N89-13319

Analog-to-Digital Conversion Techniques for Precision Photometry

Chet B. Opal
McDonald Observatory
University of Texas
Austin TX 78712

ABSTRACT

Three types of analog-to-digital converters are described: parallel, successive-approximation, and integrating. The functioning of comparators and sample-and-hold amplifiers is explained. Differential and integral linearity are defined, and good and bad examples are illustrated. The applicability and relative advantages of the three types of converters for precision astronomical photometric measurements are discussed. For most measurements integral linearity is more important than differential linearity. Successive-approximation converters should be used with multielement solid state detectors because of their high speed, but dual slope integrating converters may be superior for use with single element solid state detectors where speed of digitization is not a factor. In all cases the input signals should be tailored so that they occupy the upper part of the converter's dynamic range; this can be achieved by providing adjustable gain, or better by varying the integration time of the observation if this is possible.

Introduction

Analog circuits are noisy, don't hold calibration, are a nuisance to adjust, and are difficult to design and test. Digital circuitry has no noise, requires no calibration or adjustments, and is (by comparison) straightforward to design and test. It is not surprising that each year more functions which used to be done with analog circuitry are performed digitally. Most data are going to be fed into a digital computer for analysis anyway, so the earlier signals are digitized in the signal processing chain the better.

All photometric detectors produce low level outputs in analog form which must be amplified and filtered before digitization (a process called signal conditioning). Simple brute force analog-to-digital (A/D) conversion will seldom produce accurate results, so proper signal conditioning will always be very important. Signal conditioning itself is outside the scope of this article; instead, the emphasis is on how best to present the analog voltage to the A/D converter in order to achieve the desired accuracy.

Several types of A/D converters will be described, and their relative merits and faults discussed. Some ancillary circuitry and methods of computer interfacing are also described. First, however, some of the terms used when evaluating A/D converters will be defined and discussed.

Resolution and Linearity

An A/D converter transforms input voltages, currents, resistance ratios, or other analog signals into numbers represented by logic signals. In the following discussion we shall usually assume that the analog input is a voltage, and that the digitized values (data numbers) are expressed as binary numbers, although neither need be the case. The interval between digital steps is the resolution, which is determined by the number of bits in the digital value. A converter with 4-bit resolution thus has 2^4, or 16 steps, each with 1/16 of the total digitization interval (the conversion range).

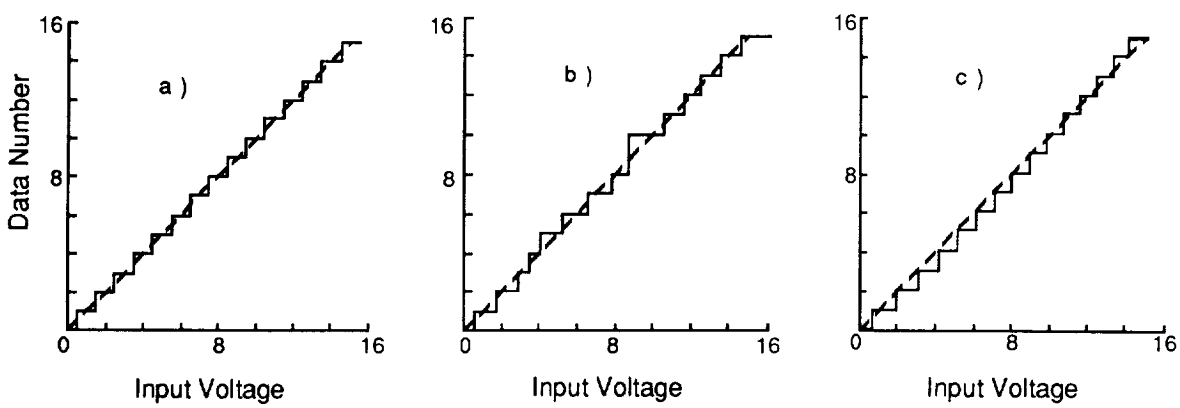

Fig. 1. Linearity and quantization error

Figure 1a shows the output value as a function of input voltage for a perfect A/D converter (the dotted line has unit slope). This hypothetical device is a 4-bit unit which converts analog signals over the range -0.5 to +15.5 V to an integer between 0 and 15. The converted value is correct everywhere to +/- 0.5 V; since a change of 1 in the least significant bit (LSB) of the integer represents 1.0 V, this converter would be said to have an accuracy of 1/2 LSB. It follows that if an A/D converter is really accurate to 1/2 LSB everywhere, it is perfect. This uncertainty in the LSB is called quantization error and is an inevitable consequence of the digitization process. Accuracy can also be expressed as a percentage of full scale (about 3.3% of full scale for a 4-bit converter). Note that the accuracy gets worse as the input voltage drops: with this converter a 1.0 V signal can only be measured to 50% accuracy. Obviously, whenever possible a converter should be operated near full scale to achieve the best accuracy.

Figure 1b shows what can happen if we relax the accuracy to 1 LSB. This is what manufacturers usually mean when they claim 1/2 LSB accuracy (it is within 1/2 LSB of perfect). In this case the input signal can cover a range of up to 2.0 V and still be digitized to the same value. In fact, certain values may not even appear in the output at all (there is no input voltage which will produce the data number 9 in the example). The output increases monotonically with the input, but this doesn't have to be the case to meet a 1 LSB accuracy specification (many manufacturers, however, are kind enough to guarantee a monotonic output with no missing codes). Because of the variation of the widths of the input bins a converter of this type is said to exhibit poor differential linearity; if it is used to measure the difference of two nearly equal voltages the errors will be larger than would be expected for the given resolution. However, this hypothetical converter does exhibit reasonably good linearity for large changes in input voltage, so it is said to exhibit good integral linearity.

Figure 1c shows a converter which also has an accuracy of about 1 LSB. Here the widths of the bins are reasonably constant, indicating good differential linearity, but the output meanders about the desired straight line input/output curve, so this converter would be said to exhibit poor integral linearity.

Whether integral or differential linearity is more important depends on the application. In a nuclear pulse height analyzer, for example, an A/D converter with good differential linearity is very important; if the digitization bins are unequal in width the spectrum will appear extremely ragged and it will be difficult or impossible to measure the positions of lines. On the other hand, if one is subtracting the detector output from the sky from that of a star, integral linearity is more important than differential linearity.

Clearly, it is advantageous to digitize signals to as many bits as possible. Sixteen bits sounds like a reasonable goal (most computer hardware handles 16-bit numbers efficiently): this corresponds to 1 part in 65,536, or about 15 parts per million (ppm). As we shall see, obtaining meaningful numbers to this accuracy is not easy.

Hardware

Comparators

The basic building block of an A/D converter is the comparator. This circuit has two analog inputs and a digital output. If the + input is greater than the - input, the output will be a logical "1," if it is less the output will be a logical "0."

Figure 2 shows a comparator circuit which can be built from parts lying around the lab. The pair of transistors and associated collector and emitter resistors form a simple differential amplifier. The V1 input to the circuit is from a carbon microphone (borrowed from the lab telephone). The bias point of the left-hand transistor is determined by the relative resistances of the microphone and its load resistor, while that of the right-hand transistor, V2, can be adjusted with a potentiometer (pot). The pot is set so that, with no input, the light emitting diode (LED) between the two collectors just goes out (goes to logical "0"), at which point V1 and V2 are equal. Sound impinging on the microphone causes its resistance to vary, and V1 will fluctuate up and down. When the resistance increases, V1 is higher than V2 and the left transistor draws more current, bringing its collector voltage down; it also robs current from the right transistor, causing its collector voltage to rise, so the LED will light (go to logical "1"). As you talk into the mike the LED will flicker: voila'--digitized speech!

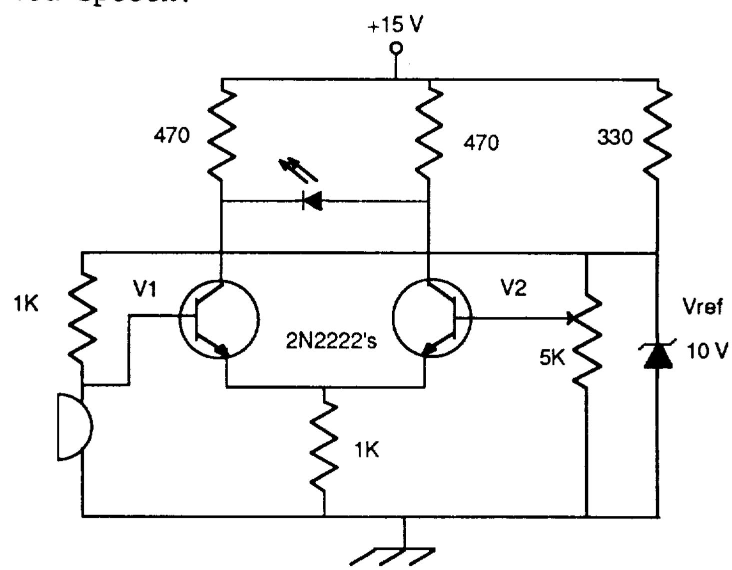

Fig. 2. A simple comparator

The above circuit has a lot of potential problems. If the two transistors are not well matched in characteristics, the point at which the circuit triggers will vary with the operating point and temperature. Usually integrated circuit (IC) comparators are used, so the transistors are next to each other on the chip, are closely matched in characteristics, and operate at nearly the same temperature. Even so, if one transistor has been driven on harder than the other for some period of time, there will be a temperature imbalance and the triggering point may change. As a result the circuit will exhibit hysteresis: the trigger point for a signal with a positive slope will be different than that for a negative slope. Most IC comparators have a much higher gain than the circuit shown, and a small amount of hysteresis is actually desirable because it produces a cleaner output: because of noise, a perfect comparator produces a rashy output as the input signal crosses the trigger point. In fact, a small amount of positive feedback is sometimes introduced to produce controlled hysteresis.

The total of the possible offsets in the triggering voltage caused by internal noise, hysteresis, temperature variations, aging, power supply tolerances, etc., in the comparator must be less than

1/2 LSB in order to preserve the accuracy of an A/D converter. Designing a comparator which can be used with a fast, 16-bit converter is not trivial.

In the above circuit, the accuracy of Vref was not important because the trigger point depended on the ratios of the resistances, not on absolute voltages. If, instead, the circuit is to be used to measure an external voltage V1 by comparing it to Vref, additional factors become important. Obviously a stable voltage divider is needed; in most cases sufficient stability can be achieved with metal film resistors of similar value deposited on the same substrate. The voltage standard Vref is another matter. For best accuracy Zener diodes with internal heaters and temperature regulators are used; even these have temperature coefficients of 1 ppm/degC So changes due to the difference in temperature between the laboratory and the dome, or even on the dome floor during the night, can exceed the 15 ppm tolerance needed for 16-bit accuracy. Even worse, most reference sources have long term drifts approaching 100 ppm/year, so they will need to be recalibrated for each observing run.

A simple comparator is all that is needed to digitize photomultiplier signals. The pot is adjusted so that the circuit is triggered only by the amplified signal pulses, and the resulting "1"/"0" pulses are fed to a counter. Accuracy is achieved by counting a lot of pulses over a period of time. An adjustable comparator used in this way is usually termed a discriminator (it "discriminates" between signal pulses and noise).

Most solid state detectors don't have internal amplification sufficient to produce countable pulses from single photon events, so their signals must be digitized to more than one bit of accuracy. Before discussing how comparators are used in various multi-bit converters, it is necessary to introduce one more building block.

Sample-and-Hold Amplifiers

A Sample-and-Hold (S/H) is used to store an analog voltage prior to digitization. Basically all it is is a capacitor which can be connected to an input source by closing a switch. The capacitor voltage follows the input voltage until the switch is opened, at which point it holds at that voltage so that it can be measured at leisure. Figure 3 shows a typical S/H circuit, which has some additional components. There is an operational amplifier (op amp) at the output to isolate the capacitor from the output load so the charge won't be drained off. There is another op amp which presents a high impedance to the input while providing a low impedance output to charge the capacitor quickly. The feedback to the input op amp comes all the way around from the output in order to compensate for voltage offsets in the switch and output amplifier.

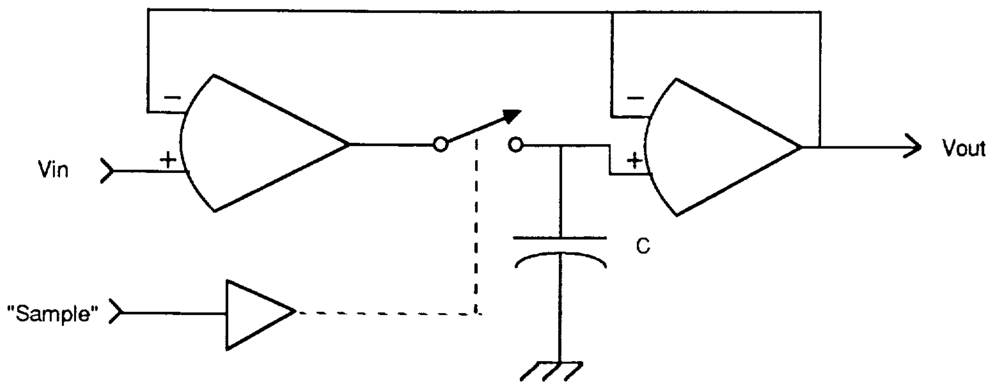

Fig. 3 Sample-and-hold Amplifier

Although the S/H circuit would appear to be simple to design, considerable care is needed to achieve a good design. There is always some resistance in the sampling switch, and the output impedance of amplifier driving the capacitor is finite, so there will be an RC time constant associated with charging the capacitor and output will lag somewhat behing the input. Note that if the output is to equal the input to 15 ppm, a settling time of more than 11 time constants is required. The time constant can be made shorter by decreasing the value of C; on the other hand, a smaller value of C makes the effects of leakage currents in the switch and offset currents in the output buffer amp larger, causing increased drift in the output. Some coupling of the switch control signal into the output is inevitable; this is minimized by increasing the value of C. So a number of compromises are necessary, depending on the application.

Multi-element detectors (CCD's, Reticons, Infrared Arrays) produce complex waveforms which must be sampled at precise times before being fed to the A/D converter; S/H elements with fast and well defined acquisition times are needed, but, because everything is done quickly, a fair amount of output drift can be tolerated. Low drift is more important in, for example, an infrared detector system where the telescope is nodded for sky chopping. A fast S/H with a droop rate of only 1 microvolt/microsec sounds very good for CCD readout electronics, but this translates to 1 V/sec which may be intolerable for other applications. If analog voltages are to be stored for more than a few milliseconds, large value Teflon-dielectric storage capacitors and chopper-stabilized (varactor) operational amplifiers should be used. In general, the requirements of fast acquisition time and low drift are incompatible.

Parallel A/D Converters

Certainly the conceptually simplest A/D converter is the parallel type, shown in Figure 4. A four-bit version of the unit shown would consist of 15 comparators, each tapped off a different resistor in a chain between Vref and ground. A 12.2 V signal, for example, will exceed the threshold of the lower twelve comparators and each will indicate a logical "1," the last 3 comparators will not be triggered and will indicate a logical "0." The resulting outputs produce a "thermometer" type indication of the input voltage; these bits must be re-coded into a conventional binary value using some logic circuitry (equivalent to that used in priority encoders).

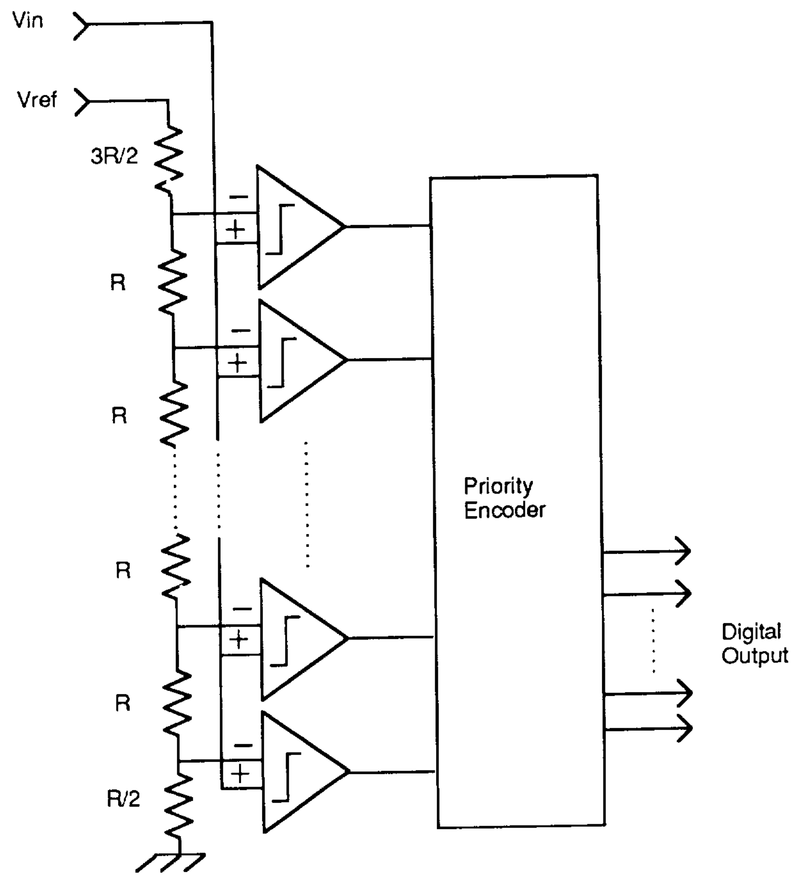

Fig. 4. Parallel-conversion A/D Converter

The advantage of the parallel A/D converter is its high speed, limited only by the frequency response of the comparators. Digitization rates of tens of megahertz can be achieved (which is why these are also called "flash" converters). Their obvious disadvantage is that the number of comparators and the complexity of the logic increase exponentially with the number of bits. Since the number of elements which can be placed on an integrated circuit has historically doubled every eighteen months, and since 8-bit "flash" A/D's were available ten years ago, we should rightfully expect 14- or 15-bit units by now, and our 16-bit problems should be solved in a few years. Alas, complexity alone is not the issue: the tolerances of the resistor ladder are tighter and the precision of the comparators must also increase as the resolution improves. So today we can only buy 9-bit flash converters; granted, they are cheaper and faster than the old 8-bit units, but we will have a long wait for 16-bit versions.

Because of the low precision achieved, parallel converters are not of much use for digitizing photometric signals. Conversion times measured in nanoseconds are mainly useful for processing TV, radar, SETI, and similar signals. For multielement photometric detectors conversion speeds of a few microseconds are adequate, and for single element detectors millisecond speeds will suffice. Fortunately there are other methods available for getting the necessary precision.

Successive-approximation A/D converters

The successive-approximation A/D converter uses a comparator, a digital-to-analog (D/A) converter, and some logic (Figure 5). The D/A converter part of the circuit consists of a switch and resistor for each bit; the currents flowing from the voltage reference are summed by an op amp and fed to the comparator. The resistors are weighted in binary fashion so that the current flowing through the LSB resistor is 1/32768'th that flowing through the most significant bit (MSB) resistor.

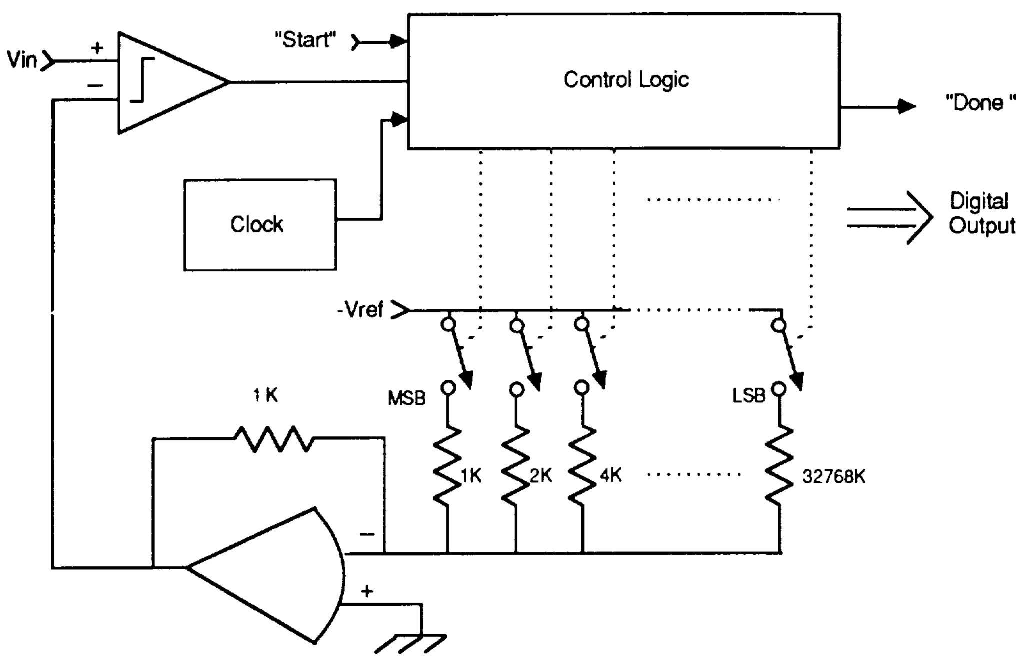

Fig. 5. Successive-approximation A/D Converter

The conversion proceeds as a binary sort. Suppose that we are converting a -0.5 to 10.5 V signal to 16-bit precision, and that that the signal is about 4.0 V. First the switch in series with the MSB resistor is thrown, which produces 5.0 V at the output of the

DAC. This voltage exceeds that of the signal, so the comparator outputs a logical "0." The logic circuitry therefore opens the MSB switch and closes the next most significant bit switch. This puts 2.5 V on the comparator, which is less than the signal, so the comparator outputs a "1." The logic therefore leaves this switch on, and tests the 1.25 V switch, which it leaves on, and so on. Only sixteen decisions are required to digitize to sixteen bits, so this procedure can be quite fast (typically on the order of 20 microsec, and 2 microsec converters are advertised). Because of the finite digitizing time, successive-approximation A/D converters are usually preceeded by a sample-and-hold amplifier to hold the voltage constant during the digitization process.

There are some problems with the circuit as shown. The tolerance of the 33 Megohm resistor can be rather lax, but that on the most significant bit is severe (it must be accurate to 1 part in 65,000). The switches when open must have a resistance that is large compared to the largest resistor (in this case 33 Megohm) yet small compared to the tolerance on the smallest resistor (1K/65000 ~ .01 ohm). The reference source must not only be stable but must be able to put up with wide variations in load depending on which resistors are switched in. In practice a different resistor switching scheme is used, known as an R/2R ladder, which uses nearly the same value of resistors everywhere but relies on diverting progressively more of the current for the less significant bits into a dummy load instead of into the op amp summing node. By using single-pole double-throw switches to dump the current in an "open" switch the current in each leg can be kept constant, minimizing load variations seen by the voltage reference.

The main disadvantage of the successive-approximation A/D converter is poor differential linearity. It is impossible to trim all the resistors values properly and to get them to track with temperature so that the digitizing intervals are uniformly spaced. These A/D converters do not age well, and need to be checked periodically for missing codes, "sticky" bits, and other signs of improper operation.

Integrating A/D Converters

One rather simple method of digitizing a signal is to feed the signal to one input of a comparator and a linear ramp to the other, and to determine the time it takes the ramp to rise from zero to the signal voltage. This time can be measured by a high speed counter fed by a stable clock, the resulting count total is proportional to the signal voltage. This type of converter, known as the Wilkinson type, is commonly used in pulse height analyzers because no resistor switching is involved and there is only one comparator; consequently the circuit exhibits good differential linearity.

For most purposes a better design, which is somewhat slower at digitizing but is more tolerant of component variations, is the dual-slope integrating A/D converter shown in Figure 6. Here the signal voltage is integrated upward from zero until the counter overflows, the input is then switched and the reference voltage is integrated downward, while the counter starts counting from zero again, until the integrated voltage falls to zero and triggers the comparator, which stops the counter. At this point the number of counts recorded is proportional to the ratio of the input voltage to the reference voltage. The precision achievable is determined by the speed of the counters and the integration interval, and can be made very high. The beauty of this circuit is that short term variations in the values of the R and C in the integrator or in the clock frequency cancel out; only the accuracy of the reference voltage matters. Because the signal is integrated over a period of time, noise variations are averaged out, and, if the signal integration period is made a multiple of the power line period, hum pickup will be cancelled out to boot.

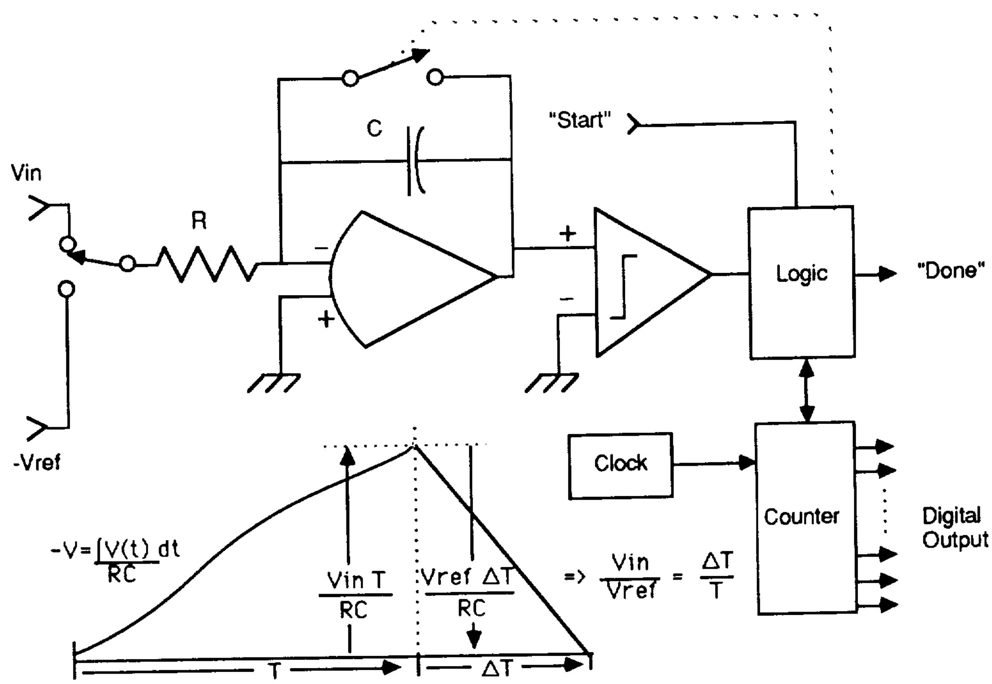

Fig. 6. Integrating A/D Converter

The main disadvantage of the integrating A/D converter is that it is very, very slow. This type of converter is really only suitable for use with single element detectors.

Improving Dynamic Range

Most photometric measurements involve the determination of ratios rather than absolute values. If a bright star and a faint star are measured for the same length of time with the same circuitry, the digitized output from the fainter star will be less accurate than that of the brighter star because of quantization error, and the resulting brightness ratio will be even more inaccurate. Some A/D converters function in ratio mode, where the reference voltage can be replaced by a stored value of the voltage measured for the brighter object. This will somewhat improve the accuracy if the objects are not too different in brightness, but is probably not worth the trouble.

It is tempting to consider digitizing the logarithms of the signals, in which case ratios become differences and quantization error is less of a problem. However, most measurements first require taking the difference of the brightness of the star and its nearby sky (or perhaps some other zero subtraction is required), and this cannot be done with logarithms. Also, a logarithmic amplifier is nonlinear and difficult to calibrate and usually has unpleasant temperature sensitivity, peculiar noise characteristics, and other problems. Similar considerations apply to square-root amplifiers, which otherwise might be a good idea to use for shot-noise-limited signals. A better method is to provide switchable gains, either through attenuators or amplifiers, so that the signals from both objects are digitized at voltages near the full scale of the A/D converter. Converters with built-in automatic ranging (also called "floating point" converters) are available; the most useful ones offer steps of factors of 2 in the gain, rather than factors of 10. Obviously the attenuators, amplifiers, and automatic gain changes must be carefully calibrated in order to preserve the accuracy of ratio measurements.

The best way to compare two objects is to use different integration times so that they produce about the same output from the detector. Time is the one variable which is easily measured to very high precision. Unfortunately, it is not always possible to vary the exposure time without influencing other variables.

Packaging

Some typical packaging schemes are shown in Figure 7.

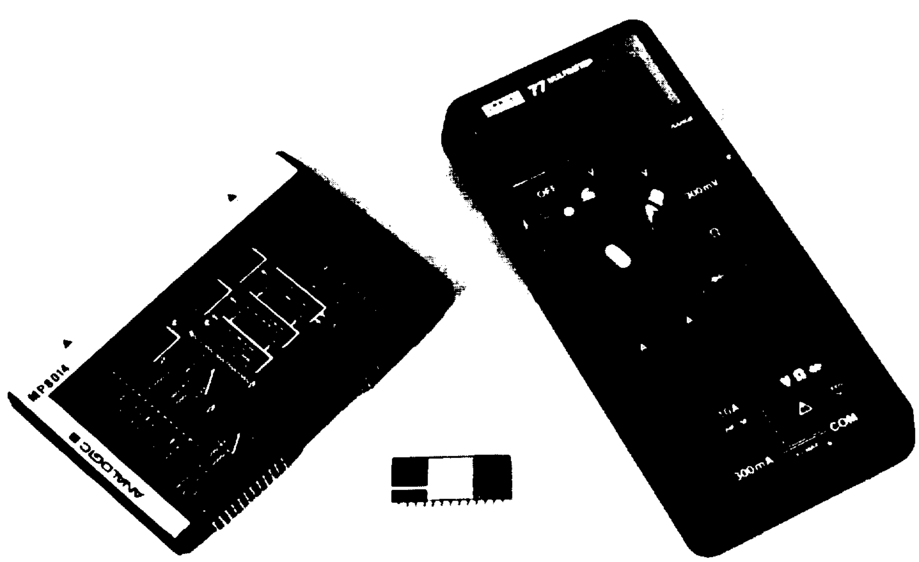

Fig. 7. Typical A/D Converter Packaging

Direct conversion A/D converters are usually monolithic IC's and require no external components. Some low-end successive-approximation converters are available as single IC's, but more commonly they are hybrids of discrete components and integrated circuits assembled into a module which can be mounted on a circuit board. Most of the more complex functions of dual slope integrators are carried out by integrated circuits, but usually several external components are required. The lowly laboratory digital voltmeter shown in the figure is actually a rather sophisticated dual slope integrator with automatic ranging and zeroing; unfortunately it lacks a computer interface.

The outputs of the direct conversion A/D are available continuously at the output; these can be read into the computer via a parallel I/O port chips can be sampled at will. The successive-approximation A/D must be commanded to start a conversion, and produces a result some time later, so is more difficult to interface. Since the successive-approximation process is serial in nature, serial outputs and a clock are usually provided, and it is straightforward to transmit the data in serial form for some distance to the computer with a minimum of wiring. The serial data can be reassembled into a word at the computer interface and then read into a parallel port. Dual-slope integrating converters must be commanded to start a conversion, and the result is available in parallel form after a somewhat arbitrary (and long!) time delay.

Plug-in circuit boards with a 12-bit successive-approximation A/D converter are available for most computer buses; these usually include a multiplexer which typically allows one of 16 analog inputs to be selected for measurement. These types of boards are designed to monitor voltages, temperatures, etc., and are not well suited to photometric measurements. To achieve accuracies in the 14- to 16-bit range with multielement detectors it will probably be necessary to design and build a custom data acquisition system. For single element detectors, it should be possible to adapt a precision laboratory dual-slope A/D converter (i.e., a digital voltmeter) which is already packaged. These units provide guarded differential inputs (usually for at least two signals so that they can be used in ratio mode), and have provisions for automatic range changing, zero offset cancellation, and other frills. The laboratory instruments often include an IEEE-488 computer interface which allows remote control of the sampling intervals and voltage ranges and serves for data acquisition. IEEE-488 boards are available for all the popular mini- and microcomputers.

Testing and Calibrating A/D Converters

The established manufacturers of A/D converters have sophisticated test equipment to evaluate their products under variationsin temperature, load, duty cycle, component aging, etc., which would be pointless to attempt to duplicate in the laboratory. Nevertheless, some tests should be performed to verify that the converter is working properly in the first place and that it works as well now as it used to. A couple of simple test procedures are described below. The ultimate test, of course, is the consistency of results at the telescope.

To test an A/D converter it will be helpful to have two simple programs available in the data acquision computer. The first should sample the converter output at regular intervals (say ten times/sec), convert the binary number to decimal, and display both the decimal number and the bit pattern on the screen. The second should sample the converter continuously at as high a rate as possible for a programmable interval and then display a histogram of the frequency of the individual values measured.

To perform a static gain and linearity test, connect a stable DC power supply and a good laboratory digital voltmeter (DVM) to the converter input. Run the program which gives the continuous display of the converter output. Do the DVM and converter numbers agree, or are they at least in constant proportion? Are any values missing or just nonsense? Check that bits change in the proper sequence as the voltage is varied to verify that the wires going from the converter to the computer are not flipped, and that there are no shorted or open wires. If an absolute voltage standard is available, and the converter has a calibration adjustment, adjust it so that the converter and DVM agree (usually at about 3/4 of full scale for optimal linearity). For more serious testing, a precision D/A converter can be interfaced to the computer and the entire procedure automated.

To perform a differential linearity test, feed the signal from an analog (not digital!) function generator to the input of the converter. Set the generator to produce a sawtooth which spans at least the digitization range of the converter at a frequency of a few hertz or less, and run the histogram program. After accumulating data long enough to get reasonable statistics, plot the histogram. The frequency distribution should be reasonably uniform except at the ends of the scale where the sawtooth exceeded the range of the converter. If the device is a successive-approximation converter, there will be regular patterns apparent in the data at intervals of 4, 8, 16, etc., data numbers. If there are huge peaks in the distribution followed by blank channels, the differential linearity is bad and the converter need to be readjusted or replaced. No differential linearity effects should be detected with an integrating converter; if these are present it indicates that there is feedback from the digital outputs back into the input signal (perhaps from the computer). Note that large scale trends in the histogram more likely reflect nonlinearities in the function generator than in the A/D converter.

Summary and Recommendations

Three types of A/D converters have been discussed: parallel, successive-approximation, and integrating. The parallel type does not yet produce the accuracy needed for photometric measurements, but could become interesting in the future. A successive-approximation converter is the proper choice if conversion times in the microseconds are required. If conversion speed is not a factor, a successive-approximation converter could still be used, but an integrating converter has better differential linearity and will probably be easier to integrate into a detector system.

For additional general discussion, the books by Bruck (1974), Jung (1979), and Sheingold (1980) are recommended. The trade magazines Electronics, Electronic Design, and EDN, often review the state of the art in A/D and D/A converters. The data sheets and product description books from manufacturers such as Analog Devices, Analogic, Burr-Brown, Datel, Hybrid Systems, Intech, and Teledyne Philbrick are very useful. Integrating A/D converter systems and calibration equipment are available from laboratory instrument manufacturers such as Beckman Instruments, Fluke, Hewlett-Packard, Keithley, and Tektronix. Read product descriptions with care!

Achieving good performance with A/D converters involves more than just buying a fast converter with a lot of bits. A 14-bit converter operated properly near the upper part of its conversion range will produce much better data than an improperly operated 16-bit converter. Computers greatly simplify data acquisition and reduction, but the performance of a detector system ultimately depends not on digital wizardry but on the care taken in conditioning and digitizing the analog signals coming from the detector.

References

Bruck, D. B., Data Conversion Handbook, Hybrid Systems, Burlington, Mass., 1974.

Jung, W. G., IC Converter Cookbook, Howard W. Sams, Indianapolis, Ind., 1979.

Sheingold, D. H. (Ed.), Analog-Digital Conversion Notes, Analog Devices, Norwood, Mass., 1980.