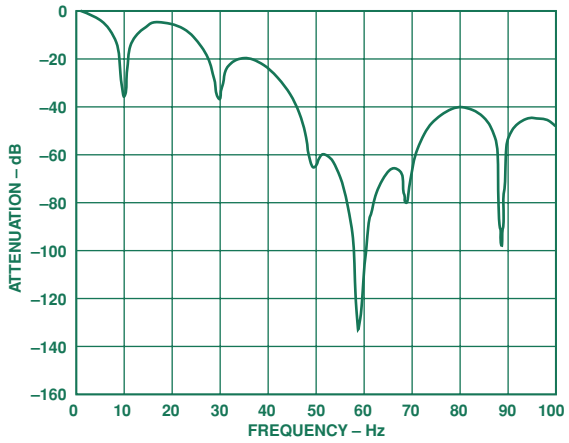


Figure 3. Filter profile showing simultaneous 50 Hz and 60 Hz rejection nulls with sinc^3 filter control word of 69.

Typical Applications

The AD7719 provides a complete analog front end for implementing low frequency measurements using temperature-, pressure-, and other transducers. In weigh-scale applications, for example, a secondary variable—such as temperature—may need to be monitored in addition to the primary variable from the bridge transducer in order to compensate for variations of bridge properties with temperature.

Traditionally, sigma-delta ADCs have used a single converter with integrated multiplexer on the front end to measure multiple input variables. This means that the end user has to switch channels on the front end to measure the secondary variable; as a consequence, measurement speed suffers from the settling time and latency associated with the digital filter when switching input sources. In systems where the sigma-delta ADC uses a second-order modulator and a third-order digital filter, the output settling time to a step input is three times the data rate in order to fully flush the digital filter of all data pertaining to the previous channel. This can greatly reduce the system throughput achievable in these applications.

The AD7719 overcomes this problem by incorporating two independent ADC channels, converting in parallel. The primary variable and the secondary variable are converted simultaneously, and output data from both measurements is available in parallel, thus avoiding the latency associated with multiplexed data acquisition systems. In addition, the on-chip current sources can be used to excite temperature sensors, such as thermistors or RTDs for temperature monitoring.

A second commonly encountered issue in low-power, battery-operated weighing systems is the unnecessary consumption of power in the front-end transducer when in standby mode. The AD7719's on-chip low-side power switches can address this issue by removing the power to the transducer when in low-power mode, thus offering substantial power savings.

Another issue with weigh-scale applications concerns calibration: when and how often should it take place? Because the AD7719 is factory-calibrated, and the signal chain uses a chopping scheme in its implementation, the gain and offset drift are reduced to a minimum, thereby eliminating the need for calibration in the field. This is a key performance advantage when the AD7719 is used in weigh-scale applications (Figure 4).

In the circuit of Figure 4, the main channel monitors the bridge by transducer, and the secondary channel monitors temperature by means of a thermistor. The bridge transducer's differential output terminals (OUT+ and OUT–), are connected to differential input terminals, AIN1 and AIN2. A typical bridge with a sensitivity of 3 mV/V will produce a rated full-scale output of 15 mV when excited with a 5 V excitation source. The excitation voltage for the bridge can be used to directly provide the reference for the ADC via a suitable resistor divider, which will allow the full dynamic range of the input to be utilized. Because this implementation is fully ratiometric, variations in the excitation voltage do not introduce errors in the system. The choice of resistance values, 20 kΩ and 12 kΩ, as shown in the diagram, gives a 1.875 V reference voltage for the AD7719 when the excitation voltage is 5 V. With the main channel programmed for a gain of 128, the full-scale 15 mV input span corresponds to the full output span from the transducer. A key requirement in weigh-scale applications is to reject ac power-mains frequency components (50 Hz and 60 Hz) as much as possible. Simultaneous 50 Hz and 60 Hz rejection can be obtained by programming the AD7719 for an output data rate of 19.8 Hz. 13-bit peak-to-peak resolution will be achieved with the AD7719 configured for a gain of 128 with a 19.8 Hz update rate. The peak- to-peak resolution can be increased by reducing the update rate or by performing extra digital filtering in the controller.

Temperature is measured using a thermistor and the AD7719's secondary channel. Thermistors, high-temperature-coefficient

electrical circuit elements formed with semiconductor material, are available with negative or positive temperature coefficients (NTC or PTC). An NTC thermistor acts like a resistor with a temperature coefficient of typically –3%/°C to –5%/°C. Thermistors offer the benefits of high stability, precision, small size, and compatibility at a competitive price in many applications. They have fast response and are among the highest-sensitivity temperature transducers available. The operating temperature range of the circuit of Figure 4 is determined by the choice of thermistor. Using a 1K7A1 thermistor from Betatherm, with a nominal resistance of 1 kΩ at 25°C, and employing the 200 μA excitation-current source, the operating temperature range is –26°C to +70°C.

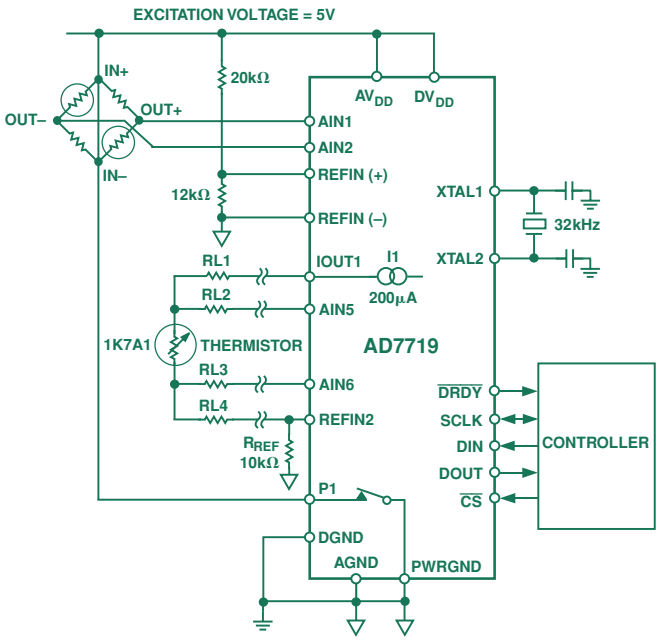


Figure 4. Weigh-scale application takes advantage of many features of the AD7719.

In this application, the same 200 μA current source that excites the thermistor also generates the reference voltage for the AD7719. As a result, variations in the excitation current do not affect performance, and the configuration provides fully ratiometric conversion. The most common wiring arrangement in these applications is a 4-wire force/sense configuration in order to reduce the effects of lead resistances on system performance. Although the lead resistance of the drive wires shifts the common-mode voltage, it does not degrade the performance of the circuit. Lead resistance of the sense wires is immaterial, as no current flows in these wires due to the high input impedance of the AD7719 analog inputs. However, the reference-setting resistor must have a low temperature coefficient to avoid errors in the reference voltage with temperature changes. By configuring the secondary channel on the AD7719 for a 19.8 Hz update rate, 16-bit peak-to-peak performance can be obtained.

Another application that makes good use of the matched current sources on the AD7719 is where a 3-wire RTD is used for precision temperature measurement, in the manner shown in Figure 5*. In 3-wire configurations, the lead resistances would cause errors if a single current source were used, as the 200 μA excitation current, flowing through RL1 and RL3, will develop a voltage drop across RL1, which adds to the RTD voltage and causes an error between AIN1 and AIN2.

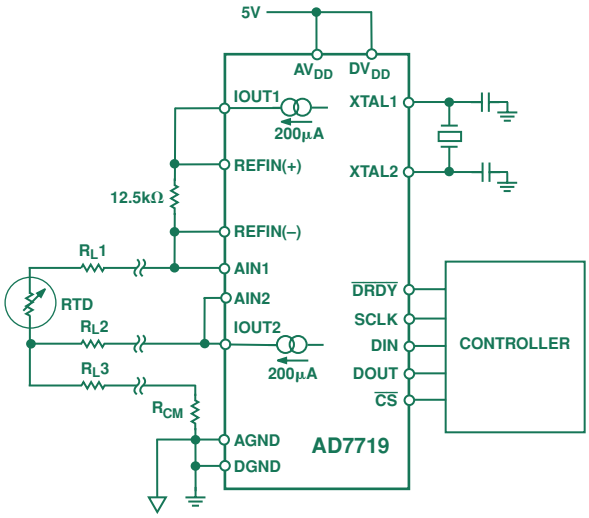


Figure 5. 3-wire RTD measurement using the AD7719.

In the scheme shown in Figure 5 however, the second current source is used to null out this error by furnishing an equal and opposite compensating current, IOUT2, through RL2, which produces an equal voltage drop in the opposite direction. This current adds to IOUT1 and flows harmlessly to ground through R3 and any common-mode resistance, producing a common-mode voltage, which is rejected by the differential inputs.

This analysis assumes that RL1 and RL2 are equal, since the leads would normally be of the same material and of equal length, and that the common-mode voltage developed by the sum of the currents is within the common-mode range of the ADC. The current from IOUT1 is also used to develop a reference voltage for the AD7719, across the 12.5 kΩ resistor, as shown, and applied to the differential reference inputs of the AD7719. This scheme ensures that the analog input voltage span remains ratiometric to the reference voltage. Any errors in the analog input voltage, due to the temperature drift of the RTD's current source, is compensated for by the variation in the reference voltage. The two RTD current sources are typically matched to better than 1%. The voltage compliance on either current source is AV_{DD} – 0.6 V.

*See the article, "Transducer/Sensor Excitation and Measurement Techniques," in *Analog Dialogue* 34-05 (2000).

The PDA Challenge—Met by the AD7873 Resistive-Touch-Screen Controller ADC

by Paul Kearney (paul.kearney@analog.com)

INTRODUCTION

Today, handheld personal digital assistants (PDAs) are popping up just about everywhere. They are finding favor, not only with busy executives, but also with mainstream consumers—and even in the classroom! Their ease of use and portability have been keys to their rapidly growing success. The most popular examples* include the Palm Pilot series, the Handspring Visor, and the Sony Clie. It is estimated that this year alone some 15 million PDA units will be sold worldwide, with annual sales expected to grow to 34 million units by 2004. A common feature of all of these PDA units is their method of entering data via a stylus and a resistive touch screen. With it the user can easily enter dates in his/her diary, send email, keep shorthand minutes of meetings, etc.

The touch screen itself is typically a 4-wire (±X and ±Y) resistive element. (There are also 5-wire resistive screens and capacitor screens on the market, but they are typically more expensive.) The touch screen interfaces with the host microprocessor via an A/D converter (ADC), such as the AD7873 and AD7843, which have special features tailored to the application. This article will discuss the common application issues faced by designers when interfacing an ADC to a resistive touch screen, and how they are solved with the AD7873.

Theory of Operation and Details of the Application

The touch screen usually consists of two layers of transparent resistive material—usually indium tin oxide (ITO) or some other form of resistive polyester material, with silver ink for electrodes. The total resistance of each layer varies from vendor to vendor, but typical screens are in the 100 to 900 Ω range. The two layers are stacked on an insulating layer of glass, separated by tiny spacer dots. They are interfaced electrically to a controller A/D converter. Figures 1 and 2 show, in simplified form, how a controller ADC might interface with a 4-wire resistive screen.

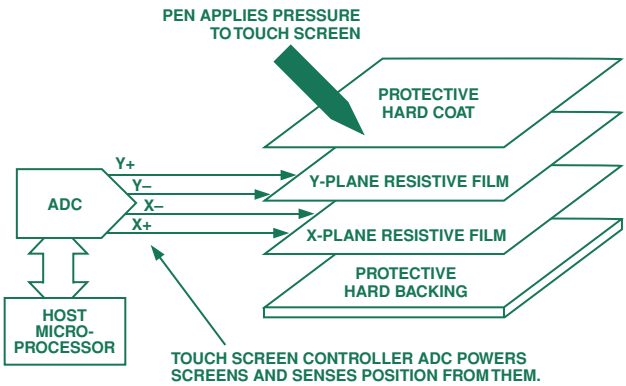


Figure 1. The touch screen uses a multilayer sandwich of resistive films and protective coatings.

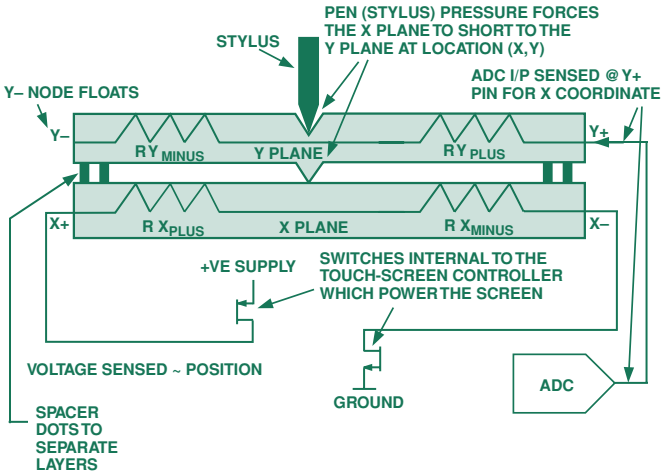


Figure 2. Stylus interfacing with a 4-wire resistive touch screen. The X-coordinate measurement is depicted.

During measurement of a given coordinate, one of the resistive planes is powered along its axis through switches on the controller ADC chip, and the other plane is used to sense the location of the coordinate on the powered plane. For X coordinate measurement, the X plane is powered. The Y plane is used to sense where the pen is located on the powered plane as follows: At the location where the pen depresses the touch screen, the planes are shorted. The voltage 'picked up' on the sensed plane is proportional to the location of the touch on the powered plane. This voltage is then converted using the controller's ADC.

For a Y-coordinate measurement, power is applied to the Y plane, the X plane is used to sense the position, and the voltage is digitized. The digital code corresponding to the X and Y coordinates is then operated on by the host microprocessor, and the command, information, or instruction intended by the stylus location is registered.

Basic AD7873 touch-screen interface

Figure 3 illustrates a simplified block diagram of the AD7873. In addition to serving as a coordinate-measurement transducer, it provides a number of peripheral functions important for the PDA's function, including measurements of temperature and battery condition, touch detection, and pressure measurement—plus an on-chip 2.5-V voltage reference.

Figure 4 illustrates a typical application diagram for the AD7873 interfacing with a 4-wire resistive touch screen. First let's discuss the basic functionality—stylus coordinate measurement. The set of touch-screen switches at the left in Figure 3 drive the resistive screen via the so-called tablet pins, X+, Y+, X–, and Y–. The tablet pins also serve as sensors of position voltage from the screen when not active, in applying the method of interfacing described briefly above.

A number of things were considered in the implementation of the design. First, the switches must be able to source and sink current to/from the low impedance screen. Say, for example, the supply is 5 V and the screen impedance is 200 Ω. The p-MOS switches to the positive supply (on the X+ and Y+ pins) must be able to source 25 mA when the screen is powered for coordinate measurement. Similarly the n-MOS switches to ground (these are

*All brand or product names mentioned are trademarks or registered trademarks of their respective holders.

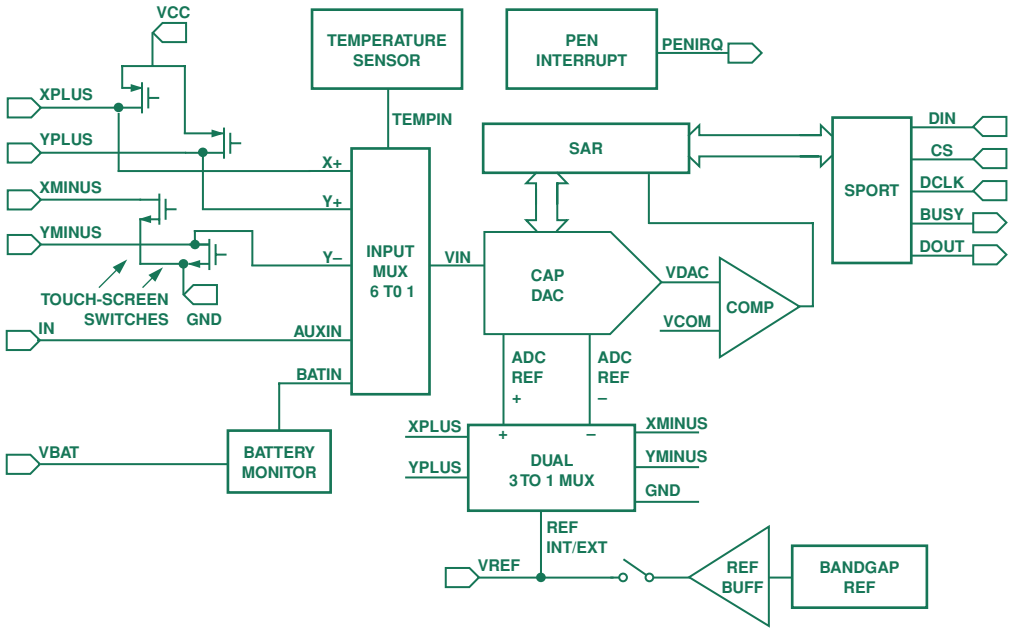


Figure 3. AD7873 block diagram.

on the Y– and X– pins) must be capable of sinking this 25 mA. The switches' drain and source connections must be carefully designed at the layout stage to avoid electromigration problems (due to the large current density in these nodes). The conservatively designed AD7873 switches are capable of supplying a 70 Ω screen, using a 5-V supply.

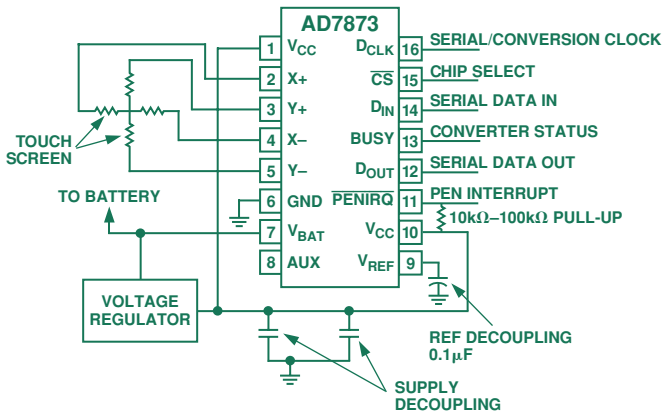


Figure 4. Typical application circuit using the AD7873.

The on-resistance of the switches, about 6 Ω, also posed a challenge. For example, when used with screen impedances as low as 100 Ω and a minimum supply voltage of 2.7 V, the voltage drop across the switches is appreciable and greatly reduces the dynamic range of signals applied to the converter input. To avoid a corresponding reduction of conversion accuracy and effective resolution, and because the ON resistance of the switches may not track the resistance of the screen over temperature and supply, the converter must be able to operate in a ratiometric mode. Figures 5 and 6 illustrate respectively how the AD7873 can be configured in both ratiometric and single-ended modes of operation for coordinate measurements.

Using the ratiometric mode, the actual ADC reference is taken from the drain nodes of the switches powering the screen. For example, if a Y coordinate is being measured, the Y plane is powered and the REF+ and REF– of the ADC are taken from the Y+ and Y– pins, respectively. The A/D conversion will be ratiometric; i.e., the result of the conversion will be equal to the ratio of touch-screen resistance at the measuring point to the total touch-screen resistance, irrespective of the voltage drop in the AD7873's switches. This is the best way of ensuring the accuracy of coordinate measurements.

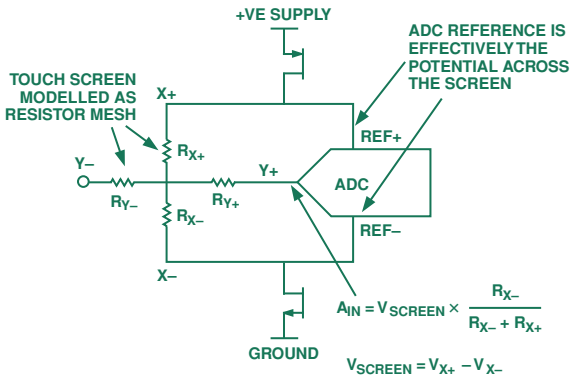


Figure 5. Ratiometric operation of the ADC is independent of switch on-resistance voltage drop.

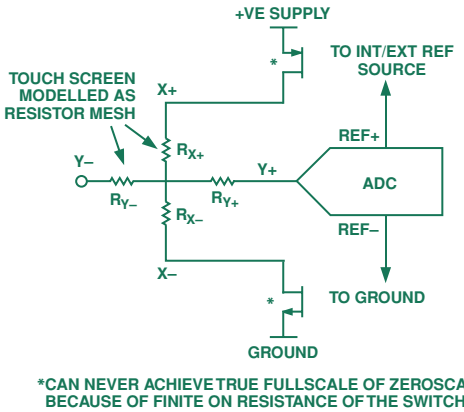


Figure 6. Single-ended operation of the ADC.

But there is a trade-off: In ratiometric mode, the screen must remain powered during the actual conversion process, in addition to the acquisition phase, since the screen voltage is used as the ADC reference. Alternatively, during a single-ended conversion, the screen need only be powered during the acquisition of the input signal. For the AD7873 this requires 3 clock cycles, at worst 20% of the overall conversion time (assuming a 15 clocks-per-conversion mode—which is the device's fastest valid throughput rate). However, used ratiometrically, at this fastest rate, the screen will be permanently powered while the device is on. If the screen resistance is 100 Ω and a 5-V supply is used, the screen will draw 50 mA—an obvious consideration in battery-operated handheld devices.

With this dilemma in mind, the AD7873 has been designed to offer the designer a choice of four power-down options to better manage the overall power consumption of the touch-screen controller, using the device's 2-bit selective power-down option control.

Power-down bit setting

00. Full power down between conversions. Both the ADC and the internal reference are powered down between conversions. If the ratiometric mode is selected, the touch-screen drivers will be turned off between conversions (for cycles longer than 15 clocks).
01. The ADC remains on between conversions, but the reference powers down. (An external reference may be used.) The touch screen power is switched off between conversions, as in mode 00.
10. The ADC powers down between conversions, while the internal reference remains on. This is useful, as the reference requires about 7 μs to power up. So when doing single-ended measurements (such as battery and temperature readings) with the internal reference, there is no need to allow a delay for the reference to power up. The touch screen power is switched off between conversions, as in mode 00.
11. The ADC and internal reference remain powered on between conversions. In this mode, the switch drivers to the touch screen will remain on until the selected input channel or power-down mode is changed, or until CS (convert start) is brought high.

Whatever option is used, the designer must realize that the bulk of the power will be dissipated in the external touch screen when it is powered. The AD7873 itself will dissipate only about 2.4 mW with the internal reference enabled, a 2-MHz data clock, throughput of 125 kSPS and with a 3.6-V supply—while a 100 Ω screen is dissipating 129.6 mW! So the converter dissipates less than 2% of the power dissipated in the screen. However, it should be noted that in the first 3 of the 4 cases listed above, the screen-drive switches can be turned off between conversions. This feature is useful for providing a considerable reduction in average screen power if the application can allow for a reduced throughput rate (with the clock frequency held at 2 MHz).

The first three power-down options above will be especially helpful. For example, using those options, a throughput rate of 20 kSPS would reduce the screen power to an average of 20 mW under the conditions described above. For character recognition from the screen, throughput rates of 20 kSPS may be acceptable without any discernible reduction in performance. In general, however, a response-speed tradeoff should be anticipated. The screens themselves may have large parasitic capacitance associated with them—10 nF is not uncommon. This may result in a non-negligible RC time constant, so that time must be allowed for settling before screen measurements can be taken. For this reason, in many circumstances, powering down the screens between conversions might not be such a good idea—unless steps are taken to account for the large screen R-C time constant (this will be discussed later).

Resolution of the converter itself needs to be in the 10- to 12-bit range for typical screens. The AD7873 provides 12 bits of resolution. The B grade guarantees 12-bit no-missing-code (NMC) performance, while the lower-cost A version provides 11-bit NMC.

Peripheral functions—battery, temperature, and pressure measurement: The AD7873 supports a number of important peripheral functions vital to the PDA:

A dedicated battery monitoring ADC channel consists of an attenuator circuit that divides down the unregulated battery supply voltage by 4 and digitizes it. A supply of up to 6 volts can be applied to the V_{BAT} input pin. The accuracy of the function is important in applications that use battery supplies with a very shallow discharge slope and a very sharp knee (e.g., Li-ion batteries), so it's important for the system to know where the battery is in its discharge profile at any given time. Usually, the microprocessor instructs the battery channel to digitize the battery voltage every few seconds. Should the battery enter the sharp discharge part of its curve (Figure 7), it could be damaged and not recover from such a 'deep discharge' event. In practice, typical PDAs will flag a condition well before the knee is reached and respond appropriately. The AD7873 battery channel will give typical accuracy performance in the 0 to ±1% range with worst case error of ±3%, using the internal reference. Where better accuracy is required, the user may need to calibrate this reading, depending on the battery's discharge characteristic.

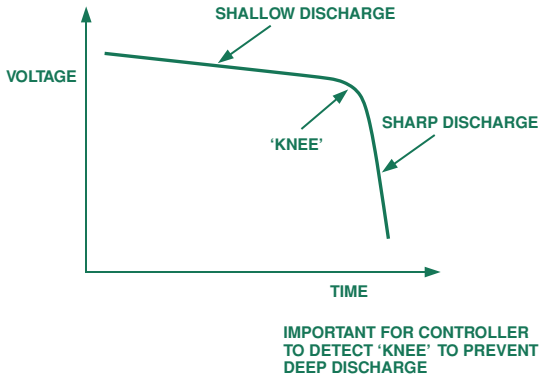


Figure 7. Typical battery discharge curve.

Temperature measurement is the second important peripheral function performed by the AD7873. Temperature within the PDA's case is an especially important parameter to monitor in devices containing rechargeable batteries. Typically, if the temperature exceeds 45°C during charging, the microprocessor needs to be flagged and provide an appropriate action to avoid permanent damage to the PDA due to overheating. Figure 8 shows the scheme used to measure temperature.

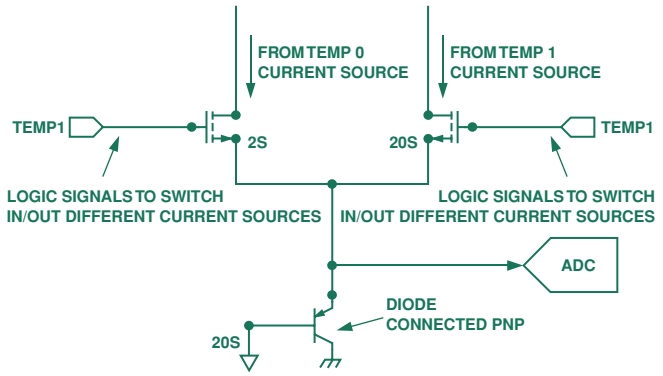


Figure 8. Temperature-measurement block diagram.

The AD7873 provides two modes of temperature measurement. The first, the single-conversion method, performs a simple conversion of the junction voltage of a diode-connected PNP transistor, biased with a constant current. The diode voltage will vary with temperature by about –2.1 mV/°C. Typically the reading would be factory-calibrated at 25°C during manufacture of the PDA. This method will typically provide a resolution of about 0.3°C with an accuracy of ±2°C.

To avoid calibration by the PDA manufacturer, a second—differential—method is provided on the AD7873. Two conversions are required, Temp 0 and Temp 1. Temp 0 is performed with a low value of diode bias current, I_0, and Temp 1 is performed with a bias current, I_1=105 I_0. From the Ebers-Moll equation and using some simple math, we can show that:

$$\Delta Vbe = (kT \mid q) \times 1n(N)$$

where

k = Boltzmann's constant, 1.38054 × 10^{-23} eV/K
q = Charge on the electron, 1.602189 × 10^{-19} C
T = Absolute temperature in kelvins
N= I_1/I_0, typically 105, as suggested for the AD7873.

Hence

$$T(°C) = 2490 \times \Delta Vbe - 273 \ \ K$$

Because **Δ**Vbe is limited to about 142 mV, the resolution of the differential method is significantly less then that of the single-conversion method. Typical resolution of about 1.6°C is achieved. The accuracy of both methods is typically the same, about ±2°C. Figure 9 plots a typical accuracy comparison for both methods over the temperature range, 0°C to 70°C. The main advantage of the differential temperature method is that it eliminates the need for calibration by the PDA manufacturer.

Pressure (or more precisely touch resistance) can be calculated via a number of simple arithmetic manipulations using the AD7873. One can determine whether the touch response is being generated by the stylus or a finger—or some other object—by measuring the contact resistance between the X and Y plates. This provides an indication of the size of the depressed area and the applied pressure. The area of the spot touched is proportional to the touch resistance. Two methods can be used. The first requires the user to know the total resistance of the X plane membrane. The second method requires that the total resistance of both X and Y membranes be known. Equations and diagrams can be found on the data sheet.

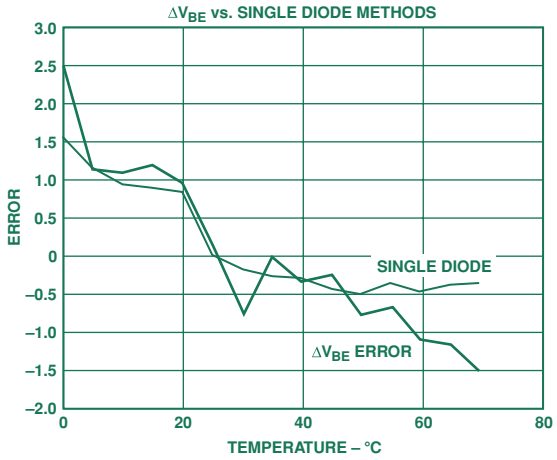


Figure 9. Comparative errors of differential- and single-conversion temperature measurement.

Waking up with a touch—The pen-interrupt function

Figure 10 illustrates the pen interrupt function, which provides an active-low open-drain output signal to the host microprocessor. When the part is powered down in a mode where the pen interrupt is enabled, a touch on the screen will cause the voltage on the PENIRQ pin to be pulled to ground. In this mode the Y– pin switch driver to ground is turned on during power down; when the two screens touch, the X+ pin is pulled to ground via the Y– pin—thus initiating an interrupt. The touch-resistance threshold is typically <30 kΩ for the AD7873. This level will ensure that spurious interrupts do not occur (if, for example, the screen were accidentally brushed against in a user's pocket). The microprocessor can use this interrupt to wake up the AD7873 and start coordinate measurements. Normally the PENIRQ pin voltage will idle at logic high. An external pull-up resistor in the range 10 kΩ to 100 kΩ is required on the PENIRQ pin for optimum operation of the function. Implementing this function using an open-drain output ensures that both the rising and falling edges of the PENIRQ signal are sharp, uninfluenced by the touch-screen capacitance (which can be as large as 10 nF) or the screen resistance itself.

Application issues when interfacing the AD7873 to a touch screen

We've already mentioned the trade off between power consumption in the screen and the time required for the screen to settle before coordinate readings can be taken—as a consequence of the large parasitic capacitance between the plates (about 10nF in some cases). Additional sources of potential error exist. The screens themselves can pick up a lot of noise from the LCD panel and backlight circuitry. The screen can also act as an antenna, picking up noise from external EMI/RFI sources. Mechanical bounce when the screen is touched carelessly is also a potential source of error. In most cases the designer will seek to minimize this noise by installing low pass filters on the tablet pins to ground. Capacitors of the order of 0.01 μF are common. (It's important to note that series resistance is not recommended for these filters, since it would lower the resolution of the converter because of the added voltage drop across the resistor). Because the filter capacitance, parasitic screen capacitance, etc., all act to increase the RC time constant of the screen, it is not advisable to make screen coordinate measurements in single-ended mode. The acquisition time of 3 clock cycles simply may not be long enough to allow the screen to settle before taking readings. For this reason, the ratiometric mode is much better. In that mode, the screen remains powered for the

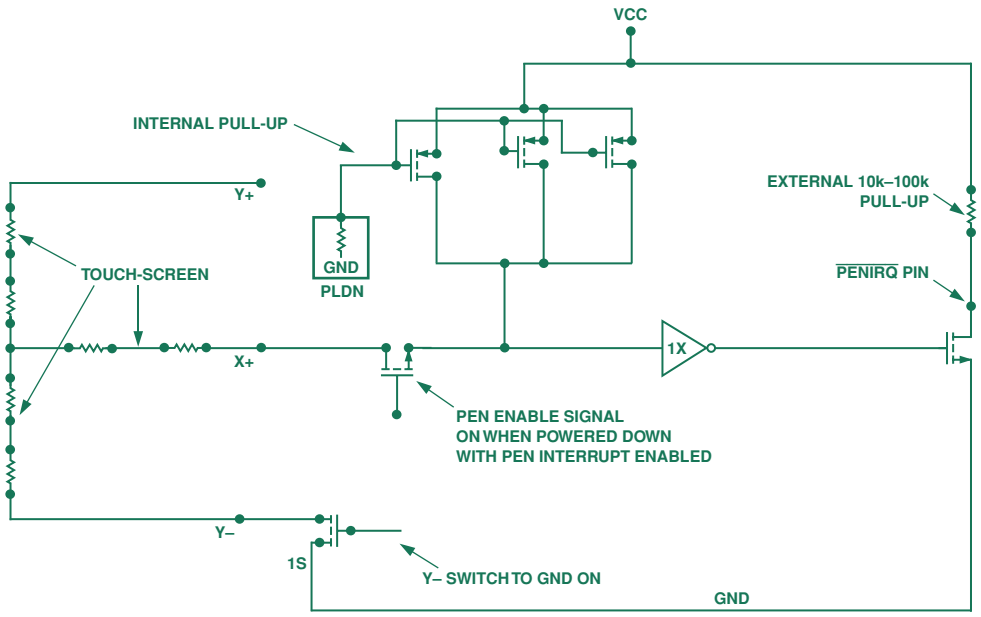


Figure 10. Pen-interrupt circuit.

full conversion cycle and, by averaging readings, an accurate result can be obtained—provided that the device is powered up in a mode where the touch-screen switch drivers remain turned on.

One can also get an accurate result without the need for averaging by simply delaying the acquisition time of the converter by including a delay between either the 6th, 7th, or 8th bits of the DIN word, as shown in Figure 11. The above discussion demonstrates the need for the designer to take account of the type of screen used with the AD7873. The converter has enough modes of operation and speed (minimum DCLK frequency is 10 kHz) to do its basic job—accurate coordinate measurement—with the most demanding of resistive touch screens and environments.

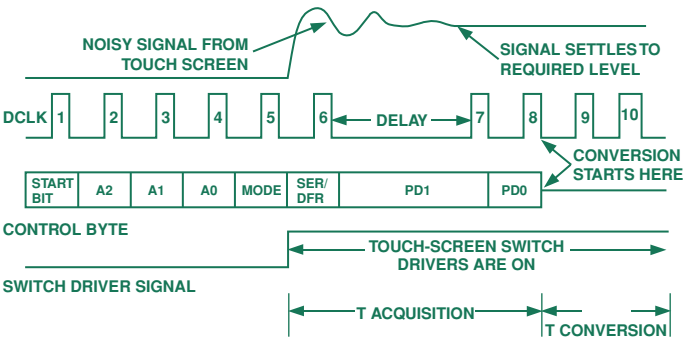


Figure 11. Delay during acquisition to allow noisy screen signal to settle.

Another problem designers may consider is failure due to high-energy voltage spikes discharged from the screen to the converter via the tablet pins. During PDA manufacture, the screen is floating and can build up considerable charge. This charge could eventually discharge through the tablet pins of the converter and permanently damage it. Then, during operation the screen is exposed to the outside world, and ESD events can occur which could damage the converter via the tablet pins—rendering the entire PDA useless. The AD7873 includes a silicon controlled rectifier (SCR)-based protection scheme at the tablet pins to make these I/Os as robust as possible against damage from ESD events. The ESD structure is capable of 15-kV protection—an attractive feature for PDA manufacturers, reducing risk of failure in the field—as well as during manufacture itself.

SUMMARY

The AD7873 facilitates coordinate measurement from a 4-wire resistive touch screen, meeting well the particular requirements of interfacing an ADC controller chip to the touch screen. Besides the basic coordinate measurement, the AD7873 provides the designer with battery monitoring, temperature sensing, touch detection, and pressure-measurement functions. We've also discussed issues designers can face when designing such systems–including power consumed by the screen, touch-screen settling issues, and potential ESD damage. The reader should now have a good insight into how the key circuitry of a PDA works, and an appreciation for the complex nature of the designer's task—and how a well-designed integrated-circuit chip can make it easier.

Versatile Programmable Amplifiers Use Digital Potentiometers with Nonvolatile Memory

by Alan Li (alan.li@analog.com)

INTRODUCTION

In concept, an op amp and a mechanical potentiometer can easily be combined to form an adjustable-gain amplifier, useful in many applications where electronic adjustments are needed. However, this combination is often unfeasible because of the potentiometer's limited resolution, poor temperature coefficient, high resistance drift over time, and the difficulties of remote adjustment. Now, the AD523x family of digital potentiometers with nonvolatile memory[1] can replace their mechanical counterparts and make these circuits practical (Figure 1).

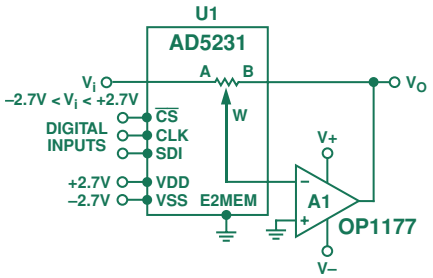


Figure 1. Programmable amplifier/attenuator.

The features offered by the AD523x digital potentiometers are extensive, but controlling them is relatively simple. A menu of sixteen instructions provides functions such as setting the "wiper" position, storing the wiper setting, reading back the wiper position, increment/decrement-by-one, and increment/decrement-by-6 dB (log taper adjustment). A write-protect pin provides memory protection, and additional nonvolatile memory is available for storing user information. Members of the AD523x family are available with one to four channels, 64- to 1024-position resolutions and 10-kΩ to 250-kΩ terminal resistance.

For operation, the user must furnish a power supply and digital controls (via 3-wire SPI serial communications) to realize the flexibility and reliability of these digital potentiometers. The parts will operate with a single supply, but dual supplies are required for ac or bipolar operation. A microcontroller or DSP is usually available to generate the control signals, but the nonvolatile memory enables users to permanently program the devices in the factory without ever having to readjust the settings in the field. This allows the devices to be controlled with personal computers[2] using programs such as Visual Basic, C, or LabView. It is unnecessary, therefore, to include a controller in these applications unless dynamic adjustment is needed. Evaluation boards for the AD5232 and AD5235 are available, making it easy to set up and perform preliminary evaluations with a personal computer.

Basic programmable amplifier

In the circuit of Figure 1, the gain (negative) is simply the ratio of the two terminal resistances, and the output voltage is:

$$V_O = -\left(\frac{R_{WB}}{R_{WA}}\right) \times V_i \quad (1)$$

$$V_O = -\left(\frac{D}{(2^N - D)}\right) \times V_i \quad (2)$$

Where:

R_{AB} = nominal end-to-end terminal resistance

R_{WB} = terminal resistance, W to B, $R_{WB} = R_{AB} \times D/2^N$

R_{WA} = terminal resistance, W to A ,
$R_{WA} = R_{AB} - R_{WB} = R_{AB} \times (1 - D/2^N)$

D = Base-10 equivalent of the binary word

N = number of bits

The gain expression implies a balanced quasi-logarithmic characteristic. This inverting configuration is useful because it makes available a wide range of gains, from very small to very large, with unity near half-scale (MSB). Because the resistors are fabricated on a single monolithic chip, resistance ratios are inherently matched, and the circuit can yield a temperature coefficient as low as 50 ppm/°C if using an AD5235. The grounded +input minimizes common-mode input errors. This circuit is a basic building block that suits many applications, especially where small signals are present and where high gain is required.

Bipolar programmable-gain amplifier with linear step adjustment

In the basic circuit of Figure 1, the output is always inverted with respect to the input, regardless of whether the circuit is providing gain or attenuation. The change in gain as the potentiometer is incremented is nonlinear. While this is useful in some cases, other applications may call for bipolar gain and/or a simple linear relationship. For example, motors need to rotate freely in both forward and reverse directions, thermal electric coolers heat or cool lasers, depending upon the direction of the current flow, and LCD panels require bipolar voltages for the contrast and brightness controls. In the most general case, to create a bipolar drive with arbitrary end points and linear step adjustment, a dual digital potentiometer, such as the AD5232, can be applied as shown in Figure 2. The output, Vo, can now be linearly programmed to amplify voltages between $+V_i$ and $-K \times V_i$, where K is the ratio of the two terminal resistances of U1 (Eq. 1). A2 provides buffered amplification for V_{W2}, minimizing the influence of the wiper resistance.

Notes
[1]The terms "nonvolatile memory" and "E2MEM" are used interchangeably
[2]Source code in Visual Basic is available to run on personal computer. Please contact the author for details

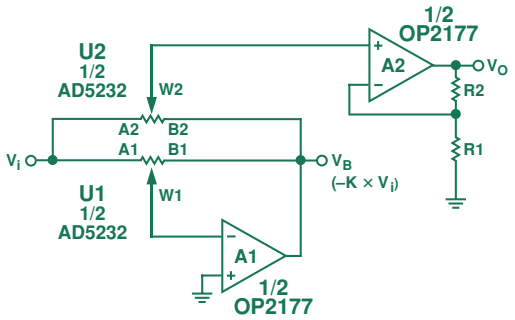


Figure 2. Bipolar programmable-gain amplifier with linear step adjustment

The transfer function in Figure 2 is:

$$\frac{V_O}{V_i} = \left(1 + \frac{R2}{R1}\right) \times \left(\frac{D2}{2^{N_2}}(1+K) - K\right) \quad (3)$$

In the simpler (and much more usual) case, where $K = 1$, a single digital potentiometer, such as the AD5231, is used in location U2, and U1 is replaced by a matched pair of resistors to apply V_i and $-V_i$ at the end terminals of the digital pot. The relationship will be

$$V_O = \left(1 + \frac{R2}{R1}\right) \times \left(\frac{2D}{2^N} - 1\right) \times V_i \quad (4)$$

Table 1 shows the result of adjusting D, with A2 configured: as a unity-gain follower ($R_1 = \infty$, R2 = 0), with a gain of 2 ($R_1 = R_2$), and with a gain of 10 ($R_2 = 9 \times R_1$). The result is a bipolar amplifier with linearly programmable gain and 256-step resolution.

Table 1. Circuit gain versus *D*

D (step)	R1 = ∞, R2 = 0	R1 = R2	R2 = 9 × R1
0	–1	–2	–10
64	–0.5	–1	–5
128	0	0	0
192	+0.5	+1	+5
255	+0.992	+1.984	+9.92

As implied in Figure 2, R1/R2 and/or R3/R4 can optionally be replaced by digital potentiometers—if tight temperature-coefficient matching and very high gains are desired. If discrete resistors are used, resistor matching is imperative.

The wiper resistance of the digital potentiometer is the on-resistance of the internal solid-state switches, typically 50 Ω to 100 Ω. This is relatively small when compared with the nominal resistance R_{AB}, but the wiper resistance approximately doubles over the operating temperature range and can become the major source of error when the device is programmed to operate at low values. The wiper terminal of the potentiometer should always be connected to a high-impedance node, such as the input terminal of an op amp, as shown in the above circuits. The OP1177 family, the fourth generation of the industry-standard OP07, was chosen for its low offset and low bias-current characteristics. This minimizes the effects of wiper resistance on the voltage divider ratio at the tap point.

The circuit in Figure 2 can also serve as a programmable gain amplifier (PGA) for A/D converters. A/D converters often need to handle dynamic input ranges from a few millivolts to several volts. The AD7723 16-bit, 1.2-MSPS sigma-delta ADC, for example, has an input voltage limit of ± 2V when using the internal reference. A PGA can be applied as a front-end signal-conditioning device that scales millivolt-level signals to the input range of the converter.

If the input voltage range is higher than ± 2.7 V, discrete resistors are needed in series with U1 due to the 5.5V maximum voltage limitation of the AD523x potentiometers. Figure 3 shows the use of series resistors in a high-voltage PGA application of the configuration of Figure 2. As noted earlier, resistor matching is essential.

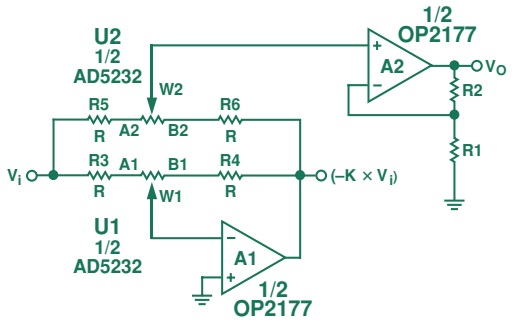


Figure 3. PGA handles high input voltage.

By configuring the AD523x nonvolatile memory digital potentiometers and op amps shown in this article, the user can easily design a versatile amplifier with high-resolution programmability, bipolar controllability, and linear/log step adjustment capability.

In addition, users may observe that digital potentiometers may replace not only some of their mechanical counterparts, but also nonprecision D/A converters in some cost-sensitive applications.

References

Sergio Franco, *Design with Operational Amplifiers and Analog Integrated Circuits*, second edition. McGraw Hill, 1998.

The engineering staff of Analog Devices, Inc., Daniel H. Sheingold, ed., *Analog-Digital Conversion Handbook*, third edition. Englewood Cliffs, NJ: Prentice-Hall, 1986.

Digital Potentiometer Selection Guide (http://www.analog.com/support/standard_linear/selection_guides/dig_pot.html)

Digital Potentiometer Product Tree (http://www.analog.com/industry/data_converters/product_trees/digitalpotentiometers.html)

Advanced LCD Driver Lowers Cost of High Performance Data Projectors

By Edmund Spence (ed.spence@analog.com) and Zoltan Frasch (zoltan.frasch@analog.com)

INTRODUCTION

Data projectors based on liquid crystal display (LCD) technology have made enormous strides in recent years—with smaller, lower cost units offering increased brightness and higher resolution. As a result, the projector has evolved from an installed large conference room display fixture to become an everyday tool for work groups and sales organizations, and a staple of home-entertainment systems. The DecDriver® IC, discussed here, is a major factor in present and future improvements to visual quality as well as cost reduction of these displays.

Today, the dominant technology used in projector display engines comprises three (RGB—red, green, blue) high-temperature polysilicon (HTPS) liquid crystal display (LCD) *microdisplays*. The three-color LCD microdisplays are presented with separate color data, and their light outputs are summed optically before being projected through the lens onto the screen. This system allows greater control over the color quality, with improved brightness and efficiency, and much less potential for 'color breakup' than field-sequential single-panel systems based on a rotating color wheel.

Recently, a newer LCD technology is emerging onto the projection market, based on lower cost silicon wafers—liquid crystal on silicon (LCOS). LCOS-based projection engines have the potential to enable low-cost home theater—as well as other large venue displays, such as public display and large rear-projection monitors or workstations. The benefits that LCOS microdisplays bring to these applications include higher resolutions than are presently achieved with HTPS LCD panels, combined with the potential—not yet materialized—for lower cost. Other LCOS advantages include higher pixel densities, smaller panel sizes, and higher aperture ratios. In home entertainment, the use of projector engines will include both front and rear projection—and will require resolutions in excess of 2 million full-color pixels (10-bit *gamma-corrected*—see Application Note AN-548).

Both LCOS and HTPS display technologies are LCD-based and require high-performance drive electronics to provide high-quality, high-resolution displays. (see Footnote, page 4).

The microdisplay interface differs from today's laptop displays, which are now most often *a-Si* (amorphous, active matrix LCD display). These laptop displays require an individual driver for each column of pixels. This interface is slow and the IC devices, which tend to integrate as many as 384 drivers per chip, are disproportionately long.

HTPS or LCOS *microdisplays*, on the other hand, have integrated a multiplexer (MUX) function to distribute the imaging signal among the pixel columns. The physical pixel itself serves as the load capacitance in a simple sample-and-hold circuit, which holds the imaging signal while the LCD material responds. The input of the MUX requires a greatly reduced number of interface channels, trading off increased speed for simpler circuitry. This reduces on-chip fanout to the microdisplay, which currently can measure as little as 0.5 inches along the diagonal. In addition, this configuration also reduces the power dissipation, area, and cost of the drive electronics. To top it off, while a high-end laptop display might include 8-bit drivers, the microdisplay used for projection engines today uses *10-bit* gamma-corrected inputs. Such higher quality images are needed because the projected image is so big, showing display artifacts not easily noticed on a 10-inch laptop display. The need for higher quality images and greater speed demands increased performance from the drive electronics.

Key figures of merit for these analog drivers now include good absolute output voltage *accuracy*, fast output voltage *settling* into capacitive loads, high *data rates*, compact *footprint*, and low power *dissipation*. Interface designs targeting home theater markets should continue to rely on analog inputs to the LCD microdisplays to avoid high-speed logic noise and enhance the quality of the signal applied to the display pixels.

Before the DecDriver chips came on the scene, the legacy drive electronics solution for microdisplays was based on sample-and-hold topologies to decimate the digital imaging data in time. Decimation is required to match the incoming high-speed data rate to the relatively limited bandwidth of the LCD pixel. A MUX function integrated onto the HTPS panel distributes the analog image signal across the pixel columns, loading a finite set of pixels on each clock cycle and continuing until a line of pixels fills. Sample-and-hold-based drivers suffer from several limitations that lead to poor image quality—large PCB area, limited ability to obtain resolutions greater than XGA (see Table 1, page 4), and high power dissipation. Errors in the sample-and-hold function—due to pedestal, droop, and settling time—limit the display quality of this architecture, affecting color match and hampering the system speed needed for higher resolutions. Whether or not the controller ASICs (responsible for timing, image signal processing, etc.) include high-speed video DACs, discrete video amplifiers are inherently needed to provide the necessary dynamic voltage range and settling time. They in turn drive the sample-and-hold amplifiers needed for decimating.

In addition, increasing the pixel counts—while keeping refresh rates fixed—dictates even faster drive electronics. This is particularly true of LCOS, because the faster response times and double frame rates of LCOS technology require panel interfaces with fewer (but much faster) channels than are required for HTPS.

Apart from speed issues, home theater quality LCOS displays have the same needs as HTPS. The analog inputs to the panels must have good channel-to-channel accuracy, wide dynamic range, and fast settling time.

Digitally decimated architecture

The AD8380 DecDriver (decimating driver) IC is a monolithic silicon solution that delivers all the performance needed to drive 10-bit gamma-corrected images directly onto high-resolution HTPS and LCOS panels with high speed and accuracy. It effectively replaces the sample-and-hold function by latching high-speed digital data from the controller, then transferring all channels on a given signal—thus accomplishing decimation-in-time by converting the data to parallel analog imaging signals.

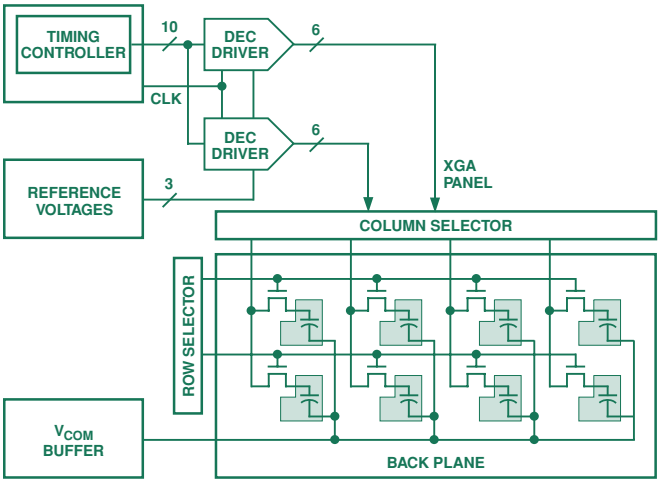


Figure 1. Simplified view of an XGA microdisplay panel driven by two DecDriver ICs.

The DecDriver is designed to optimize settling time and power dissipation by fabricating the DACs and drive amplifiers on the same chip, using a new high-density 26-V fast bipolar process. By integrating the high-voltage output drive amplifiers with the fast bipolar DACs, they can be factory-trimmed together to meet required absolute accuracy specifications.

This complete solution, designed for high output precision, also allows complete control of the imaging signal—including contrast, brightness, signal inversion, and output VCOM levels—with no sacrifice in accuracy. The combined speed, flexible logic control, and laser-trimmed output accuracy permit modular design using multiple DecDriver devices interchangeably in XGA, SXGA, and higher resolution systems.

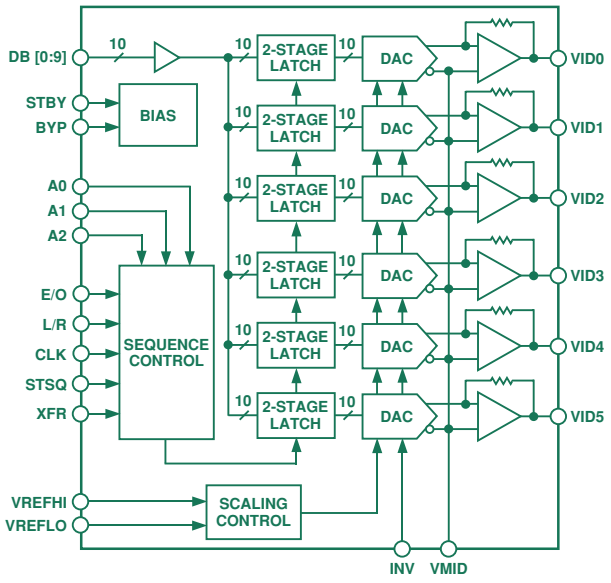


Figure 2. AD8380 DecDriver IC functional block diagram. Fast 10-bit input is latched, then passed through the DACs to output video amplifiers on XFR pulse. Output voltage levels are controlled with VREF, INV, and VMID controls.

Front- and rear-projection systems require scans in opposite directions from one another, so L/R controls are provided to determine the direction in which data is latched, making it easier to design both front- and rear-projection systems. E/O provides the ability to latch on alternate clock edges, simplifying the demultiplexing of dual high-speed data paths.

Display performance

Higher resolution displays call for higher pixel clock rates, and consequently faster drive electronics. Table 1 (page 4) shows the pixel clock and system clock rates of common video (VESA and SMPTE) formats.

Maximum assigned frequencies are 7 MHz for HTPS panels and 40 MHz for LCOS panels.

Table 2 tabulates the corresponding maximum settling times and number of required input channels of HTPS and LCOS panels with common video formats. Input data for LCOS panels is assumed frame-doubled since pixel density is too high to use column or line inversion without suffering crosstalk.

Inadequate driver operating frequency results in incompatibility with certain video formats or types of panels. A compatible driver must be able to operate at the system CLK frequency and provide the required output channels accurately within the required settling time. Excessive settling times or unmatched output channel accuracy will cause ghosting or mismatched column-to-column voltage levels, resulting in vertical lines on the image.

Accuracy

Light intensity of each color component of a given pixel depends on the driver output level, which depends on the digitally coded amplitude. Errors in D/A conversion and driver amplification in a given channel and from channel to channel and color to color lead to errors in intensity and color value; systematic errors degrade the display by producing annoying visual effects. *Absolute error* in driver outputs is proportional to rms error in the output of each driver. The absolute error specifications of the DecDriver include *all* errors: i.e., DAC nonlinearity, full-scale error, offset error, amplifier offset, and channel-matching errors.

To best correlate image artifacts and driver errors, the rms or differential error voltage, V_{DE}, (Figure 3) is defined as:

$$V_{DE}(n) = \tfrac{1}{2}\left[V_{OUTN}(n) - V_{OUTP}(n)\right] - \left[V_{FS} \times (1 - n/1023)\right]$$

where

$V_{OUTN}(n)$ is the output voltage when INV is driven high

$V_{OUTP}(n)$ is the output voltage when INV is driven low

$1/2[V_{OUTN}(n) - V_{OUTP}(n)]$ is the rms value of the output

$(V_{FS} \times (1 - n/1023))$ is the rms value of the ideal output

n is one of 2^{10} input code values

V_{FS} is the full-scale output voltage

A common-mode error (Figure 4) which shifts the transfer function away from the midpoint, V_{MID}, is defined as:

$$V_{CME}(n) = \tfrac{1}{2}\{\tfrac{1}{2}\left[V_{OUTN}(n) + V_{OUTP}(n)\right] - V_{MID}\}$$

where

$$\tfrac{1}{2}\left[V_{OUTN}(n) + V_{OUTP}(n)\right]$$

is the dc average value of the output.

Common-mode errors result in increasing crosstalk as pixel density increases (i.e., as pitch decreases).

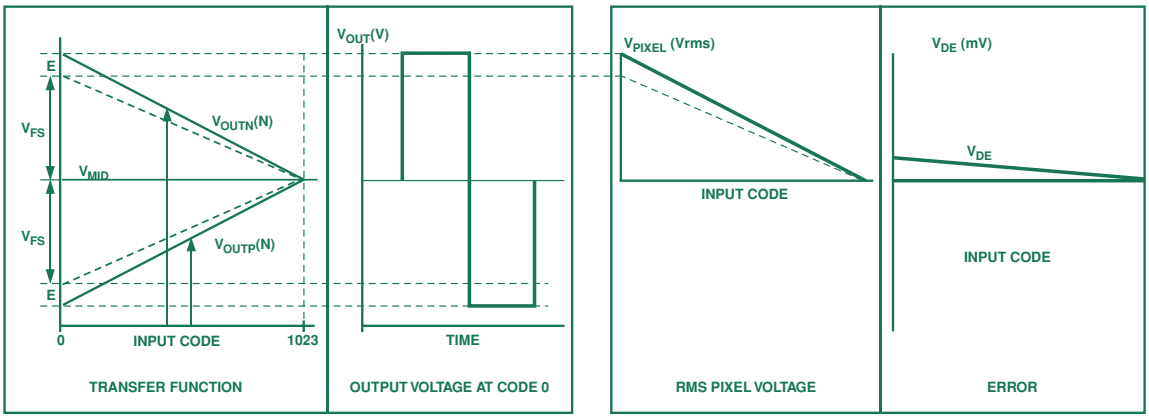


Figure 3. A typical case of V_{DE} or differential error. From left to right, plots show transfer function, time-domain rms signal at code zero, rms voltage as seen by the pixel, and the error vs. DAC code.

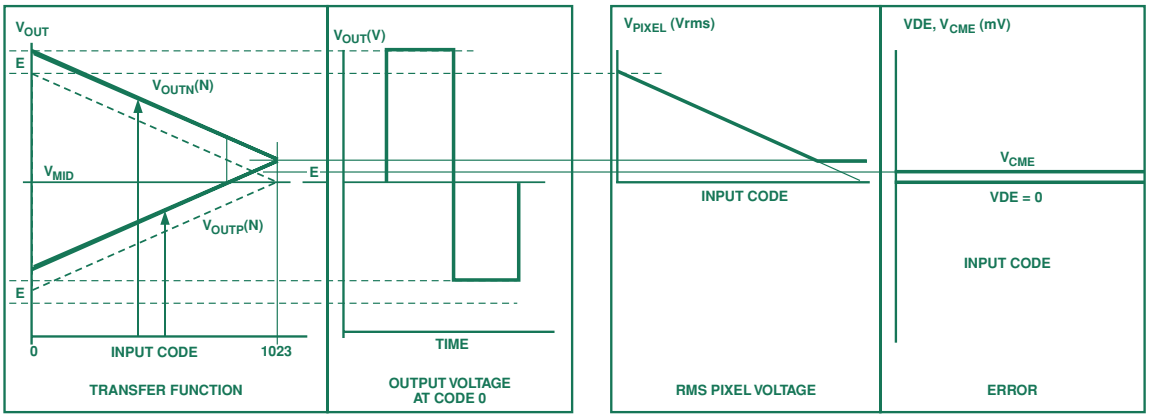


Figure 4. A typical example of V_{CME} or common-mode (offset) error. As in Figure 3, AD8380 transfer function, time-domain rms signal at code zero, rms voltage seen by the pixel, and the V_{CME} error as a function of DAC code.

Device characteristics

The AD8380, a 10-bit, 6-channel device, is the first implementation of the DecDriver architecture and is currently available in production quantity. Its analog controls adjust output reference and full-scale levels; its digital controls include Invert, Right/Left loading, Even/Odd data, and sequential- and addressed-latch loading.

It dissipates 550 mW from 15-V analog and 3.3-V digital supplies, or less than 100 mW per analog channel.

Reference inputs and logic controls foster modular design with common inputs and controls. SXGA and higher resolution projection systems have been constructed using multiple devices per color panel. Operation over the rated temperature range has been achieved with input data clock rates as high as 150 MSPS.

Maximum V_{DE} error is less than ±7.5 mV (or 1.5 gray-scale levels). As noted above, this includes all errors due to DAC nonlinearity, offset and full-scale errors, as well as amplifier gain errors. Common-mode or V_{CME} errors are less than ±3.5 mV (0.7 gray-scale levels). See Figure 5.

Output amplifier settling time (Figure 6) is typically 35 ns to 0.25% for a 5-V step into a 150-pF load. Table 2 shows that this is fast enough to drive LCOS panels with HDTV resolution.

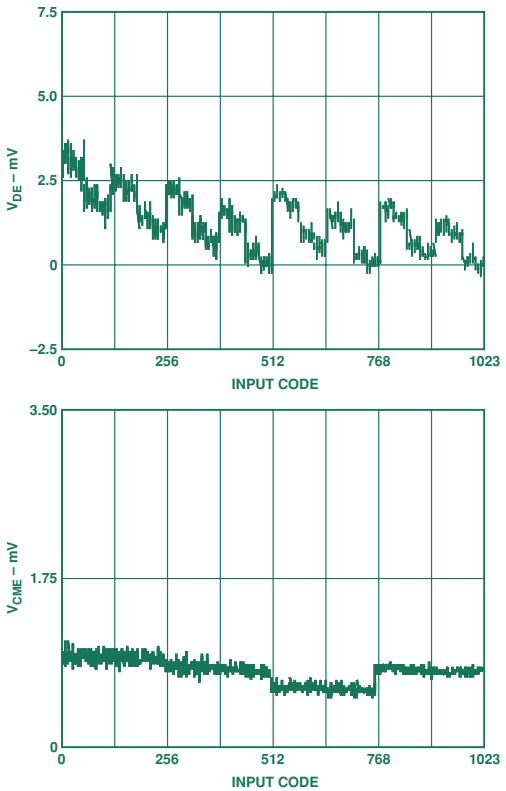


Figure 5. Typical V_{DE} and V_{CME} as functions of input code.

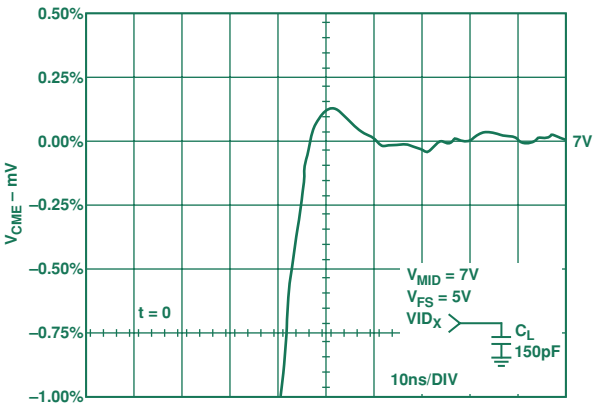


Figure 6. Typical DecDriver output settling time.

SUMMARY

In a projected image of a commercially available 12-channel XGA system using HTPS panels, no vertical lines or other image artifacts are visible. Superb color matching can be observed on such a projection system with 1000-lumen output, a significant improvement over legacy systems. The circuit footprint is reduced by elimination of external trim components between DACs integrated on the CMOS controller, gain amplifiers, and sample-holds used for de-multiplexing. Integrated control of both input logic options and output image signal levels makes the DecDriver IC a complete drive electronics subsystem, directly interfacing to both the CMOS controller and the LCD panel, with a further saving in board area. Superior dynamic performance of the integrated output drive amplifiers, coupled with fast logic inputs and output accuracy, means that multiple DecDriver ICs can be used to implement higher resolution systems for SXGA or UXGA projectors. Similar results have been demonstrated with SXGA LCOS panels—and higher resolution systems using the DecDriver architecture are in evaluation.

FOOTNOTE

About pixels, resolution, formats, clock speed, etc.

A digital video display originates with a grid of small areas each of which is briefly stimulated to produce, transmit, or reflect an individual electrically determined light intensity, which is retained during a frame's display. These areas are known as picture elements or *pixels*. A digital color display system involves three sources of pixels (red, green, blue, or RGB), whose intensities combine to illuminate the area in color. In the simplest form of *monochrome* display, the pixel areas in the top line are illuminated in horizontal sequence, then the pixels in the next line are illuminated in sequence, and so on, until the last pixel in the last line has been illuminated; then the scan starts a new *frame*.

The number of pixels per line, multiplied by the number of lines, is called the *resolution*, and the ratio of horizontal to vertical pixels is called the *aspect ratio*. In the real world, the electrical signals that control the display need time for instructions at the end of each line and at the end of each frame, so some time is allowed for additional samples. Thus, for example, in a standard computer *format* called *SVGA*, the visible display area resolution is 800 pixels per line with 600 lines per picture—but the theoretical number of possible samples is 1056 per line, and there are in effect 632 lines per frame.

To the average human visual apparatus, in order for the display to appear to provide a constant picture without flickering and to show continuous motion, the number of frames shown per second should be at least 50, and is more usually about 60. A simple calculation for this example shows that the electrical signal defining the picture must be capable of displaying with correct intensity $800 \times 600 \times 60 = 28.8$ million pixels per second and handling up to $1056 \times 632 \times 60 = 40.04$ million samples per second. This huge number is known as the *pixel clock rate*.

Table 1 compares the resolution and speed numbers for a variety of standard formats used in digital television, digitized analog TV, and computers. It also shows a key difference between computer formats and some TV formats: *interlacing*. In standard analog TV, the complete picture is displayed by two interlaced frames with half as many lines per frame; the lines of the alternate frames are interlaced to fill the display space, relying for continuity on persistence of vision. Formats retaining those vestiges require a pixel clock only half as fast as the system clock.

Table 1. Pixel clock rates for common video formats.

VIDEO DISPLAY FORMATS AND REQUIRED NUMBER OF HTPS AND LCOS CHANNELS BASED ON THE MAXIMUM DATA UPDATE FREQUENCY SUPPORTED BY THE PANEL.										INTERLACED VIDEO LINE DOUBLED
NAME	ASPECT RATIO	RESOLUTION PIXELS	LINES	SAMPLES/ LINES	LINES/ FRAME	V. RATE (Hz)	H. RATE (kHz)	PIXELS CLOCK (MHz)	SYSTEM CLOCK (MHz)
DIGITAL TV FORMAT									
HDTV 1080i	16:9	1920	1080	2200	1125	60.00	33.75	74.25	148.50
HDTC 720p	16:9	1280	720	1650	750	60.00	45.00	74.25	74.25
SDTV 480p	16:9/4:3	704	480	858	525	59.94	31.47	27.00	27.00
SDTV 480i	16:9/4:3	704	480	858	525	59.94	15.73	13.50	27.00
DIGITIZED ANALOG TV FORMATS									
NTSC 525i	4:3	640	480	858	525	59.94	15.73	13.50	27.00
PAL 625i	4:3	768	576	960	625	50.00	15.63	15.00	30.00
COMPUTER FORMATS									
VGA	4:3	640	480	800	525	60.00	31.50	25.20	25.20
SVGA	4:3	800	600	1056	632	60.00	37.92	40.04	40.04
XGA	4:3	1024	768	1343	807	60.00	48.42	65.03	65.03
SXGA	5:4	1280	1024	1688	1067	60.00	64.02	108.07	108.07
UXGA	4:3	1600	1200	2160	1250	60.00	75.00	162.00	162.00

Panels and pixels

The number of D/A conversion and driver channels required depends on the type of display and the video format. Both HTPS and LCOS require more than a single channel per color, but the slower HTPS displays require more channels, from six to as many as 24 for HDTV, SXGA, and UXGA, while LCOS displays require from two to eight channels. The 6-channel AD8380 DecDriver is designed to meet this need for driving multiple channels of 10-bit data with adequate settling time, as well as to provide many of the control functions. Table 2 indicates the channel and settling time requirements for each of the formats in Table 1.

Table 2. Maximum allowable settling times for HTPS and LCOS panels, with assumed number of channels.

VIDEO DISPLAY FORMATS	CHANNELS REQUIRED		CHANNELS USED		REQUIRED SETTLING TIME IN ns AT CHANNELS USED	
	HTPS	LCOS	HTPS	LCOS		
	f_{MAX} (MHz)		f_{MAX} (MHz)		HTPS	LCOS
NAME	7	40	7	40		
DIGITAL TV						
HDTV 1080i	21.2	7.4	24	8	162	27
HDTC 720p	10.6	3.7	12	4	162	27
SDTV 480p	3.9	1.3	6	2	222	37
SDTV 480i	3.9	1.3	6	2	222	37
DIGITIZED ANALOG TV						
NTSC 525i	3.9	1.3	6	2	222	37
PAL 625i	4.3	1.5	6	2	200	33
COMPUTER FORMATS						
VGA	3.6	1.3	6	40	238	40
SVGA	5.7	2.0	6	25	150	25
XGA	9.3	3.3	12	31	182	31
SXGA	15.4	5.4	24	37	222	37
UXGA	23.1	8.1	24	25	148	25

Sensing, Analyzing, and Acting in the First Moments of an Earthquake

by Giuseppe Olivadoti

Earthquakes and seismic activity have always been a hot issue. Attention is focused time and again by disasters such as the earthquakes in Turkey, Taiwan, and India. It has become apparent that the power of an earthquake is not something we are currently prepared to handle. Much of the problem is that the damaging earthquake waves seem to come out of nowhere without warning. However, this may not truly be the case. Earthquakes, if properly analyzed, can actually give warning of their incipient occurrence, even if only moments before the ground starts shaking. A critical objective is to quickly identify the precursors of the destructive waves of the earthquake in time to initiate an alarm and a shutdown of vulnerable facilities.

When feasible, early detection can potentially be very valuable. Consider for instance the devastating damage and loss reported from the earthquake that occurred in Turkey on Tuesday, August 17, 1999. The 45-second earthquake, of Richter magnitude 7.4, had an epicenter approximately 7 miles (11 km) southeast of Izmit, an industrial city roughly 56 miles (90 km) east of Istanbul. The earthquake was felt over a large area—as far east as Ankara, which is about 200 miles (320 km) away. Unofficial estimates place the death toll between 30,000 and 40,000.

Although the collapse of commercial and residential buildings caused most of the deaths and injuries, a widely publicized and spectacular tank explosion—which occurred at the massive Tüpras refinery in Korfez—caused significant deaths and injuries due to the fires that followed in its wake. Fire in one of the tank farms quickly spread to other tank farms through pipelines and distribution systems and burned out of control for several days, prompting an evacuation within a three-mile radius. Some loss of life and property at the Tüpras refinery might have been prevented if valves controlling pipelines and distribution systems carrying highly flammable material had been shut off. A few extra moments to react to the earliest warnings in such a system might have allowed valves to be shut off on pipelines and distribution systems, and an alarm to be sounded.

In many quake-prone locations, safety codes, even for homes, require acceleration-sensitive shut-off valves. While undoubtedly quite useful, they respond only upon arrival of the surface wave, and then only (in many cases) for vibration in a single plane. Also, they may respond with false alarms to vibrations caused by large vehicles and to other non-earthquake shocks, requiring wasteful resetting procedures.

How might earthquakes be analyzed to give advance warning?

When an earthquake occurs, energy radiates outward in all directions. The energy travels through and around the earth as three types of seismic waves called *primary*, *secondary*, and *surface* waves.

The energy of primary waves (or P waves) travels through the earth as a sequence of back-and-forth vibrations in a plane (x- and y-axis) parallel to the direction of propagation of the seismic wave. The wave's passage through the earth causes the pushing (compression) and pulling (dilation) of particles in its path, and it can travel through solids or liquids. P waves are the fastest of the three types of seismic waves. Figure 1 shows the passage of P waves through the earth.

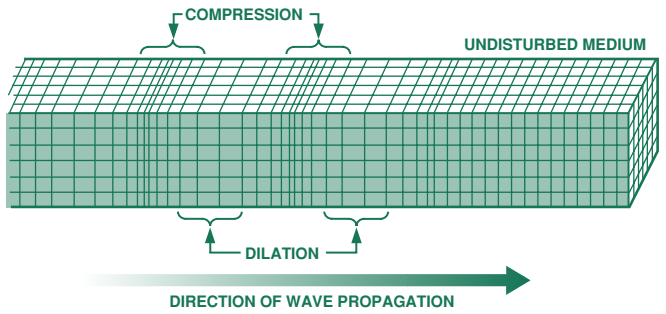


Figure 1. Passage of P waves through the earth's crust*.

Secondary waves (or S waves, see Figure 2), also referred to as *shear* waves, can travel through solids, but unlike *P* waves cannot travel through liquids. The energy of *S* waves travels through the earth as a sequence of up-and-down vibrations perpendicular to the surface of the earth. Its passage causes particles to vibrate in all directions, North-South and East-West. Its velocity is between that of *P* waves and that of surface waves.

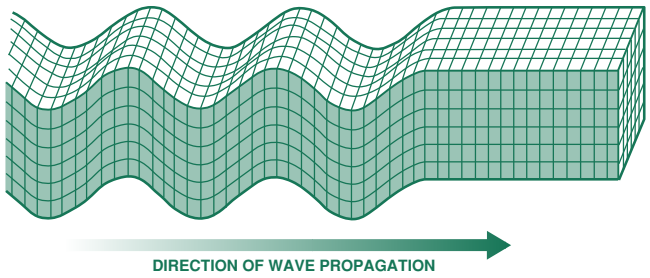
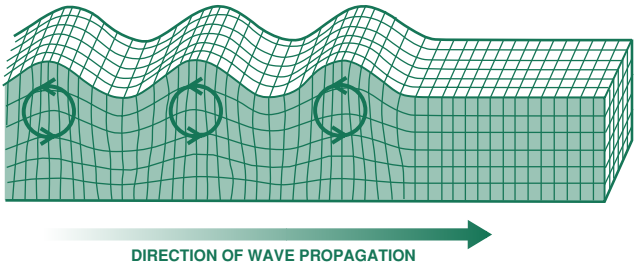


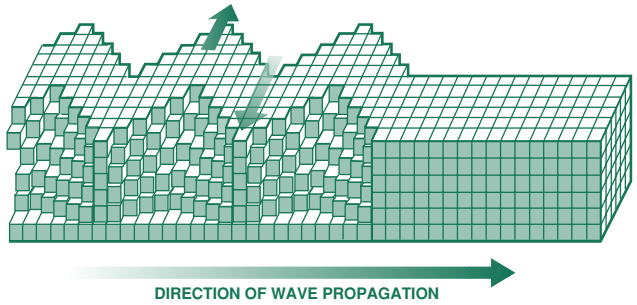
Figure 2. Passage of S waves through the earth's crust*.

Surface waves are the slowest and by far the most destructive of the three types of seismic waves. Surface waves travel along the surface of the earth as two types of waves: Rayleigh waves have a horizontal shearing motion similar to S waves, while Love waves have a rolling motion in the vertical plane much like water waves. Figure 3 shows the passage of both the Rayleigh and Love waves through the earth.

**Nuclear Explosions and Earthquakes: The Parted Veil.* by Bruce A. Bolt © 1976 by W. H. Freeman and Company. Used with Permission.



a. Passage of Rayleigh waves through earth's crust.



b. Passage of Love waves through earth's crust.

Figure 3. Surface Waves*

P waves typically travel 1.68 times faster than S waves and 2 to 3 times faster than surface waves, which typically travel at about 3.7 km/s. Thus there is typically a one-second separation between the *P* and *S* waves for every 8 km traveled. S waves travel about 4 km/s faster than surface waves, so every 4 km away from the epicenter typically adds one extra second of delay between the P-S complex and the arrival of the surface waves.

The various types of earthquake waves follow this pattern. At a given distance from the epicenter, the P wave arrives first, then the S wave, both of which have such small energies that they are not threatening. Finally, the surface waves arrive with all of their damaging energies (Figure 4). It is predominantly the surface waves that we would notice as the earthquake. This knowledge, that preceding any surface or destructive earthquake waves there are tell-tale body waves, can be used to help predict the arrival time of the damaging surface waves.

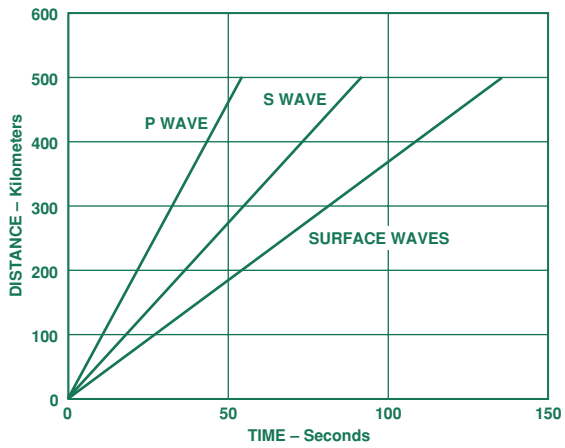


Figure 4. Comparison of arrival times at a given distance from the epicenter. Calculations are based on short distance approximations of Surface and Body wave velocities.

**Nuclear Explosions and Earthquakes: The Parted Veil.* by Bruce A. Bolt © 1976 by W. H. Freeman and Company. Used with Permission.

For example, at a distance of 7 miles (11 km), nearly 3 seconds would elapse during which the P-S complex could be sensed and identified, and alarms and valve closures initiated. Rapidly responding low-cost sensors and digital signal processors could render a rapid decision and allow nearly 100% of this time for mechanical operations.

The elements of a detection system, shown in Figure 5, would include 3-dimensional sensing of earth motion, filtering, analysis, and actuation of alarms, valve closures, etc.

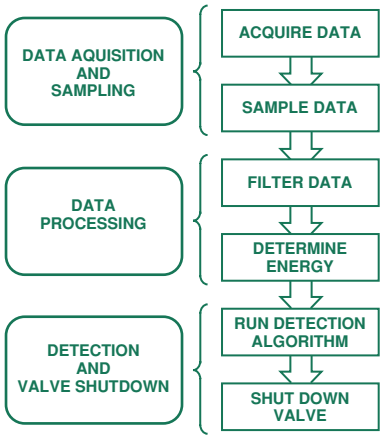


Figure 5. Functional diagram of detection system.

The simplest of approaches is shown in the flow diagram (Figure 6). While continuously monitoring the X, Y, and Z accelerometer outputs, an unusually large acceleration disturbance is identified, and labeled as a possible P (compression) wave. The system stands by, waiting to identify a following transverse disturbance as an S wave. If none occurs within a time corresponding to (say) 500 km of P-wave travel, the system shrugs it off as a false alarm (or a distant phenomenon). If, on the other hand, a transverse disturbance does show up, the detection system actuates an alert, which includes a surface-wave ETA (estimated time of arrival), and possibly seeks confirmation from similar systems in the locality before issuing the alarm. If they are in constant contact, this could all occur within less than a few milliseconds after the presence of an S wave is verified.

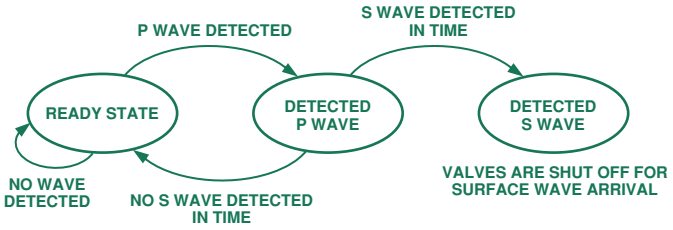


Figure 6. Detection and verification process flow chart.

Applying signal processing to solve this problem

In addressing this problem the motion in all three axes—X, Y, and Z—must be measured promptly and accurately. An accelerometer would be a perfect example of a device that is capable of sensing this sort of seismic movement. An example could be the fast-responding Analog Devices ADXL202 high-sensitivity and very-low-cost, two-axis accelerometer family. The data produced by the array of accelerometers would then need to be processed

continuously. The data would need to be filtered so that random noise would be removed and any received signal compared with earthquake signatures using energy-detection algorithms to identify and predict the arrival time of an earthquake at the sensing point (and elsewhere)—and the likely energy level of the surface waves. An example of a device for this sort of processing is a DSP, such as the low-cost, high-performance floating-point ADSP-21161N. It offers 32-bit precision, on-chip memory, and a host of other advanced architectural features that make it suitable for filtering and analysis, as well as decision-making.

The key advantages of doing this type of detection:

This type of detection could be very useful in high-risk earthquake areas. Exploiting the differences in arrival time of an earthquake could offer a few seconds advance notice that a destructive wave will be present shortly. This time could be used for a number of things, such as stopping the flow of hazardous or flammable waste from flowing through pipes, and halting production lines of hazardous or flammable materials. This type of detection can also offer fairly reliable earthquake identification by exploiting its characteristic multiple wave system. It is because of the knowledge that seismic waves travel in groups of three that the detection system can filter out extraneous noise such as might be produced by large trucks, rock blasting, etc.

The disadvantages of relying on this type of detection:

This approach (and most others) will be ineffective if the epicenter of the earthquake is too close to the sensing device. There will generally not be enough time for mechanical devices to react, even if the entire analysis process requires only a few milliseconds or less. Since the time difference between seismic waves is due to their propagation velocity relative to each other, there needs to be some minimum distance away from the epicenter (see Figure 4) to establish an identification with reasonable confidence. That minimum distance will depend on the application. For instance, in protecting a gas or oil line, one must take into account the speed with which a valve can seal the pipe.

CONCLUSION

It is reasonable to consider that a low-cost seismic detection system can be designed using tried-and-true high-performance digital signal processing plus the kinds of motion sensors that are now used in collision detection for airbags in millions of automobiles. An example of an experimental system embodying these principles, using an Analog Devices DSP and accelerometer, can be seen in the *SHARC 2000 International DSP Conference* proceedings. A seismic detection system offers the possibility of having a few critical moments of advance notice before the arrival of a destructive surface earthquake wave. Perhaps if these ideas plant the seed of designs for useful devices that can undergo extensive testing and be produced in high volume, some aspects of seismic disasters could be contained, saving human works and lives.

References

Bolt, Bruce. *Nuclear Explosions and Earthquakes: The Parted Veil.* Figure 3.5, p. 48. Copyright 1976. W. H. Freeman and Company.

Carlson, Shawn. "The New Backyard Seismology." *Scientific American* April, 1996: 100-102.

Lemaire, Christophe. "Low-Cost, iMEMS Accelerometers Sense Strong Motion Seismic Activity for Hazard Mitigation Applications." *Analog Devices Products and Datasheets*: 4 pp. 16 July 1999.
http://www.analog.com/publications/whitepapers/products/Seismart.html.

McEvilly, Thomas. "US4516206: Post-processing of seismic parameter data based on valid seismic event determination." *IBM Patent Network*.
http://www.delphion.com/details? pn=US04516206.

"Operating Principles." *QuakeAlert Earthquake Detection System*: 1 p. 16 July 1999.
http://www. quakeguard.com/func.html.

SHARC® International DSP Conference 2000 proceedings. Norwood, MA: Analog Devices, Inc. pp. 64–72 (ISBN 0-916550-23-0)

"Travel Time Curve." Earthquakes.: 1 p. 25 Jan 2001.
http://www.geo.arizona.edu/saso/Earthquakes/Recent/ttc.html.

A Useful Role for the ADXL202 Dual-Axis Accelerometer in Speedometer-Independent Car-Navigation Systems

by Peter Shih (peter.shih@analog.com) and
Harvey Weinberg (harvey.weinberg@analog.com)

INTRODUCTION

Vehicle navigation using the Global Positioning System (GPS) has been of increasing interest over the past decade; GPS navigation is frequently installed in today's high-end luxury cars and in many commercial vehicles. Because they rely on high-frequency radio signals from satellites, vehicles with GPS navigation systems can travel into situations where they lose the GPS signal for short intervals. With appropriate algorithms, GPS can be integrated with other course-plotting techniques to provide users with continuously accurate navigation information.

Dead-(deductive) reckoning is one method widely used in vehicle navigation. It utilizes three distinct inputs to predict position: a set of starting coordinates, the direction of travel, and the speed of travel. Its accuracy is limited due to its relative positioning scheme; the absolute position error grows in proportion to the distance traveled. Other methods of non-GPS vehicle navigation include map-matching, inertial navigation, and Delta A measurement. Map-matching is based on the principle that if you are traveling near a road or parallel to it, there's a good chance that you are indeed on that road (may not work well in heavily populated areas). Inertial navigation relies on an accelerometer to derive velocity as the integral of acceleration. In Delta A measurement, the GPS signals subsequently recovered from the receiver are correlated with inputs from non-GPS systems. This method can correct for the inaccuracies of an accelerometer, such as noise and zero-*g* offset over temperature/time. More than one of these techniques can be used in conjunction with the GPS to display the position more accurately.

While it would be helpful to use the car's installed speedometer, certain difficulties arise. In general, speedometer information is unavailable because it is not bussed out to systems outside the engine/ABS/stability control computer(s). Because it is considered "safety-critical" in many cases, the speedometer output will not be connected to anything that could take down the bus. Besides, GPS systems are usually built by third parties who may want to build a generic product with a wide potential market. It is in applications of this sort that the ADXL202 dual-axis accelerometer can be used to develop accurate speed estimates for the navigation system. A digital compass or gyroscope is used in conjunction with an accelerometer to determine the approximate direction of travel. The information is then translated by the navigation system (in conjunction with other methods mentioned above) to determine location relative to the point at which the signal was lost.

The method for determining velocity to be described here uses an accelerometer to sense the time interval for both front and back wheels to encounter a bump in the road (while moving straight ahead). Whether one is driving on a local road or a highway, there will always be imperfections in the road. These imperfections translate into bumps and jolts sensed immediately by the car's wheels, and ultimately by its passengers. In order to track the speed by sensing these bumps, an accelerometer is used to identify their magnitudes and timing. Thus, for a car with a given wheelbase (W), the interval (T_1) for both axles to encounter a bump can be used to compute the speed at which the car is traveling, using the following equation: (See Figure 1.)

Speed [mi/hr] = (W [ft]/T_1 [s]) × (3600 s/hr)/(5280 ft/mi)[1].

Sample data log:

M_{XX} = Magnitude of bumps (duty cycle %)
t_{XX} = Instantaneous time of bumps (sec)
T_X = Duration between two correlating bumps (sec)
S_0 = Previous valid speed (mph)
S_1 = Current calculated speed (mph)

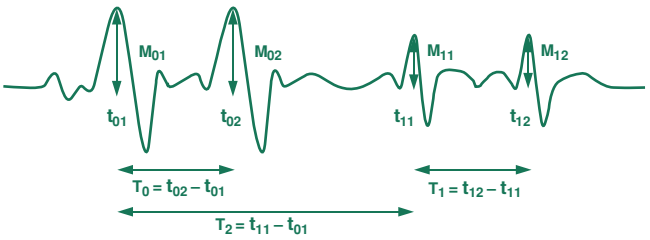


Figure 1. Event timing for speed measurement.

When logging data during a typical drive on the local roads to develop experimental information, it is not easy for the accelerometer to discriminate between bounces and vibrations in the car's suspension system and the spike-pairs caused by irregularities in the road. Thus a filtering system is needed to isolate the bumps. The ADXL202EB-232 Evaluation Board[2] has internal software that allows data to be smoothed with low-pass filtering. This provides a better chance of recognizing the road bumps and using them in calculations. This problem having been solved, a correlation problem arises—for example, if there are two similar bumps less than a car-length apart, it is difficult to make some sense of the four total bumps experienced by the car within a short time. Hence the necessity of coming up with an algorithm to cleanly translate the data points into valid speedometer readings.

If the accelerometer is placed halfway between the axles with the X-axis parallel to the Earth's surface and aimed straight ahead, and the Y-axis perpendicular to the Earth's surface, the magnitudes of the bump impulses produced by the front and rear wheels are approximately equal (depending somewhat on the vehicle's suspension system). To identify bump pairs, it is necessary to make magnitude comparisons to match the bumps originating with the front and rear axles. At the same time, the current tabulated speed must be compared with the last valid speed to determine whether or not the current calculated speed is feasible. For instance, if a

[1]Speed (km/hr) = (W [m]/T_1 [s]) × (3600 s/hr)/(1000 m/km)
[2]http://www.analog.com/techsupt/eb/ADXL202EB-232A_c.pdf

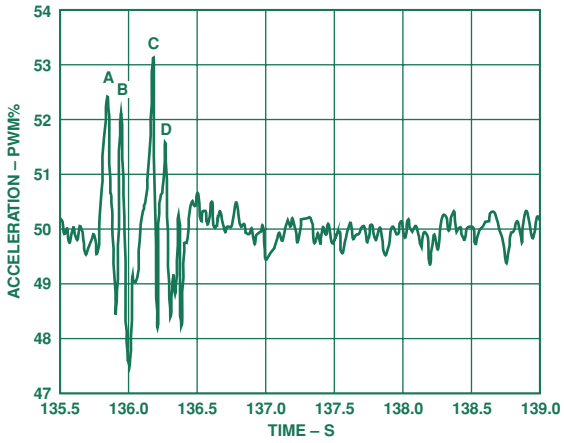
vehicle was going at 25 mph about one second ago, it is extremely unlikely for the current speed to be 45 mph or more. So by using timing and speed comparisons, any outputs that don't make sense will be rationalized or ignored.

Data analysis

The digital outputs of the ADXL202 are duty-cycle modulated; on-time is proportional to acceleration. A 50% duty cycle (square-wave output) represents a nominal 0-*g* acceleration; the scale factor is ±12.5% duty cycle change per *g* of acceleration. These nominal values are affected by the initial tolerances of the device, including zero-*g* offset error and sensitivity error.

In the application described here, a 50% duty cycle output corresponds to a perfectly smooth ride—no bumps or vibration detected by the accelerometer. In general, because of its suspension dynamics, a vehicle is more responsive to bumps at lower speeds. So for lower speeds there needs to be less sensitivity to the magnitude of the bumps (M_{XX}) and the threshold level can be higher. Magnitudes less than the threshold level will be considered invalid data, while those above the threshold (valid magnitudes) will pass on to the next phase of filtering.

The purpose of the next stage is to block out unfeasible speeds implied by two adjacent bumps that are closer together than the wheelbase of the vehicle. To tackle this problem, if S_1 (defined below) is beyond a general acceleration limit of 20 mph/s when compared to S_0, the set of data is invalid. However, the configuration of the four bumps can be translated into legitimate speeds by pairing the first and third bumps—and the second and fourth bumps (Figure 2).



A: M_{01} = 52.41%, t_{01} = 135.862s
B: M_{11} = 52.15%, t_{11} = 135.938s
C: M_{02} = 53.08%, t_{02} = 136.179s
D: M_{12} = 51.66%, t_{12} = 136.242s

Figure 2. X-axis (forward-back), parallel to Earth's surface

Figure 2 shows recent data logged with the ADXL202EB in a car traveling at a constant speed of 20 mph. At first glance, it seems that A and B are two correlating bumps, as well as C and D. However, $t_{11} - t_{01}$ = 0.076 seconds, which translates into a speed of about 81 mph. This would be compared with the last valid speed, using the Delta A measurement method, and would negate the correlation of A & B. Then A and C are paired: $t_{02} - t_{01}$ = 0.317 seconds, as are B and D (0.304 s), which translate into respective speeds of about 19.4 mph and 20.2 mph, respectively. From Equation 1 and a 9-foot wheelbase, T_X for 20 mph is equal to 0.307 seconds. The results here show differences of 3.2% and 1% respectively.

This simple solution deals with the most common source of false-readings. But of course there are many other bump configurations that could cause faulty speed-readings. Many of them can be dealt with by increasingly clever algorithms and signal conditioning, but in the end, one must realize that this computation is a part of a system to substitute for temporary loss of GPS signals, intended to maintain reasonable accuracy over short intervals.

Which axis?

One can consider using either the X- or the Y-axis (or both) for acceleration data to measure the bumps on the road. The Y-(vertical) axis measures the actual magnitude, as modified—and frequently muddied—by the car's suspension-system dynamics, while the X-axis measures the magnitude of the front-back acceleration component (accelerometer cross-axis) as the car passes over the bump.

The first approach (Figure 3) measures the Y-axis acceleration (perpendicular to Earth's surface). In a no-bump situation the measurement will be 1 *g*, established by Earth's static gravitational force. This nominal 62.5% output (50% + 12.5%/*g*) can be offset to read 50%, as a starting point for positive or negative vertical deflection forces.

The second approach (seen in Figure 2) uses the X-axis (parallel to Earth's surface) to measure forward-and-back acceleration. In a no-bump situation the measurement will be 0 *g*. As the car's motion is influenced by the bump, the accelerometer picks up a Y-axis acceleration spike, strongly filtered by the car's suspension system. At the same time, the X-axis picks up a smaller but "cleaner" forward component of that acceleration spike due to front-and-back motion induced by the bump (and the accelerometer's cross-axis sensitivity). During trial runs, this latter method (Figure 2) gave better results.

This approach helps in filtering out unwanted noise. Furthermore, one can see in Figure 3 that upon hitting a bump, the vertical component of acceleration tends to show a rather low damping factor. Relying on forward-back motion can overcome these complications.

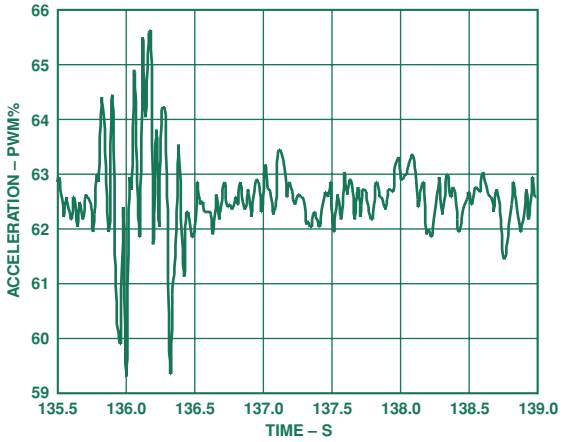


Figure 3. Y-(vertical) axis, perpendicular to Earth's surface.

APPENDIX

The trials described in this article were performed using the ADXL202EB-232 Evaluation Board and Crossbow software on a personal computer. Here is a step-by-step description of the process and its flow chart (Figure 4).

1. Connect ADXL202EB-232 to serial cable, then to RS-232 port of computer.
2. Open the software program X-Analyze provided by Crossbow.
3. Click on "Add a connection" and select 'ADXL202-EB-232A on COM1'.
4. Click on "Configure a connection"; then click on "Calibrate".
5. To calibrate, hold the board with the XY plane perpendicular to the ground and rotate the board 360° about that plane.
6. Select update rate to be as fast as possible and the logging rate to be 50 Hz; Select filtering rate to be 1.
7. Select logging folder; this is where the .txt logged files will be saved.
8. Click on "Save & Exit".
9. Mount/Attach board onto vehicle with Y-axis facing either towards the bottom of the vehicle and X-axis straight ahead. Make sure that the board is securely mounted so that it cannot move relative to the car's body when going over bumps.
10. Click on "Log all connections" when ready to log data.
11. Once logging data, the same button is used to "Stop logging".
12. Open .txt files and copy and paste onto Excel to create charts/graphs.

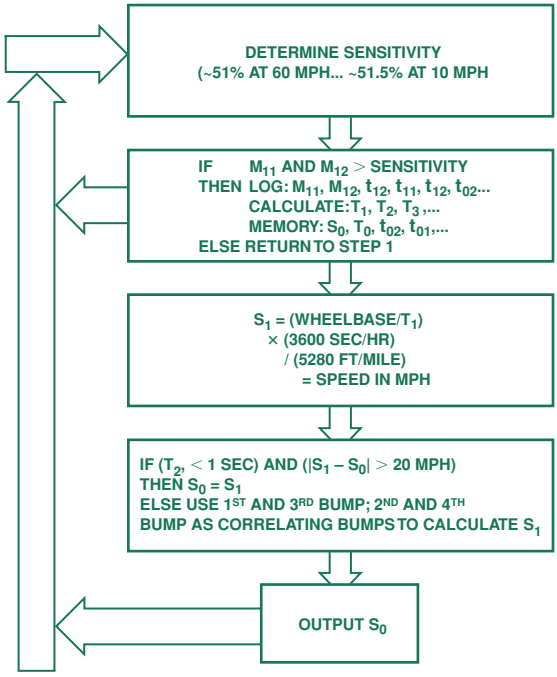


Figure 4. Computation flow chart.

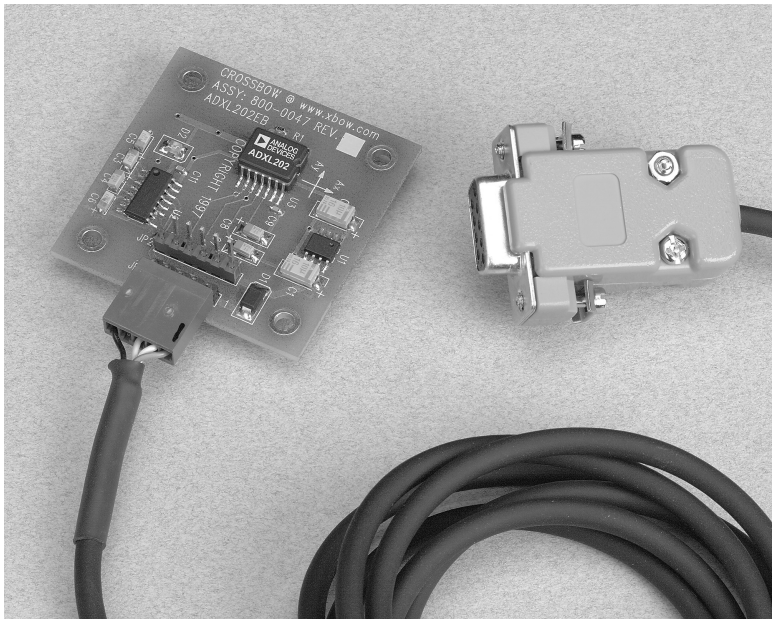


Figure 5. Evaluation board.

Ask The Applications Engineer—24

by Steve Guinta*

RESISTANCE

Q. I'd like to understand the differences between available resistor types and how to select the right one for a particular application.

A. Sure, let's talk first about the familiar "discrete" or axial-lead type resistors we're used to working with in the lab; then we'll compare cost and performance tradeoffs of the discretes and thin- or thick-film networks.

Axial Lead Types: The three most common types of axial-lead resistors we'll talk about are *carbon composition*, or carbon film, *metal film* and *wirewound*:

- *carbon composition* or carbon film-type resistors are used in general-purpose circuits where initial accuracy and stability with variations of temperature aren't deemed critical. Typical applications include their use as a collector or emitter load, in transistor/FET biasing networks, as a discharge path for charged capacitors, and as pull-up and/or pull-down elements in digital logic circuits.

  Carbon-type resistors are assigned a series of standard values (Table 1) in a quasi-logarithmic sequence, from 1 ohm to 22 megohms, with tolerances from 2% (carbon film) to 5% up to 20% (carbon composition). Power dissipation ratings range from 1/8 watt up to 2 watts. The 1/4-watt and 1/2-watt, 5% and 10% types tend to be the most popular.

  Carbon-type resistors have a poor temperature coefficient (typically 5,000 ppm/°C); so they are not well suited for precision applications requiring little resistance change over temperature, but they are inexpensive—as little as 3 cents [USD 0.03] each in 1,000 quantities.

  Table 1 lists a decade (10:1 range) of standard resistance values for 2% and 5% tolerances, spaced 10% apart. The smaller subset in lightface denote the only values available with 10% or 20% tolerances; they are spaced 20% apart.

Table 1. Standard resistor values: 2%, 5% and 10%

10	**16**	27	**43**	68
11	18	**30**	47	**75**
12	**20**	33	**51**	82
13	22	**36**	56	**91**
15	**24**	39	**62**	100

Carbon-type resistors use color-coded bands to identify the resistor's ohmic value and tolerance:

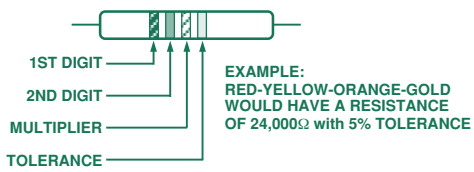


Table 2. Color code for carbon-type resistors

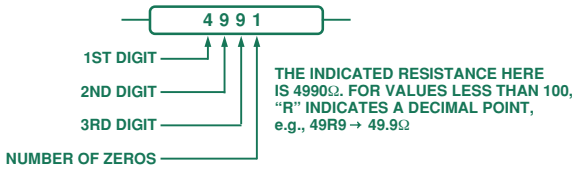
digit	color	multiple	# of zeros	tolerance
–	silver	0.01	–2	10%
–	gold	0.10	–1	5%
0	black	1	0	–
1	brown	10	1	–
2	red	100	2	2%
3	orange	1 k	3	–
4	yellow	10 k	4	–
5	green	100 k	5	–
6	blue	1 M	6	–
7	violet	10 M	7	–
8	gray	–	–	–
9	white	–	–	–
–	none	–	–	20%

- *Metal film* resistors are chosen for precision applications where initial accuracy, low temperature coefficient, and lower noise are required. Metal film resistors are generally composed of Nichrome, tin oxide or tantalum nitride, and are available in either a hermetically sealed or molded phenolic body. Typical applications include bridge circuits, RC oscillators and active filters. Initial accuracies range from 0.1 to 1.0 %, with temperature coefficients ranging between 10 and 100 ppm/°C. Standard values range from 10.0 Ω to 301 kΩ in discrete increments of 2% (for 0.5% and 1% rated tolerances).

Table 3. Standard values for film-type resistors

1.00	1.29	1.68	2.17	2.81	3.64	4.70	6.08	7.87
1.02	1.32	1.71	2.22	2.87	3.71	4.80	6.21	8.03
1.04	1.35	1.74	2.26	2.92	3.78	4.89	6.33	8.19
1.06	1.37	1.78	2.31	2.98	3.86	4.99	6.46	8.35
1.08	1.40	1.82	2.35	3.04	3.94	5.09	6.59	8.52
1.10	1.43	1.85	2.40	3.10	4.01	5.19	6.72	8.69
1.13	1.46	1.89	2.45	3.17	4.09	5.30	6.85	8.86
1.15	1.49	1.93	2.50	3.23	4.18	5.40	6.99	9.04
1.17	1.52	1.96	2.55	3.29	4.26	5.51	7.13	9.22
1.20	1.55	2.00	2.60	3.36	4.34	5.62	7.27	9.41
1.22	1.58	2.04	2.65	3.43	4.43	5.73	7.42	9.59
1.24	1.61	2.09	2.70	3.49	4.52	5.85	7.56	9.79
1.27	1.64	2.13	2.76	3.56	4.61	5.96	7.72	9.98

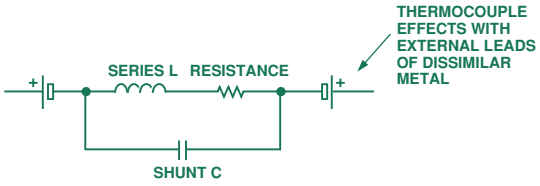
Metal film resistors use a 4 digit numbering sequence to identify the resistor value instead of the color band scheme used for carbon types:



*His photo and a brief biography appear in Analog Dialogue 30-2 (1996), page 2.

• *Wirewound* precision resistors are extremely accurate and stable (0.05%, <10 ppm/°C); they are used in demanding applications, such as tuning networks and precision attenuator circuits. Typical resistance values run from 0.1 Ω to 1.2 MΩ.

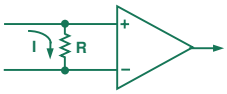
High Frequency Effects: Unlike its "ideal" counterpart, a "real" resistor, like a real capacitor (*Analog Dialogue* 30-2), suffers from parasitics. (Actually, any two-terminal element may look like a resistor, capacitor, inductor, or damped resonant circuit, depending on the frequency it's tested at.)



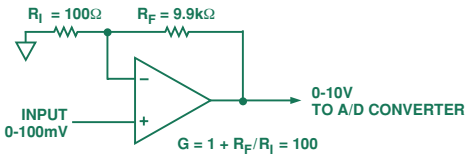
Factors such as resistor base material and the ratio of length to cross-sectional area determine the extent to which the parasitic L and C affect the constancy of a resistor's effective dc resistance at high frequencies. Film type resistors generally have excellent high-frequency response; the best maintain their accuracy to about 100 MHz. Carbon types are useful to about 1 MHz. Wirewound resistors have the highest inductance, and hence the poorest frequency response. Even if they are non-inductively wound, they tend to have high capacitance and are likely to be unsuitable for use above 50 kHz.

Q. *What about temperature effects? Should I always use resistors with the lowest temperature coefficients (TCRs)?*

A. Not necessarily. A lot depends on the application. For the single resistor shown here, measuring current in a loop, the current produces a voltage across the resistor equal to I × R. In this application, the absolute accuracy of resistance at any temperature would be critical to the accuracy of the current measurement, so a resistor with a very low TC would be used.



A different example is the behavior of gain-setting resistors in a gain-of-100 op amp circuit, shown below. In this type of application, where gain accuracy depends on the ratio of resistances (a ratiometric configuration), resistance matching, and the tracking of the resistance temperature coefficients (TCRs), is more critical than absolute accuracy.



Here are a couple of examples that make the point.

1. Assume both resistors have an actual TC of 100 ppm/°C (i.e., 0.01%/°C). The resistance following a temperature change, ΔT, is

$$R = R_0(1+ TC\ \Delta T)$$

For a 10°C temperature rise, both R_f and R_i increase by 0.01%/°C × 10°C = 0.1%. Op amp gains are [to a very good approximation] 1 + R_F/R_I. Since both resistance values, though quite different (99:1), have increased by the same *percentage*, their ratio—hence the gain—is unchanged. Note that the gain accuracy depends just on the resistance *ratio*, independently of the absolute values.

2. Assume that R_I has a TC of 100 ppm/°C, but R_F's TC is only 75 ppm/°C. For a 10°C change, R_I increases by 0.1% to 1.001 times its initial value, and R_F increases by 0.075% to 1.00075 times its initial value. The new value of gain is

$$(1.00075\ R_F)/(1.001\ R_I) = 0.99975\ R_F/R_I$$

For an ambient temperature change of 10°C, the amplifier circuit's gain has decreased by 0.025% (equivalent to 1 LSB in a 12-bit system).

Another parameter that's not often understood is the self-heating effect in a resistor.

Q. *What's that?*

A. Self-heating causes a change in resistance because of the increase in temperature when the dissipated power increases. Most manufacturers' data sheets will include a specification called "thermal resistance" or "thermal derating", expressed in degrees C per watt (°C/W). For a 1/4-watt resistor of typical size, the thermal resistance is about 125°C/W. Let's apply this to the example of the above op amp circuit for full-scale input:

Power dissipated by R_I is

E^2/R = (100 mV)^2/100 Ω = 100 μW, leading to a temperature change of 100 μW × 125°C/W = 0.0125°C, and a negligible 1-ppm resistance change (0.00012%).

Power dissipated by R_F is

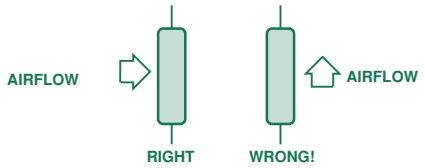
$^2/R$ = (9.9 V)^2/9900 Ω = 9.9 mW, leading to a temperature change of 0.0099 W × 125°C/W = 1.24°C, and a resistance change of 0.0124%, which translates directly into a 0.012% gain change.

Thermocouple Effects: Wirewound precision resistors have another problem. The junction of the resistance wire and the resistor lead forms a thermocouple which has a thermoelectric EMF of 42 μV/°C for the standard "Alloy 180"/Nichrome junction of an ordinary wirewound resistor. If a resistor is chosen with the [more expensive] copper/nichrome junction, the value is 2.5 μV/°C. ("Alloy 180" is the standard component lead alloy of 77% copper and 23% nickel.)

Such thermocouple effects are unimportant in ac applications, and they cancel out when both ends of the resistor are at the same temperature; however if one end is warmer than the other, either because of the power being dissipated in the resistor, or

its location with respect to heat sources, the net thermoelectric EMF will introduce an erroneous dc voltage into the circuit. With an ordinary wirewound resistor, a temperature differential of only 4°C will introduce a dc error of 168 μV—which is greater than 1 LSB in a 10-V/16-bit system!

This problem can be fixed by mounting wirewound resistors so as to insure that temperature differentials are minimized. This may be done by keeping both leads of equal length, to equalize thermal conduction through them, by insuring that any airflow (whether forced or natural convection) is normal to the resistor body, and by taking care that both ends of the resistor are at the same thermal distance (i.e., receive equal heat flow) from any heat source on the PC board.



Q. What are the differences between "thin-film" and "thick-film" networks, and what are the advantages/disadvantages of using a resistor network over discrete parts?

A. Besides the obvious advantage of taking up considerably less real estate, resistor networks—whether as a separate entity, or part of a monolithic IC—offer the advantages of high accuracy via laser trimming, tight TC matching, and good temperature tracking. Typical applications for discrete networks are in precision attenuators and gain setting stages. Thin film networks are also used in the design of monolithic (IC) and hybrid instrumentation amplifiers, and in CMOS D/A and A/D converters that employ an R-2R Ladder network topology.

Thick film resistors are the lowest-cost type—they have fair matching (<0.1%), but poor TC performance (>100 ppm/°C) and tracking (>10 ppm/°C). They are produced by screening or electroplating the resistive element onto a substrate material, such as glass or ceramic.

Thin film networks are moderately priced and offer good matching (0.01%), plus good TC (<100 ppm/°C) and tracking (<10 ppm/°C). All are laser trimmable. Thin film networks are manufactured using vapor deposition.

Tables 4 compares the advantages/disadvantages of a thick film and several types of thin-film resistor networks. Table 5 compares substrate materials.

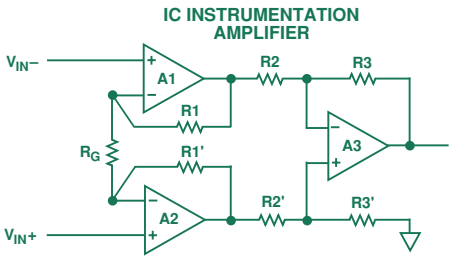
Table 4. Resistor Networks

Type	Advantages	Disadvantages
Thick film	Low cost	Fair matching (0.1%)
	High power	Poor TC (>100 ppm/°C)
	Laser-trimmable	Poor tracking TC
	Readily available	(10 ppm/°C)
Thin film on glass	Good matching (<0.01%)	Delicate
	Good TC (<100 ppm/°C)	Often large geometry
	Good tracking TC (2 ppm/°C)	Low power
	Moderate cost	
	Laser-trimmable	
	Low capacitance	
Thin film on ceramic	Good matching (<0.01%)	Often large geometry
	Good TC (<100 ppm/°C)	
	Good tracking TC (2 ppm/°C)	
	Moderate cost	
	Laser-trimmable	
	Low capacitance	
	Suitable for hybrid IC substrate	
Thin film on silicon	Good matching (<0.01%)	
	Good TC (<100 ppm/°C)	
	Good tracking TC (2 ppm/°C)	
	Moderate cost	
	Laser-trimmable	
	Low capacitance	
	Suitable for hybrid IC substrate	

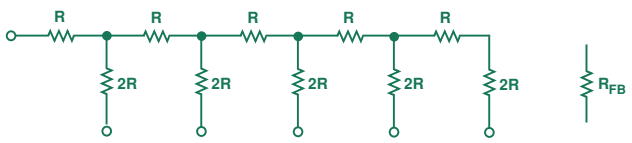
Table 5. Substrate Materials

Substrate	Advantages	Disadvantages
Glass	Low capacitance	Delicate
		Low power
		Large geometry
Ceramic	Low capacitance	Large geometry
	Suitable for hybrid IC substrate	
Silicon	Suitable for monolithic construction	Low power
		Capacitance to substrate
Sapphire	Low capacitance	Low power
		Higher cost

In the example of the IC instrumentation amplifier shown below, tight matching between resistors R1-R1', R2-R2', R3-R3' insures high common-mode rejection (as much as 120 dB, dc to 60 Hz). While it is possible to achieve higher common-mode rejection using discrete op amps and resistors, the arduous task of matching the resistor elements is undesirable in a production environment.



Matching, rather than absolute accuracy, is also important in R-2R ladder networks (including the feedback resistor) of the type used in CMOS D/A converters. To achieve *n*-bit performance, the resistors have to be matched to within 1/2*n*, which is easily achieved through laser trimming. Absolute accuracy error, however, can be as much as ±20%. Shown here is a typical R-2R ladder network used in a CMOS digital-analog converter.



Ask The Applications Engineer—25

by Grayson King*

OP AMPS DRIVING CAPACITIVE LOADS

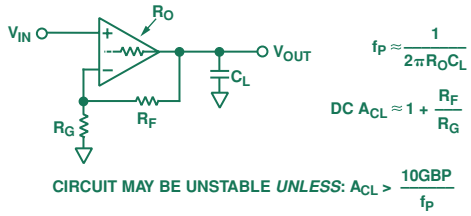
Q. *Why would I want to drive a capacitive load?*

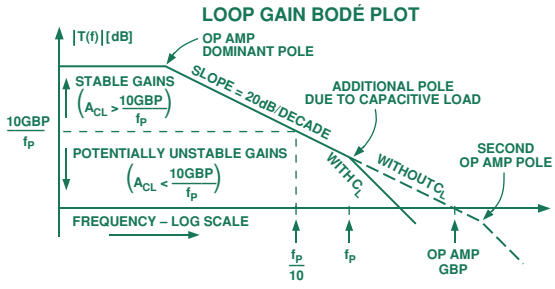
A. It's usually not a matter of choice. In most cases, the load capacitance is not from a capacitor you've added intentionally; most often it's an unwanted parasitic, such as the capacitance of a length of coaxial cable. However, situations do arise where it's desirable to decouple a dc voltage at the output of an op amp—for example, when an op amp is used to invert a reference voltage and drive a dynamic load. In this case, you might want to place bypass capacitors directly on the output of an op amp. Either way, a capacitive load affects the op amp's performance.

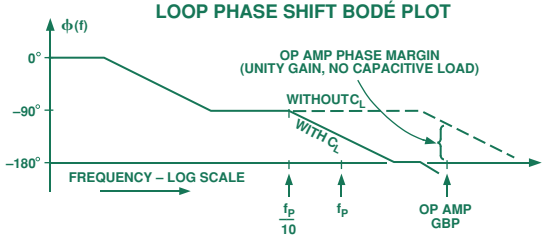
Q. *How does capacitive loading affect op amp performance?*

A. To put it simply, it can turn your amplifier into an oscillator. Here's how:

Op amps have an inherent output resistance, R_o, which, in conjunction with a capacitive load, forms an additional pole in the amplifier's transfer function. As the Bode plot shows, at each pole the amplitude slope becomes more negative by 20 dB/decade. Notice how each pole adds as much as –90° of phase shift. We can view instability from either of two perspectives. Looking at amplitude response on the log plot, circuit instability occurs when the sum of open-loop gain and feedback attenuation is greater than unity. Similarly, looking at phase response, an op amp will tend to oscillate at a frequency where loop phase shift exceeds –180°, if this frequency is below the closed-loop bandwidth. The closed-loop bandwidth of a voltage-feedback op amp circuit is equal to the op amp's gain-bandwidth product (GBP, or unity-gain frequency), divided by the circuit's closed loop gain (A_{CL}).







Phase margin of an op amp circuit can be thought of as the amount of additional phase shift at the closed loop bandwidth required to make the circuit unstable (i.e., phase shift + phase margin = –180°). As phase margin approaches zero, the loop phase shift approaches –180° and the op amp circuit approaches instability. Typically, values of phase margin much less than 45° can cause problems such as "peaking" in frequency response, and overshoot or "ringing" in step response. In order to maintain conservative phase margin, the pole generated by capacitive loading should be at least a decade above the circuit's closed loop bandwidth. When it is not, consider the possibility of instability.

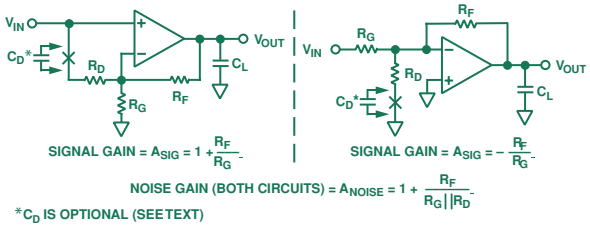
Q. *So how do I deal with a capacitive load?*

A. First of all you should determine whether the op amp can safely drive the load on its own. Many op amp data sheets specify a "capacitive load drive capability". Others provide typical data on "small-signal overshoot vs. capacitive load". In looking at these figures, you'll see that the overshoot increases exponentially with added load capacitance. As it approaches 100%, the op amp approaches instability. If possible, keep it well away from this limit. Also notice that this graph is for a specified gain. For a voltage feedback op amp, capacitive load drive capability increases proportionally with gain. So a VF op amp that can safely drive a 100-pF capacitance at unity gain should be able to drive a 1000-pF capacitance at a gain of 10.

A few op amp data sheets specify the open loop output resistance (R_o), from which you can calculate the frequency of the added pole as described above. The circuit will be stable if the frequency of the added pole (f_P) is more than a decade above the circuit's bandwidth.

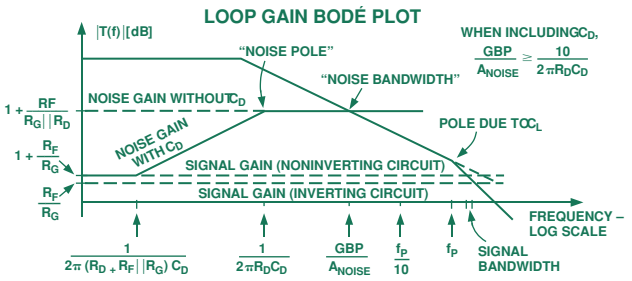
If the op amp's data sheet doesn't specify capacitive load drive or open loop output resistance, and has no graph of overshoot versus capacitive load, then to assure stability you must assume that any load capacitance will require some sort of compensation technique. There are many approaches to stabilizing standard op amp circuits to drive capacitive loads. Here are a few:

Noise-gain manipulation: A powerful way to maintain stability in low-frequency applications—often overlooked by designers—involves increasing the circuit's closed-loop gain (a/k/a "noise gain") without changing signal gain, thus reducing the frequency at which the product of open-loop gain and feedback attenuation goes to unity. Some circuits to achieve this, by connecting R_D between the op amp inputs, are shown below. The "noise gain" of these circuits can be arrived at by the given equation.



*His photo and a brief biography appear in Analog Dialogue 31-2, page 2.

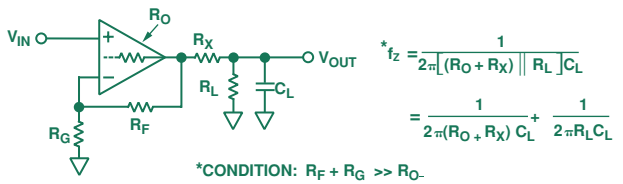
Since stability is governed by noise gain rather than by signal gain, the above circuits allow increased stability without affecting signal gain. Simply keep the "noise bandwidth" (GBP/A_{NOISE}) at least a decade below the load generated pole to guarantee stability.

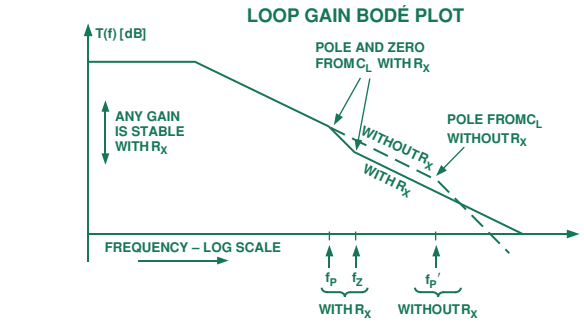


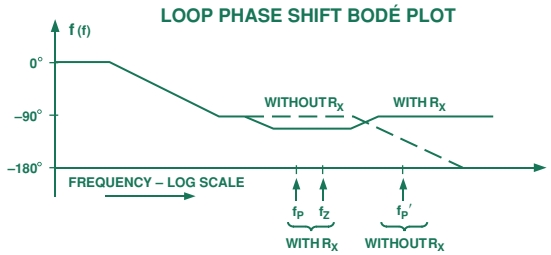
One disadvantage of this method of stabilization is the additional output noise and offset voltage caused by increased amplification of input-referred voltage noise and input offset voltage. The added dc offset can be eliminated by including C_D in series with R_D, but the added noise is inherent with this technique. The effective noise gain of these circuits with and without C_D are shown in the figure.

C_D, when used, should be as large as feasible; its minimum value should be 10 $A_{NOISE}/(2\pi R_D GBP)$ to keep the "noise pole" at least a decade below the "noise bandwidth".

Out-of-loop compensation: Another way to stabilize an op amp for capacitive load drive is by adding a resistor, R_X, between the op amp's output terminal and the load capacitance, as shown below. Though apparently outside the feedback loop, it acts with the load capacitor to introduce a zero into the transfer function of the feedback network, thereby reducing the loop phase shift at high frequencies.



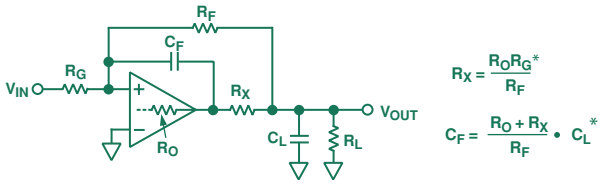




To ensure stability, the value of R_X should be such that the added zero (f_Z) is at least a decade below the closed loop bandwidth of the op amp circuit. With the addition of R_X, circuit performance will not suffer the increased output noise of the first method, but the output impedance as seen by the load will increase. This can decrease signal gain, due to the resistor divider formed by R_X and R_L. If R_L is known and reasonably constant, the results of gain loss can be offset by increasing the gain of the op amp circuit.

This method is very effective in driving transmission lines. The values of R_L and R_X must equal the characteristic impedance of the cable (often 50 Ω or 75 Ω) in order to avoid standing waves. So R_X is pre-determined, and all that remains is to double the gain of the amplifier in order to offset the signal loss from the resistor divider. Problem solved.

In-loop compensation: If R_L is either unknown or dynamic, the effective output resistance of the gain stage must be kept low. In this circumstance, it may be useful to connect R_X inside the overall feedback loop, as shown below. With this configuration, dc and low-frequency feedback comes from the load itself, allowing the signal gain from input to load to remain unaffected by the voltage divider, R_X and R_L.



The added capacitor, C_F, in this circuit allows cancellation of the pole and zero contributed by C_L. To put it simply, the zero from C_F is coincident with the pole from C_L, and the pole from C_F with the zero from C_L. Therefore, the overall transfer function and phase response are exactly as if there were no capacitance at all. In order to assure cancellation of both pole/zero combinations, the above equations must be solved accurately. Also note the conditions; they are easily met if the load resistance is relatively large.

Calculation is difficult when R_O is unknown. In this case, the design procedure turns into a guessing game—and a prototyping nightmare. A word of caution about SPICE: SPICE models of op amps don't accurately model open-loop output resistance (R_O); so they cannot fully replace empirical design of the compensation network.

It is also important to note that C_L must be of a known (and constant) value in order for this technique to be applicable. In many applications, the amplifier is driving a load "outside the box," and C_L can vary significantly from one load to the next. It is best to use the above circuit only when C_L is part of a closed system.

One such application involves the buffering or inverting of a reference voltage, driving a large decoupling capacitor. Here, C_L is a fixed value, allowing accurate cancellation of pole/zero combinations. The low dc output impedance and low noise of this method (compared to the previous two) can be very beneficial. Furthermore, the large amount of capacitance likely to decouple a reference voltage (often many microfarads) is impractical to compensate by any other method.

All three of the above compensation techniques have advantages and disadvantages. You should know enough by now to decide which is best for your application. All three are intended to be applied to "standard", unity gain stable, voltage feedback op amps. Read on to find out about some techniques using special purpose amplifiers.

Q. *My op amp has a "compensation" pin. Can I overcompensate the op amp so that it will remain stable when driving a capacitive load?*

A. Yes. This is the easiest way of all to compensate for load capacitance. Most op amps today are internally compensated for unity-gain stability and therefore do not offer the option to "overcompensate". But many devices still exist with inherent stability only at very high noise gains. These op amps have a pin to which an external capacitor can be connected in order to reduce the frequency of the dominant pole. To operate stably at lower gains, increased capacitance must be tied to this pin to reduce the gain-bandwidth product. When a capacitive load must be driven, a further increase (overcompensation) can increase stability—but at the expense of bandwidth.

Q. *So far you've only discussed voltage feedback op amps exclusively, right? Do current feedback (CF) op amps behave similarly with capacitive loading? Can I use any of the compensation techniques discussed here?*

A. Some characteristics of current feedback architectures require special attention when driving capacitive loads, but the overall effect on the circuit is the same. The added pole, in conjunction with op-amp output resistance, increases phase shift and reduces phase margin, potentially causing peaking, ringing, or even oscillation. However, since a CF op amp can't be said to have a "gain-bandwidth product" (bandwidth is much less dependent on gain), stability can't be substantially increased simply by increasing the noise gain. This makes the first method impractical. Also, a capacitor (C_F) should NEVER be put in the feedback loop of a CF op amp, nullifying the third method. The most direct way to compensate a current feedback op amp to drive a capacitive load is the addition of an "out of loop" series resistor at the amplifier output as in method 2.

Q. *This has been informative, but I'd rather not deal with any of these equations. Besides, my board is already laid out, and I don't want to scrap this production run. Are there any op amps that are inherently stable when driving capacitive loads?*

A. Yes. Analog Devices makes a handful of op amps that drive "unlimited" load capacitance while retaining excellent phase margin. They are listed in the table, along with some other op amps that can drive capacitive loads up to specified values. About the "unlimited" cap load drive devices: don't expect to get the same slew rate when driving 10 µF as you do when driving purely resistive loads. Read the data sheets for details.

Part Number	Ch	BW MHz	SR V/µs	v_n nV/$\sqrt{Hz}$	i_n fA/$\sqrt{Hz}$	V_{OS} mV	I_b nA	Supply Voltage Range [V]	I_Q mA	R_O Ω	Cap Load Drive [pF]	Notes
AD817	1	50	350	15	1500	0.5	3000	5-36	7	8	**unlim**	
AD826	2	50	350	15	1500	0.5	3000	5-36	6.8	8	**unlim**	
AD827	2	50	300	15	1500	0.5	3000	9-36	5.25	15	**unlim**	
AD847	1	50	300	15	1500	0.5	3000	9-36	4.8	15	**unlim**	
AD848	1	35	200	5	1500	0.5	3000	9-36	5.1	15	**unlim**	G_{MIN} = 5
AD849	1	29	200	3	1500	0.3	3000	9-36	5.1	15	**unlim**	G_{MIN} = 25
AD704	4	0.8	0.15	15	50	0.03	0.1	4-36	0.375		**10000**	
AD705	1	0.8	0.15	15	50	0.03	0.06	4-36	0.38		**10000**	
AD706	2	0.8	0.15	15	50	0.03	0.05	4-36	0.375		**10000**	
OP97	1	0.9	0.2	14	20	0.03	0.03	4-40	0.38		**10000**	
OP279	2	5	3	22	1000	4	300	4.5-12	2	22	**10000**	
OP400	4	0.5	0.15	11	600	0.08	0.75	6-40	0.6		**10000**	
AD549	1	1	3	35	0.22	0.5	0.00015	10-36	0.6		**4000**	
OP200	2	0.5	0.15	11	400	0.08	0.1	6-40	0.57		**2000**	
OP467	4	28	170	6	8000	0.2	150	9-36	2		**1600**	
AD744	1	13	75	16	10	0.3	0.03	9-36	3.5		**1000**	comp.term
AD8013	3	140	1000	3.5	12000	2	3000	4.5-13	3.4		**1000**	current fb
AD8532	2	3	5	30	50	25	0.005	3-6	1.4		**1000**	
AD8534	4	3	5	30	50	25	0.005	3-6	1.4		**1000**	
OP27	1	8	2.8	3.2	1700	0.03	15	8-44	6.7	70	**1000**	
OP37	1	12	17	3.2	1700	0.03	15	8-44	6.7	70	**1000**	G_{MIN} = 5
OP270	2	5	2.4	3.2	1100	0.05	15	9-36	2		**1000**	
OP470	4	6	2	3.2	1700	0.4	25	9-36	2.25		**1000**	
OP275	2	9	22	6	1500	1	100	9-44	2		**1000**	
OP184	1	4.25	4	3.9	400	0.18	80	4-36	2		**1000**	
OP284	2	4.25	4	3.9	400	0.18	80	4-36	2		**1000**	
OP484	4	4.25	4	3.9	400	0.25	80	4-36	2		**1000**	
OP193	1	0.04	15	65	50	0.15	20	3-36	0.03		**1000**	
OP293	2	0.04	15	65	50	0.25	20	3-36	0.03		**1000**	
OP493	4	0.04	15	65	50	0.28	20	3-36	0.03		**1000**	
OP297	2	0.5	0.15	17	20	0.08	0.05	4-40	0.525		**1000**	
OP497	4	0.5	0.15	25	20	0.08	0.06	4-40	0.525		**1000**	

REFERENCES

Practical Analog Design Techniques, Analog Devices 1995 seminar notes. Cap load drive information can be found in section 2, "High-speed op amps" (Walt Jung and Walt Kester). **Available on our Web site: www.analog.com or see the book purchase card**

Application Note AN-257: "Careful design tames high-speed op amps," by Joe Buxton, in ADI's *Applications Reference Manual* (1993). A detailed examination of the "in-loop compensation" method. Free. **Circle 8**

"Current-feedback amplifiers," Part 1 and Part 2", by Erik Barnes, *Analog Dialogue* 30-3 and 30-4 (1996), now consolidated in *Ask The Applications Engineer* (1997). **Available on our Web site.**

Ask The Applications Engineer—26

by Mary McCarthy & Anthony Collins*

SWITCHES AND MULTIPLEXERS

Q. *Analog Devices doesn't specify the bandwidth of its ADG series switches and multiplexers. Is there a reason?*

A. The ADG series switches and multiplexers have very high input bandwidths, in the hundreds of megahertz. However, the bandwidth specification by itself is not very meaningful, because at these high frequencies, the off-isolation and crosstalk will be significantly degraded. For example, at 1 MHz, a switch typically has off-isolation of 70 dB and crosstalk of –85 dB. Both off-isolation and crosstalk degrade by 20 dB per decade. This means that at 10 MHz, the off-isolation is reduced to 50 dB and the crosstalk increases to –65 dB. At 100 MHz, the off-isolation will be down to 30 dB while the crosstalk will have increased to –45 dB. So it is not sufficient to consider bandwidth alone—the off-isolation and crosstalk must be considered to determine if the application can tolerate the degradation of these specifications at the required high frequency.

Q. *Which switches and multiplexers can be operated with power supplies less than those specified in the data sheet?*

A. All of the ADG series switches and multiplexers operate with power supplies down to +5 V or ±5 V. The specifications affected by power-supply voltage are timing, on resistance, supply current and leakage current. Lowering power supply voltage reduces power supply current and leakage current. For example, the ADG411's $I_{S(OFF)}$ and $I_{D(OFF)}$ are ±20 nA, and $I_{D(ON)}$ is ±40 nA, at +125°C with a ±15-V power supply. When the supply voltage is reduced to ±5 V, $I_{S(OFF)}$ and $I_{D(OFF)}$ drop to ±2.5 nA, while $I_{D(ON)}$ is reduced to ±5 nA at +125°C. The supply currents, I_{DD}, I_{SS} and I_L, are 5 µA maximum at +125°C with a ±15-V power supply. When a ±5-V power supply is used, the supply currents are reduced to 1 µA maximum. The on-resistance and timing increase as the power supply is reduced. The Figures below show how the timing and on-resistance of the ADG408 vary as a function of power supply voltage.

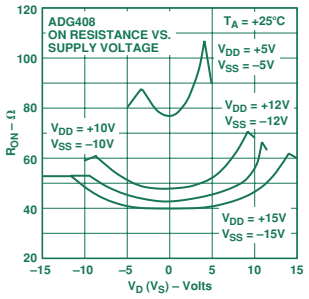


Figure 1. On-Resistance vs Power Supply.

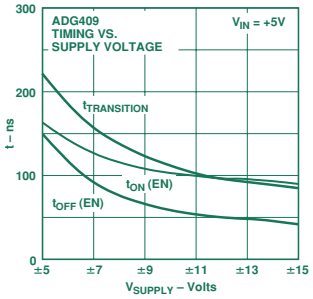


Figure 2. Timing vs Power Supply.

Q. *Some of the ADG series switches are fabricated on the DI process. What is it?*

A. DI is short for dielectric isolation. On the DI process, an insulating layer (trench) is placed between the NMOS and PMOS transistors of each CMOS switch. Parasitic junctions, which occur between the transistors in standard switches, are eliminated, resulting in a completely latchup-proof switch. In junction isolation (no trench used), the N and P wells of the PMOS and NMOS transistors form a diode which is reverse-biased in normal operation. However, during overvoltage or power-off conditions, when the analog input exceeds the power supplies, the diode is forward biased, forming a silicon controlled rectifier (SCR)-like circuit with the two transistors, causing the current to be amplified significantly, leading eventually to latch up. This diode doesn't exist in dielectrically isolated switches, making the part latchup proof.

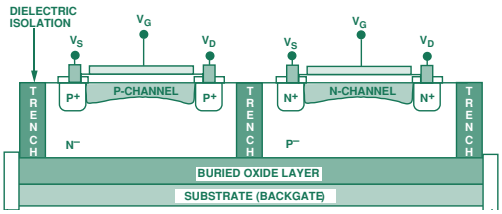


Figure 3. Dielectric Isolation.

Q. *How do the fault-protected multiplexers and channel protectors work?*

A. A channel of a fault-protected multiplexer or channel protector consists of two NMOS and two PMOS transistors. One of the PMOS transistors does not lie in the direct signal path but, is used to connect the source of the second PMOS to its backgate. This has the effect of lowering the threshold voltage, which increases the input signal range for normal operation. The source and backgate of the NMOS devices are connected for the same reason. During normal operation, the fault-protected parts operate as a standard multiplexer. When a fault condition occurs on the input to a channel, this means that the input has exceeded some threshold voltage which is set by the supply rail voltages. The threshold voltages are related to the supply rails as follows: for a positive overvoltage, the threshold voltage is given by $V_{DD} - V_{TN}$ where V_{TN} is the threshold voltage of the NMOS transistor (typically 1.5 V). For a negative overvoltage, the threshold voltage is given by $V_{SS} - V_{TP}$, where V_{TP} is the threshold voltage of the PMOS device (typically 2 V). When the input voltage exceeds these threshold voltages, with no load on the channel, the output of the channel is clamped at the threshold voltage.

Q. *How do the parts operate when an overvoltage exists?*

A. The next two figures show the operating conditions of the signal path transistors during overvoltage conditions. This one demonstrates how the series N, P and N transistors operate when a *positive* overvoltage is applied to the channel. The first NMOS transistor goes into saturation mode as the voltage on its drain exceeds $(V_{DD} - V_{TN})$. The potential at the source of the NMOS device is equal to $(V_{DD} - V_{TN})$. The other MOS devices are in a non-saturated mode of operation.

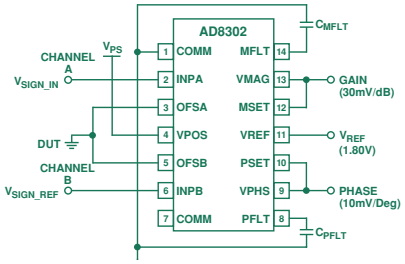


Figure 4. Positive Overvoltage on the Channel.

*Their biographies, and a photo of Mary McCarthy, appear in Analog Dialogue 31-3, page 23.

When a *negative* overvoltage is applied to a channel, the PMOS transistor enters a saturated mode of operation as the drain voltage exceeds ($V_{SS} - V_{TP}$). As with a positive overvoltage, the other MOS devices are non-saturated.

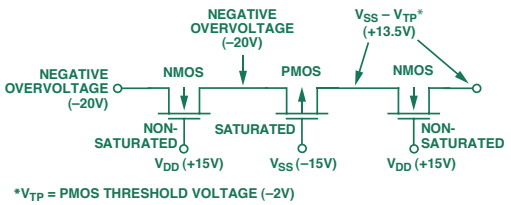


Figure 5. Negative Overvoltage on the Channel.

Q. *How does loading affect the clamping voltage?*

A. When the channel is loaded, the channel output will clamp at a value of voltage between the thresholds. For example, with a load of 1 kΩ, V_{DD} = 15 V, and a positive overvoltage, the output will clamp at $V_{DD} - V_{TN} - \Delta V$, where ΔV is due to the IR voltage drop across the channels of the non-saturated MOS devices. In the example shown below the voltage at the output of the clamped NMOS is 13.5 V. The on-resistance of the two remaining MOS devices is typically 100 Ω. Therefore, the current is 13.5 V/(1 kΩ + 100 Ω) = 12.27 mA. This produces a voltage drop of 1.2 V across the NMOS and PMOS resulting in a clamp voltage of 12.3 V. The current during a fault condition is determined by the load on the output, i.e., V_{CLAMP}/R_L.

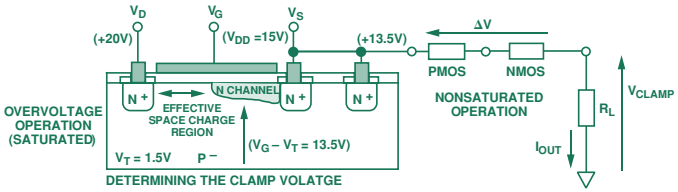


Figure 6. Determining the clamping point.

Q. *Do the fault-protected multiplexers and channel protectors function when the power supply is absent.*

A. Yes. These devices remain functional when the supply rails are down or momentarily disconnected. When V_{DD} and V_{SS} equal 0 V, all the transistors are off, as shown, and the current is limited to sub nanoampere levels.

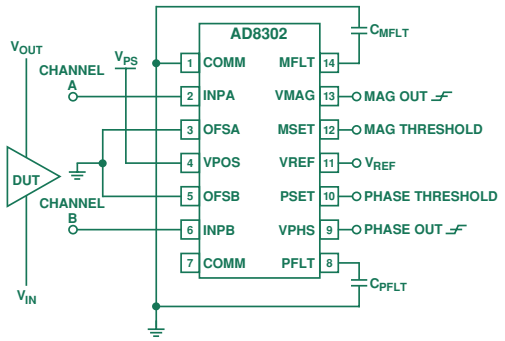


Figure 7. Power Supplies Absent.

Q. *What is "charge injection"?*

A. Charge injection in analog switches and multiplexers is a level change caused by stray capacitance associated with the NMOS and PMOS transistors that make up the analog switch. The Figure below models the structure of an analog switch and the stray capacitance associated with such an implementation. The structure basically consists of an NMOS and PMOS device in parallel. This arrangement produces the familiar "bathtub" resistance profile for bipolar input signals. The equivalent circuit shows the main parasitic capacitances that contribute to the charge injection effect, C_{GDN} (NMOS gate to drain) and C_{GDP} (PMOS gate to drain). The gate-drain capacitance associated with the PMOS device is about twice that of the NMOS device, because for both devices to have the same *on*-resistance, the PMOS device has about twice the area of the NMOS. Hence the associated stray capacitance is approximately twice that of the NMOS device for typical switches found in the marketplace.

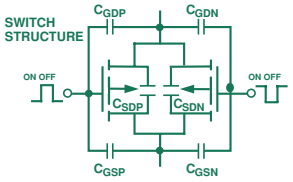


Figure 8

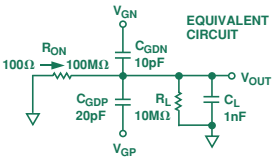


Figure 9

Figure 8. CMOS Switch Structure showing parasitic capacitance.
Figure 9. Equivalent circuit showing the main parasitics which contribute to charge injection.

When the switch is turned on, a positive voltage is applied to the gate of the NMOS and a negative voltage is applied to the gate of the PMOS. Because the stray gate-to-drain capacitances are mismatched, unequal amounts of positive and negative charge are injected onto the drain. The result is a removal of charge from the output of the switch, manifested as a negative-going voltage spike. Because the analog switch is now turned on this negative charge is quickly discharged through the on resistance of the switch (100 Ω). This can be seen in the simulation plot at 5 μs. Then when the switch is turned off, a negative voltage is applied to the gate of the NMOS and a positive voltage is applied to the gate of the PMOS. The result is charge added to the output of the switch. Because the analog switch is now off, the discharge path for this injected positive charge is a high impedance (100 MΩ). The result is that the load capacitance stores this charge until the switch is turned on again. The simulation plot clearly shows this with the voltage on C_L (as a result of charge injection) remaining constant at 170 mV until the switch is again turned on at 25 μs. At this point an equivalent amount of negative charge is injected onto the output, reducing the voltage on C_L to 0 V. At 35 μs the switch is turned on again and the process continues in this cyclic fashion.

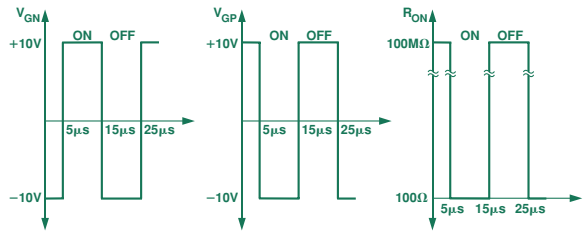


Figure 10. Timing used for simulation in Figure 11.

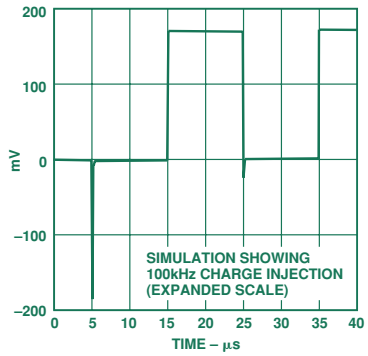


Figure 11. Output of simulation to show the effect of charge injection switching at 100 kHz.

At lower switching frequencies and load resistance, the switch output would contain both positive and negative glitches as the injected charge leaks away before the next switch transition.

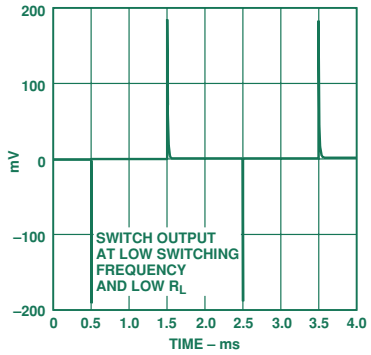


Figure 12. Switch output at low switching frequencies and low resistive loads.

Q. *What can be done to improve the charge injection performance of an analog switch?*

A. As noted above, the charge injection effect is caused by a mismatch in the parasitic gate-to-drain capacitance of the NMOS and PMOS devices. So if these parasitics can be matched there will be little if any charge injection effect. This is precisely what is done in Analog Devices CMOS switches and multiplexers. The matching is accomplished by introducing a dummy capacitor between the gate and drain of the NMOS device.

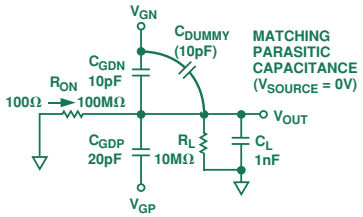


Figure 13. Matching parasitics at V_{SOURCE} = 0 V (ground).

Unfortunately the matching is only accomplished under a specific set of conditions, i.e., when the voltage on the Source of both devices is 0 V. The reason for this is that the parasitic capacitances, C_{GDN} and C_{GDP}, are not constant; they vary with the Source voltage. When the Source voltage of the NMOS and PMOS is varied, their channel depths vary, and with them, C_{GDN} and C_{GDP}. As a consequence of this matching at V_{SOURCE} = 0 V the charge injection effect will be noticeable for other values of V_{SOURCE}.

NOTE: Charge injection is usually specified on the data sheet under these matched conditions, i.e., V_{SOURCE} = 0 V. Under these conditions, the charge injection of most switches is usually quite good in the order of 2 to 3 pC max. However the charge injection will increase for other values of V_{SOURCE}, to an extent depending on the individual switch. Many data sheets will show a graph of charge injection as a function of Source voltage.

Q. *How do I minimize these effects in my application?*

A. The effect of charge injection is a voltage glitch on the output of the switch due to the injection of a fixed amount of charge. The glitch amplitude is a function of the load capacitance on the switch output and also the turn on and turn off times of the switch. The larger the load capacitance, the smaller will be the voltage glitch on the output, i.e., $Q = C \times V$, or $V = Q/C$, and Q is fixed. Naturally, it may not always be possible to increase the load capacitance, because it would reduce the bandwidth of the channel. However, for audio applications, increasing the load capacitance is an effective means of reducing those unwanted "pops" and "clicks".

Choosing a switch with a slow turn on and turn off time is also an effective means of reducing the glitch amplitude on the switch output. The same fixed amount of charge is injected over a longer time period and hence has a longer time period in which to leak away. The result is a wider glitch but much reduced in amplitude. This technique is used quite effectively in some of the audio switch products, such as the SSM-2402/SSM-2412, where the turn on time is designed to be of the order of 10 ms.

Another point worth mentioning is that the charge injection performance is directly related to the *on*-resistance of the switch. In general the lower the R_{ON}, the poorer the charge injection performance. The reason for this is purely due to the associated geometry, because R_{ON} is decreased by increasing the area of the NMOS and PMOS devices, thus increasing C_{GDN} and C_{GDP}. So trading off R_{ON} for reduced charge injection may also be an option in many applications.

Q. *How can I evaluate the charge injection performance of an analog switch or multiplexer?*

A. The most efficient way to evaluate a switch's charge injection performance is to use a setup similar to the one shown below. By turning the switch on and off at a relatively high frequency (>10 kHz) and observing the switch output on an oscilloscope (using a high impedance probe), a trace similar to that shown in Figure 11 will be observed. The amount of charge injected into the load is given by $\Delta V_{OUT} \times C_L$. Where ΔV_{OUT} is the output pulse amplitude.

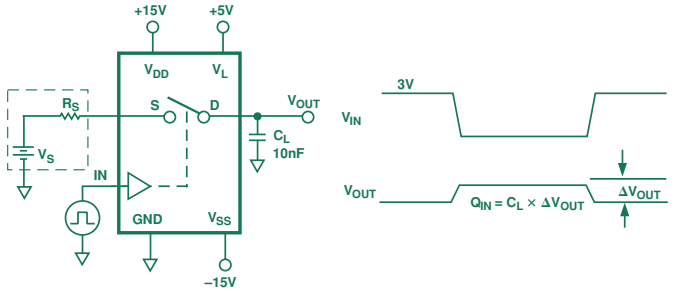


Figure 14. Evaluating the charge injection performance of an analog Switch or Multiplexer.

Ask the Applications Engineer—27

By Bill Englemann*

SIGNAL CORRUPTION IN INDUSTRIAL MEASUREMENT

Q. *What problems am I most likely to run into when instrumenting an industrial system?*

A. The five kinds of problems most frequently reported by customers of our *I/O Subsystems (IOS) Division* are:

1. GROUND LOOPS

Ground loops are the bane of instrumentation engineers and technicians. They cause many lost hours troubleshooting obscure and hard-to-diagnose measurement problems. Do these symptoms sound familiar?

- Readings slowly drift even though you know the sensor is not changing.
- Readings shift when another piece of equipment is turned on.
- Measurements differ when a calibration device is connected at the end of an instrument cable instead of directly at the input.
- A 60-Hz sine wave is superimposed upon your dc measurement input.
- There are unexplained measurement equipment failures.

Any of these problems can be caused by ground loops—inadvertent flows of current through "ground," "common" and "reference" paths connected to points at nominally the same potential. And all of these problems can be eliminated by *isolation*, the key signal-conditioning attribute we offer in all our signal conditioning series.

Sometimes separate grounding of two pieces of equipment introduces a potential difference and causes current to flow through signal lines. Why would this happen if they were both grounded? Because the earth and metal structures are actually relatively poor conductors of electricity when compared with the copper wires that carry power and signals. This inherent resistance to current flow varies with the weather and time of year and causes current to flow through any wires that are connecting the two devices. Many factory and plant buildings experience potentials of several tens or hundreds of volts. Appropriate signal conditioning eliminates the possibility of ground loops by electrically *isolating* the equipment. Signal conditioning will also protect equipment, rejecting potentially damaging voltage levels before entering the sensitive measurement system.

Isolation provides a completely floating input and output port, where there is no electrical path from field input to output and to power. Hence, there is no path for current to flow, and no possibility of ground loops.

Q. *How is this possible? How can we provide a path for the signal from input to output, without any path for current to flow?*

A. It's done by magnetic isolation. A representation of the signal is passed through a transformer, which creates a magnetic—not a galvanic—connection. We have perfected the use of transformers for accurate, reliable low-level signal isolation. This approach employs a modulator and demodulator to transmit the signal across the transformer barrier, and can achieve isolation levels of 2500 volts ac.

One of the most frequently encountered application problems involves measuring a low-level sensor such as a thermocouple in the presence of as much as hundreds of volts of ground potential. This potential is known as *common-mode voltage*. The ability of a high-quality signal conditioner to reject errors caused by common-mode voltage, while still accurately amplifying low-level signals is known as *common mode rejection (CMR)*. Our *5B*, *6B* and *7B* Series signal conditioning subsystems provide sufficient common-mode rejection to reduce the impact of these errors by a factor of 100 million to 1!

2. MISWIRING AND OVERVOLTAGE

You know what happens when a cable from a sensitive data acquisition board is routed into another cabinet, or another part of the building—the input and output wiring terminals are grouped among hundreds of other terminals carrying diverse signals and levels: dc signals, ac signals, milli-voltage, thermocouples, dc power, ac power, proximity switches, relay circuits, etc. It's not difficult to imagine even a well trained technician or electrician connecting a wire to the wrong terminal. Wiring diagrams are often updated in real time with a red pen, as system needs change. Equipment gets replaced with "equivalents." Sometimes power supplies fail and excess voltages are applied inadvertently. What can you do to protect your measurement system?

The answer lies in using rugged signal conditioning on every analog signal lead. This inexpensive insurance policy provides protection against miswiring and overvoltage on each input and output signal line. For example, the use of a *5B Series* signal conditioner will provide 240 -V ac of protection, even on input lines used to measure sensitive thermocouple signals, with levels in the millivolt range. You can literally connect a 240-V ac line across the same input lines used to measure the thermocouple, without any damage. The use of signal conditioning to interface with field I/O will protect all measurement and data acquisition equipment on the system side.

3. LOSS OF RESOLUTION

Resolution is the smallest change in the measurement that the analog-digital converter (ADC) system can detect and respond to. For example, if a temperature reading steps from 100.00° to 100.29° to 100.58°, as the actual temperature gradually increases through this range, the resolution (least-significant-bit value) is 0.29°. This would occur if you had a signal conditioner measuring a thermocouple with a range of 0° to +1200° and a 12-bit ADC. There are two ways to improve this (make the resolution smaller) and detect smaller changes - use a higher resolution ADC or use a smaller measurement range.

For example, a 15-bit plus sign ADC of the type used in our *6B Series* would offer resolution of 0.037° on the 0 to 1200° range example, 8 times smaller! On the other hand, if you knew that most of the time the temperature would be in the vicinity of 100°, you could order from Analog Devices a thermocouple signal conditioner with a custom range, calibrated for the exact thermocouple type and temperature measurement range. For example, a custom-ranged signal conditioner with a span of +50° to +150° would offer resolution of 0.024° with a 12-bit ADC, a big improvement over the 0 to 1200° range.

**His photo and a brief biography appear in Analog Dialogue 33, page 71.*

4. MULTIPLE SIGNALS DON'T ALL HAVE THE SAME PROPERTIES

This can pose quite a challenge to traditional industrial measurement approaches where 4, 8 or even 16 channels are dedicated to interfacing to the same signal type. For example, let's say you need to measure two J thermocouples, one 0 to +10 V signal, four 4-20 mA signals and two platinum RTDs (resistance temperature detector). You can either buy individual transmitters for each channel and then wire them all into a common 4-20 mA input board, or use a signal conditioning solution from Analog Devices that is configured channel-by-channel, but is also integrated into a simple backplane subsystem.

These subsystems incorporate all connections for input, output and field wiring, as well as simple connections for a dc power supply. They offer a choice of output options: 0 to +5 V, 0 to +10 V, 4-20 mA and RS-232/485, and more! Input and output modules are mix-and-match compatible on a per-channel basis and hot-swappable for the ultimate flexibility.

5. ELECTRICAL INTERFERENCE

Today's industrial factories and plants contain all kinds of interference sources: engines and motors, fluorescent lights, two-way radios, generators, etc. Each of these can radiate electro-magnetic noise that can be picked up by wiring, circuit boards and measurement modules. Even with the best shielding and grounding practices, this interference can show up as noise on the signal measurement. How can this be eliminated? By providing high noise rejection in the signal conditioning subsystem.

Lower-frequency noise can be eliminated by choosing signal-conditioning subsystems with excellent common mode and normal mode rejection. Common mode noise present on both the plus and minus inputs can be seen when measuring either the plus or minus input with respect to a common point like ground. Normal-mode noise is measured in the difference between the plus and minus inputs. A typical common mode rejection specification on our signal conditioning subsystems is 160 dB. This log scale measurement means that the effect of any common mode voltage noise is reduced relative to signal by a factor of 10^8, or 100 million to 1!

Very high frequency noise in the radio frequency bands can cause dc offsets due to rectification. It requires other approaches, including careful circuit layout and the use of RFI filters such as ferrite beads. The performance measures are indicated by our compliance with the EN certifications for electromagnetic susceptibility popularized by the CE mark requirements of the European community. A typical application where this is important would be where a two-way radio is used within a few feet of the input wiring and signal conditioning subsystem. It is necessary to reject measurement errors whenever the radios are transmitting. Good panel layout practice and the use of signal conditioning will ensure the best accuracy in these noisy environments.

CONCLUSION

Q. *What are some good installation and wiring practices?*

A. Here are a few suggestions. You may also want to take a look at *"Design Tools"* and the Analog Devices book, *Practical Analog Design Techniques*, available for sale in hard copy and free on the Web.

- Avoid installing sensitive measuring equipment, or wire carrying low level signals, near sources of electrical and magnetic noise, such as breakers, transformers, motors, SCR drives, welders, fluorescent lamp controllers, or relays.
- Use twisted pair wiring to reduce magnetic noise pickup. Look for 10 to 12 twists per foot.
- Use shielded cable with the shield connected to circuit common at the input end only.
- Never run signal-carrying wires in the same conduit that carries power lines, relay contact leads or other high-level voltages or currents.
- In extremely high interference environments, mount signal conditioning and measurement equipment inside grounded and closed metal cabinets.

Ask the Applications Engineer—28

By Eamon Nash*

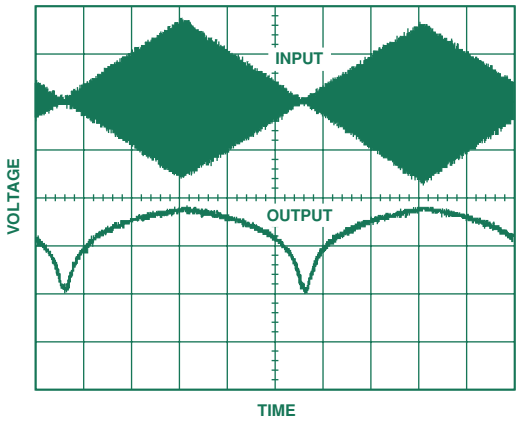
LOGARITHMIC AMPLIFIERS EXPLAINED

Q. *I've just been reading data sheets of some recently released Analog Devices log amps and I'm still a little confused about what exactly a log amp does.*

A. You're not alone. Over the years, I have had to deal with lots of inquiries about the changing emphasis on functions that log amps perform and radically different design concepts. Let me start by asking you, what do you expect to see at the output of a log amp?

Q. *Well, I suppose that I would expect to see an output proportional to the logarithm of the input voltage or current, as you describe in the* Nonlinear Circuits Handbook| <*http://www.analog.com/publications/magazines/Dialogue/Anniversary/books.html*> *and the* Linear Design Seminar Notes| <*http://www.analog.com/publications/press/misc/press_123094.html*>.

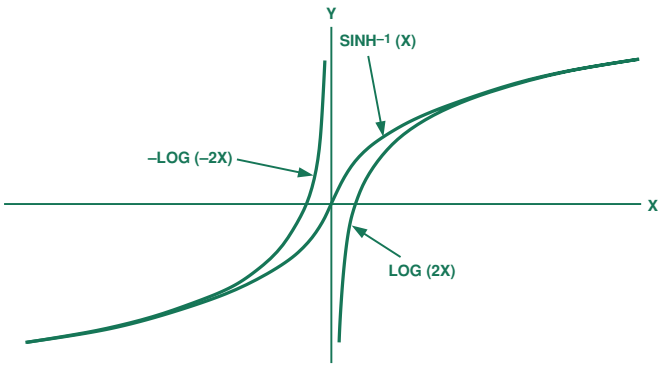
A. Well, that's a good start but we need to be more specific. The term log amp, as it is generally understood in communications technology, refers to a device which calculates the log of an input signal's envelope. What does that mean in practice? Take a look a the scope photo below. This shows a 10-MHz sine wave modulated by a 100-kHz triangular wave and the gross logarithmic response of the AD8307, a 500-MHz 90-dB log amp. Note that the input signal on the scope photo consists of many cycles of the 10-MHz signal, compressed together, using the time/div knob of the oscilloscope. We do this to show the envelope of the signal, with its much slower repetition frequency of 100 kHz. As the envelope of the signal increases linearly, we can see the characteristic "log (x)" form in the output response of the log amp. In contrast, if our measurement device were a linear envelope detector (a filtered rectified output), the output would simply be a tri-wave.



Q. *So I don't see the log of the instantaneous signal?*

A. That's correct, and it's the source of much of the confusion. The log amp gives an indication of the instant-by-instant low-frequency changes in the envelope, or *amplitude*, of the signal in the log domain in the same way that a digital voltmeter, set to "ac volts," gives a steady (linear) reading when the input is connected to a constant amplitude sine wave and follows any adjustments to the amplitude. A device that calculates the instantaneous log of the input signal is quite different, especially for bipolar signals.

On that point, let's digress for a moment to consider such a device. Think about what would happen when an ac input signal crosses zero and goes negative. Remember, the mathematical function, log x, is undefined for x real and less than or equal to zero, or $-x$ greater than or equal to zero (see figure).



However, as the figure shows, the inverse hyperbolic sine, $\sinh^{-1} x$, which passes symmetrically through zero, is a good approximation to the combination of log $2x$ and minus log $(-2x)$, especially for large values of $|x|$. And yes, it is possible to build such a log amp; in fact, Analog Devices many years ago manufactured and sold Model 752 N & P temperature-compensated log diode modules, which—in complementary feedback pairs—performed that function. Such devices, which calculate the instantaneous log of the input signal are called *baseband log amps* (the term "true log amp" is also used). The focus of this discussion, however, is on envelope-detecting log amps, also referred to as demodulating log amps, which have interesting applications in RF and IF circuitry for communications.

Q. *But, from what you have just said, I would imagine that a log amp is generally not used to demodulate signals?*

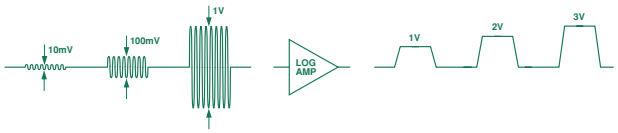
A. Yes, that is correct. The term demodulating came to be applied to this type of device because a log amp recovers the log of the envelope of a signal in a process somewhat like that of demodulating AM signals.

In general, the principal application of log amps is to measure signal *strength*, as opposed to detecting signal *content*. The log amp's output signal, which can represent a many-decade dynamic range of high-frequency input signal amplitudes by a relatively narrow range, is typically used to regulate gain. The classic example of this is using a log amp in an automatic gain control loop, to regulate the gain of a variable-gain amplifier. The receiver of a cellular base station, for example, might use the signal from a log amp to regulate the receiver gain. In transmitters, log amps are also used to measure and regulate transmitted power.

However, there are some applications where a log amp is used to demodulate a signal. The figure shows a received signal that has been modulated using *amplitude shift keying* (ASK). This simple modulation scheme, similar to early transmissions of radar pulses, conveys digital information by transmitting a series of RF bursts (logic 1 = burst, logic 0 = no burst). When this

*His photo and a brief biography appear in Analog Dialogue 33, page 71.

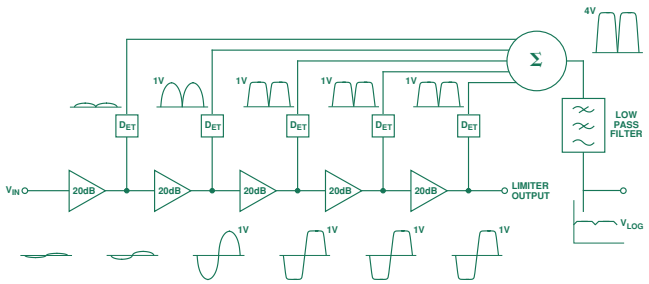
signal is applied to a log amp, the output is a pulse train which can be applied to a comparator to give a digital output. Notice that the actual amplitude of the burst is of little importance; we only want to detect its presence or absence. Indeed, it is the log amp's ability to convert a signal which varies over a large dynamic range (10 mV to 1 V in this case) into one that varies over a much smaller range (1 V to 3 V) that makes the use of a log amp so appealing in this application.



Q. Can you explain briefly how a log amp works?

A. The figure shows a simplified block diagram of a log amp. The core of the device is a cascaded chain of amplifiers. These amplifiers have linear gain, usually somewhere between 10 dB and 20 dB. For simplicity of explanation, in this example, we have chosen a chain of 5 amplifiers, each with a gain of 20 dB, or 10¥. Now imagine a small sine wave being fed into the first amplifier in the chain. The first amplifier will amplify the signal by a factor of 10 before it is applied to the second amplifier. So as the signal passes through each subsequent stage, it is amplified by an additional 20 dB.

Now, as the signal progresses down the gain chain, it will at some stage get so big that it will begin to clip (the term *limit* is also used) as shown. In the simplified example, this clipping level (a desired effect) has been set at 1 V peak. The amplifiers in the gain chain would be designed to limit at this same precise level.



After the signal has gone into limiting in one of the stages (this happens at the output of the third stage in the figure), the limited signal continues down the signal chain, clipping at each stage and maintaining its 1 V peak amplitude as it goes.

The signal at the output of each amplifier is also fed into a full wave rectifier. The outputs of these rectifiers are summed together as shown and applied to a low-pass filter, which removes the ripple of the full-wave rectified signal. Note that the contributions of the earliest stages are so small as to be negligible. This yields an output (often referred to as the "video" output), which will be a steady-state quasi-logarithmic dc output for a steady-state ac input signal. The actual devices contain innovations in circuit design that shape the gain and limiting functions to produce smooth and accurate logarithmic behavior between the decade breaks, with the limiter output sum comparable to the *characteristic*, and the contribution of the less-than-limited terms to the *mantissa*.

To understand how this signal transformation yields the log of the input signal's envelope, consider what happens if the input signal is reduced by 20 dB. As it stands in the figure, the unfiltered output of the summer is about 4 V peak (from three stages that are limiting and a fourth that is just about to limit). If the input signal is reduced by a factor of 10, the output of one stage at the input end of the chain will become negligible, and there will be one less stage in limiting. Because of the voltage lost from this stage, the summed output will drop to approximately 3 V. If the input signal is reduced by a further 20 dB, the summed output will drop to about 2 V.

So the output is changing by 1 V for each factor-of-10 (20-dB) amplitude change at the input. We can describe the log amp then as having a slope of 50 mV/dB.

Q. O.K. I understand the logarithmic transformation. Now can you explain what the Intercept is?

A. The slope and intercept are the two specifications that define the transfer function of the log amp, that is, the relationship between output voltage and input signal level. The figure shows the transfer function at 900 MHz, and over temperature, of the AD8313, a 100-MHz-to-2.5-GHz 65-dB log amp. You can see that the output voltage changes by about 180 mV for a 10 dB change at the input. From this we can deduce that the slope of the transfer function is 18 mV/dB.

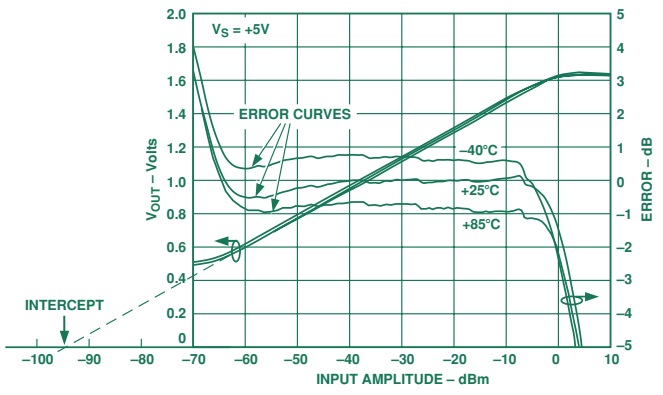
As the input signal drops down below about –65 dBm, the response begins to flatten out at the bottom of the device's range (at around 0.5 V, in this case). However, if the linear part of the transfer function is extrapolated until it crosses the horizontal axis (0 V theoretical output), it passes through a point called the *intercept* (at about –93 dBm in this case). Once the slope and intercept of a particular device are known (these will always be given in the data sheet), we can predict the nominal output voltage of the log amp for any input level within the linear range of the device (about –65 dBm to 0 dBm in this case) using the simple equation:

$$V_{OUT} = Slope \times (P_{IN} - Intercept)$$

For example, if the input signal is –40 dBm the output voltage will be equal to

$$18\ mV/dB \times (-40\ dBm - (-93\ dBm)) = 0.95\ V$$

It is worth noting that an increase in the intercept's value *decreases* the output voltage.



The figure also shows plots of deviations from the ideal, i.e., *log conformance*, at –40°C, +25°C, and +85°C. For example, at +25°C, the log conformance is to within at least ±1 dB for an input in the range –2 dBm to –67 dBm (over a smaller range, the log conformance is even better). For this reason, we call the AD8313 a 65-dB log amp. We could just as easily say that the AD8313 has a dynamic range of 73 dB for log conformance within 3 dB.

Q. *In doing some measurements, I've found that the output level at which the output voltage flattens out is higher than specified in the data sheet. This is costing me dynamic range at the low end. What is causing this?*

A. I come across this quite a bit. This is usually caused by the input picking up and measuring an external noise. Remember that our log amps can have an input bandwidth of as much as 2.5 GHz! The log amp does not know the difference between the wanted signal and the noise. This happens quite a lot in laboratory environments, where multiple signal sources may be present. Remember, in the case of a wide-range log amp, a –60-dBm noise signal, coming from your colleague who is testing his new cellular phone at the next lab bench, can wipe out the bottom 20-dB of your dynamic range.

A good test is to ground both differential inputs of the log amp. Because log amps are generally ac-coupled, you should do this by connecting the inputs to ground through coupling capacitors.

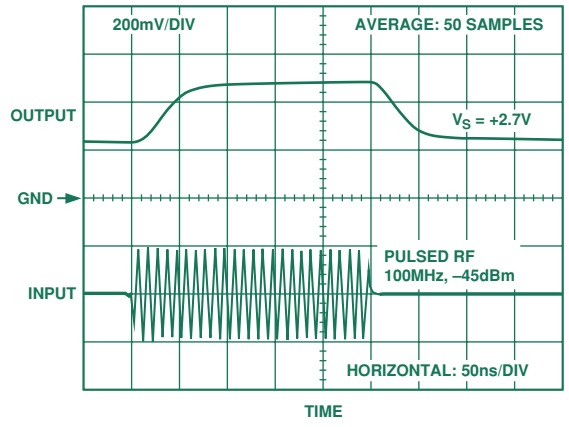
Solving the problem of noise pickup generally requires some kind of filtering. This is also achieved indirectly by using a matching network at the input. A narrow-band matching network will have a filter characteristic and will also provide some gain for the wanted signal. Matching networks are discussed in more detail in data sheets for the AD8307, AD8309, and AD8313.

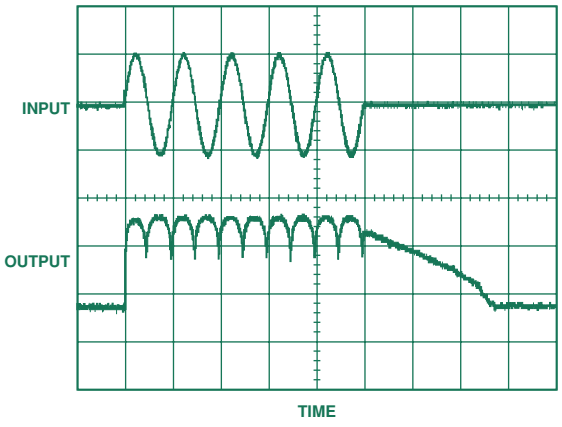
Q. *What corner frequency is typically chosen for the output stage's low-pass filter?*

A. There is a design trade-off here. The corner frequency of the on-chip low-pass filter must be set low enough to adequately remove the ripple of the full-wave rectified signal at the output of the summer. This ripple will be at a frequency 2 times the input signal frequency. However the RC time constant of the low-pass filter determines the maximum rise time of the output. Setting the corner frequency too low will result in the log amp having a sluggish response to a fast-changing input envelope.

The ability of a log amp to respond to fast changing signals is critical in applications where short RF bursts are being detected. In addition to the ASK example discussed earlier, another good example of this is RADAR. The figure on the left shows the response of the AD8313 to a short 100 MHz burst. In general, the log-amp's response time is characterized by the metric 10% to 90% rise time. The table below compares the rise times and other important specifications of different Analog Devices log amps.

Now take a look at the figure on the right. This shows you what will happen if the frequency of the input signal is lower than the corner frequency of the output filter. As might be expected, the full wave rectified signal appears unfiltered at the output. However this situation can easily be improved by adding additional low-pass filtering at the output.





Part Number	Input Bandwidth	10%–90% Rise Time	Dynamic Range	Log Conformance	Limiter Output
AD606	50 MHz	360 ns	80 dB	±1.5 dB	Yes
AD640	120 MHz	6 ns	50 dB	±1 dB	Yes
AD641	250 MHz	6 ns	44 dB	±2 dB	Yes
AD8306	500 MHz	67 ns	95 dB	±0.4 dB	Yes
AD8307	500 MHz	500 ns	92 dB	±1 dB	No
AD8309	500 MHz	67 ns	100 dB	±1 dB	Yes
AD8313	2500 MHz	45 ns	65 dB	±1 dB	No

Q. *I notice that there is an unusual tail on the output signal at the right. What is causing that?*

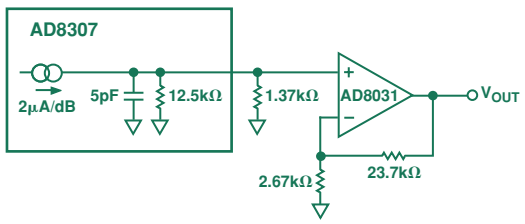
A. That is an interesting effect that results from the nature of the log transformation that is taking place. Looking again at transfer function plot (i.e. voltage out vs. input level), we can see that at low input levels, small changes in the input signal have a significant effect on the output voltage. For example a change in the input level from 7 mV to 700 μV (or about –30 dBm to –50 dBm) has the same effect as a change in input level from 70 mV to 7 mV. That is what is expected from a logarithmic amplifier. However, looking at the input signal (i.e., the RF burst) with the naked eye, we do not see small changes in the mV range. What's happening in the figure is that the burst does not turn off instantly but drops to some level and then decays exponentially to zero. And the log of a decaying exponential signal is a straight line similar to the tail in the plot.

Q. *Is there a way to speed up the rise time of the log amp's output?*

A. This is not possible if the internal low-pass filter is buffered, which is the case in most devices. However the figure shows one exception: the unbuffered output stage of the AD8307 is here represented by a current source of 2 μA/dB, which is looking at an internal load of 12.5 kΩ. The current source and the resistance combine to give a nominal slope of 25 mV/dB. The 5-pF capacitance in parallel with the 12.5-kΩ resistance combines to yield a low-pass corner frequency of 2.5 MHz. The associated 10%-90% rise time is about 500 ns.

In the figure, an external 1.37-kΩ shunt resistor has been added. Now, the overall load resistance is reduced to around 1.25 kΩ. This will *decrease* the rise time ten-fold. However the overall logarithmic slope has also decreased ten-fold. As a result, external gain is required to get back to a slope of 25 mV/dB.

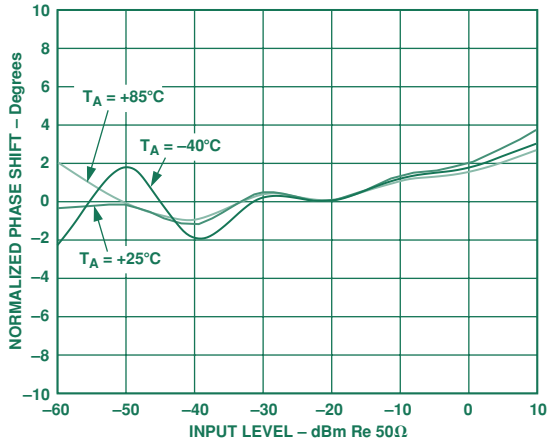
You may also want to take a look at the Application Note AN-405. This shows how to improve the response time of the AD606.



Q. *Returning to the architecture of a typical log amp, is the heavily clipped signal at the end of the gain chain in any way useful?*

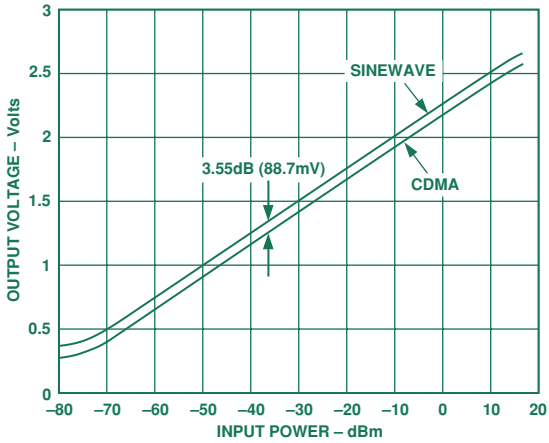
A. The signal at the end of the linear gain chain has the property that its amplitude is constant for all signal levels within the dynamic range of the log amp. This type of signal is very useful in phase- or frequency demodulation applications. Remember that in a phase-modulation scheme (e.g. QPSK or broadcast FM), there is no useful information contained in the signal's amplitude; all the information is contained in the phase. Indeed, amplitude variations in the signal can make the demodulation process quite a bit more difficult. So the signal at the output of the linear gain chain is often made available to give a limiter output. This signal can then be applied to a phase or frequency demodulator.

The degree to which the phase of the output signal changes as the input level changes is called *phase skew*. Remember, the phase between input and output is generally not important. It is more important to know that the phase from input to output stays constant as the input signal is swept over its dynamic range. The figure shows the phase skew of the AD8309's limiter output, measured at 100 MHz. As you can see, the phase varies by about 6° over the device's dynamic range and over temperature.



Q. *I noticed that something strange happens when I drive the log amp with a square wave.*

A. Log amps are generally specified for a sine wave input. The effect of differing signal waveforms is to shift the effective value of the log amp's intercept upwards or downwards. Graphically, this looks like a vertical shift in the log amp's transfer function (see figure), without affecting the logarithmic slope. The figure shows the transfer function of the AD8307 when alternately fed by an unmodulated sine wave and by a CDMA channel (9 channels on) of the same rms power. The output voltage will differ by the equivalent of 3.55 dB (88.7 mV) over the complete dynamic range of the device.



The table shows the correction factors that should be applied to measure the rms signal strength of various signal types with a logarithmic amplifier which has been characterized using a sine wave input. So, to measure the rms power of a squarewave, for example, the mV equivalent of the dB value given in the table (–3.01 dB, which corresponds to 75.25 mV in the case of the AD8307) should be subtracted from the output voltage of the log amp.

Signal Type	Correction Factor (Add to Output Reading)
Sine Wave	0 dB
Square Wave or DC	–3.01 dB
Triangular Wave	+0.9 dB
GSM Channel (All Time Slots On)	+0.55 dB
CDMA Forward Link (Nine Channels On)	+3.55 dB
Reverse CDMA Channel	0.5 dB
PDC Channel (All Time Slots On)	+0.58 dB
Gaussian Noise	+2.51 dB

Q. In your data sheets you sometimes give input levels in dBm and sometimes in dBV. Can you explain why?

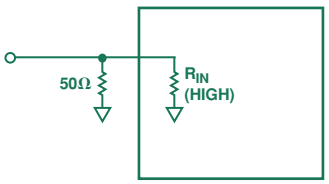
A. Signal levels in communications applications are usually specified in dBm. The dBm unit is defined as the power in dB relative to 1 mW i.e.,

$$Power\ (\text{dBm}) = 10 \log_{10} (Power/1\ \text{mW})$$

Since power in watts is equal to the rms voltage squared, divided by the load impedance, we can also write this as

$$Power\ (\text{dBm}) = 10 \log_{10} ((V_{rms}^2/R)/1\ \text{mW})$$

It follows that 0 dBm occurs at 1 mW, 10 dBm corresponds to 10 mW, +30 dBm corresponds to to 1 W, etc. Because impedance is a component of this equation, it is always necessary to specify load impedance when talking about dBm levels.



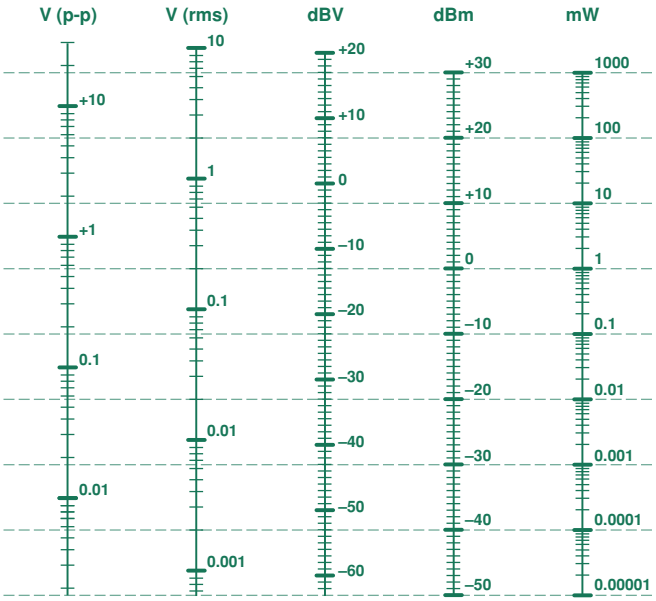
Log amps, however fundamentally respond to voltage, not to power. The input to a log amp is usually terminated with an external 50-Ω resistor to give an overall input impedance of approximately 50 Ω, as shown in the figure (the log amp has a relatively high input impedance, typically in the 300 Ω to 1000 Ω range). If the log amp is driven with a 200-Ω signal and the input is terminated in 200 Ω, the output voltage of the log amp will be higher compared to the same amount of power from a 50-Ω input signal. As a result, it is more useful to work with the *voltage* at the log amp's input. An appropriate unit, therefore, would be dBV, defined as the voltage level in dB relative to 1 V, i.e.,

$$Voltage\ (\text{dBV}) = 20 \log_{10} (V_{rms}/1\ \text{V})$$

However, there is disagreement in the industry as to whether the 1-V reference is 1 V peak (i.e., amplitude) or 1 V rms. Most lab instruments (e.g., signal generators, spectrum analyzers) use 1 V rms as their reference. Based upon this, dBV readings are converted to dBm by adding 13 dB. So –13 dBV is equal to 0 dBm.

As a practical matter, the industry will continue to talk about input levels to log amps in terms of dBm power levels, with the implicit assumption that it is based on a 50 Ω impedance, even if it is not completely correct to do so. As a result it is prudent to provide specifications in *both* dBm and dBV in data sheets.

The figure shows how mV, dBV, dBm and mW relate to each other for a load impedance of 50 Ω. If the load impedance were 20 Ω, for example, the V (rms), V (p-p) and dBV scales will be shifted downward relative to the dBm and mW scales. Also, the V (p-p) scale will shift relative to the V (rms) scale if the peak to rms ratio (also called crest factor) is something other than √2 (the peak to rms ratio of a sine wave).



Ask the Applications Engineer—29

By Harvey Weinberg*

ACCELEROMETERS—FANTASY and REALITY

Q. *As an Applications Engineer for the Micromachined Products Division supporting the compact, low-cost, gravity-sensitive Analog Devices line of integrated accelerometers[1] you must get to hear lots of creative ideas from prospective customers about how to employ accelerometers in useful ways.*

A. Yes, but sometimes what they are suggesting violates physical laws! We've come to rate them in various categories on an informal "reality" scale:

- ***Real*** – A real application that actually works today and is currently in production.
- ***Possible*** – An application that is technically feasible, but not (to our knowledge) in production.
- ***Fantasy*** – An application that could be possible if we had much better technology.
- ***Dream Land*** – An application that would be nice, but any practical implementation we can think of would violate some physical laws.

Would you like to see some of these ideas (the ones we're free to mention), hear about their potential benefits, think about them, and guess which category they belong in?

Q. *Sure, go ahead.*

A. O.K. You can find (and link to) the answers, with commentary, below.

Washing Machine Load Balancing
Unbalanced loads during the high-speed spin-cycle cause washing machines to shake, and in some cases, to even "walk" across the floor if unrestrained. An accelerometer senses the accelerations present during the unbalanced spin. If such an imbalance is present, the washing machine redistributes the load by jogging the drum back and forth a few times until the load is balanced. *Real or fantasy*? See answer A.

Machine Health Monitors
Many industries change or overhaul mechanical equipment using a calendar-based preventive maintenance schedule. This is especially true in applications where one cannot afford or tolerate unscheduled down-time. So machinery with plenty of service life left is often prematurely rebuilt at a cost of millions of dollars across many industries. By embedding accelerometers in bearings, or other rotating equipment, service life can be extended without additional risk of sudden failure.

The accelerometer senses the vibration of bearings or other rotating equipment to determine their condition. *Real or fantasy*? See answer B.

Automatic Leveling
Accelerometers measure the absolute inclination of an object (e.g. a large machine, a mobile home, etc.). The tilt information is used by a microcontroller to automatically level it. *Real or fantasy*? See answer C.

Scroll Control for PDAs and Handyphones
The accelerometer allows the microcontroller to recognize gestures allowing the user one handed control of mobile devices. *Real or fantasy*? See answer D.

Car Alarm
Here the accelerometer senses if a car is being jacked up or being picked up by a tow truck, and sets off the alarm. *Real or fantasy*? See answer E.

Ski Bindings
The accelerometer measures the total shock energy and signature to determine if the binding should release. *Real or fantasy*? See answer F.

Personal Navigation
In this application, position is determined by dead reckoning (double integration of acceleration over time to determine actual position). *Real or fantasy*? See answer G.

Subwoofer Servo Control
An accelerometer is mounted on the cone of the subwoofer to provide positional feedback used to servo out distortion. *Real or fantasy*? See answer H.

Neuromuscular Stimulator
This application helps people, who have lost control of their lower leg muscles, to walk—by stimulating muscles at the appropriate time. *Real or fantasy*? See answer I.

Car-Noise Cancellation
The accelerometer senses low-frequency vibration in the passenger compartment, and the noise-cancellation system nulls it out, using the speakers in the car stereo system. *Real or fantasy*? See answer J.

ANSWERS

A. *Washing-machine load balancing:* **Real**. This application is currently in production. With better load balance, faster spin rates can be used to wring more water out of clothing, making the drying process more efficient. Fewer mechanical components are required for damping the drum motion, making the overall system lighter, and less expensive. In addition, transmission and bearing service life may be extended because of lower peak loads present on the motor. See http://www.analog.com/industry/*i*MEMS/markets/industrial/washing_machine.html for more information about washing machines. *Go to next question.*

B. *Machine health monitoring:* **Possible.** This application has been demonstrated but is not yet in production.

Using the vibrational "signature" of bearings to determine their condition is a well proven and industry-accepted method of equipment maintenance, However, the cost of accelerometers and the associated signal conditioning equipment has been traditionally too high. The ADXL105[2] offers a complete vibration measurement and signal conditioning solution on a chip at very low cost. See http://www.analog.com/industry/*i*MEMS/markets/industrial/machine.html for more information about machine health. *Go to next question.*

[1]http://www.analog.com/industry/iMEMS.
[2]http://products.analog.com/products/info.asp?product=ADXL105.

**His photo and a brief biography appear in Analog Dialogue 33, page 72.*

C. *Automatic leveling:* **Real to Fantasy** (depending on the application). There are some applications where this is practical and currently in production. Others are too demanding for current products.

Self-leveling is a very demanding application, as absolute precision is required. Surface micromachined accelerometers have impressive resolution, but absolute tilt measurement with high accuracy (to better than 1%) requires temperature stability and hysteresis performance that today's surface-micromachined accelerometers cannot achieve. Applications needing absolute accuracy to within ±3° or more are currently possible and a few such applications are in production. *Go to next question.*

D. *Scroll control:* **Real**. This application is currently (or almost, depending on press time) in production.

A PDA (like the 3com Palm Pilot) is incredibly handy, but almost impossible to use one-handed. Like when you're driving, or on the phone. Adding an accelerometer lets the PDA accept gesture inputs, like tilting up or down, to control the cursor or page up/down control. See http://www.analog.com/industry/*i*MEMS/markets/consumer/peripherals.html if you are interested in game controllers or are a Palm Pilot user and want to see how to add a tilt function to your PDA. *Go to next question.*

E. *Car alarm:* **Real**. This application is currently in production in OEM and after-market automotive anti-theft systems.

One of the most popular methods of auto theft is where the car is stolen by simply towing it away. Conventional car alarms do not protect against this. Shock sensors cannot measure changes in inclination, and ignition-disabling systems are ineffectual. Here is an application where the high-resolution capabilities of the ADXL202 are used to advantage. The accelerometer measures if the car's inclination is changing by more than 0.5° per minute. If so, the alarm is sounded, hopefully scaring off the would-be thief. Absolute stability is not required here (unlike automatic leveling systems) as temperature does not change significantly in a minute or less. See http://www.analog.com/industry/*i*MEMS/markets/automotive/car_alarms.html if you're interested in car alarms. *Go to next question.*

F. *Ski bindings:* **Fantasy**. This is a practical accelerometer application, but current battery technology (particularly low temperature performance) makes it impractical.

All mechanical ski bindings are highly evolved, but limited in their performance. Measuring the actual shock experienced by the skier would be a much more accurate way to determine if a binding should release. Intelligent systems could be developed that could take each individual's capability and physiology into account. Smaller and lighter batteries that perform well at low temperatures will, eventually, enable this application. *Go to next questio*n.

G. *Personal navigation*: **Dream Land.** Long term integration results in the accumulation of error due to small dc errors in the accelerometer, integrator input circuitry, wiring thermocouples, etc. Double integration compounds the errors (t^2). Without some way of "resetting" the actual position from time to time, huge errors result. This is analogous to building an op-amp integrator by simply putting a capacitor across it. Even if the accelerometer's accuracy is improved by ten or one hundred times better than currently available, huge errors would still eventually result. It would just take longer to happen.

Accelerometers can be used in conjunction with a GPS system when the GPS signals are briefly unavailable. Integration over a short time (a minute or so) can give satisfactory results. See http://www.analog.com/industry/*i*MEMS/markets/consumer/car_nav.html for more information about navigation. *Go to next question.*

H. *Subwoofer servo:* **Real**. Several active subwoofers with servo control are on the market today.

Servo control of subwoofers has several advantages. Harmonic distortion, as well as power compression, can be greatly reduced. In addition servo control can also electronically lower the *Q* of the speaker/enclosure system, enabling the use of smaller enclosures[3]. The ADXL190[4] is small and light; its mass, added to that of the loudspeaker cone, does not change the overall acoustic characteristics significantly. See http://www.analog.com/industry/*i*MEMS/markets/consumer/subwoofers.html for more information about active subwoofer applications, and http://www.analog.com/industry/*i*MEMS/markets/consumer/subwoofers/Subwoof.html#cir for circuits. *Go to next question.*

I. *Neuromuscular stimulator:* **Real.** This application is very near (if not already in) production.

When walking, the forefoot is normally raised when moving the leg forward, then lowered when pushing the leg backward. The accelerometer is worn somewhere on the lower leg or foot, where it senses the position of the leg. The appropriate muscles are then electronically stimulated to flex the foot as required.

This is a classic example of how micromachined accelerometers have made a product feasible. Earlier models used a liquid tilt sensor or a moving ball bearing (acting as a switch) to determine the leg position. Liquid tilt sensors had problems because of sloshing of the liquid, so only slow walking was possible. Ball-bearing switches were easily confused when walking on hills. Using an accelerometer, the differential between leg back and leg forward is measured, so hills do not fool the system and there are no liquid slosh problems. The low power consumption of the accelerometer allows the system to work with a small lithium battery, making the overall package unobtrusive. *Go to next question.*

J. *Car-noise cancellation:* **Dream Land.** While the accelerometer has no trouble picking up the vibration in the passenger compartment, noise cancellation is highly phase-dependent. So while we may cancel the noise in one location (say around the head of the driver), it will probably be increased at other locations.

CONCLUSION

Because of their sensitivity, compactness, low cost, ruggedness, and ability to measure both static and dynamic acceleration forces, surface micromachined accelerometers have made numerous new applications possible. Many of them were not anticipated because they were not thought of as classic accelerometer applications. The imagination of designers now seems to be the limiting factor in the scope of potential applications—but sometimes designers can become too imaginative! While performance improvements continue to enable more applications, it's wise to try to stay away from "solutions" that violate laws of physics.

[3]See R. A. Greiner and T. M. Sims, Jr., Loudspeaker Distortion Reduction, *Journal of the Audio Engineering Society*, Vol. 32, No. 12.
[4]http://products.analog.com/products/info.asp?product=ADXL190.

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

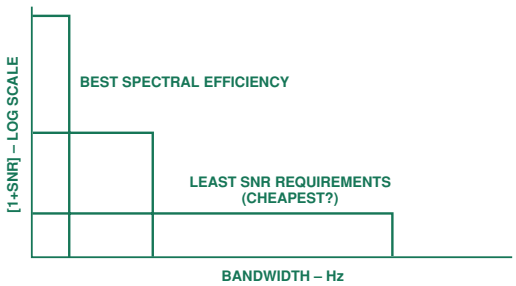


Figure 1. Shannon's capacity limit: equal theoretical capacity.

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.

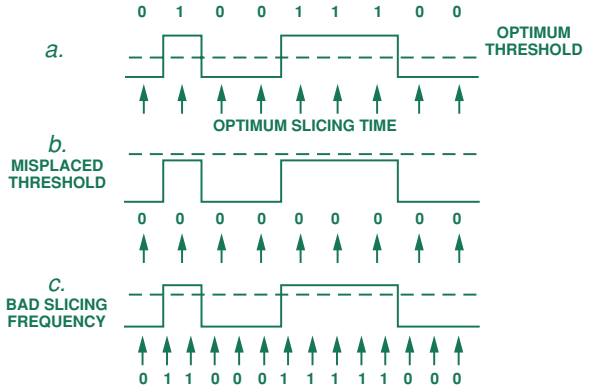


Figure 2. Simplified bit voltage transmission (PCM).

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.

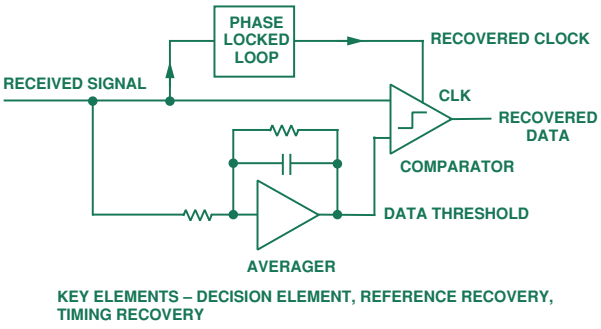


Figure 3. Idealized PCM.

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

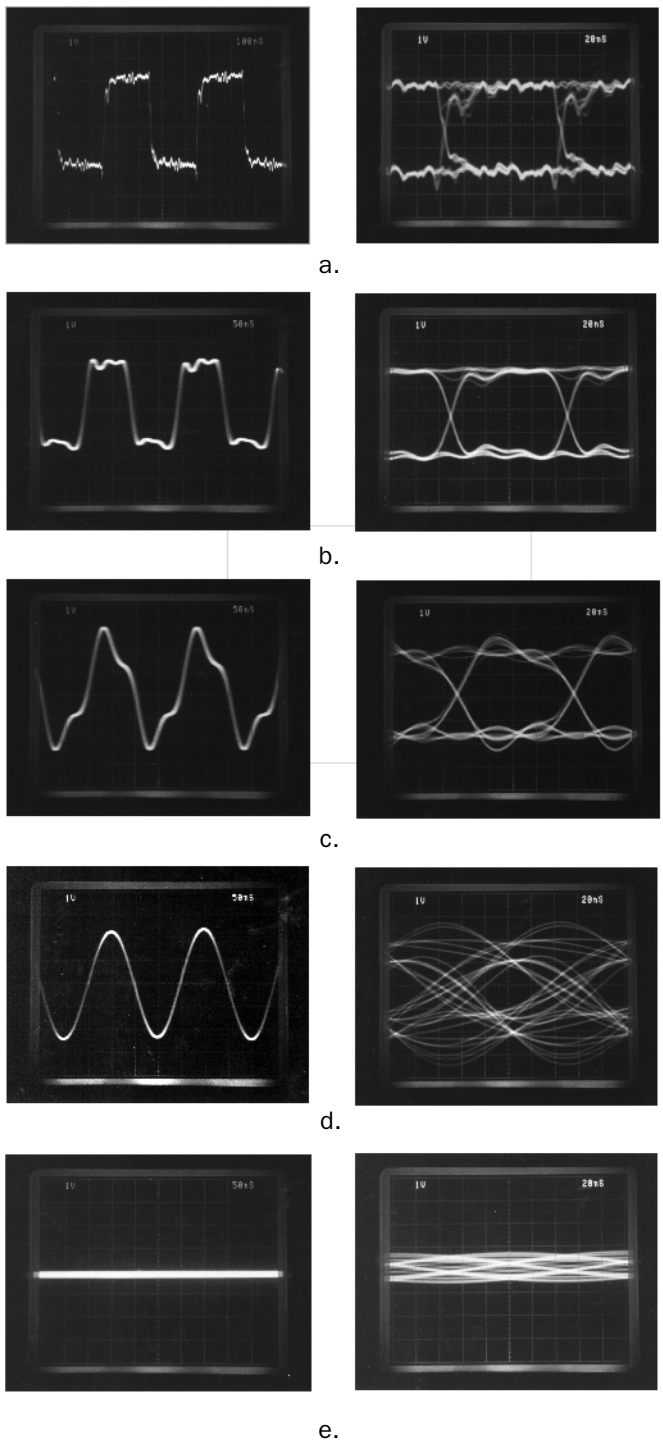


Figure 4. Scope waveforms vs. time (L) and eye diagrams (R).

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]

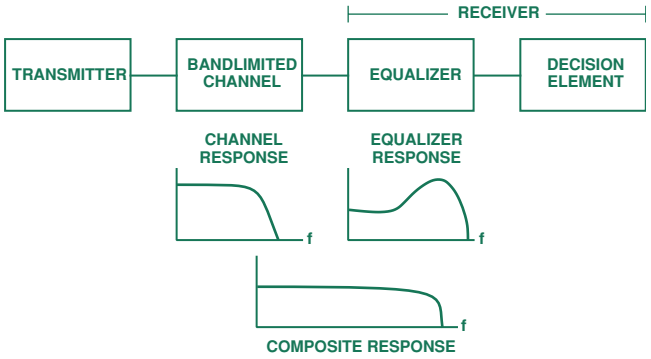


Figure 5. Channel equalization.

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

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

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.

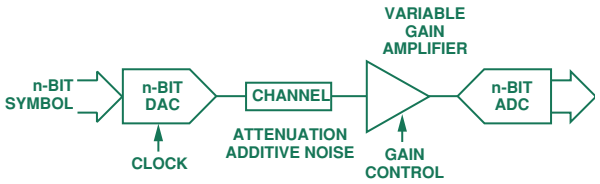


Figure 6. Simplified PAM transmitter/receiver.

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.

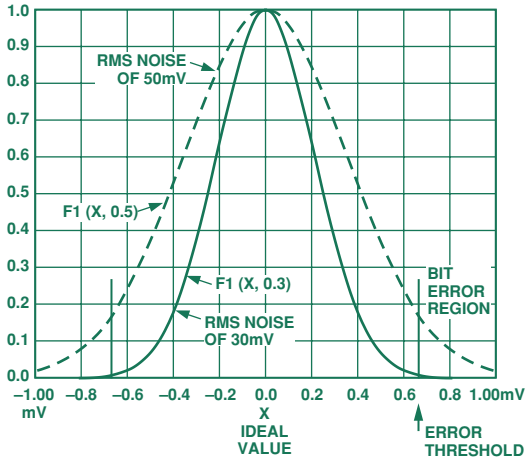


Figure 7. Ideal signal plus noise vs. error threshold: Threshold at 2 σ, and threshold at 1 σ.

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.

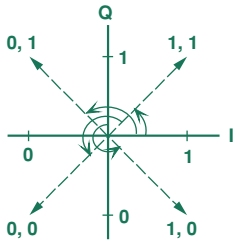


Figure 1. 2-bit QPSK phases.

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.

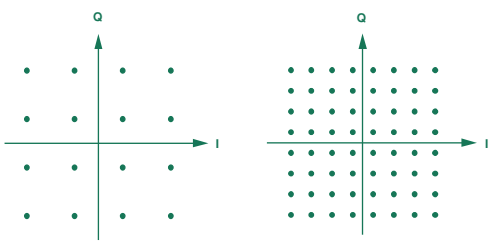


Figure 2. QAM constellations. a) 4 bits: 2-bit I and 2-bit Q. b) 6 bits: 3-bit I and 3-bit Q.

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.

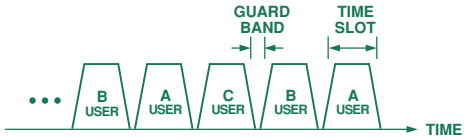


Figure 1. Illustration of time multiplexing, showing guard bands.

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)

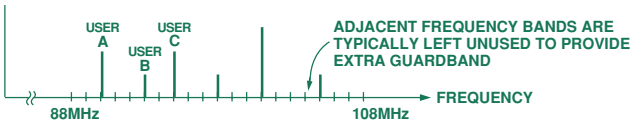


Figure 2. Illustration of frequency multiplexing, showing unused frequency bands to provide separation.

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.

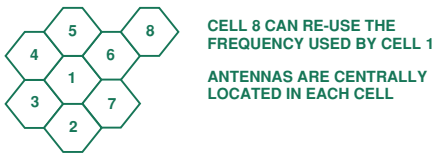


Figure 3. Illustration of geographical multiplexing, showing honeycomb of cells with base-station antennas at centers.

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.)

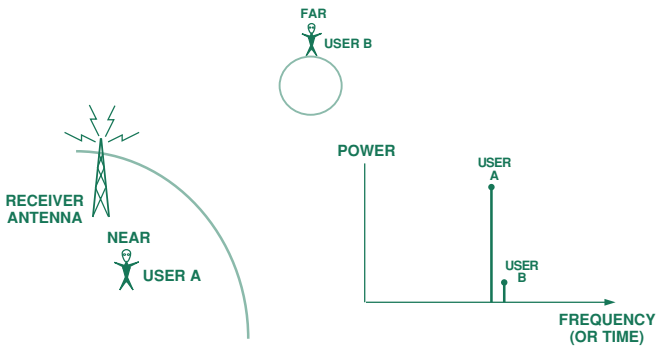


Figure 4. Near-far effect calls for the ability to handle wide dynamic range between adjacent channels.

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.

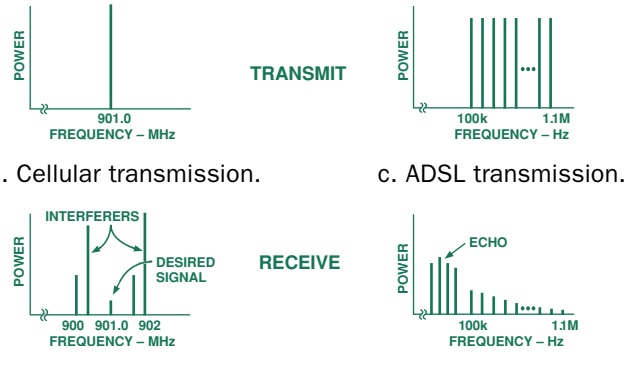
Selecting Mixed Signal Component for Digital Communications Systems

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).

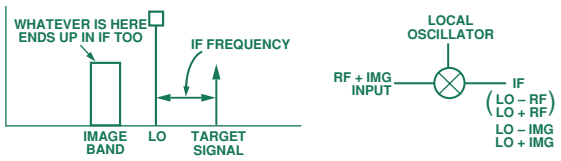


Figure 2. Mixing—the image problem.

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.

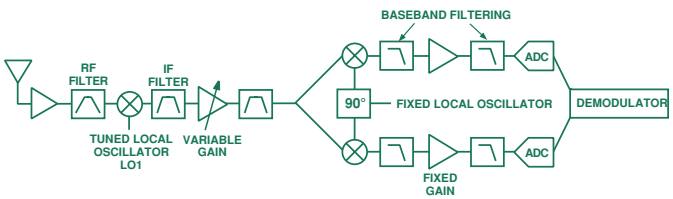
Conventional Receiver Design: Figure 3a illustrates a possible architecture for a GSM receiver path, and Figure 3b illustrates that of an ADSL modem. As noted earlier, the task of the receive

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).

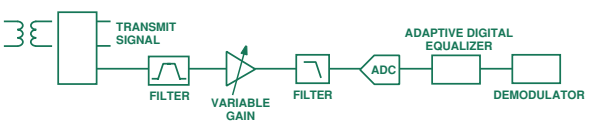
The GSM receiver signal path shown in Figure 3a illustrates the use of alternating gain and filter stages to provide the required selectivity and sensitivity. Channel selection, or tuning, is accomplished by varying the frequency of the first local oscillator, LO1. Variable gain and more filtering is 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 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.



a. GSM receiver.



b. ADSL modem receiver.

Figure 3. Typical receiver architectures.

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<signal<$(N+1/2)F_s$ will be translated down intact, signals in the bands $(N-1/2)F_s$<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.

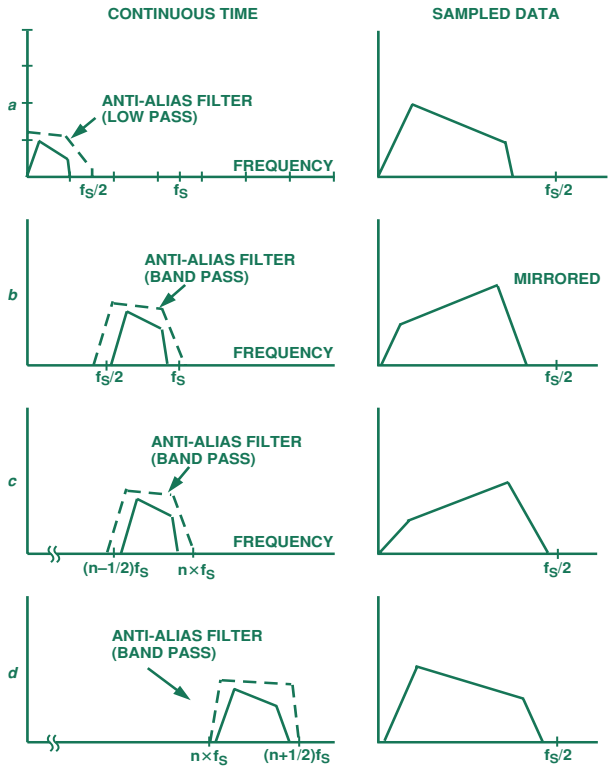


Figure 1. Aliasing, and frequency translation through undersampling.

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.

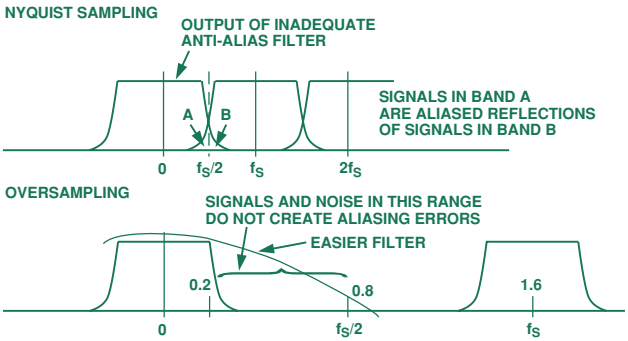


Figure 2. Oversampling makes filtering easier.

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.

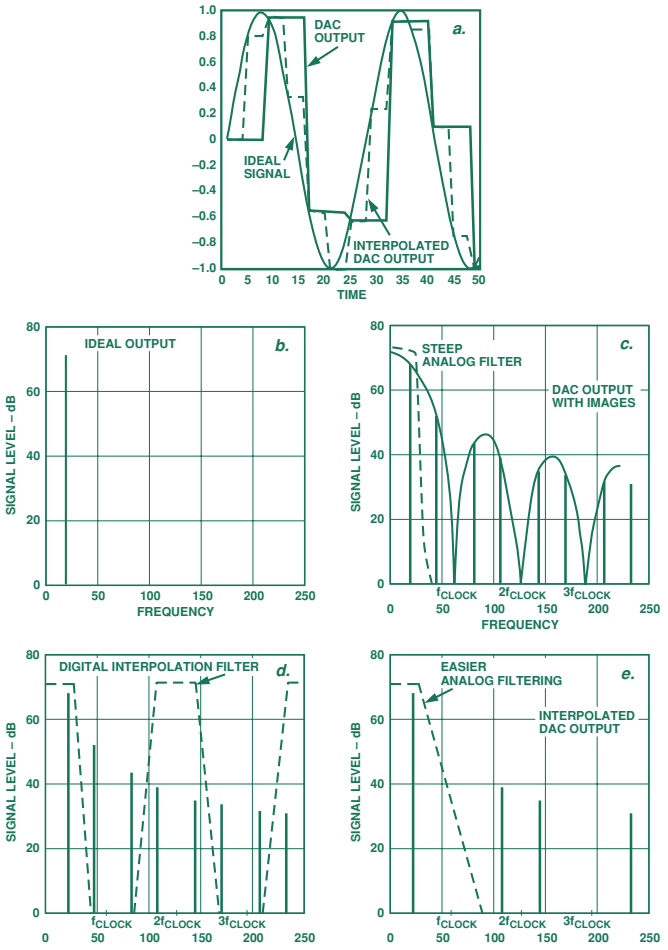


Figure 3. Time domain and frequency domain representation of continuous time and discrete sampled sine wave, and an interpolated discrete sampled sine wave.

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 ($2 \times 1.2 - 1.1$) and 1.0 MHz ($2 \times 1.1 - 1.2$). 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.

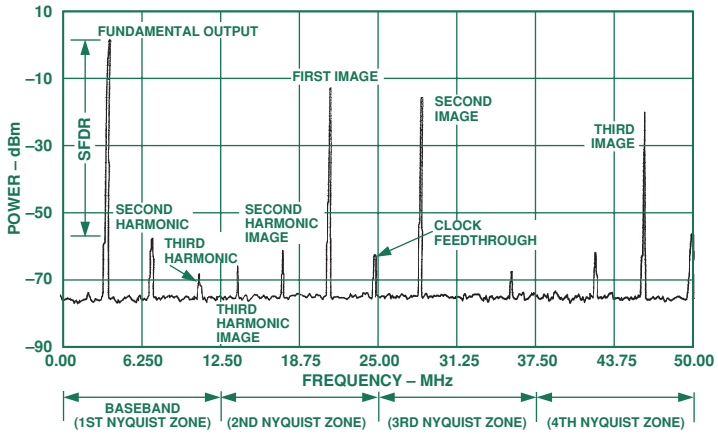


Figure 4. Different error effects in the output spectrum of a DAC.

Worldwide Sales Directory

North American Sales Offices and Representatives

Alabama: 256-721-1788
Alaska: 425-822-1655*; 714-641-9391*
Arizona: 714-641-9391*; 602-675-9252*
Arkansas: 972-907-4300*
California: 714-641-9391* (South), 408-382-3280* (North)
Colorado: 714-641-9391*
Connecticut: 203-748-6883
Delaware: 410-782-7980*; 800-392-3780*
Florida: 770-497-4440*
Georgia: 770-497-4440*
Hawaii: 714-641-9391*
Idaho: 714-641-9391*; 425-822-1655
Illinois: 847-252-5400*
Indiana: 219-434-0949*; 219-642-4585, 317-859-7886
Iowa: 319-364-5582
Kansas: 913-829-2800
Kentucky: 317-859-7886 (West) 513-742-8447 (East)
Louisiana: 972-907-4300*
Maine: 781-461-3000*
Maryland: 410-782-7980*; 800-392-3780*
Massachusetts: 781-461-3000*
Michigan: 248-615-4000
Minnesota: 736-792-4220*; 952-893-1400
Mississippi: 256-721-1788
Missouri: 636-946-1001
Montana: 425-822-1655*; 714-641-9391*
Nebraska: 913-829-2800
Nevada: 408-559-2037*; 714-641-9391*
New Hampshire: 781-461-3000*
New Jersey: 732-303-0775 (North), 410-782-7980* (South), 800-392-3780*
New Mexico: 714-641-9391*; 602-675-9252*
New York: 781-461-3000*; 631-563-7676
North Carolina: 919-834-0505
North Dakota: 952-893-1400
Ohio: 440-247-0060, 513-742-8447, 614-487-1882
Oklahoma: 972-907-4300*
Oregon: 425-822-1655*; 714-641-9391*
Pennsylvania: 410-782-7980*; 215-643-7790* (East), 440-247-0060 (West)
Rhode Island: 781-461-3000*
South Carolina: 919-834-0505*
South Dakota: 952-893-1400
Tennessee: 770-638-8998 (East), 256-721-1788 (West)
Texas: 972-907-4300*
Utah: 303-666-8874*
Vermont: 781-461-3000*
Virginia: 410-782-7980*
Washington: 425-822-1655*
West Virginia: 410-782-7980*
Wisconsin: 414-543-5777*; 920-668-8460
Wyoming: 714-641-9391*
Puerto Rico: 787-273-8535
Canada: 905-821-7800, 613-225-3878, 514-672-0801, 604-465-6892

*Analog Devices, Inc., Direct Sales Offices

North American Distributors

Avnet 800-332-8638
Future Electronics –U.S. 800-388-8731, –Canada 800-361-6747
Newark Electronics 800-463-9275
Pioneer 800-657-0168
Semi Dice 508-238-8344

International Direct Sales Offices

Australia: 03 5986 7755
Austria: 1 888 55 04 0
Brazil: 11 5561 2686
Denmark: 44 845800
France: 01 46 74 45 00
Germany: 089/76 903 551, 0221/68 929 3, 0711/139 908 0
Hong Kong: 852 2506 9336
India: 91 11 623 0565, 91 11 668 2565
Ireland: 61 229 011
Israel: 9 7744461
Italy: 02 210401 1
Japan: 813 5402 8200
Korea: 2 554 3301
Mexico: 528 365 2020
Netherlands: 76 5233200
New Zealand: 613 5986 7755
People's Republic of China: 755 378 1959
Shanghai Office: 21 6426 7009
Singapore: Tel; 6887 0650
Sweden: 0 8733 3800
Taiwan: 8862 2719 5612
United Kingdom: 01932 266 000, 01256 472111
(Firefly Technology LTD MIL/Aero Representative)

International Distributors

Argentina: Elko Components Electronics S.A. 54 11 4372 1101; Semak S.A. 54 11 4381 2108

Australia: Adilam (Systems Only) 613 9761 4466; Future Electronics 613 9899 7944; Lockhart Pty. Ltd. 618 9245 9101; Unique Technology 61 3 9760 4277

Austria: Silica/Avnet EMG 43 2236 3782780; Kristl Seibt & Co_Ges.m.b.H 43 316 5995 78; RS Components (Systems Only) 43 2852 505; Spoerle Electronic 43 1 360460

Belarus: SHPAT & Service 375 17213435

Benelux: ACAL Belgium nv/sa 32 2 720 59 83; ACAL Nederland bv 31 40 250 26 02; Spoerle Electronic Belgium 32 2 725 46 60; Spoerle Electronic Nederland 31 30 639 12 34

Brazil: Alphatech Servicoes E Com. Ext 55 11 6976 3433; Avnet do Brasil Ltda. 55 11 5079 2150

Bulgaria: K1 Electronics 35 92 953 1550

Croatia: Alma Electronic 385 1 3666 793

Czechia: Amtek 420 5 471255 55

Denmark: Arrow Nordic Components 70 10 22 11; Impact Memec 44 44 66 66

Estonia: Adimir Ltd., 372 6654260; Arrow Electronics Estonia Oü 372 6774 250

Finland: Arrow Nordic Components 358 9 476660

France: Silica/Avnet EMG 1 64 47 00 84; CCI Electronique 1 41 80 70 00; Impact Memec 1 49 78 03 24; RS Composants 3 44 10 15 00

Germany: Silicia/Avnet SEI 49 89 45110 0; Impact 49 2162 5039 0; Sasco Semiconductor 49 89 46 11 0; Semitron W Rock GMBH 49 7742 8001 0; Spoerle Gmbh 49 6103 304223; Spoerle Lager 49 3103 304 122

Greece: Micrelec 30 10 5695042

Hong Kong: Arrow 2484 2685; Chinatronic 2376 2075; Future Electronics (HK) Ltd. 2420 6238; General Engineers Hong Kong 2503 2212

Hungary: SMD Technology 36 1 370 7994

India: BBS Electronics Pte., Ltd. (Bangalore) 91 80 530 6611, (Mumbai) 91 22 824 2374, (New Delhi) 91 11 647 4978; (Pune) 91 20 746 3965, (Thiruvanan Thapuram) 91 471 331376; Excelpoint Systems (Pte) Ltd. (Bangalore) 91 80 527 4211, (Mumbai) 91 22 857 6600, (New Delhi) 91 11 645 2808; Future Electronics (Bangalore) 91 80 558 7885, (Mumbai) 91 22 570 2745, (New Delhi) 91 11 646 9745

Ireland: Abacus Polar 353 614 80166; Arrow Electronics 353 18307207; Avnet Lyco 353 1 8901000

Italy: Silica/Avnet Nordstar 39 02 660921; Eurolink Systems srl 06 612401; La Tecnika Due s.r.l. 011 2425905 Lasi Elettronica 02 661431; Tecnika Due s.a.s. 0438 943211; Tecnika Due s.a.s. 39 02 9595931

Korea: Analog World Co. Ltd 2701 5993; Future Electronics Inc. 2 555 6736; Lite On Korea Ltd. 2 650 9700; Tong Baek Electronics Co. Ltd. 2 715 6623

Latvia: CARO Ltd. 371 7210526

Lithuania: Elgerta Ltd. 370 2 65 26 83

Malaysia: Excelpoint Systems (PTE) Ltd. 603 244 8929

Mexico: Future Electronics Mexico SA de CV (Ixcalli) 525 893 5761, (Jalisco) 523 122 0043, Latin Rep 523 647 4250

Netherlands: Acal Nederland 040 250 2602; Spoerle Electronic 030 6391234

New Zealand: Future Electronics 64 3 366 2887, Memec EBV (NZ) Ltd. 64 3 379 3889

Norway: Impact Memec A/S 47 66 981370

People's Republic of China:
Arrow China Electronics Ltd.: Beijing 10 8809 1638; Chengdu 28 6626 536; Fuzhou 591 7848 456; Nanjing 25 4507 458; Qingdao 532 5919 810; Shanghai 21 5306 1871; Shenzhen 755 359 2920

ChinaTronic: Beijing 10 6848 6895; Chengdu 28 627 0793; Fuzhou 591 782 8575; Guangzhou 20 711 1007; Shanghai 21 6259 4939; Shenzhen 755 332 5655; Tian Jin 22 2731 3965

Excelpoint Systems (PTE) Ltd.: Beijing 10 6837 3894; Chengdu 28 662 4475; Shanghai 21 6482 2280; Shenzhen 755 377 9492

Future Advanced Electronics Ltd.: Beijing 10 6874 8672; Guangzhou 20 8381 4927; Nanjing 25 482 9241; Shanghai 21 6443 7045; Shenzhen 755 219 0525

Wuhan P & S Electronics Co., Ltd.: 86 27 8749 3500

Phillippines: Future Electronics Tel: 63 2 807 5092,

Poland: ADDIS 48 32 330 46 90; P.E.P. Alfine 48 61 8205811

Portugal: SEI/Selco 351 297 38 21

Romania: Top 9+ ELECTRONIC SYSTEM 401 2320424

Russia: AO ARGUSSOFT 7 095 216 5855 (Moscow), 7 812 310 6234 (St. Peterburg); AO Autex 70 95 334 7741; Eltech Ltd. 7812 327 9090; Prosoft (Systems Only) 70 95 2340 636

Singapore: Excelpoint Systems (PTE) 741 8966; BBS Electronics Pte Ltd. 65 748 8400; Future Electronics Tel. 65 479 1300

Slovakia: Dialogue 421 33 7744220

Slovenia: Alma Electronic S.P.O. 286 1 2561381

South Africa: CST Electronics Pty. Ltd. 27 11 452 0706; Parsec Components 27 12 349 2282

Spain: Silicia/Avnet 91 372 72 00

Sweden: Arrow Nordic Components 46 856 265700; Impact Memec (IE Komponenter) 08 80 46 85; Impact Norway A/S 47 66 981370

Switzerland: Sasco Semiconductor 0041 24 447 0100, 0041 1 817 6262; Spoerle Electronic GMBH 0041 24 447 0100, 0041 1 817 6662

Taiwan: Answer Technology Co., Ltd. 8862 2698 2526; Arrow Strong Electronics Co., Ltd. 8862 2917 9917; AV Tech Inc. 8862 2696 2700; Elitetron Electronic Co., Ltd. 8862 2789 3300; EPCO Technology Co., Ltd. 8862 8797 2627

Thailand: Excelpoint Systems Supplies Co., Ltd. 662 5761877/8; Future Electronics Tel. 66 2 361 8400,

Turkey: ELEKTRO Elektronik A.S. 90 216 4610700; Arrow Electronics 9021 64645090

Ukraine: VD Mais 380 44 2272262

United Kingdom: Abacus-Polar Holdings 01525 858000; Arrow Electronics UK Ltd. 01234 270027, 01279 441144; Silica/Avnet EMG 01438 788300; Farnell Electronic Components 0870 1200200; Impact Memec 01844 261188; RS Components Ltd. 01536 2012

05/02

Analog Devices GmbH
Am Westpark 1–3, D-81373 München, Germany; Tel: 49 89 76903 0; Fax: 49 89 76903 157

Japan Headquarters
New Pier Takeshiba, South Tower Building, 1-16-1 Kaigan, Minato-ku, Tokyo 105-6891, Japan; Tel: 813 5402 8200; Fax: 813 5402 1063

Southeast Asia Headquarters
4501 Nat West Tower, Times Square, 1 Matheson Street Causeway Bay, Hong Kong, PRC; Tel: 852 2506 9336; Fax: 852 2506 4755

Worldwide Headquarters
One Technology Way, P.O. Box 9106, Norwood, MA 02062 9106, U.S.A.
Tel: 781 329 4700, (1 800 262 5643, U.S.A. only), Fax: 781 326 8703
www.analog.com

ANALOG DEVICES

WORLDWIDE HEADQUARTERS
One Technology Way, P.O. Box 9106
Norwood, MA 02062-9106, U.S.A.
Tel: 781 329 4700, (800 262 5643 U.S.A. only)
Fax: 781 326 8703
www.analog.com

ANALOG DEVICES GmbH
Am Westpark 1–3 D-81373
München, Germany
Tel: 49 89 76903 0; Fax: 49 89 76903 157

JAPAN HEADQUARTERS
New Pier Takeshiba, South Tower Building
1-16-1 Kaigan, Minato-ku, Tokyo 105-6891, Japan
Tel: 813 5402 8200; Fax: 813 5402 1063

SOUTHEAST ASIA HEADQUARTERS
4501 Nat'West Tower, Times Square
1 Matheson Street, Causeway Bay, Hong Kong, PRC
Tel: 852 2 506 9336; Fax: 852 2 506 4755

PRINTED IN U.S.A.
AD35–166–5/02