Chapter 5
Interface Electronic Circuits

Engineers like to solve problems.
If there are no problems handily available,
they will create their own problems.

-Scott Adams

5.1 Input Characteristics of Interface Circuits

A system designer is rarely able to connect a sensor directly to processing, monitoring, or recording instruments, unless a sensor has a built-in electronic circuit with an appropriate output format. When a sensor generates an electric signal, that signal often is either too weak, or too noisy, or it contains undesirable components. Besides, the sensor output may be not compatible with the input requirements of a data acquisition system, that is, it may have a wrong output format. To mate a sensor and a processing device, they either must share a "common value" or some kind of a "mating" device is required in-between. In other words, signal from a sensor usually has to be *conditioned* before it is fed into a processing device (a load). Such a load usually requires either voltage or current as its input signal.

An interface or a *signal conditioning* circuit has a specific purpose: to bring signal from the sensor up to the format that is compatible with the load device. Figure 5.1 shows a stimulus that acts on a sensor, which is connected to a load through an interface circuit. To do its job effectively, an interface circuit must be a faithful slave of two masters: the sensor and the load device. Its input characteristics must be matched to the output characteristics of the sensor and its output must be interfaceable with the load. This book, however, focuses on the sensors, therefore, below we will discuss only the front stages of the interface circuits. Also, we will discuss some typical excitation circuits that are required for active sensors, that is, for the sensors which need electrical signals to produce electrical outputs.

The input part of an interface circuit may be specified through several standard numbers. These numbers are useful for calculating how accurately the circuit can

J. Fraden, *Handbook of Modern Sensors*,
DOI 10.1007/978-1-4419-6466-3_5, © Springer Science+Business Media, LLC 2010

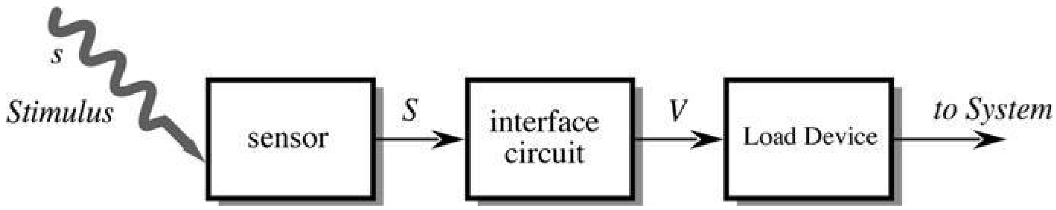


Fig. 5.1 Interface circuit matches the signal formats of a sensor and a load device

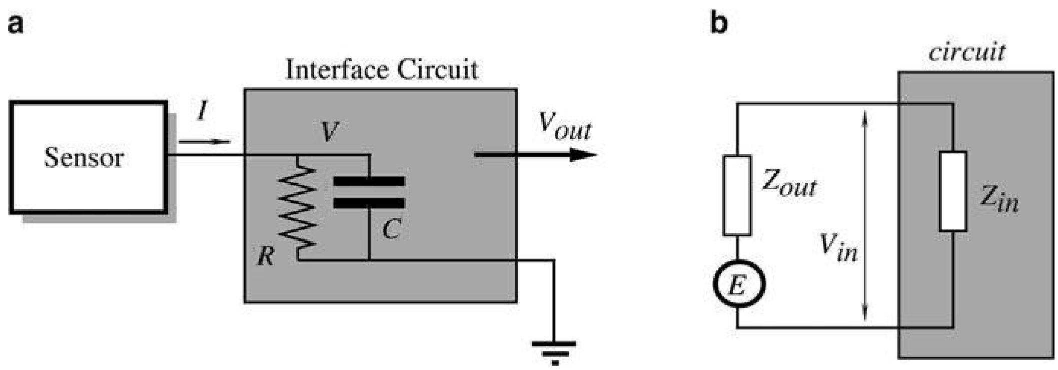


Fig. 5.2 Complex input impedance of an interface circuit (**a**), and equivalent circuit of a voltage generating sensor (**b**)

process the sensor's output signal and what would be the circuit's contribution to a total error budget?

The input impedance; shows by how much the circuit loads the sensor. The impedance may be expressed in a complex form as:

$$\mathbf{Z} = \frac{\mathbf{V}}{\mathbf{I}}, \tag{5.1}$$

where **V** and **I** are complex notations for the voltage and the current across the input impedance. For example, if the input of a circuit is modeled as a parallel connection of input resistance, R and input capacitance, C (Fig. 5.2a), the complex input impedance may be represented as

$$\mathbf{Z} = \frac{R}{1 + j\omega RC}, \tag{5.2}$$

where ω is the circular frequency and $j = \sqrt{-1}$ is the imaginary unity. At very low frequencies, a circuit having a relatively low input capacitance and resistance has an input impedance, which is almost equal to the input resistance: $\mathbf{Z} \approx R$. Relatively low, here it means that the reactive part of the above equation becomes small, i.e., the following holds:

$$RC << \frac{1}{\omega}. \tag{5.3}$$

Whenever an input impedance of a circuit is considered, the output impedance of the sensor must be taken into account. For example, if the sensor is of a capacitive nature, to define a frequency response of the input stage, sensor's capacitance must be connected in parallel with the circuit's input capacitance. Formula (5.2) suggests that the input impedance is function of the signal frequency. With an increase in the signal rate of change, the input impedance becomes lower.

Figure 5.2b shows an equivalent circuit for a voltage generating sensor. The circuit is comprised of the sensor output, $\mathbf{Z}_{out}$, and the circuit input, $\mathbf{Z}_{in}$, impedances. The output signal from the sensor is represented by a voltage source, e, which is connected in series with the output impedance. Instead of a voltage source, for some sensors it is more convenient to represent the output signal as outgoing from a current source, which would be connected in parallel with the sensor output impedance. Both representations are equivalent to one another, so we will use voltage. Accounting for both impedances, the circuit input voltage, V_{in} is represented as

$$V_{in} = e\frac{Z_{in}}{Z_{in} + Z_{out}}. \tag{5.4}$$

In any particular case, an equivalent circuit of a sensor should be defined. This helps to analyze the frequency response and the phase lag of the sensor-interface combination. For instance, a capacitive detector may be modeled as a pure capacitance connected in parallel with the input impedance. Another example is a piezoelectric sensor which can be represented by a very high resistance (on the order of $10^{11}\ \Omega$) shunted by a capacitance (in the order of 10pF).

To illustrate the importance of the input impedance characteristics, let us consider a purely resistive sensor connected to the input impedance as shown in Fig. 5.2. The circuit's input voltage as function of frequency, f, can be expressed by a formula

$$V = \frac{E}{\sqrt{1 + \left(\frac{f}{f_c}\right)^2}} \tag{5.5}$$

where $f_c = (2\pi RC)^{-1}$ is the corner frequency, (i.e., the frequency where the amplitude drops by 3 dB). If we assume that a 1% accuracy in the amplitude detection is required, then we can calculate the maximum stimulus frequency that can be processed by the circuit:

$$f_{max} \approx 0.14 f_c, \tag{5.6}$$

or $f_c \approx 7f_{max}$; that is, the impedance must be selected in such a way as to assure a sufficiently high corner frequency. For example, if the stimulus' highest frequency is 100 Hz, the corner frequency must be selected at least at 700 Hz. In practice, f_c is selected even higher, because of the additional frequency limitations in the subsequent circuits.

One should not overlook a speed response of the front stage of the interface circuit. Operational amplifiers, which are the most often used building blocks of interface circuits, usually have limited frequency bandwidths. There are the so-called programmable operational amplifiers, which allow the user to control (to program) the bias current and, therefore, the first stage frequency response. The higher the current, the faster would be the response.

Figure 5.3 is a more detailed equivalent circuit of the input properties of an interface circuit, for instance, an amplifier or an A/D converter. The circuit is characterized by the input impedance $\mathbf{Z}_{in}$ and several generators. They represent voltages and currents that are generated by the circuit itself. These signals are spurious and may pose substantial problems if not handled properly. All these interfering signals are temperature-dependent.

Voltage e_o is called the input *offset voltage*. If the input terminals of the circuit are shorted together, that voltage would simulate a presence of an input dc signal having a value of e_o. It should be noted that the offset voltage source is connected in series with the input and its resulting error is independent of the output impedance of the sensor.

The input *bias current* i_o is also internally generated by the circuit. Its value is quite high for many bipolar transistors, much smaller for the JFET s, and even more lower for the CMOS circuits. This current may present a serious problem when a circuit or sensor employs high-impedance components. The bias current passes through the input impedance of the circuit and the output impedance of the sensor, resulting in a spurious voltage drop. This voltage may be of a significant magnitude. For instance, if a piezoelectric sensor is connected to a circuit having an input resistance of 1 GΩ (10^9 Ω) and the input bias current of 1 nA (10^{-9}A), the voltage drop at the input becomes equal to 1 GΩ·1 nA=1 V, a very high value indeed. In contrast to the offset voltage, the error resulting from bias current is proportional to the output impedance of the sensor. This error is negligibly small for the sensors having low output resistances. For instance, an inductive detector is not sensitive to a magnitude or variations in the bias current.

A circuit board *leakage current* may be a source of errors while working with high-impedance circuits. This current may be the result of lower surface resistance in the printed circuit board (PCB). Possible causes for that are: poor quality PCB material, surface contamination with solder flux residue (a poorly washed board),

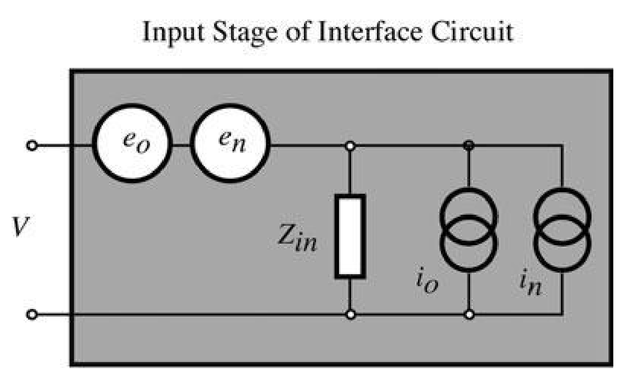


Fig. 5.3 Equivalent circuit of electrical noise sources at an input stage

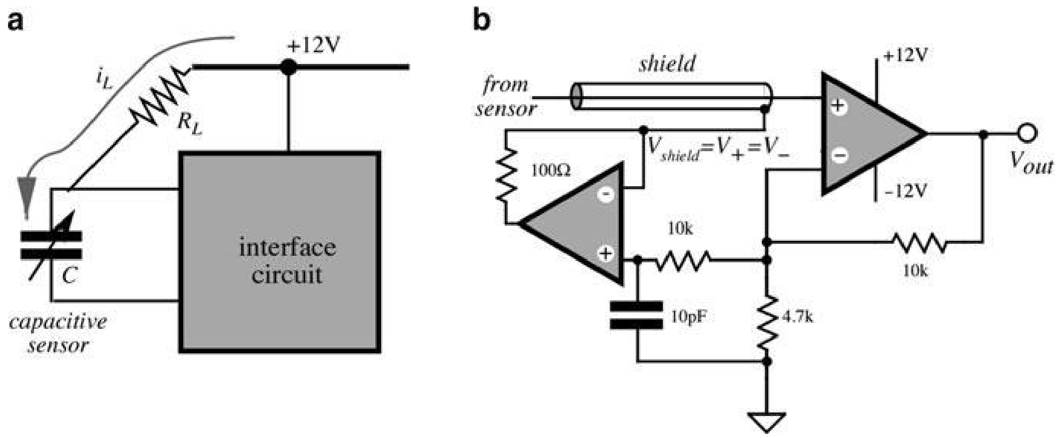


Fig. 5.4 Circuit board leakage affects input stage (**a**); driven shield of the input stage (**b**)

moisture, and degraded conformal coating. Figure 5.4a shows that a power supply bus and the board resistance, R_L, may cause leakage current, i_L, through the sensor's output impedance. If the sensor is capacitive, its output capacitance will be very quickly charged by the leakage current. This will not only cause an error, but may even lead to the sensor's destruction, especially if the sensor uses some chemical compound (e.g., a resistive moisture sensor).

There are several techniques known to minimize the board leakage current effect. One is a careful board layout to keep higher voltage conductors away from the high-impedance components. A leakage through the board thickness in multi-layer boards should not be overlooked. Another method is electrical guarding, which is an old trick. The so-called driven shield is also highly effective. Here, the input circuit is surrounded by a conductive trace that is connected to a low-impedance point at the same potential as the input. The guard absorbs the leakage from other points on the board, drastically reducing currents that may reach the input terminal. To be completely effective, there should be guard rings on both sides of the printed circuit board. As an example, an amplifier is shown with a guard ring, driven by a relatively low impedance of the amplifier's inverting input.

It is highly advisable to locate the high-impedance interface circuits as close as possible to the sensors. However, sometimes connecting lines can not be avoided. Coaxial shielded cables with good isolation are recommended [1]. Polyethylene or virgin (not reconstructed) Teflon is best for the critical applications. In addition to potential insulation problems, even short cable runs can reduce bandwidth unacceptably with high source resistances. These problems can be largely avoided by bootstrapping the cable's shield. Figure 5.4b shows a voltage follower connected to the inverting input of an amplifier. The follower drives the shield of the cable, thus reducing the cable capacitance, the leakage and spurious voltages resulting from cable flexing. A small capacitance at the follower's noninverting input improves its stability.

Another problem that must be avoided is connecting to the input of an amplifier any components, besides a sensor, that potentially may cause problems. An example of such a “troublemaker” is a ceramic capacitor. In a hope to filter out high-frequency transmitted noise at the input, a designer quite frequently uses filter capacitors either at the input, or in the feedback circuit, of an input stage. If for a cost-saving or space saving reason she selects a ceramic capacitor, she may get what is not expecting. Many capacitors possess the so-called dielectric absorption properties, which are manifested as a memory effect. If such a capacitor is subjected to a charge spike either from a sensor, or from a power supply, or just from any external noise source, the charge will alter the capacitor’s dielectric properties in such a way as a capacitor now behaves like a small battery. That “battery” may take a long time to lose its charge: from few seconds to many hours. The voltage generated by that “battery” is added to the sensor’s signal and may cause significant errors. If a capacitor must be employed at the input stage, a film capacitor should be used instead of ceramic.

5.2 Amplifiers

Most passive sensors produce weak output signals. The magnitudes of these signals may be on the order of microvolts (μV) or picoamperes (pA). On the other hand, standard electronic data processors, such as A/D converters, frequency modulators, data recorders, etc. require input signals of sizable magnitudes on the order of volts (V) and milliamperes (mA). Therefore, an amplification of the sensor output signals has to be made with a voltage gain up to 10,000 and a current gain up to 1 million. Amplification is part of a signal conditioning. There are several standard configurations of amplifiers that might be useful for the amplifying signals from various sensors. These amplifiers may be built of discrete components, such as semiconductors, resistors, capacitors, and inductors. Alternatively, the amplifiers are frequently composed of standard building blocks, such as operational amplifiers and various discrete components.

It should be clearly understood that a purpose of an amplifier is much broader than just increasing the signal magnitude. An amplifier may be also an impedance-matching device, an enhancer of a signal-to-noise ratio, a filter, and an isolator between input and output.

5.2.1 Operational Amplifiers

One of the principle building blocks for the amplifiers is the so-called *operational amplifier* (OPAM), which is either an integrated (monolithic) or hybrid (a combination of monolithic and discrete parts) circuit. An integrated OPAM may contain hundreds of transistors, as well as resistors and capacitors. An analog circuit designer, by arranging around the OPAM discrete components (resistors, capacitors, inductors, etc.), may create an infinite number of useful circuits, not only the amplifiers, but many others circuits as well. Operational amplifiers are also used as cells in custom-made integrated circuits of the analog or mixed technology types. These circuits are called *application-specific integrated circuits* or ASICs for short. Below, we will describe some typical circuits with OPAM, which are often used in conjunction with various sensors.

As a building block, a good operational amplifier has the following properties (a schematic representation of OPAM is shown in Fig. 5.5):

– Two inputs: one is inverting (−) and the other is noninverting (+);
– A high input resistance (on the order of hundreds of MΩ or even GΩ);
– A low output resistance (a fraction of Ω);
– An ability to drive capacitive loads;
– A low input offset voltage e_o (few mV or even μV);
– A low input bias current i_o (few pA or even less);
– a very high *open loop gain* A_{OL} (at least 10^4 and preferably over 10^6). That is, the OPAM must be able to magnify (amplify) a voltage difference V_{in}, between its two inputs by a factor of A_{OL};
– a high common mode rejection ratio (CMRR). That is, the amplifier suppresses the in-phase equal magnitude input signals (common-mode signals) V_{CM} applied to its both inputs;
– low intrinsic noise;
– a broad operating frequency range;
– a low sensitivity to variations in the power supply voltage;
– a high environmental stability of its own characteristics.

For the detailed information and the application guidance the user should refer to data sheets and catalogues published by the respective manufacturers. Such catalogues usually contain the selection guides for every important feature of an OPAM. For instance, OPAMs are grouped by such criteria as low offset voltages, low bias currents, low noise, bandwidth, etc.

Figure 5.5a depicts an operational amplifier without any feedback components. Therefore, it operates under the so-called open-loop conditions. An open loop gain,

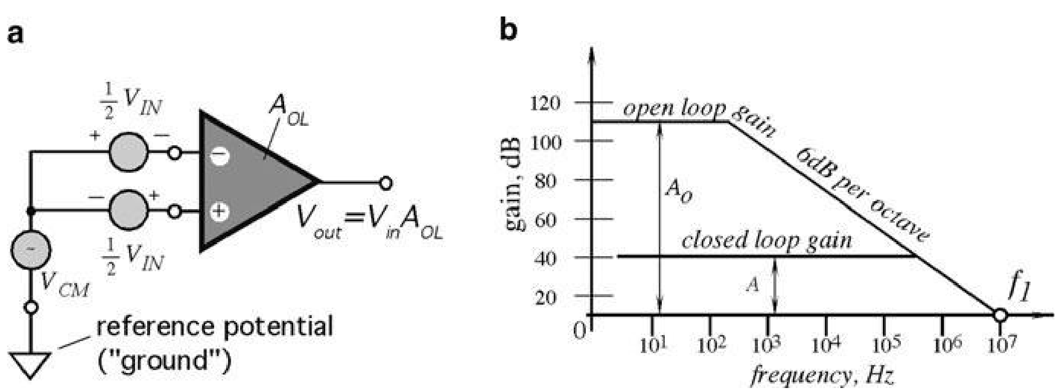


Fig. 5.5 General symbol of an operational amplifier (**a**), and gain/frequency characteristic of an OPAM (**b**)

A_{OL}, of an OPAM is always specified but is not a very stable parameter. Its frequency dependence may be approximated by a graph of Fig. 5.5b. The A_{OL} changes with the load resistance, temperature, and the power supply fluctuations. Many amplifiers have an open loop gain temperature coefficient on the order of 0.2 to 1%/°C and the power supply gain sensitivity on the order of 1%/%. An OPAM is very rarely used with an open loop (without the feedback components) because the high open-loop gain may result in circuit instability, a strong temperature drift, noise, etc. For instance, if the open-loop gain is 10^5, the input voltage drift of 10 μV would cause the output drifts by about 1 V.

The ability of an OPAM to amplify small-magnitude high-frequency signals is specified by the gain-bandwidth product (GBW), which is equal to the frequency f_1 where the amplifier gain becomes equal to unity. In other words, above the f_1 frequency, the amplifier cannot amplify. Figure 5.6a depicts a noninverting amplifier where resistors R_1 and R_2 define the feedback loop. The resulting gain $A = 1 + R_2/R_1$ is a closed-loop gain. It may be considered constant over a much broader frequency range [see Fig. 5.5b], however, f_1 is the frequency limiting factor regardless of the feedback. A linearity, gain stability, the output impedance, and gain accuracy are all improved by the amount of feedback and now depend mainly on characteristics of the feedback components. As a general rule for moderate accuracy, the open loop gain of an OPAM should be at least 100 times greater than the closed loop gain at the highest frequency of interest. For even higher accuracy, the ratio of the open and closed loop gains should be 1,000 or more.

A typical data sheet for an OPAM specifies the bias and offset voltages. Due to limitations in manufacturing technologies, any OPAM acts not only as a pure amplifier, but as a generator of voltages and currents, which may be related to its input [Fig. 5.6b]. Since these spurious signals are virtually applied to the input terminals, they are amplified along with the useful signals.

Because of offset voltages and bias currents, an interface circuit does not produce zero output when zero input signal is applied. In dc-coupled circuits, these undesirable input signals may be indistinguishable from the useful signal. If the input offset voltage is still too large for the desired accuracy, it can be trimmed

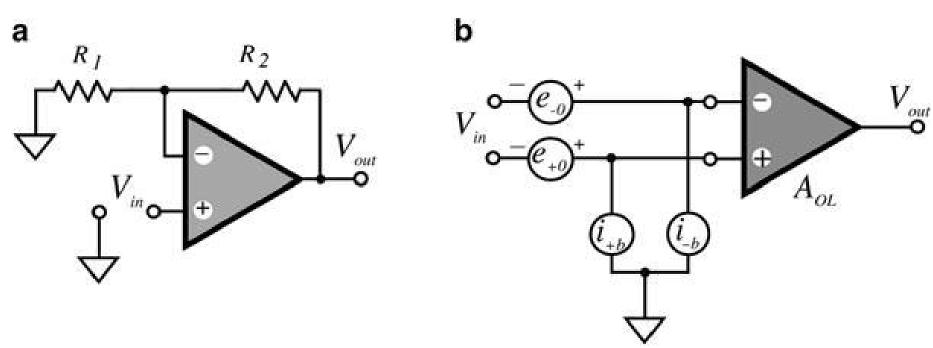


Fig. 5.6 Noninverting amplifier (**a**); offset voltages and bias currents in an operational amplifier are represented by generators connected to its inputs (**b**)

out either directly at the amplifier (if the amplifier has dedicated trimming terminals) or in the independent offset compensation circuit.

An application engineer should be concerned with, the output offset voltage, which can be derived from formula:

$$V_o = A(e_o + i_o R_{eqv}) \tag{5.7}$$

where R_{eqv} is the equivalent resistance at the input (a combination of the sensor's output resistance and the input resistance of the amplifier), e_o is the input offset voltage, and i_o is the input bias current. The offset is temperature-dependent. In circuits where the amplifier has high gain, the output voltage offset may be a source of substantial error. There are several ways to handle this difficulty. Among them is selecting an amplifier with low bias current, high input resistance, and low offset voltage. Chopper-stabilized amplifiers are especially efficient for reduction of offset voltages.

5.2.2 Voltage Follower

A voltage follower (Fig. 5.7) is a an electronic circuit that provides impedance conversion from a high to low level. A typical follower has high input impedance (the high input resistance and the low input capacitance) and low output resistance (the output capacitance makes no difference). A good follower has a voltage gain very close to unity (typically, 0.999 at lower frequencies) and a high current gain. In essence, it is a current amplifier and impedance converter. Its high input and low output impedances make it indispensable for interfacing between many sensors and signal processing devices.

A follower, when connected to a sensor, makes very little effect on the latter's performance, thus providing a buffering function between the sensor and the load. When designing a follower, these tips might be useful:

- For the current-generating sensors, the input bias current of the follower must be at least 100 times smaller than the sensor's current.
- The input offset voltage must be either trimmable or smaller than the required LSB.
- The temperature coefficient of the bias current and the offset voltage should not result in errors of more than 1 LSB over an entire temperature range.

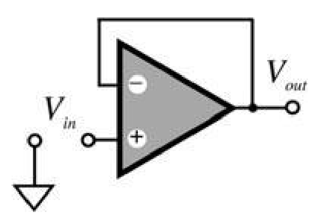


Fig. 5.7 Voltage follower with an operational amplifier

5.2.3 Instrumentation Amplifier

An instrumentation amplifier (IA) has two inputs and one output (Fig. 5.8). It is distinguished from an operational amplifier by its finite gain (which is usually no more than 100) and the availability of both inputs for connecting to the signal sources. The latter feature means that all necessary feedback components are connected to other parts of the instrumentation amplifier, rather than to its non-inverting and inverting inputs. The main function of the IA is to produce an output signal which is proportional to the difference in voltages between its two inputs:

$$V_{out} = a(V_+ - V_-) = a\Delta V, \tag{5.8}$$

where V_+ and V_- are the input voltages at noninverting and inverting inputs respectively, and a is the gain. It is important to assure high input resistances for both inputs, so that the amplifier can be used in a true differential form. A differential input of the amplifier is very important for rejection of common mode interferences having an additive nature (see Sect. 5.10). Thus, the IA should have a high common-mode rejection ratio (CMRR), that is, its output signal should be insensitive to the value of V_+ or V_- but responsive only to their difference.

An instrumentation amplifier can be either built from several discrete OPAMs, in a monolithic or hybrid forms. Presently, instrumentation amplifiers are readily available from many manufacturers in form of monolithic chips. An example of a high-quality monolithic instrumentation amplifier is INA118 from Burr-Brown/ Texas Instruments (www.ti.com). It offers low offset voltage of 50 μV (the input signal produced by the IA when both inputs are connected together) and high CMRR (110 dB). The gain is programmed by a single resistor. Also, many integrated circuits, such as microcontrollers or DSP (digital signal processors) have built-in input instrumentation amplifiers for direct interface between the input sensors and the internal A/D (analog-to-digital) converters.

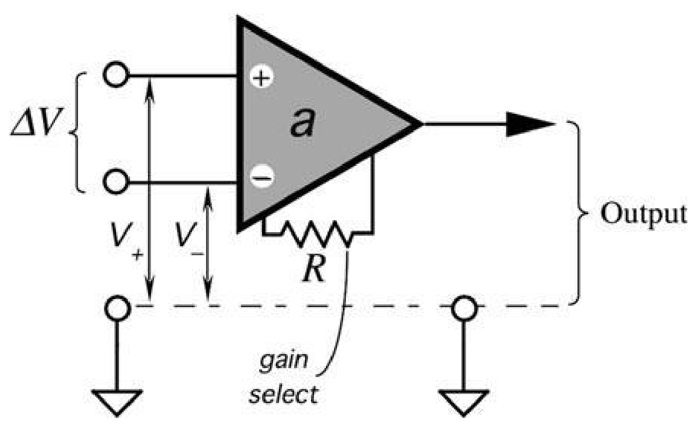


Fig. 5.8 Instrumentation amplifier

5.2.4 Charge Amplifiers

Charge amplifiers (CA) is a very special class of circuits, which must have extremely low bias currents. These amplifiers are employed to convert to voltage signals from the capacitive sensors, quantum detectors, pyroelectric sensors, and other devices, which generate very small charges (on the order of pico-coulombs, pC) or currents (on the order of pico-amperes). A basic circuit of a charge-to-voltage converter is shown in Fig. 5.9a. A capacitor, C, is connected into a feedback network of an OPAM. Its leakage resistance r must be substantially larger than the impedance of the capacitor at the lowest operating frequency. A good film capacitor is usually recommended along with a good quality printed circuit board where the components are coated with conformal coating.

A transfer function of the converter is

$$V_{out} = -\frac{\Delta Q}{C}. \tag{5.9}$$

Special integrated charge sensitive preamplifiers are commercially available for precision applications.

Many sensors can be modeled as capacitors. Some capacitive sensors are active, that is, they require an excitation signal. Examples are the microphones, capacitive force, and pressure transducers and humidity detectors. Other capacitive sensors are passive, that is they directly convert a stimulus into an electric charge or current. Examples are the piezoelectric and pyroelectric detectors. There are also noncapacitive sensors that still can be considered as current generators. An example is a photodiode.

A current generating sensor is modeled by a leakage resistance, r, connected in parallel with a current generator that has an infinitely high internal resistance (Fig. 5.10). The sensor generates current, i, which has two ways to outflow: to the sensors leakage resistance, r, as current, i_o, and the other, i_{out}, toward the interface circuit input impedance, Z_L. Naturally, current i_o is useless and to minimize the

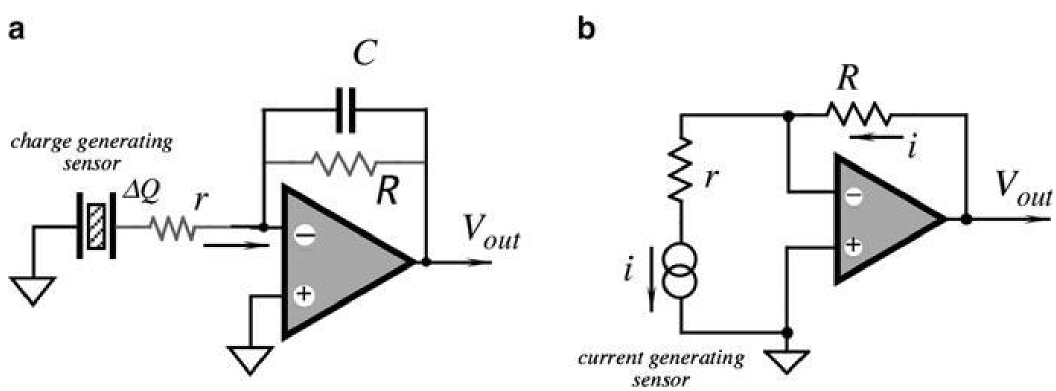


Fig. 5.9 Charge-to-voltage (**a**) and current-to-voltage (**b**) converters

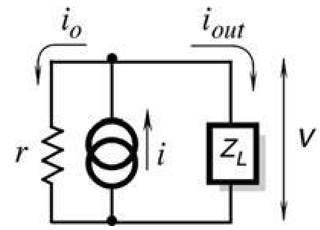


Fig. 5.10 An equivalent circuit of a current-generating sensor

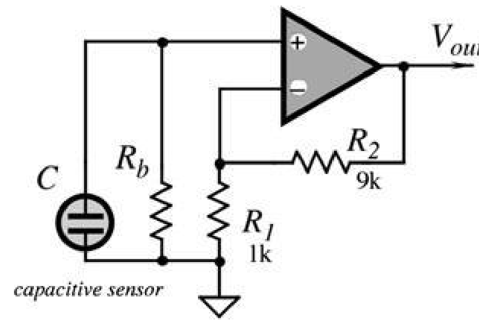


Fig. 5.11 Noninverting current-to-voltage converter

error of the current-to-voltage conversion, the leakage resistance of the sensor must be much larger than the impedance of the interface circuit (Fig. 5.11).

Ohm's law suggests that to convert electric current i_{out} into voltage, current should pass through an appropriate impedance and the voltage drop across that impedance is proportional to the magnitude of the current. Figure 5.9b shows a basic current-to-voltage converter where the current-generating sensor is connected to the inverting input of an OPAM, which serves as a virtual ground. In other words, voltage at the inverting input is almost equal to that at the noninverting input, which is grounded. The sensor operates at nearly zero voltage across its terminals and its current defines the output voltage of the OPAM:

$$V_{out} = -iR \tag{5.10}$$

Resistor, $r << R$ is often required for the circuit stability because at high frequencies, without that resistor the OPAM would operate near the open loop gain, which may result in oscillations. This is especially true than the sensor has small internal resistance. The advantage of the virtual ground is that the output signal does not depend on the sensor's capacitance. The circuit produces voltage whose phase is shifted by 180° with respect to the current. A noninverting circuit shown in Fig. 5.12a can convert and amplify the signal, however, its speed response depends on both the sensor's capacitance and the converting resistor R_1. Thus, the response to a step function in a time domain can be described by

$$V_{out} = iR_b\left(1 + \frac{R_2}{R_1}\right)(1 - e^{-\frac{t}{rC}}). \tag{5.11}$$

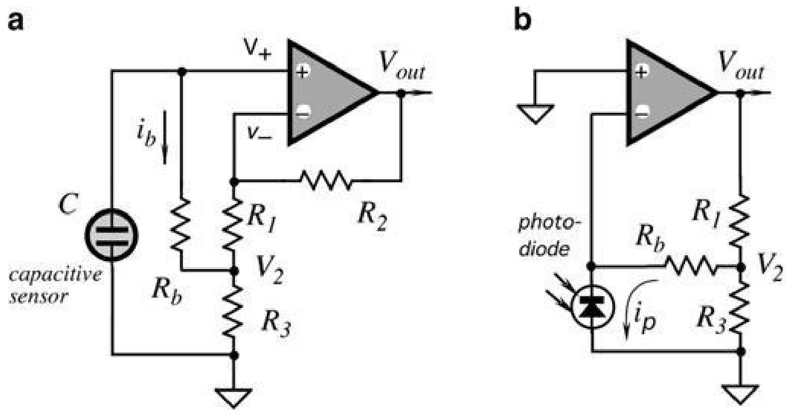


Fig. 5.12 Resistance multipliers (**a**) and (**b**)

When converting currents from such sensors as piezo and pyroelectrics, the resistor R_b [R in circuit 5.9(B)] may be required on the order of tens or even hundreds of gigohms. In many cases, resistors of such high values may be not available or impractical to use due to poor environmental stability. A high ohmic resistor can be simulated by a circuit, which is known as a resistance multiplier. For a positive input of an amplifier is implemented by adding a positive feedback around the amplifier. Figure 5.12a shows that R_1 and R_3 form a resistive divider. Due to a high open loop gain of the OPAM, voltages at noninverting and inverting inputs are almost equal to one another: $V_+ \approx V_-$. As a result, voltage, V_2, at the divider is

$$V_2 = V_- \frac{R_3}{R_1 + R_3} \approx V_+ \frac{R_3}{R_1 + R_3}, \tag{5.12}$$

and current through the resistor is defined trough the voltage drop:

$$i_b = \frac{\Delta V}{R_b} = \frac{V_+ - V_2}{R_b} = \frac{V_+}{R_b} \frac{R_1}{R_1 + R_3}. \tag{5.13}$$

From this equation, the input voltage can be found as a function of the input current and the resistive network:

$$V_+ = i_b R_b \left(1 + \frac{R_3}{R_1}\right). \tag{5.14}$$

It's seen that the resistor R_b is multiplied by a factor of $(1 + R_3/R_1)$. For example, if the highest resistor you may consider is 10 MΩ, by selecting the multiplication factor in parenthesis of say 5, you'll get a virtual resistance of 50 MΩ. Resistance multiplication, while being a powerful trick, should be used with caution. Specifically, noise, bias current and offset voltage, all of them are also multiplied by the same factor ($1 + R_3/R_1$), which may be undesirable in some applications. Further, since the network forms a positive feedback, it may cause circuit instability. Therefore, in practical circuits, a resistance multiplication should be limited to a factor of 10. If a resistance multiplication is required for a negative input of an amplifier, the circuit of Fig. 5.12b comes in handy. the resistance multiplication is governed by the same (5.14).

5.3 Light-to-Voltage Converters

Light-to-voltage converters are based on combination of photosensors and current-to-voltage converter circuits. For detecting extremely low-intensity light, typically one or several photons, the photomultipliers are generally employed (Chap. 15), however, for less demanding applications three types of a photosensor are available: a photodiode, phototransistor, and photoresistor (Chap. 14). They employ a photoeffect that was discovered by Albert Einstein and won him the Nobel Prize. The difference between a photodiode and a phototransistor is in construction of the semiconductor chip. A photodiode has one p-n junction, while a phototransistor has two junctions where the base of the transistor may be floating or may have a separate terminal. The base current is a photo-induced current that is multiplied by the transistor's β to produce the collector current. Thus, a phototransistor is equivalent to a photodiode with a built-in current amplifier.

From the electrical point of view, a photodiode can be represented by an equivalent circuit shown in Fig. 5.13a. It consists of a current generator (internal input impedance is infinitely large), a parallel regular diode (like a rectifier diode), resistance of the junction R_j, capacitance of the junction C_j, and a serial resistance R_s. The current generator generates a photocurrent proportional to the photon flux. This current flows in the direction from the cathode (−) to the anode (+) of the photodiode. Note that for very strong illuminations, the photocurrent will start flowing through a nonlinear rectifier diode, which will degrade a linearity.

A photodiode can be used in voltaic or current modes. In the voltaic mode, a photodiode is connected to a very high resistor (10^7–10^9 Ω) and a good voltage

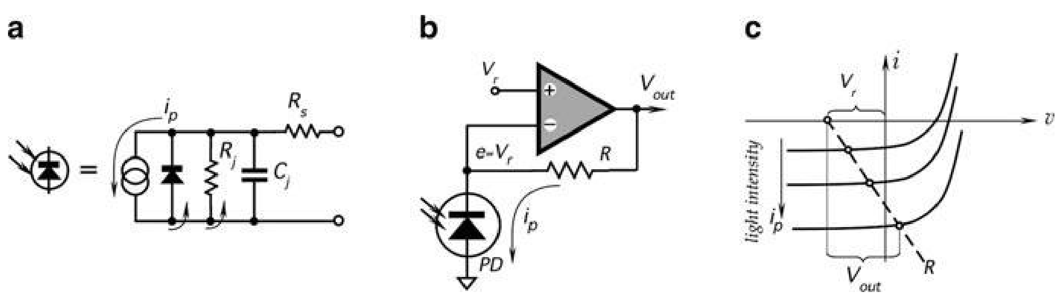


Fig. 5.13 Equivalent circuit of a photo-diode (**a**) A reverse-biased photo-diode with a current-to voltage converter; (**b**) load diagram of the circuit (**c**)

amplifier. The diode will work like a battery with voltage proportional to the light intensity. This voltage is the result of a photocurrent i_p passing through the internal junction resistance R_j. In a current mode, the photodiode is virtually shorted (a voltage across the diode is zero) and current i_p is drawn to the current-to-voltage converter as described below. This node is more popular, especially for applications where a high-speed response is required.

A circuit with an operational amplifier is shown in Fig. 5.13b. Note that the reference voltage V_r creates a constant reverse bias across the photodiode. Figure 5.13c shows the operating points for a load resistor R. Features of the circuits used with a reverse-biased photodiode are high-speed response and wide-proportional-range of output. Therefore, this circuit is generally used. Another circuit with an operational amplifier has a zero-bias across the photodiode as shown in Fig. 5.14a. This arrangement provides a near-ideal short-circuit current in a wide operating range. The output voltage (V_{OUT}) is given as $V_{OUT} = i_p R$. Figure 5.14b shows the output voltage vs. radiant intensity (a transfer function). An arrangement with no bias and high-impedance loading to the photodiode provides less influence by dark current and a wide linear range of the photocurrent relative to the radiant intensity. It should be noted that for rather small illuminations to get a meaningful output (few hundred millivolts), a value of resistor R should be quite large on the order of 100 MΩ or even several GΩ. If such resistors are not available, a resistance multiplication circuit as shown in Fig. 5.12c may be cautiously employed. The multiplication factor should not be greater than 10 as all the bad things in the circuit are also multiplied: the offset voltage, bias current, and noise. When using the high-Ohmic resistors, the photosensor and interface circuit should be electrically shielded. Even a minute capacitive coupling of the environment to such resistors brings in a lot of interferences, especially from power lines (60 or 50 Hz).

The interface circuits for a phototransistor are similar, except that they have to provide a voltage across the collector-emitter terminals as shown in Fig. 5.15a. The transfer function of this circuit is shown in Fig. 5.15b. A phototransistor circuit is more sensitive to light but for the price of higher nonlinearity at stronger irradiances.

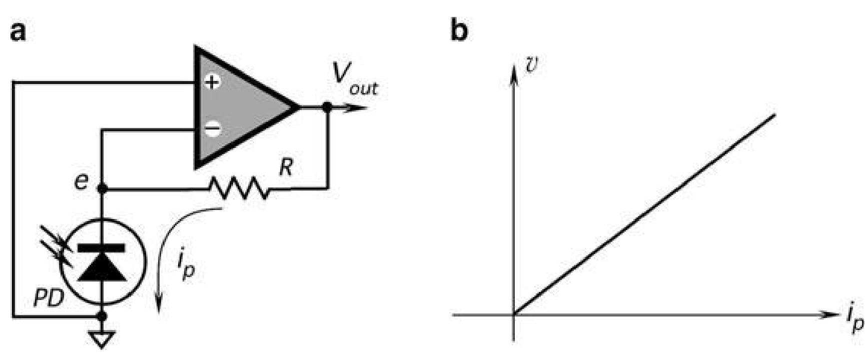


Fig. 5.14 A zero-biased photodiode with a current-to-voltage converter (**a**) and a transfer function (**b**)

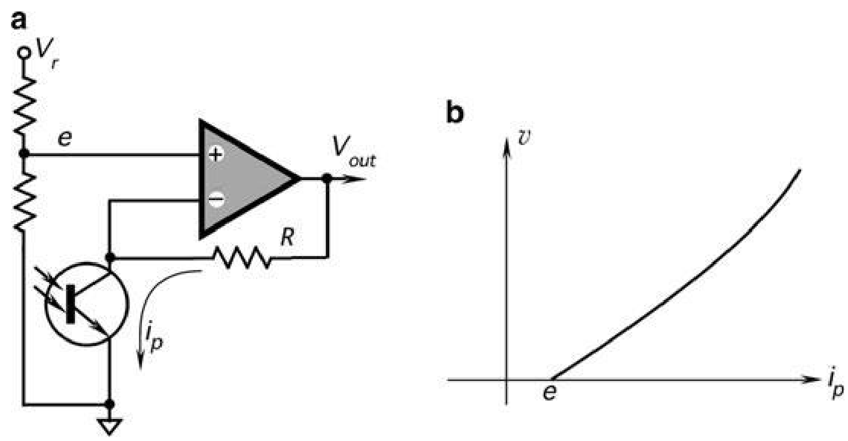


Fig. 5.15 Light-to-voltage converted with a photo-transistor (**a**); transfer function (**b**)

We are not describing here interface circuits for photoresistors because any suitable resistance measurement circuits can be used for the purpose. An example is the Wheatstone bridge circuits which we will discuss below.

5.4 Excitation Circuits

External power is required for operation of the *active* sensors. Examples are: temperature sensors (thermistors and RTDs), pressure sensors (piezoresistive and capacitive), and displacement (electromagnetic and resistive). The power may be delivered to a sensor in different forms. It can be a constant voltage, constant current, and sinusoidal or pulsing currents. It may even be delivered in the form of light or ionizing radiation. The name for that external power is an excitation signal. In many cases, stability and precision of the excitation signal directly relates to the sensor's accuracy and stability. Hence, it is imperative to generate the signal with such accuracy that the overall performance of the sensing system is not degraded. Below, we review several electronic circuits that feed sensors with appropriate excitation signals.

5.4.1 Current Generators

Current generators are often used as excitation circuits to feed sensors with predetermined currents that, within limits, are independent of the sensor properties, stimulus value, or environmental factors. In general terms, a current generator (current pump) is a device which produces electric current independent of the load impedance. That is, within the capabilities of the generator, amplitude of its output current must remain substantially independent of any changes in the impedance of the load. It is said that an ideal current source (generator) has infinitely high output resistance, so any series load value will not change anything. When generating current for a variable load, according to Ohm's law, the corresponding voltage must change in synch.

Usefulness of the current generators for the sensor interfaces is in their ability to produce excitation currents of precisely controlled magnitude and shape. Hence, a current generator should not only produce current which is load independent, but it also must be controllable from an external signal source (a wave-form generator), which in most cases has a voltage output. A good current generator must produce current that follows the control signal with high fidelity and is independent of the load over a broad range of impedances.

There are two main characteristics of a current generator: the output resistance and the voltage compliance. The output resistance should be as high as practical. A voltage compliance is the highest voltage that can be developed across the load without affecting the output current. For a high resistive load, according to Ohm's law, a higher voltage is required for a given current. For instance, if the required excitation current is $i = 10$ mA and the highest load impedance at any given frequency is Z_L=10 kΩ, a voltage compliance of at least $iZ_L = 100$ V would be needed. Below, we cover some useful circuits with increased voltage compliance where the output currents can be controlled by external signals.

A unipolar current generator is called either a current source (generates the outflowing current), or a current sink (generates the in-flowing currents). Here, unipolar means that it can produce currents flowing in one direction only, usually toward the ground. Many of such generators utilize current-to-voltage characteristics of transistors. A voltage-controlled current source or sink may include an operational amplifier (Fig. 5.16). In such a circuit, a precision and stable resistor R_1 defines the output current, i_{out}. The circuit contains a feedback loop through the OPAM that keeps voltage across resistor R_1 constant and thus the constant current. To deliver a higher current at a maximum voltage compliance, a voltage drop across the sensing resistor R_1 should be as little as possible. In effect, that current is equal to V_1/R_1. For better performance, the current through the base of the output transistor should be minimized, hence, a field-effect rather than bipolar transistor is often used as an output current delivering device.

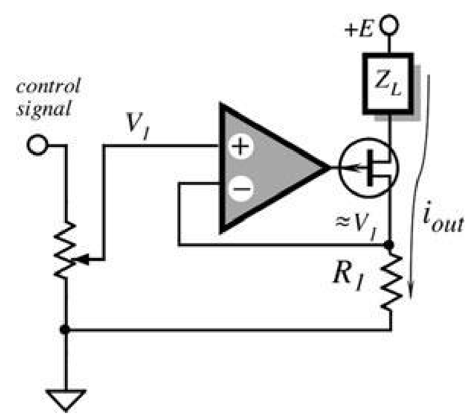


Fig. 5.16 Current sources with OPAM

It is well known hat the transistor's collector current is very little dependent on collector voltages. This feature was employed by the so-called current mirrors. A current mirror has one current input and at least one (may be several) current output. Therefore, the output current is controlled by the input current. The input current is supplied from an external source (like a voltage source plus a resistor) and should be of a known value. The so-called Wilson current mirrors have the control currents of about the same magnitude as the output currents, that is, they have an input-output ratio 1:1. Mirrors with other ratios also were designed, such as 1:2 and 1:4. Commercially available current mirrors may have a very broad current range, such as from 3 nA to 3 mA as in the integrated current mirror ADL5315 from Analog Devices.

For many sensors, bipolar current generators may be required. Such a generator provides a sensor with the excitation current, which may flow in both directions (in- and out-flowing). Figure 5.17 shows a noninverting (B) and inverting (A) circuits with an operational amplifier where the load is connected as a feedback. Current though the load Z_L, is equal to V_I/R_I and is load-independent. The load current follows V_I within the operating limits of the amplifier. An obvious limitation of the circuit is that the load is "floating," i.e., it is not connected to a ground bus or any other reference potential. For some applications, this is quite all right, however, many sensors need to be grounded or otherwise referenced. A circuit shown in Fig. 5.17b keeps one side of the load impedance near the ground potential, because a noninverting input of the OPAM is a virtual ground. Nevertheless, even in this circuit, the load is still fully isolated from the ground. One negative implication of this isolation may be an increased pick up of various kinds of transmitted noise.

In cases where the sensor must be grounded, a current pump invented by Brad Howland at MIT may be used [Fig. 5.18a]. The pump operation is based on utilizing both negative and positive feedbacks around the operational amplifier. The load is connected to the positive loop [2]. Current through the load is defined by

$$i_{out} = \frac{R_2}{R_1}\frac{(V_1 - V_2)}{R_5} \tag{5.15}$$

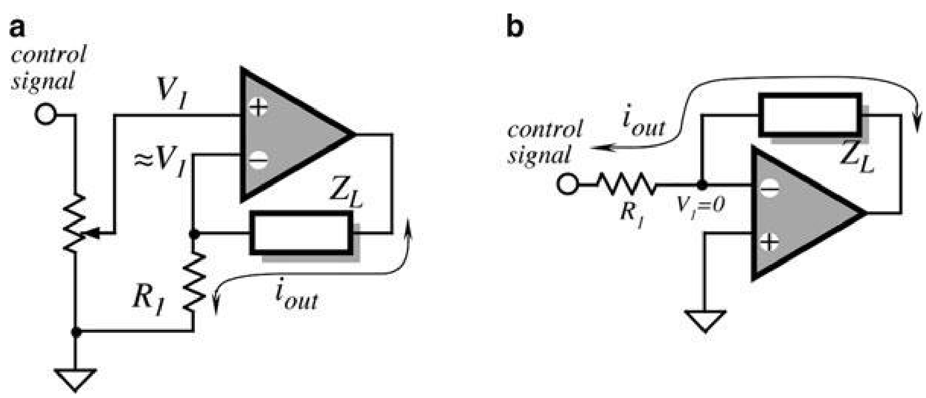


Fig. 5.17 Bipolar current generators with floating loads noninverting circuit (**a**); circuit with a virtual ground (**b**)

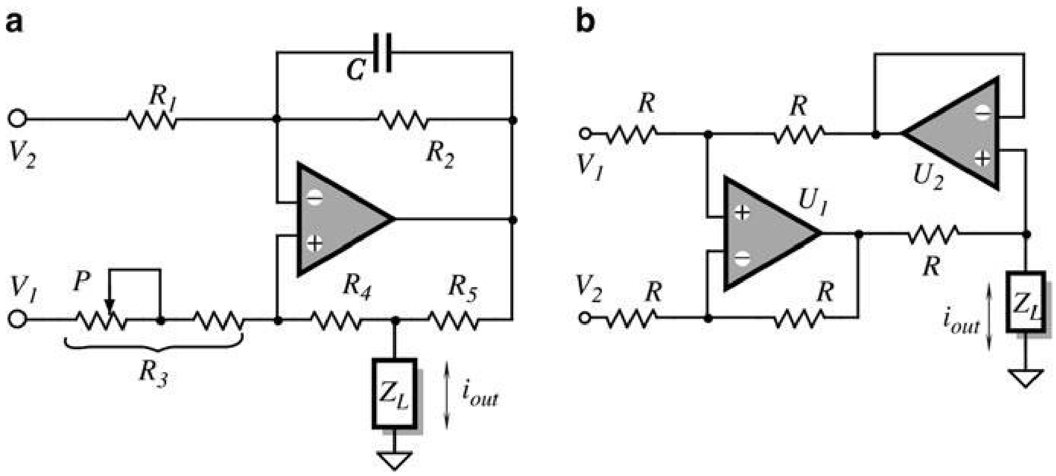


Fig. 5.18 Current generators with ground referenced loads Howland current pump (**a**); current pump with two OPAMs (**b**)

A trimming resistor, P, must be adjusted to assure that

$$R_3 = R_1 \frac{R_4 + R_5}{R_2} \tag{5.16}$$

In that circuit, each resistor may have a relatively high value (100 kΩ or higher), but the value for R_5 should be relatively small. This condition improves efficiency of the Howland current pump, as smaller voltage is wasted across R_5 and smaller current is wasted through R_4 and R_3. The circuit is stable for most of the resistive loads, however, to insure stability, a few picofarad capacitor C may be added in a negative feedback or/and from the positive input of an operational amplifier to ground. When the load is inductive, an infinitely large compliance voltage would be required to deliver the set current when a fast transient control signal is applied. Therefore, the current pump will produce a limited rising slope of the output current. The flowing current will generate an inductive spike across the output terminal, which may be fatal to the operational amplifier. It is advisable, for the large inductive load to clamp the load with diodes to the power supply buses.

An efficient current pump with four matched resistors and two operational amplifiers is shown in Fig. 5.18b. Its output current is defined by the equation

$$i_{out} = \frac{(V_1 - V_2)}{R_s}. \tag{5.17}$$

The advantage of this circuit is that resistors R may be selected with a relatively high value and housed in the same thermally homogeneous packaging for better thermal tracking.

5.4.2 Voltage References

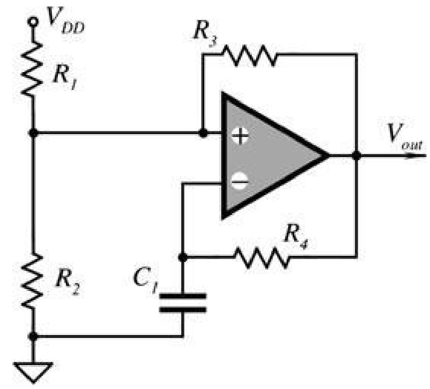
A voltage reference is an electronic device that generates constant voltage that is affected little by variations in power supply, temperature, load, aging, and other factors. There are several techniques known for generation of such voltages. Many voltage references are available in monolithic form for a large variety of voltages. Most of them operate with the so-called band gap references.

5.4.3 Oscillators

Oscillators are generators of variable electrical signals. In many applications that employ microprocessors or microcontrollers, square-wave pulses may be generated at one of the I/O ports. When no such port is available, free-standing oscillators may be developed. Any oscillator is essentially comprised of a circuit with a gain stage and some nonlinearity, and a certain amount of positive feedback. By definition, an oscillator is an unstable circuit (as opposed to an amplifier which better be stable!) whose timing characteristics should be either steady or changeable according to a predetermined functional dependence. The latter is called a *modulation*. Generally, there are three types of electronic oscillators classified according to the time-keeping components: the *RC*, *LC*, and crystal oscillators. In the *RC* oscillators, the operating frequency is defined by capacitors (*C*) and resistors (*R*), in the *LC*-oscillators by the capacitive (*C*) and inductive (*L*) components. In the crystal oscillators, operating frequency is defined by a mechanical resonant in specific cuts of piezoelectric crystals, usually quartz and ceramics. There is a great variety of oscillation circuits, coverage of which is beyond the scope of this book. Below, we briefly describe some practical circuits.

Many various multivibrators can be built with logic circuits, for instance with NOR, NAND gates, or binary inverters. Also, many multivibrators can be designed with comparators or operational amplifiers having a high open loop gain. In all these oscillators, a combination of a capacitor and a resistor is time-keeping combination. These circuits are called relaxation oscillators. A voltage across a charged capacitor is compared with another voltage, which is either constant or changing with a different rate. The moment when both voltages are equal is detected by a comparator. The indication of a comparison is fed back to the *RC*-network to alter the capacitor charging in the opposite direction, which is discharging. Recharging in a new direction goes on until the next moment of comparison. This basic principle essentially requires the following minimum components: a capacitor, a charging circuit, and a threshold device (a comparator). Several monolithic relaxation oscillators are available from many manufacturers, for instance a very popular timer, type 555, which can operate in either monostable or actable modes. For the illustration, below we describe just two discrete-component square-wave oscillators, however, there is a great variety of such circuits, which the reader can find in many books on operational amplifiers and digital systems, for instance [3].

Fig. 5.19 Square-wave oscillator with OPAM



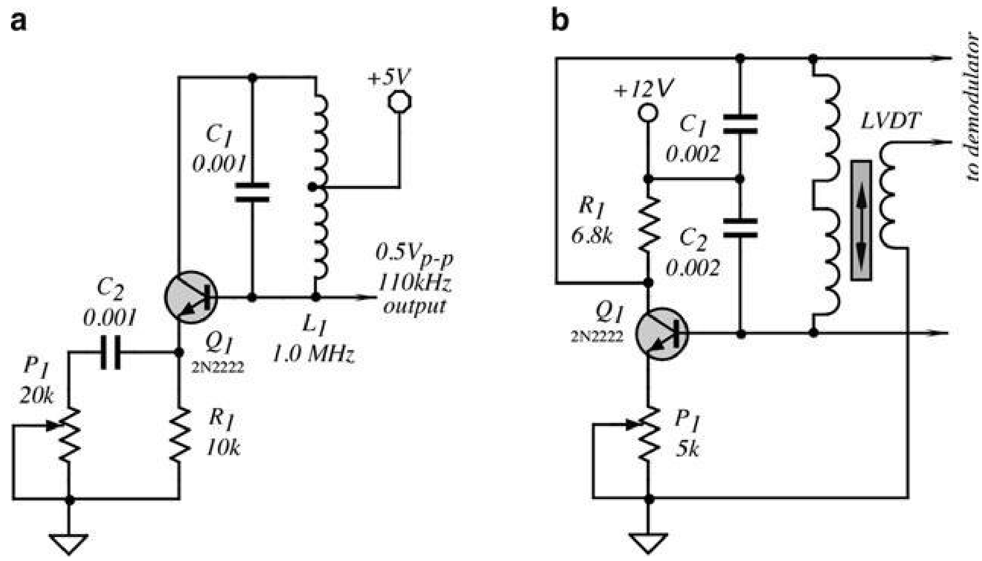


Fig. 5.20 *LC* sine-wave oscillators

A very popular square-wave oscillator (Fig. 5.19) can be built with one OPAM or a voltage comparator.[1] The amplifier is surrounded by two feedback loops: one is negative (to an inverting input) and the other is positive (to a noninverting input). The positive feedback (via R_3) controls the threshold level, while the negative loop charges and discharges timing capacitor C_1, through the resistor R_4. The frequency of this oscillator can be determined from

$$f = \frac{1}{R_4 C_1}\left[\ln\left(1 + \frac{R_1 || R_2}{R_3}\right)\right]^{-1}, \tag{5.18}$$

where $R_1||R_2$ is an equivalent resistance of parallel-connected R_1 and R_2.

---

[1]A voltage comparator differs from an operational amplifier by its faster speed response and the output circuit which is easier interfaceable with digital circuits.

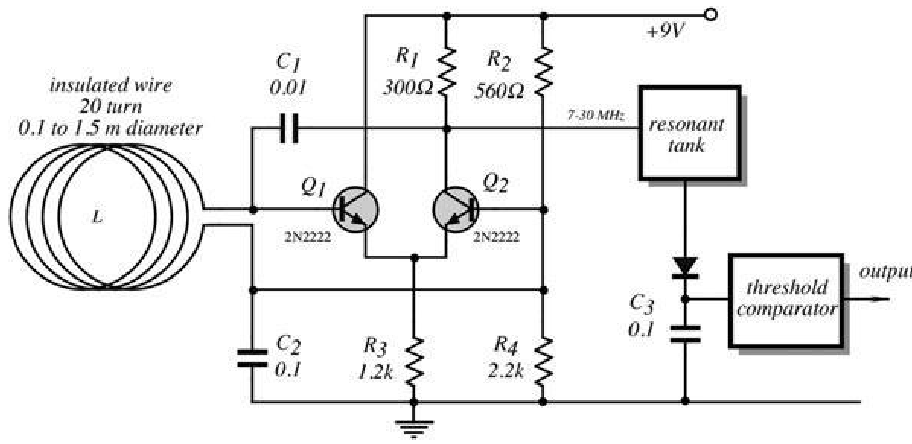


Fig. 5.21 *LC* radio-frequency oscillator as a capacitive occupancy detector

Two circuits shown in Fig. 5.20 can generate sine wave signals. They use the *npn*-transistors as amplifiers and the *LC*-networks to set the oscillating frequency. The (b) circuit is especially useful for driving the LVDT position sensors, as the sensor's transformer becomes a part of the oscillating circuit.

A radiofrequency oscillator can be used as part of a capacitive occupancy detector to detect presence of people in the vicinity of its antenna (Fig. 5.21).[2] The antenna is a coil, which together with the C_2 capacitor, determines the oscillating frequency. The antenna is coupled to the environment through its distributed capacitance, which somewhat reduces the frequency of the oscillator. When a people move into vicinity of the antenna, they bring in an additional capacitance that lowers the oscillator frequency even further. The output of the oscillator is coupled to a resonant tank (typically, an *LC*-network) which is tuned to a baseline frequency (near 30 MHz).

5.4.4 *Drivers*

As opposed to current generators, voltage drivers must produce output voltages, which over broad ranges of the loads and operating frequencies are independent of the output currents. Sometimes, the voltage drivers are called *hard voltage sources*. Usually, when the sensor, which has to be driven is purely resistive, a driver can be a simple output stage, which can deliver sufficient current. However, when the load contains capacitances or inductances, that it, the load is reactive, the output stage becomes a more complex device.

In many instances, when the load is purely resistive, there still can be some capacitance associated with it. This may happen when the load is connected though lengthy wires or coaxial cables. A coaxial cable behaves as a capacitor connected from its central conductor to its shield if the length of the cable is less than 1/4 of the

[2]See Sect. 6.3.

wavelength in the cable at the frequency of interest f. For a coaxial cable, this maximum length is given by

$$L \leq 0.0165\frac{c}{f}, \tag{5.19}$$

where c is the velocity of light in a coaxial cable dielectric.

For instance, if $f = 100$ kHz, $L \leq 0.0165\frac{3\times10^8}{10^5} = 49.5$, that is, a cable less than 49.5 m (162.4 ft) long will behave as a capacitor connected in parallel with the load [Fig. 5.22a]. For an R6-58A/U cable, the capacitance is 95 pF/m. This capacitance must be considered for two reasons: for the speed and stability of the circuits. The instability results from the phase shift produced by the output resistance of the driver R_o and the loading capacitance C_L:

$$\varphi = \text{arctang}(2\pi f R_o C_L). \tag{5.20}$$

For example, if $R_o = 100\ \Omega$ and $C_L = 1{,}000$ pF, at $f = 1$ MHz, the phase shift $\varphi \approx 32°$. This shift significantly reduces the phase margin in a feedback network, which may cause a substantial degradation of the response and a reduced ability to drive capacitive loads. The instability may be either overall, when an entire system oscillates, or localized when the driver alone becomes unstable. The local instabilities often can be cured by large bypass capacitors (on the order of 10 μF) across the power supply or the so-called Q-spoilers consisting of a serial connection of 3–10 Ω resistor and a disk ceramic capacitor connected from the power supply pins of the driver chip to ground.

To make a driver stage more tolerant to capacitive loads, it can be isolated by a small serial resistor as it is shown in Fig. 5.22b. A small capacitive feedback (C_f) to the inverting input of the amplifier, and a 10 Ω resistor may allow to drive loads as large as 0.5 μF. However, in any particular case, it is recommended to find the best values for the resistor and the capacitor experimentally.

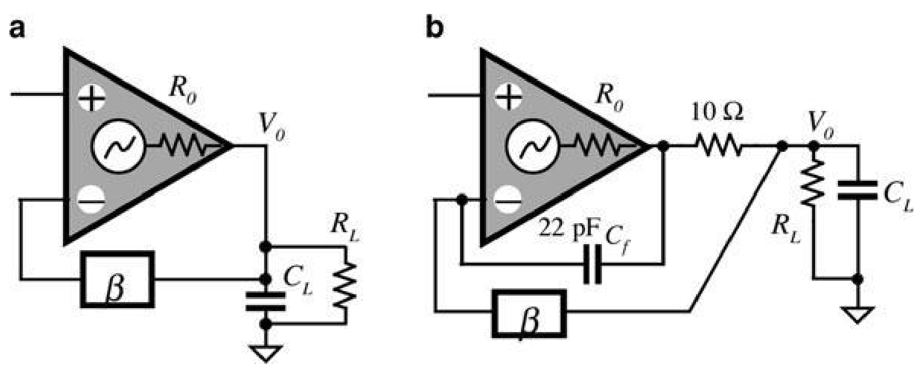


Fig. 5.22 Driving a capacitive load load capacitor is coupled to the driver's input through a feedback (**a**); decoupling of a capacitive load (**b**)

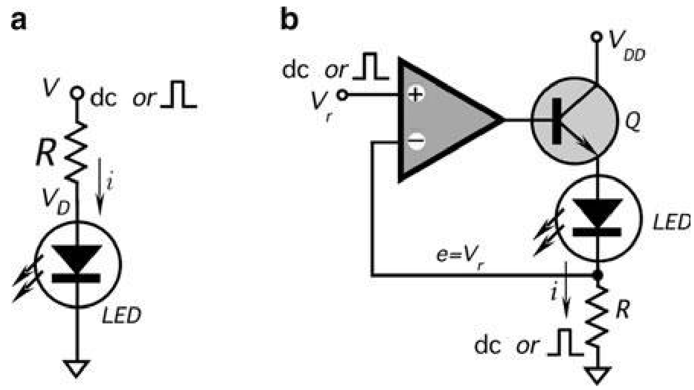


Fig. 5.23 Resistive (**a**) and currents source (**b**) drivers for LED

5.4.5 Optical Drivers

In same applications, optical sensors receive light from natural sources, like the celestial objects or fire, or from human-made sources like scintillating materials, incandescent or fluorescent light. In many other cases, a special light source must be provided. An example is a light detector in a TV set remote control receiver. This sensor produces no output till it receives a train of light (near infrared) pulses from the remote control transmitter. The transmitter must contain a light emitter with the matching spectral characteristics. The most popular light emitters are light emitting diodes (LED) that operate from the UV to near infrared spectral range. The LED produces light whose light intensity is nearly proportional to the current passing through the diode. The simplest drive circuit for LED is shown in Fig. 5.23a. It contains either a dc voltage source or a pulsing voltage source and a current limiting resistor R. Current i is defined by formula

$$i = \frac{V - V_D}{R}, \tag{5.21}$$

where V_D is voltage across the LED (typically from 1.5 to 2.0 V). This voltage depends on the current and temperature and, subsequently the intensity of light generated by the LED also will be the current- and temperature-dependent. For precision applications, the current through LED should be maintained constant. Thus, instead of a current limiting resistor, an electronic current source should be employed as shown in Fig. 5.23b. In this driver, the current is solely defined by the drive control voltage and sensing resistor R and thus is temperature-independent.

5.5 Analog-to-Digital Converters

5.5.1 Basic Concepts

The analog-to-digital converters (abbreviated as A/D, or ADC, or A2D, or A-to-D) range from discrete circuits, to monolithic ICs (integrated circuits), to high-performance hybrid circuits, modules, and even boxes. Also, the converters are available as standard cells for custom and semicustom application-specific integrated circuits (ASIC). The A/D converters transform analog data, usually voltage, into an equivalent digital format, compatible with digital data processing devices. Key characteristics of the A/D converters include absolute and relative accuracy, linearity, no-missing codes, resolution, conversion speed, stability, and price. Quite often, when price is of a major concern, the monolithic IC versions (integrated circuits) are the most efficient. The most popular A/D converters are based on a successive approximation technique because of an inherently good compromise between speed and accuracy. However, other popular techniques are used in a large variety of applications, especially when no high conversion speed is required. These include dual-ramp, quad-slope, pulse-width modulators (PWM), voltage-to-frequency (V/F) converters, and resistance-to-frequency (R/F) converters. The art of an A/D conversion is well developed. Here, we briefly review some popular architectures of the converters, however, for detailed descriptions the reader should refer to specialized texts, such as [4].

The best known digital code is *binary* (base 2). Binary codes are most familiar in representing integers, i.e., in a natural binary integer code having n bits, the least significant bit (LSB) has a weight of 2^0 (i.e., 1), the next bit has a weight of 2^1 (i.e., 2), and so on up to MSB (most significant bit), which has a weight of 2^{n-1} (i.e., $2^n/2$). The value of a binary number is obtained by adding up the weights of all nonzero bits. When the weighted bits are added up, they form a unique number having any value from 0 to 2^{n-1}. Each additional trailing zero bit, if present, essentially doubles the size of the number.

When converting signals from analog sensors, because full scale is independent of the number of bits of resolution, a more useful coding is fractional binary [4], which is always normalized to full scale. Integer binary can be interpreted as fractional binary if all integer values are divided by 2^n. For example, the MSB has a weight of 1/2 (i.e., $2^{n-1}/2^n = 2^{-1}$), the next bit has a weight of 1/4 (i.e., 2^{-2}), and so forth down to the LSB, which has a weight of $1/2^n$ (i.e., 2^{-n}). When the weighted bits are added up, they form a number with any of 2^n values, from 0 to ($1–2^{-n}$) of full scale. Additional bits simply provide more fine structure without affecting the full-scale range. To illustrate these relationships, Table 5.1 lists 16 permutations of 5-bit's worth of 1's and 0's, with their binary weights, and the equivalent numbers expressed as both decimal and binary integers and fractions.

When all bits are "1" in natural binary, the fractional number value is $1 - 2^{-n}$, or normalized full-scale less 1 LSB (1-1/16=15/16 in the example). Strictly speaking, the number that is represented, written with an "integer point," is 0.1111 (=1 − 0.0001). However, it is almost universal practice to write the code simply as the integer 1111 (i.e., "15") with the fractional nature of the corresponding number understood: "1111" → 1111/(1111 + 1), or 15/16.

For convenience, Table 5.2 lists bit weights in binary for numbers having up to 20 bits. However, the practical range for the vast majority of sensors rarely exceeds 16 bits.

Table 5.1 Integer and fractional binary codes

Decimal fraction	Binary fraction	MSB x1/2	Bit2 x1/4	Bit3 x1/6	Bit4 x1/16	Binary integer	Decimal integer
0	0.0000	0	0	0	0	0000	0
1/16 (LSB)	0.0001	0	0	0	1	0001	1
2/16 = 1/8	0.0010	0	0	1	0	0010	2
3/16 = 1/8 + 1/16	0.0011	0	0	1	1	0011	3
4/16 = 1/4	0.0100	0	1	0	0	0100	4
5/16 = 1/4 + 1/16	0.0101	0	1	0	1	0101	5
6/16 = 1/4 + 1/8	0.0110	0	1	1	0	0110	6
7/16 = 1/4 + 1/8 + 1/16	0.0111	0	1	1	1	0111	7
8/16 = 1/2 (MSB)	0.1000	1	0	0	0	1000	8
9/16 = 1/2 + 1/16	0.1001	1	0	0	1	1001	9
10/16 = 1/2 + 1/8	0.1010	1	0	1	0	1010	10
11/16 = 1/2 + 1/8 + 1/16	0.1011	1	0	1	1	1011	11
12/16 = 1/2 + 1/4	0.1100	1	1	0	0	1100	12
13/16 = 1/2 + 1/4 + 1/16	0.1101	1	1	0	1	1101	13
14/16 = 1/2 + 1/4 + 1/8	0.1110	1	1	1	0	1110	14
15/16 = 1/2 + 1/4 + 1/8 + 1/16	0.1111	1	1	1	1	1111	15

The weight assigned to the LSB is the resolution of numbers having *n* bits. The dB column represents the logarithm (base 10) of the ratio of the LSB value to unity (full scale), multiplied by 20. Each successive power of 2 represents a change of 6.02 dB [i.e., 20 $\log_{10}(2)$] or "6 dB/octave."

5.5.2 V/F Converters

A voltage-to-frequency (V/F) converter can provide a high-resolution conversion, and such is useful for sensor's special features as a long-term integration (from seconds to years), a digital-to-frequency conversion (together with a D/A converter), a frequency modulation, a voltage isolation, and an arbitrary frequency division and multiplication. The converter accepts an analog output from the sensor, which can be either voltage or current (in latter case, of course, it should be called a current-to-voltage converter). In some cases, a sensor may become a part of an A/D converter as it is illustrated in Sect. 5.6. Here, however, we will discuss only the conversion of voltage to frequency, or, in other words, to a number of square pulses per unit of time. The frequency is a digital format because pulses can be gated (selected for a given interval of time) and then counted, resulting in a binary number. All V/F converters are of the integrating type because the number of pulses per second, or *frequency*, is proportional to the *average* value of the input voltage.

By using a V/F converter, an A/D can be performed in the most simple and economical manner. The time required to convert an analog voltage into a digital number is related to the full-scale frequency of the V/F converter and the required resolution. Generally, the V/F converters are relatively slow, as compared with

Table 5.2 Binary bit weights and resolutions

BIT	2-*n*	1/2*n* fraction	dB	1/2*n* decimal	%	ppm
FS	20	1	0	1.0	100	1,000,000
MSB	2−1	1/2	−6	0.5	50	500,000
2	2−2	1/4	−12	0.25	25	250,000
3	2−3	1/8	−18.1	0.125	12.5	125,000
4	2−4	1/16	−24.1	0.0625	6.2	62,500
5	2−5	1/32	−30.1	0.03125	3.1	31,250
6	2−6	1/64	−36.1	0.015625	1.6	15,625
7	2−7	1/128	−42.1	0.007812	0.8	7,812
8	2−8	1/256	−48.2	0.003906	0.4	3,906
9	2−9	1/512	−54.2	0.001953	0.2	1,953
10	2−10	1/1,024	−60.2	0.0009766	0.1	977
11	2−11	1/2,048	−66.2	0.00048828	0.05	488
12	2−12	1/4,096	−72.2	0.00024414	0.024	244
13	2−13	1/8,192	−78.3	0.00012207	0.012	122
14	2−14	1/16,384	−84.3	0.000061035	0.006	61
15	2−15	1/32,768	−90.3	0.0000305176	0.003	31
16	2−16	1/65,536	−96.3	0.0000152588	0.0015	15
17	2−17	1/131,072	−102.3	0.00000762939	0.0008	7.6
18	2−18	1/262,144	−108.4	0.000003814697	0.0004	3.8
19	2−19	1/524,288	−114.4	0.000001907349	0.0002	1.9
20	2−20	1/1,048,576	−120.4	0.0000009536743	0.0001	0.95

successive approximation devices, however, they are quite appropriate for the vast majority of sensor applications. When acting as an A/D converter, the V/F converter is coupled to a counter, which is clocked with the required sampling rate. For instance, if a full-scale frequency of the converter is 32 kHz, and the counter is clocked 8 times per second, the highest number of pulses, which can be accumulated every counting cycle is 4,000, which approximately corresponds to a resolution of 12 bit (see Table 5.2). By using the same combination of components the V/F converter and the counter, an integrator can be built for the applications, where the stimulus needs to be integrated over a certain time. The counter accumulates pulses over the gated interval rather than as an average number of pulses per counting cycle.

Another useful feature of a V/F converter is that its pulses can be easily transmitted through communication lines. The pulsed signal is much less susceptible to noisy environment than a high-resolution analog signal. In the ideal case, the output frequency f_{out} of the converter is proportional to the input voltage V_{in}:

$$\frac{f_{out}}{f_{FS}} = \frac{V_{in}}{V_{FS}}, \tag{5.22}$$

where f_{FS} and V_{FS} are the full-scale frequency and input voltage, respectively. For a given linear converter, ratio $f_{FS}/V_{FS} = G$ is constant and is called a conversion factor, then

$$f_{out} = GV_{in}. \tag{5.23}$$

There are several known types of V/F converters. The most popular of them are the multivibrator and the charge-balance circuit.

A multivibrator V/F converter employs a free-running square-wave oscillator where charge-discharge currents of a timing capacitor are controlled by the input signal (Fig. 5.24). Input voltage V_{in} is amplified by a differential amplifier (for instance, an instrumentation amplifier) whose output signal controls two voltage-to-current converters (transistors U_1 and U_2). A precision multivibrator alternatively connects timing capacitor C to both current converters. The capacitor is charged for a half of period through transistor U_1 by the current i_a. During the second half of the timing period, it is discharged by the current i_b through transistor U_2. Since currents i_a and i_b are controlled by the input signal, the capacitor charging and discharging slopes vary accordingly, thus changing the oscillating frequency. An apparent advantage of this circuit is its simplicity and potentially very low power consumption, however, its ability to reject high-frequency noise in the input signal is not as good as in the charge-balance architecture.

The charge-balance type of converter employs an analog integrator and a voltage comparator as shown in Fig. 5.25. This circuit has such advantages as high speed, high linearity, and good noise rejection. The circuit is available in an integral form from several manufacturers, for instance, ADVFC32 and AD650 from Analog Devices, LM331 from National Semiconductors. The converter operates as follows. Input voltage V_{in} is applied to an integrator through the input resistor R_{in}. The integrating capacitor is connected as a negative feedback loop to the operational amplifier whose output voltage is compared with a small negative threshold of -0.6 V. The integrator generates a saw-tooth voltage (Fig. 5.27) that at the moment of comparison with the threshold results in a transient at the comparator's output. That transient enables a one-shot generator, which produces a square pulse of a fixed duration t_{os}. A precision current source generates constant current i, which is alternatively applied either to the summing node of the integrator, or to its output. The switch S_1 is controlled by the one-shot pulses. When the current source

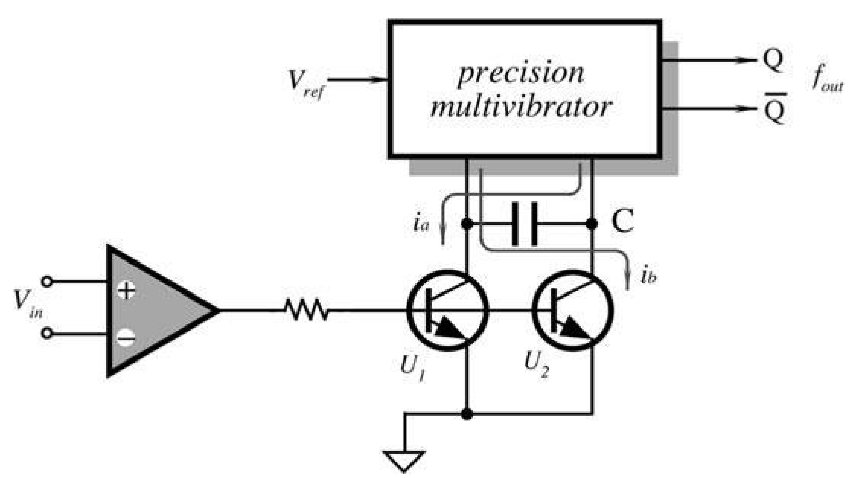


Fig. 5.24 Multivibrator type of a voltage-to-frequency converter

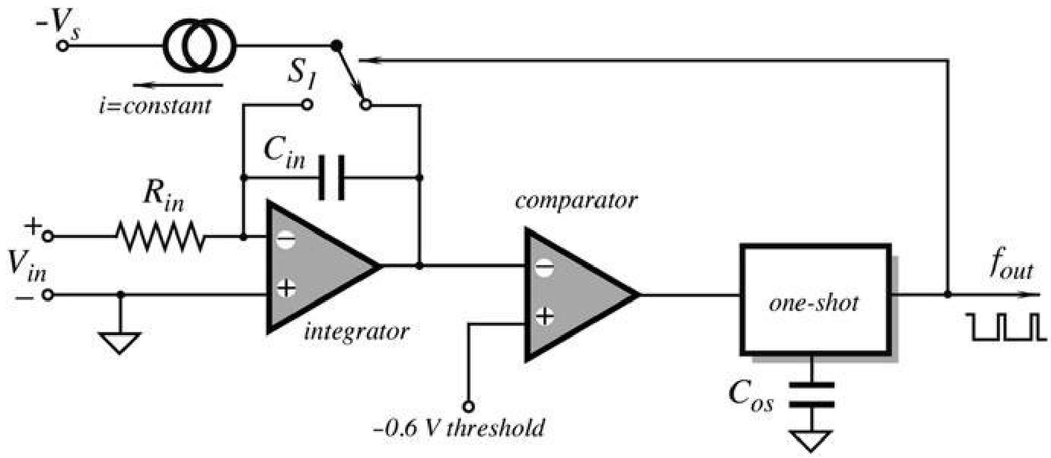


Fig. 5.25 Charge-balance V/F converter

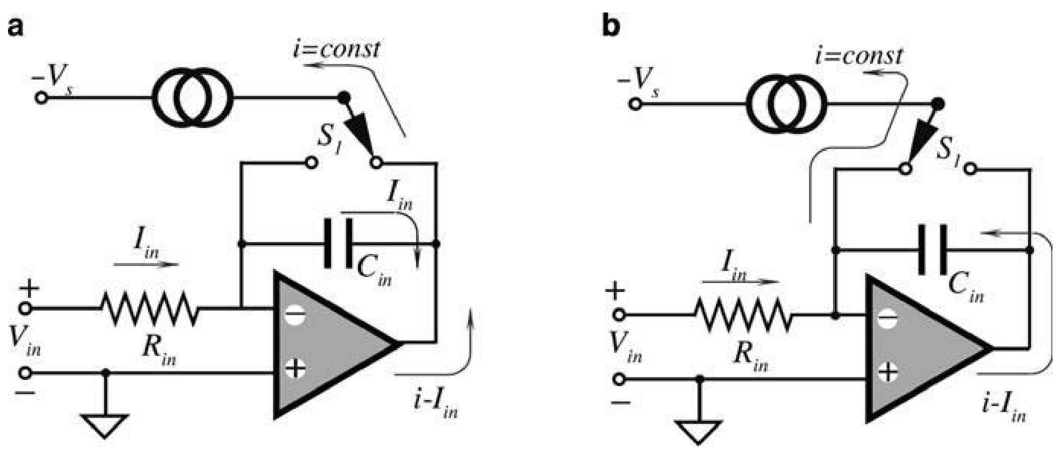


Fig. 5.26 Integrate and deintegrate (reset) phases in a charge-balance converter

is connected to the summing node, it delivers a precisely defined packet of charge $\Delta Q = it_{os}$ to the integrating capacitor. The same summing node also receives an input charge through the resistor R_{in}, thus the net charge is accumulated on the integrating capacitor C_{in}.

When the threshold is reached, the one-shot is triggered and the switch S_1 changes its state to high, thus initiating a reset period (Fig. 5.26b). During the reset period, the current source delivers its current to the summing node of the integrator. The input current charges the integrating capacitor upward. The total voltage between the threshold value and the end of the de-integration is determined by the duration of a one-shot pulse:

$$\Delta V = t_{os}\frac{\mathrm{d}V}{\mathrm{d}t} = t_{os}\frac{i - I_{in}}{C_{in}}. \tag{5.24}$$

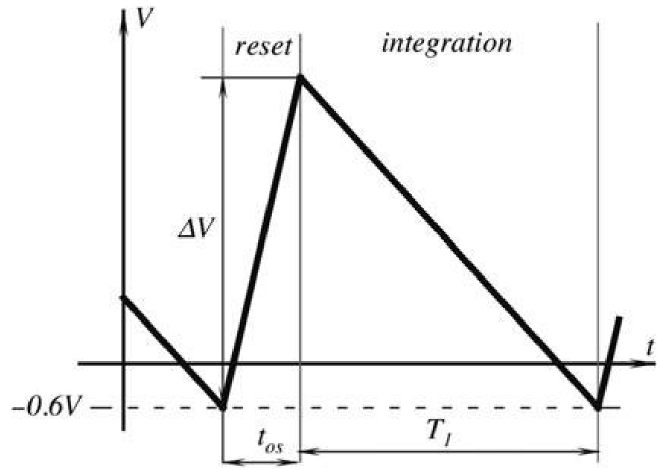


Fig. 5.27 Integrator output in a charge-balance converter

When the output signal of the one-shot circuit goes low, switch S_1 diverts current i to the output terminal of an integrator, which makes no effect on the state of the integrating capacitor C_{in}. That is, the current source sinks a portion of the output current from the operational amplifier. This time is called the integration period (Figs. 5.26a and 5.27). During the integration, the positive input voltage delivers current $I_{in} = V_{in}/R_{in}$ to the capacitor C_{in}. This causes the integrator to ramp down from its positive voltage with the rate proportional to V_{in}. The amount of time required to reach the comparator's threshold is:

$$T_1 = \frac{\Delta V}{dV/dt} = t_{os}\frac{i - I_{in}}{C_{in}}\frac{1}{I_{in}/C_{in}} = t_{os}\frac{i - I_{in}}{I_{in}}. \tag{5.25}$$

It is seen that the capacitor value does not effect duration of the integration period. The output frequency is determined by:

$$f_{out} = \frac{1}{t_{os} + T_1} = \frac{I_{in}}{t_{os}i} = \frac{V_{in}}{R_{in}}\frac{1}{t_{os}i}. \tag{5.26}$$

Therefore, the frequency of one-shot pulses is proportional to the input voltage. It depends also on quality of the integrating resistor, stability of the current generator, and a one-shot circuit. With a careful design, this type of a V/F converter may reach nonlinearity error below 100 ppm and can generate frequencies from 1 Hz to 1 MHz.

A major advantage of the integrating-type converters, such as a charge-balanced V/F converter, is the ability to reject large amounts of additive noise. By integrating of the measurement, noise is reduced or even totally eliminated. Pulses from the converter are accumulated for a gated period T in a counter. Then, the counter behaves like a filter having a transfer function in the form

$$H(f) = \frac{\sin \pi f T}{\pi f T}, \tag{5.27}$$

where f is the frequency of pulses. For low frequencies, value of this transfer function $H(f)$ is close to unity, meaning that the converter and the counter make correct measurements. However, for a frequency $1/T$ the transfer function $H(1/T)$ is zero, meaning that these frequencies are completely rejected. For example, if gating time $T = 16.67$ ms, which corresponds to a frequency of 60 Hz, the power line frequency, which is a source of substantial noise in many sensors, then the 60 Hz noise will be rejected. Moreover, the multiple frequencies (120, 180, 240 Hz, and so on) will also be rejected.

5.5.3 Dual-Slope Converters

A dual-slope converter has been very popular; it was used nearly universally in handheld digital voltmeters and other portable instruments where a fast conversion was not required. This type of converter performs an indirect conversion of the input voltage. First, it converts V_{in} into a function of time, then the time function is converted into a digital number by a pulse counter. The dual slope has the same advantage as the charge-balanced converter; they both reject frequencies $1/T$ corresponding to the integrate timing.

Dual-slope converters are often implemented as a combination of analog components (OPAMs, switches, resistors, and capacitors) and a microcontroller, which handles the functions of timing, control logic, and counting.

5.5.4 Successive Approximation Converter

These converters are widely used in a monolithic form thanks to their high speed (to 1 MHz throughput rates)and high resolution (16 bit and higher). Conversion time is fixed and independent of the input signal. Each conversion is unique, as the internal logic and registers are cleared after each conversion, thus making these A/D converters suitable for the multichannel multiplexing. The converter (Fig. 5.28) consists of a precision voltage comparator, a module comprising shifter registers and a control logic, and a digital-to-analog converter (D/A) that serves as a feedback from the digital outputs to the input analog comparator.

The conversion technique consists of comparing the unknown input, V_{in}, against a precise voltage, V_a, or current generated by a D/A converter. The conversion technique is similar to a weighing process using a balance, with a set of n binary weights (for instance, 1/2 kg, 1/4 kg, 1/8 kg, 1/16 kg, etc. up to total of 1 kg). Before the conversion cycles, all the registers must be cleared and the comparator's output is HIGH. The D/A converter has MSB (1/2 scale) at its inputs and generates an

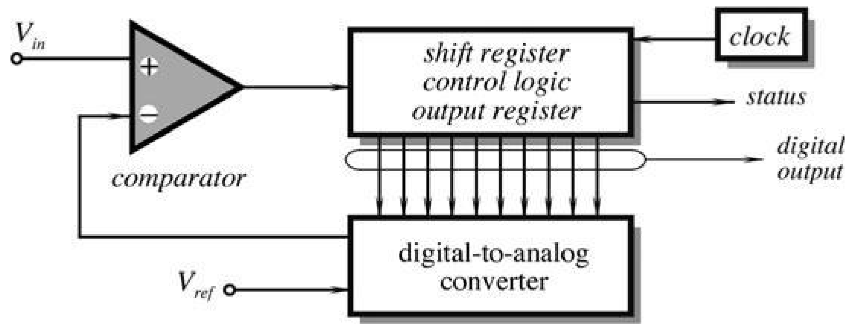


Fig. 5.28 Block-Diagram of successive-approximation A/D converter

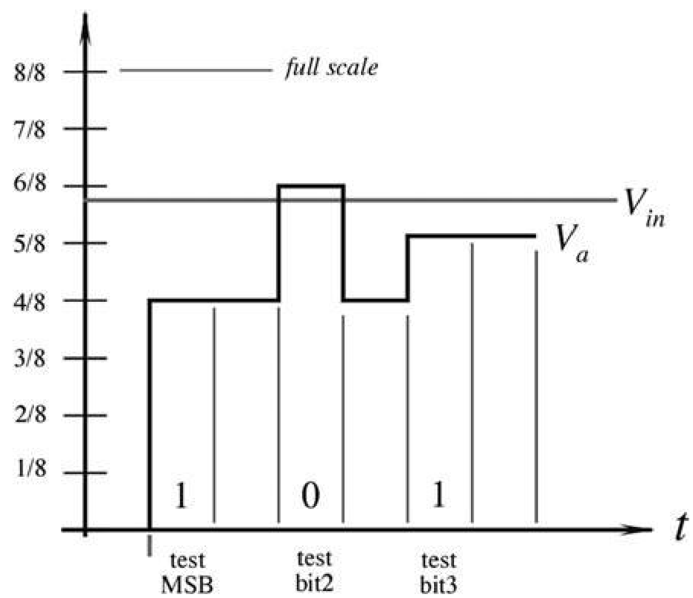


Fig. 5.29 3-bit weighing in successive approximation A/D

appropriate analog voltage, V_a, equal to 1/2 of a full scale input signal. If the input is still greater than the D/A voltage (Fig. 5.29), the comparator remains HIGH, causing “1” at the register’s output. Then, the next bit (2/8 = 1/4 of FS) is tried. If the second bit does not add enough weight to exceed the input, the comparator remains HIGH (“1” at the output), and the third bit is tried. However, if the second bit tips the scale too far, the comparator goes LOW resulting in “0” in the register, and the third bit is tried. The process continues in order of descending bit weight until the last bit has been tried. After the completion, the status line indicates the end of conversion and data can be read from the register as a valid number corresponding to the input signal.

To make the conversion valid, the input signal V_{in} must not change until all the bits are tried, otherwise, the digital reading may be erroneous. To avoid any problems with the changing input, a successive approximation converter usually is supplied with a sample-and-hold (S&H) circuit. This circuit is a short-time analog memory, which samples the input signal and stores it as a dc voltage during an entire conversion cycle.

5.5.5 Resolution Extension

In a typical data acquisition system, a monolithic microcontroller often contains an analog-to-digital converter, whose maximum resolution typically is limited to 10 bits. When resolution of a built-in converter is higher, 12 or even 14 bits, the cost may become prohibitively high. In most applications, 10 bits may be not nearly enough for correct representation of a minimum increment of a stimulus (input resolution R_o). There are several ways to resolve this problem. One is to use an amplifier before the A/D converter. For example, an amplifier of gain 4 will effectively increase the input resolution by two bits, say from 10 to 12. Of course, the price to pay is an uncertainty in the amplifier's characteristics. Another method of achieving higher resolution is use a dual-slope A/D converter whose resolution limited only by the available counter rate and the speed response of a comparator.[3] And another method is to use a 10-bit A/D converter (for instance, of a successive approximation type) with a resolution extension circuit. Such a circuit can boost the resolution by several bits, for instance from 10 to 14. A block-diagram of the circuit is shown in Fig. 5.30a. In addition to a conventional 10-bit A/D converter,

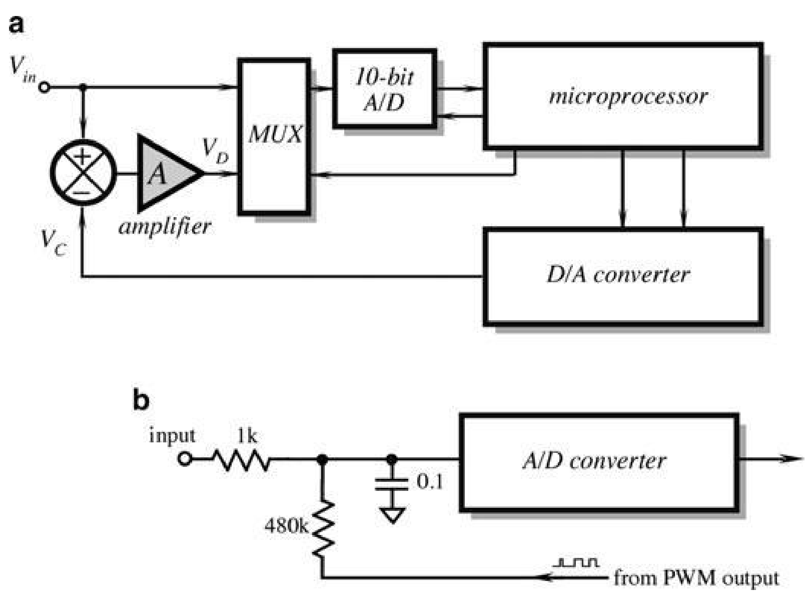


Fig. 5.30 Resolution enhancement circuit with D/A converter (**a**); adding artificial noise to the input signal for oversampling (**b**)

[3]A resolution should not be confused with accuracy.

it includes a digital-to-analog (D/A) converter, a subtraction circuit and an amplifier having gain A. In the ASIC or discrete circuits, a D/A converter may be shared with an A/D part (see Fig. 5.28).

The input signal V_m has a full-scale value E, thus for an 8-bit converter, the initial resolution will be

$$R_o = E/(2^{10} - 1) = E/1023, \tag{5.28}$$

which is expressed in volts per bit. For instance, for a 5 V full scale, the 10-bit resolution is 4.89 mV/bit. Initially, the multiplexer (MUX) connects the input signal to the A/D converter, which produces the output digital value, M, which is expressed in bits. Then, the microprocessor outputs that value to a D/A converter, which produces output analog voltage V_c, which is an approximation of the input signal. This voltage is subtracted from the input signal and amplified by the amplifier to value

$$V_D = (V_m - V_c)A \tag{5.29}$$

The voltage V_D is an amplified error between the actual and digitally represented input signals. For a full scale input signal, the maximum error $(V_m - V_c)$ is equal to a resolution of an A/D converter, therefore, for an 10-bit conversion $V_D = 4.89A$ mV. The multiplexer connects that voltage to the A/D converter which converts V_D to a digital value C:

$$C = \frac{V_D}{R_0} = (V_m - V_c)\frac{A}{R_0}. \tag{5.30}$$

As a result, the microprocessor combines two digital values: M and C, where C represents the high resolution bits. If A=255, then for the 5 V full scale, LSB ≈ 19.25 μV, which corresponds to a total resolution of 18 bit. In practice, it is hard to achieve such a high resolution because of the errors originated in the D/A converter, reference voltage, amplifier's drift, noise, etc. Nevertheless, the method is quite efficient when a modest resolution of 12 or 13 bit is deemed to be sufficient.

Another powerful method of a resolution extension is based on the so-called oversampling [18]. The idea works only if the input analog signal is changing between the sampling points. For example, if the A/D conversion steps are at 50, 70, 90 mV, etc. while the input signal is steady 62 mV, the digital number will indicate 70 mV, thus producing a digitization error of 8 mV and no oversampling would make any difference. If the input signal changes with the maximum spectral frequency f_m, according to Nyquist theorem[4] the sampling frequency $f_s > 2f_m$.

[4]A fundamental theorem of the information theory. It is also known as Nyquist-Shannon-Kotelnikov theorem. It states that the minimum sampling must be twice as fast as the highest frequency of the signal.

The oversampling requires a much higher sampling frequency than defined by Nyquist. Specifically, it is based on the formula

$$f_{os} > 2^{2+n} f_m, \tag{5.31}$$

where n is a number of the extension bits. For example, if we have a 10-bit A/D and would like to generate with it a number of 12 bits ($n = 2$), the sampling rate must be at least 16 times higher than f_m. The oversampling allows exchanging a resolution of an A/D conversion for the maximum converted frequency. Thus, this method is useful for converting relatively slow changing signals as compared with the maximum sampling rate of an A/D converter.

As it was said above, the method requires the signal to change between the samplings. If the analog signal does not include natural variations or inherent noise, an artificial noise can be added to the input signal or the A/D reference voltage to jitter signal between the samples. A practical method of adding artificial noise is shown in Fig. 5.30b. The microcontroller generates pulse width modulated (PWM) random pulses that are smoothed by a capacitor and added to the analog input signal. The magnitude of jittering must correspond to at least 0.5 LSB of the original resolution but preferably should be about 2 LSB. After sampling, to get an increased resolution, 2^{2+n} samples from the A/D are added and the result is right-shifted n times. For the above example, 16 sequential 10-bit numbers are added and then right-shifted 2 times, resulting in a 12-bit output number.

5.6 Direct Digitization

Most sensors produce low-level signals. To bring these signals to levels compatible with data processing devices, amplifiers are generally required. Unfortunately, amplifiers and connecting cables and wires may introduce additional errors, add cost to the instrument and increase complexity. Some emerging trends in the sensor-based systems are causing use of the signal conditioning amplifiers to be reevaluated (at least for some transducers) [5]. In particular, many industrial sensor-fed systems are employing digital transmission and processing equipment. These trends point toward direct digitization of sensor outputs, right in the sensor, a difficult task. It is especially true when a sensor-circuit integration on a single chip is considered.

Classical A/D conversion techniques emphasize high-level input ranges. This allows LSB step size to be as large as possible, minimizing offset and noise error. For this reason, a minimum LSB signal is always selected to be at least 100–200 μV. Therefore, a direct connection of many sensors, for instance, RTD temperature transducers or piezoresistive strain gauge s, is unrealistic. In such transducers, a full-scale (FS) output may be limited by several millivolts, meaning that a 10-bit A/D converter must have about 1 μV LSB.

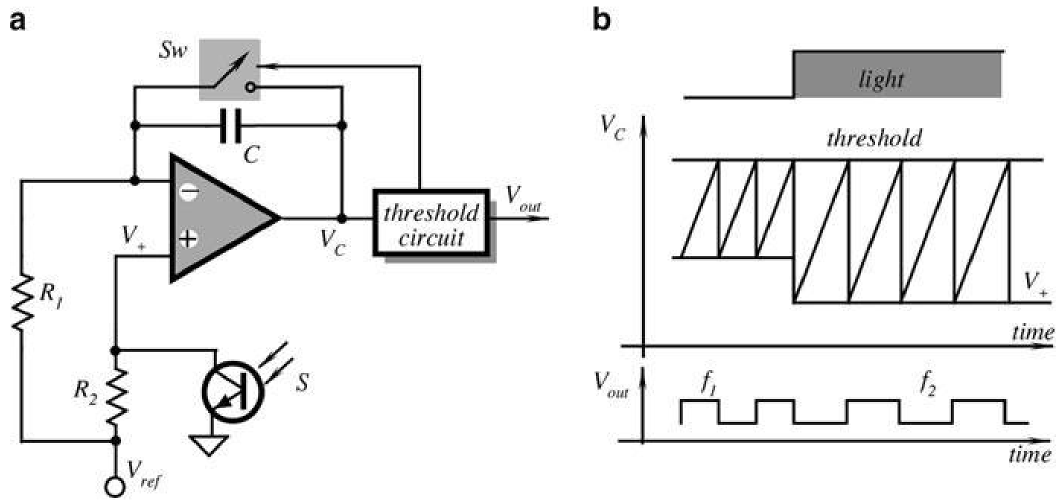


Fig. 5.31 Simplified schematic (**a**) and voltages (**b**) of a light-modulated oscillator

Direct digitization of transducers eliminates a dc gain stage and may yield a better performance without sacrificing accuracy. The main idea behind a direct digitization is to incorporate a sensor into a signal converter, for instance, an A/D converter or an impedance-to-frequency converter. All such converters perform a modulation process and, therefore, are nonlinear devices. Hence, they have some kind of nonlinear circuit, often a threshold comparator. Shifting the threshold level, for instance, may modulate the output signal, which is a desirable effect.

Figure 5.31a shows a simplified circuit diagram of a modulating oscillator. It is comprised of an integrator built with an operational amplifier and a threshold circuit. The voltage across capacitor, C, is an integral of the current whose value is proportional to voltage in the noninverting input of the operational amplifier. When that voltage reaches the threshold, switch *SW* closes, thus fully discharging the capacitor. The capacitor starts integrating the current again until the cycle repeats. The operating point of the amplifier is defined by the resistor R_2, a phototransistor S, and the reference voltage V_{ref}. A change in light flux, which is incident on the base of the transistor, changes its collector current, thus shifting the operation point. A similar circuit may be used for direct conversion of a resistive transducer, for instance, a thermistor. The circuit can be further modified for the accuracy enhancement, such as for the compensation of the amplifier's offset voltage or bias current, temperature drift, etc.

5.7 Capacitance-to-Voltage Converters

Capacitive sensors are very popular in many applications. Nowadays, micro-machining technology allows fabrication of small monolithic capacitive sensors. Capacitive pressure transducers employ a thin silicon diaphragm as a movable plate of the variable-gap capacitor, which is completed by a metal electrode on the opposing plate. The principle problem in these capacitors is a relatively low capacitance value per unit area (about 2 pF/mm^2) and resulting large die sizes. A typical device offers a zero pressure capacitance on the order of few picofarads, so that a 10-bit resolution requires the detection of capacitive shifts on the order of 15 fF or less (1 femtofarad$=10^{-15}$ F). It is obvious that any external measurement circuit will be totally impractical, as parasitic capacitance of connecting conductors at best can be on the order of 1 pF: too much with respect to the capacitance of the sensor. Therefore, the only way to make such a sensor practical is to build an interface circuit as an integral part of the sensor itself. One quite effective way of designing such a circuit is to use a switched capacitor technique. The technique is based on charge transfer from one capacitor to another by means of solid-state analog switches.

Figure 5.32a shows a simplified circuit diagram of a switched-capacitor converter [6], where variable capacitance C_x and reference capacitance C_r are parts of a symmetrical silicon pressure sensor. Monolithic MOS switches (1–4) are driven by opposite phase clock pulses, ϕ_1 and ϕ_2. When the clocks switch, a charge appears at the common capacitance node. The charge is provided by the constant voltage source, V_{PM}, and is proportional to $(C_x - C_r)$ and, therefore to applied pressure to the sensor. This charge is applied to a charge-to-voltage converter, which includes an operational amplifier, integrating capacitor C_f, and MOS discharge (reset) switch 5. The output signal is variable-amplitude pulses (Fig. 5.32b) which can be transmitted through the communication line and either demodulated to produce linear signal or can be further converted into digital data. So long as the open loop gain of the integrating OPAM is high, the output voltage is insensitive to stray input capacitance, offset voltage, and temperature drift. The minimum detectable signal (noise floor) is determined by the component noise and temperature drifts of the components. The circuit analysis shows that the minimum noise power occurs when

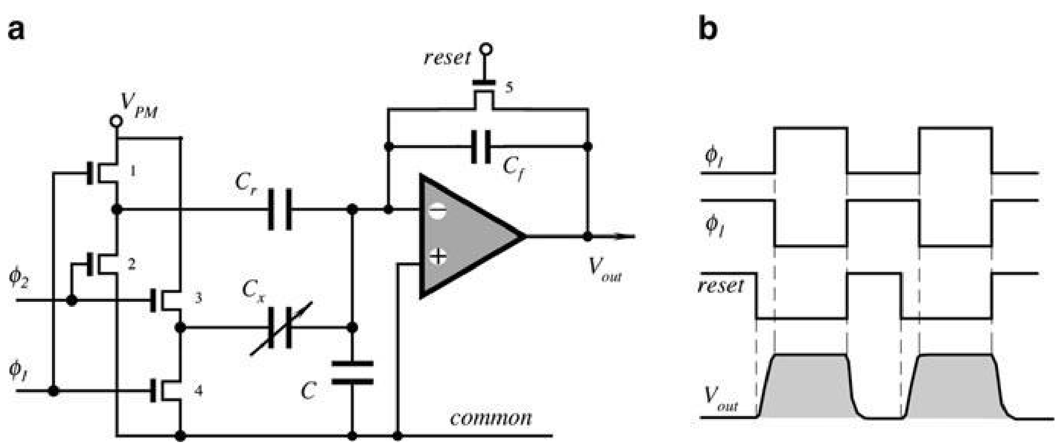


Fig. 5.32 Simplified schematic (**a**) and timing diagrams (**b**) of a differential capacitance-to-voltage converter

the integration capacitor C_f is approximately equal to the frequency compensation capacitor of the OPAM.

When the MOS reset switch goes from the on-state to the off-state, the switching signal injects some charge from the gate of the reset transistor to the input summing node of the OPAM (inverting input). This charge propagated through the gate-to-channel capacitance of the MOS transistor 5. An injection charge results in an offset voltage at the output. This error can be compensated for by a charge-canceling device [7], which can improve the signal-to-noise ratio by two orders of magnitude of the uncompensated charge.

5.8 Integrated Interfaces

A modern trend in the sensor signal conditioning is to integrate in a single silicon chip the amplifiers, multiplexers, A/D converter, and other circuits. Here are two examples of such integration. Figure 5.33 illustrates a signal conditioning circuit ZMD21013 from ZMD (www.zmd.buz). It is optimized for the low-voltage and low-power multiple resistive bridge sensor applications, such as battery-operated consumer or industrial products. This integrated circuit provides programmable amplification and A/D conversion of the sensor signals with up to three resistive bridges or two bridges and one thermocouple, thermopile or any other low-voltage generating sensor. The applied sensor will be switched on only during the sampling

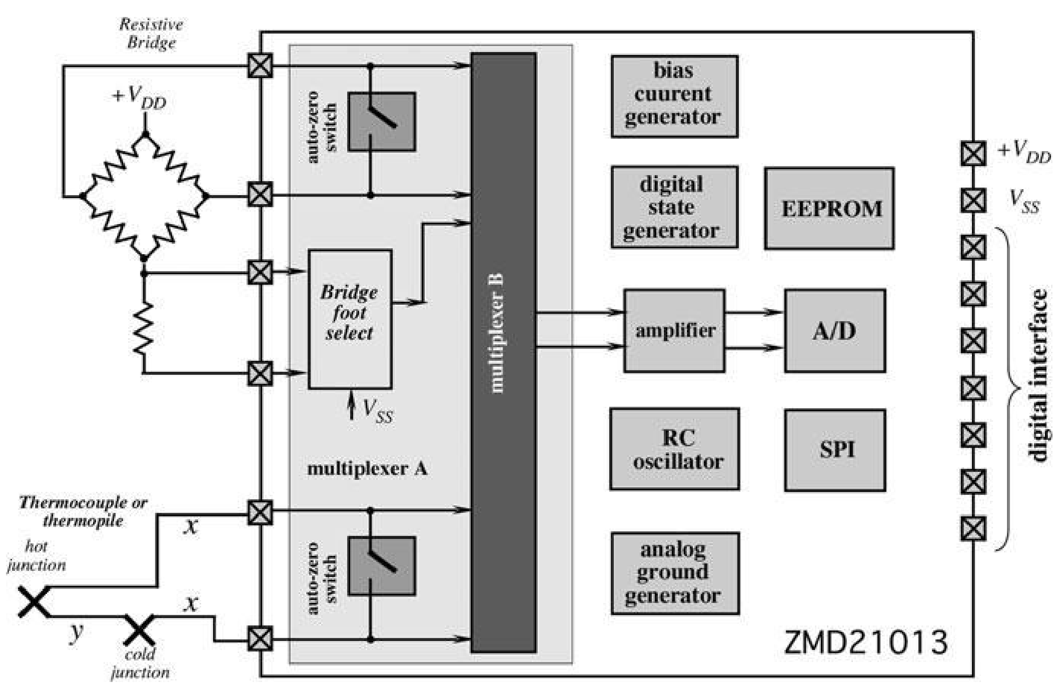


Fig. 5.33 Integrated signal conditioner

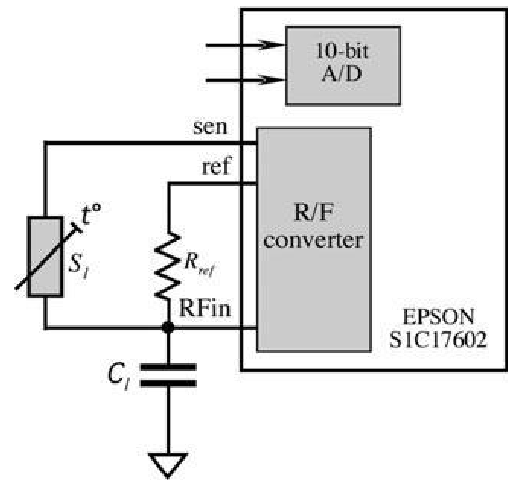


Fig. 5.34 Front end of Epson microcontroller

time, making the circuit suitable for low-power applications. The auto-zero, A/D resolution (10- to 16-bit), sample rate, input range, sensitivity, and measurement mode they all are programmable. This circuit is quite efficient for interfacing a microcontroller with the following resistive and voltage generating sensors: acceleration, pressure, force, flow, and temperature.

Another example of an integrated interface is a 16-bit microcontroller from EPSON having ultralow power consumption. It is specifically adapted for ease of interfacing with various sensors. Figure 5.34 shows that its front end has two types of an A/D converter: a successive approximation 10-bit A/D and an R/F converter with a 24-bit counter. The R/F converter is useful for monitoring temperature and humidity. This microcontroller has various serial interfaces, which enable connection with sensors that are incorporated in healthcare equipment to monitor body temperature, blood pressure, body composition, etc. In addition, this product contains a segment LCD driver for displaying alphanumeric characters and icons. It also contains a built-in MAC (multiply and accumulate) unit and dividing unit enable high-speed processing of the sensor signals. The R/F converter operates with a reference resistor R_{ref} that sets the operating point of the conversion. During the conversion cycle, the capacitor C_1 is charged through a resistive sensor and discharged through R_{ref}, thus value of the capacitor and its stability make no effect on the R/F converter's accuracy.

5.9 Ratiometric Circuits

A powerful method of improving accuracy of a sensor is a *ratiometric* technique, which is one of the most popular ways of signal conditioning. It should be emphasized, however, that the method is useful only if a source of error has a multiplicative nature but not additive. That is, the technique is useless for reduction of, for instance, thermal noise. On the other hand, it is quite potent to solve such problems as dependence of a sensor's sensitivity to such factors as power supply instability, ambient temperature, humidity, pressure, effects of aging, etc. The technique essentially requires the use of two sensors where one is the acting sensor, which responds to an external stimulus and the other is a compensating sensor, which is either shielded from that stimulus or is insensitive to it. Both sensors must be exposed to all other external effects, which may multiplicatively change their performance. The second sensor, which is often called *reference*, must be subjected to a reference stimulus, which is ultimately stable during the life time of the product. In many practical systems, the reference sensor must not necessarily be exactly similar to the acting sensor, however, its physical properties, which are subject to instabilities should be the same. For example, Fig. 5.35a shows a simple temperature detector where the acting sensor is a negative temperature coefficient (NTC) thermistor R_T. A reference resistor R_o has a value equal to the resistance of the thermistor at some reference temperature, for instance at 25°C. Both are connected via an analog multiplexer to an amplifier with a feedback resistor R. Let us assume that there is a some drift in the sensor value that can be described by a function of time $a(t)$ so that the sensor's resistance becomes $R_T(t) = a(t)R_T$. Property of the resistor R_0 is such that it also changes with the same function, so $R_o(t) = a(t)R_o$. The output signals of the amplifier produced by the sensor and the reference resistor respectively are:

$$V_N = -\frac{ER}{a(t)R_T} = -\frac{ER}{a(t)R_T}, \tag{5.32}$$

$$V_D = -\frac{ER}{a(t)R_o} = -\frac{ER}{a(t)R_o}.$$

It is seen that both voltages are functions of a power supply voltage E and the circuit gain, which is defined by resistor R. They also functions of the drift $a(t)$. The multiplexing switch causes two voltages V_N and V_D to appear sequentially at the

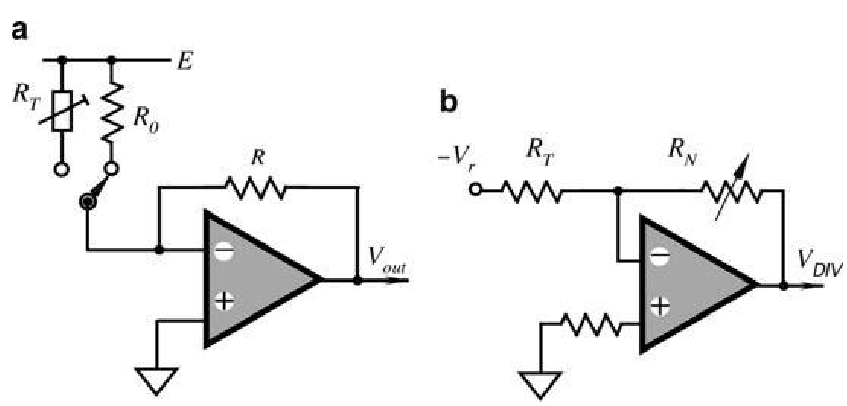


Fig. 5.35 Ratiometric temperature detector (**a**) and analog divider of resistive values (**b**)

amplifier's output. If these voltages are fed into a divider circuit, the resulting signal may be expressed as

$$V_{\mathrm{DIV}} = k\frac{V_{\mathrm{N}}}{V_{\mathrm{D}}} = k\frac{R_{\mathrm{o}}}{R_{\mathrm{T}}}, \tag{5.33}$$

where k is the divider's gain. Therefore, the divider's output signal is subject of neither power supply voltage nor the amplifier gain. It also not subject of the multiplicative drift $a(t)$. The voltage depends only on the sensor and its reference resistor. This is true only if spurious variables, such as function $a(t)$, the power supply or amplifier's gain, do not change rapidly. That is, they should not change appreciably during the multiplexing period. This requirement determines the rate of multiplexing.

A ratiometric technique essentially requires the use of a division. It can be performed by two standard methods: digital and analog. In a digital form, output signals from both the acting and the reference sensors are multiplexed and converted into binary codes in an analog-to-digital (A/D) converter. Subsequently, a computer or a microprocessor performs an operation of a division. In an analog form, a divider may be part of a signal conditioner or the interface circuit. A "divider" [Fig. 5.36a] produces an output voltage or current proportional to the ratio of two input voltages or currents:

$$V_{DIV} = k\frac{V_N}{V_D}, \tag{5.34}$$

where the numerator is denoted as V_N, the denominator V_D and k is equal to the output voltage, when $V_N = V_D$. The operating ranges of the variables (quadrants of operation) is defined by the polarity and magnitude ranges of the numerator and denominator inputs, and of the output. For instance, if V_N and V_D are both either positive or negative, the divider is of a 1-quadrant type. If the numerator is bipolar, the divider is 2-quadrant. Generally, the denominator is restricted to a single polarity,

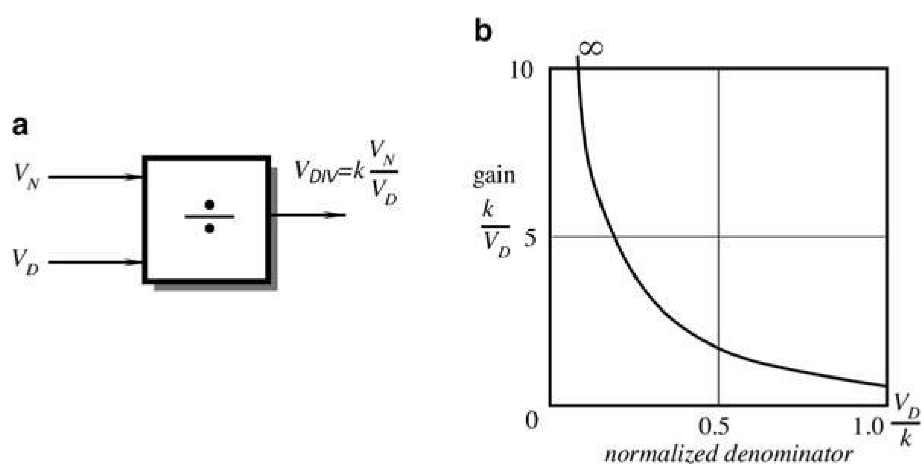


Fig. 5.36 Symbol of a divider (**a**) and gain of a divider as function of a denominator (**b**)

since the transition from one polarity to another would require the denominator to pass through zero, which would call for an infinite output (unless the numerator is also zero). In practice, the denominator is a signal from a reference sensor, which usually is of a relatively constant value.

Division has long been the most difficult of the four arithmetic functions to implement with analog circuits. This difficulty stems primarily from the nature of division: the magnitude of a ratio becomes quite large, approaching infinity, for a denominator that is approaching zero (and a nonzero numerator). Thus, an ideal divider must have a potentially infinite gain and infinite dynamic range. For a real divider, both of these factors are limited by the magnification of drift and noise at low values of V_D. That is, the gain of a divider for a numerator is inversely dependent on the value of the denominator [Fig. 5.36b]. Thus, the overall error is the net effect of several factors, such as gain dependence of denominator, numerator, and denominator input errors, like offsets, noise and drift (which must be much smaller than the smallest values of input signals). Besides, the output of the divider must be constant for constant ratios of numerator and denominator, independent of their magnitudes. For example, $10/10 = 0.01/0.01 = 1$ and $1/10 = 0.001/0.01 = 0.1$. In practice, some simple division circuits are used quite extensively. An example is an amplifier of Fig. 5.35b whose output signal is function of the resistor ratio (note that the reference voltage V_r is negative):

$$V_{DIV} = V_r \frac{R_N}{R_T}, \tag{5.35}$$

The most popular and efficient ratiometric circuits are based on Wheatstone bridge designs which are covered below.

5.10 Differential Circuits

Beside multiplicative interferences, the additive interferences are very common and pose a serious problem for low-level output signals. Consider for example a pyroelectric sensor [Fig. 14.22a] where a heat flow sensitive ceramic plate is supported inside a metal can. Since a pyroelectric is also a piezoelectric, besides heat flow the sensor is susceptible to mechanical stress interferences. Even a slight vibration will generate a spurious piezoelectric signal that may be several orders of magnitude higher than a pyroelectric current. The solution is to fabricate a sensor with dual electrodes deposited on the same ceramic substrate as shown in Fig. 14.22b. This essentially creates two identical sensors on the same ceramic plate. Both sensors respond to all stimuli nearly identically. Since they are oppositely connected and assuming that V_{pyro} and V_{piezo} from one sensor are, respectively, equal to those of the other sensor, the resulting output voltage is essentially zero:

$$V_{out} = (V_{pyro_1} + V_{piezo_1}) - (V_{pyro_2} + V_{piezo_2}) = 0 \tag{5.36}$$

If one of the sensors is blocked from receiving thermal radiation ($V_{pyro_2} = 0$), then $V_{out} = V_{pyro_1}$. In other words, thanks to subtraction ($V_{piezo_1} = V_{piezo_2}$) are cancelling each other), the combined sensor now is insensitive to a piezoelectric effect. A differential method where a sensor is fabricated is a symmetrical form and connected to a symmetrical interface circuit (e.g., differential amplifier) so that one signal is subtracted from another and is a very powerful way of noise and drift reductions. Yet, this method is effective only if a dual sensor is fully symmetrical. An asymmetry will produce a proportional loss of noise cancelation. For example, if asymmetry is 5%, the noise will be cancelled by no more than 95%.

5.11 Bridge Circuits

5.11.1 General Concept

The Wheatstone bridge circuits are popular and very effective implementations of the ratiometric technique (division) technique on a sensor level. A basic circuit is shown in Fig. 5.37. Impedances **Z** may be either active or reactive, that is they may be either simple resistances, like in the piezoresistive gauges, or capacitors, or inductors, or combinations of the above. For a pure resistor, the impedance is R, for an ideal capacitor, the magnitude of its impedance is equal to $1/(2\pi fC)$ and for an inductor, it is $2\pi fL$, where f is the frequency of the current passing through the element. The bridge output voltage is represented by

$$V_{out} = \left(\frac{\mathbf{Z_1}}{\mathbf{Z_1} + \mathbf{Z_2}} - \frac{\mathbf{Z_3}}{\mathbf{Z_3} + \mathbf{Z_4}}\right) V_{ref}, \tag{5.37}$$

The bridge is considered to be in a balanced state when the following condition is met:

$$\frac{\mathbf{Z_1}}{\mathbf{Z_2}} = \frac{\mathbf{Z_3}}{\mathbf{Z_4}}. \tag{5.38}$$

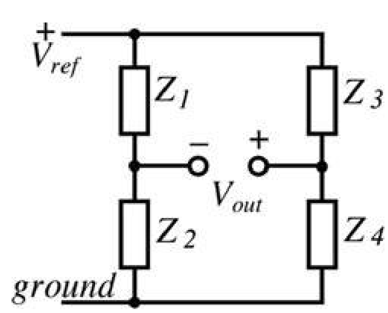


Fig. 5.37 General circuit of Wheatstone bridge

Under the balanced condition, the output voltage is zero. When at least one impedance in the bridge changes, the bridge becomes imbalanced and the output voltage goes either in a positive or negative direction, depending on the direction of the impedance change. To determine the bridge sensitivity with respect to each impedance partial derivatives may be obtained from (5.37):

$$\begin{aligned}
\frac{\partial V_{out}}{\partial \mathbf{Z_1}} &= \frac{\mathbf{Z}_2}{(\mathbf{Z_1}+\mathbf{Z_2})^2}V_{ref}\\
\frac{\partial V_{out}}{\partial \mathbf{Z_2}} &= -\frac{\mathbf{Z}_1}{(\mathbf{Z_1}+\mathbf{Z_2})^2}V_{ref}\\
\frac{\partial V_{out}}{\partial \mathbf{Z_3}} &= -\frac{\mathbf{Z}_4}{(\mathbf{Z_3}+\mathbf{Z_4})^2}V_{ref}\\
\frac{\partial V_{out}}{\partial \mathbf{Z_4}} &= \frac{\mathbf{Z}_3}{(\mathbf{Z_3}+\mathbf{Z_4})^2}V_{ref}
\end{aligned} \tag{5.39}$$

By summing these equations, we obtain the bridge sensitivity:

$$\frac{\delta V_{out}}{V_{ref}} = \frac{\mathbf{Z}_2\delta\mathbf{Z_1} - \mathbf{Z_1}\delta Z_2}{(\mathbf{Z_1}+\mathbf{Z_2})^2} - \frac{\mathbf{Z}_4\delta\mathbf{Z_3} - \mathbf{Z_3}\delta\mathbf{Z_4}}{(\mathbf{Z_3}+\mathbf{Z_4})^2}, \tag{5.40}$$

A closer examination of (5.40) shows that only the adjacent pairs of the impedances (i.e., $\mathbf{Z_1}$ and $\mathbf{Z_2}$, $\mathbf{Z_3}$, and $\mathbf{Z_4}$) have to be identical in order to achieve the ratiometric compensation (such as the temperature stability, drift, etc.). It should be noted that impedances in the balanced bridge do not have to be equal, as long as a balance of the ratio (5.38) is satisfied. In many practical circuits, only one impedance is used as a sensor, thus for $\mathbf{Z}_1$ as a sensor, the bridge sensitivity becomes

$$\frac{\delta V_{\mathrm{out}}}{V_{\mathrm{ref}}} = \frac{\delta\mathbf{Z_1}}{4\mathbf{Z_1}}. \tag{5.41}$$

The resistive bridge circuits are commonly used with strain gauges, piezoresistive pressure transducers, thermistor thermometers, hygristors, and other sensors when immunity against environmental factors is required. Similar arrangements are used with the capacitive and magnetic sensors for measuring force, displacement, moisture, etc.

5.11.2 Disbalanced Bridge

A basic Wheatstone bridge circuit [Fig. 5.38a] generally operates with a disbalanced bridge. This is called the *deflection* method of measurement. It is based on a detecting the voltage across the bridge diagonal. The bridge output voltage is

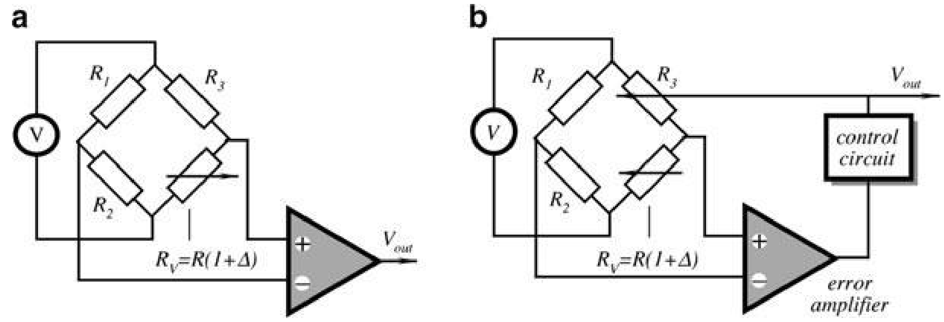


Fig. 5.38 Two methods of using a bridge circuit: Disbalanced bridge (**a**) and Null-balanced bridge with a feedback control (**b**)

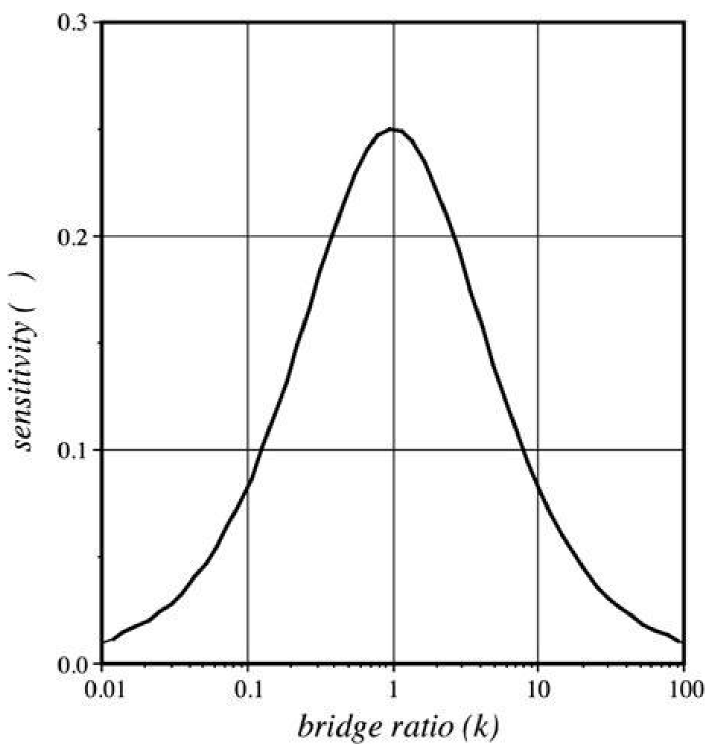


Fig. 5.39 Sensitivity of a disbalanced bridge as function of the impedance ratio

a nonlinear function of a disbalance Δ, where the sensor's resistance $R_v = R(1 + \Delta)$. However, for a small change ($\Delta < 0.05$), which often is the case, the bridge output may be considered quasi-linear. The bridge maximum sensitivity is obtained when $R_1 = R_2$ and $R_3 = R$. When $R_1 >> R_2$ or $R_2 >> R_1$, the bridge output voltage is decreased. Assuming that $k = R_1/R_2$, the bridge sensitivity may be expressed as:

$$\alpha = \frac{V}{R}\frac{k}{(k+1)^2} \tag{5.42}$$

A normalized graph calculated according to this equation is shown in Fig. 5.39. It indicates that the maximum sensitivity is achieved at $k = 1$, however, the sensitivity drops relatively little for the range where $0.5 < k < 2$. If the bridge is fed by a current source, rather by a voltage source, its output voltage for small Δ and a single variable component is represented by

$$V_{out} = i\frac{k\Delta}{2(k+1)}, \tag{5.43}$$

where i is the excitation current.

5.11.3 Null-Balanced Bridge

Another method of using a bridge circuit is called a *null-balance*. The method overcomes the limitation of small changes (Δ) in the bridge arm to achieve a good linearity. The null-balance essentially requires that the bridge is *always* maintained at the balanced state. To satisfy the requirement for a bridge balance (5.37) another arm of the bridge should vary along with the sensing arm. Figure 5.38b illustrates this concept. A control circuit modifies the value of R_3 on a command from the error amplifier. The sensor's output voltage may be obtained from the control signal of the balancing arm R_3. For example, both R_v and R_3 may be photoresistors. The R_3-photoresistor could be interfaced with a light emitting diode (LED), which is controlled by the error amplifier. Current through the LED becomes a measure of resistance R_v, and, subsequently, of the light intensity detected by the sensor.

5.11.4 Bridge Amplifiers

The bridge amplifiers for resistive sensors are probably the most frequently used sensor interface circuits. They may be of various configurations, depending on the required bridge grounding and availability of either grounded or floating reference voltages. Figure 5.40a shows the so-called active bridge, where a variable resistor (the sensor) is floating, i.e., isolated from ground, and is connected into a feedback of the OPAM. If a resistive sensor's transfer function can be modeled by a first order function:

$$R_x \approx R_o(1+\alpha), \tag{5.44}$$

where α is the input stimulus, a transfer function of this circuit is

$$V_{out} = \frac{V}{2} - \frac{1}{2}\alpha V. \tag{5.45}$$

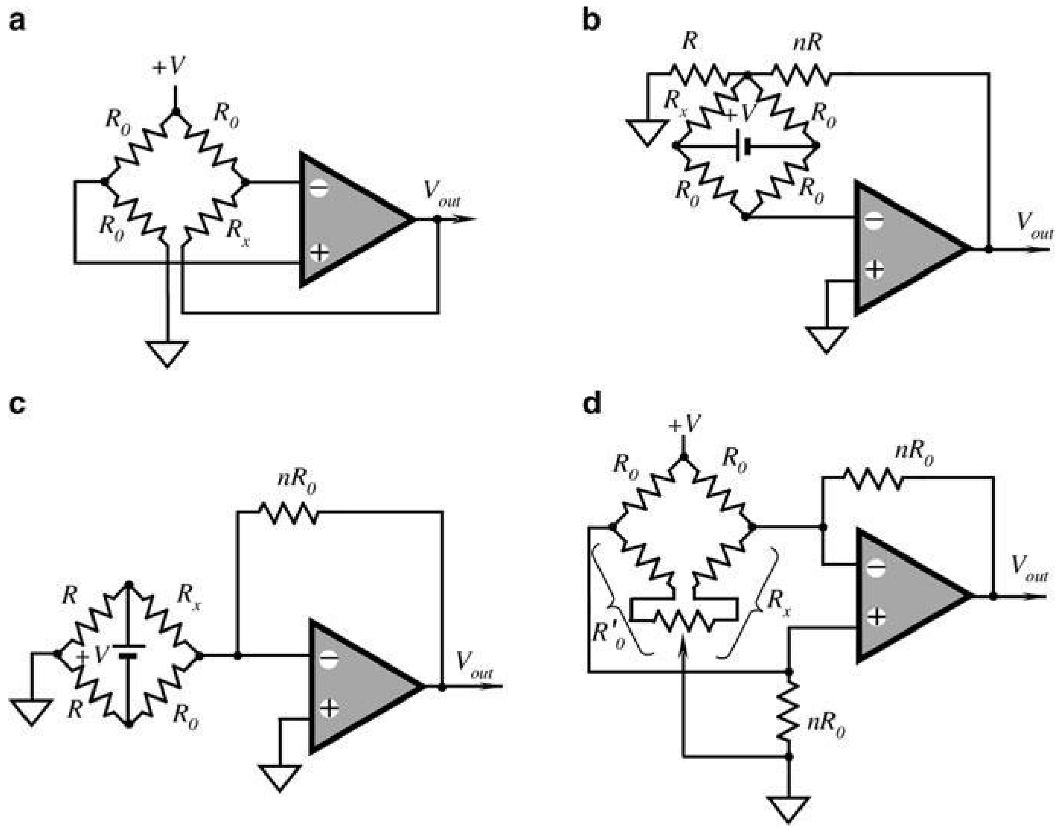


Fig. 5.40 Connection of operational amplifiers to resistive bridge circuits (disbalanced mode)

A circuit with a floating bridge and floating reference voltage source V is shown in Fig. 5.40b. This circuit may provide gain which is determined by a feedback resistor whose value is nR_o:

$$V_{out} = \frac{V}{2} + (1+n)\alpha\frac{V}{4}\frac{1}{1+\frac{\alpha}{2}} \approx \frac{V}{2}\left(1+\frac{(1+n)\alpha}{2}\right). \tag{5.46}$$

A bridge with the asymmetrical resistors ($R \neq R_o$) may be used with the circuit shown in Fig. 5.40c. It requires a floating reference voltage source V:

$$V_{\text{out}} = \frac{V}{2} + n\alpha\frac{V}{4}\frac{1}{1+\frac{\alpha}{2}} \approx \frac{V}{2}\left(1+\frac{n\alpha}{2}\right) \tag{5.47}$$

When a resistive sensor is grounded and a gain from the interface circuit is desirable, a schematic shown in Fig. 5.40d may be employed. Its transfer function is determined from

$$V_{out} = \frac{V}{2} - \frac{n}{2}\frac{V}{1+\frac{1}{2n}}\frac{\alpha}{1+\alpha} \approx 0.5V\left(1-\frac{n}{1+\frac{1}{2n}}\alpha\right) \tag{5.48}$$

Note that the circuit may contain a balancing potentiometer whose resistance sectors should be included into the corresponding arms of the bridge. The potentiometer is used to adjust the bridge component tolerances or offset the bridge balance by some fixed bias. When the bridge is perfectly balanced, its output voltage V_{out} is equal to a half of the bridge excitation voltage $+V$. To better utilize the operational amplifier open loop gain, the value of n should not exceed 100.

5.12 Data Transmission

Signal from a sensor may be transmitted to a receiving end of the system either in a digital format or analog. In most cases, a digital format essentially requires use of an analog-to-digital converter at the sensor's site. Transmission in a digital format has several advantages, the most important of which is noise immunity. Since transmission of digital information is beyond the scope of this book we will not discuss it further. In many cases, however, digital transmission can not be done for several reasons. Then, the sensor's output signal is transmitted to the receiving site in an analog form. Depending on connection, the transmission methods can be divided into a 2, 4, and 6-wire methods.

5.12.1 Two-Wire Transmission

Two-wire analog transmitters are used to couple sensors to control and monitoring devices in the process industry [10]. When, for example, a temperature measurement is taken within a process, a 2-wire transmitter relays that measurement to the control room or interfaces the analog signal directly to a process controller. Two wires can be used to transmit either voltage or current, however, current was accepted as an industry standard. The current carried by the wires varies in the range from 4 to 20 mA, which represents the entire span of an input stimuli. Zero stimulus corresponds to 4 mA while the maximum is at 20 mA. There are two advantages of using current rather than voltage as it is illustrated in Fig. 5.41. Two wires link the controller site to the sensor site. On the sensor site, there is a sensor which is connected to the so-called *two-wire transmitter*. The transmitter may be a voltage-to-current converter. That is, it converts the sensor signal into a variable current. On the controller site, there is a voltage source that can deliver current up to 20 mA. The two wires form a current loop, which at the sensor's side has the sensor and a transmitter, while at the controller side it has a load resistor and a power supply, which are connected in series. When the sensor signal varies, the transmitter's output resistance varies accordingly, thus

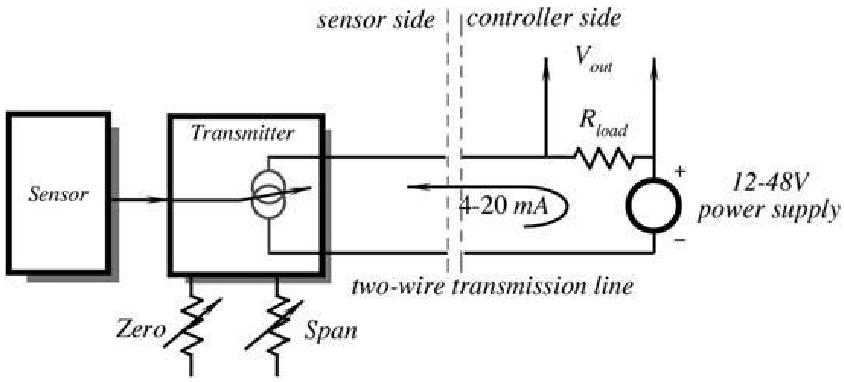


Fig. 5.41 Two-wire 20 mA analog data transmission

modulating the current in the range between 4 and 20 mA. The same current, which carries information, is also used by the transmitter and the sensor to provide their operating power. Obviously, even for the lowest output signal, which produces 4 mA current, the produced voltage must be sufficient to power the transmitting side of the loop. The loop current causes a voltage drop across the load resistor at the controller side. This voltage is a received signal, which is suitable for further processing by the electronic circuits. An advantage of the two-wire method is that the transmitting current is independent of the connecting wires resistance (as long as they do not change) and thus of the transmission line length, obviously, within the limits.

5.12.2 Four-Wire Sensing

Sometimes, it is desirable to connect a resistive sensor to a remotely located interface circuit. When such a sensor has a relatively low resistance (for instance, it is normal for the piezoresistors or RTDs to have resistances on the order of 100 Ω), the connecting wire resistances pose a serious problem since they alter the excitation voltage across the sensor. The problem can be solved by using the so-called 4-wire method (Fig. 5.42a). It allows measuring the resistance of a remote resistor without accounting for resistances of the connecting conductors. A sensor, which is the subject of measurement, is connected to the interface circuit through four rather than two wires. Two wires are connected to a current source and two others to a voltmeter or amplifier. A constant current source (current pump) has a very high output resistance, therefore the current, which it pushes through the loop is almost independent of any resistances r in that loop. An input impedance of a voltmeter or amplifier is very high, hence no current is diverted from the current loop to the voltmeter. A voltage drop across the resistor R_x is

$$V_x = R_x i_o, \tag{5.49}$$

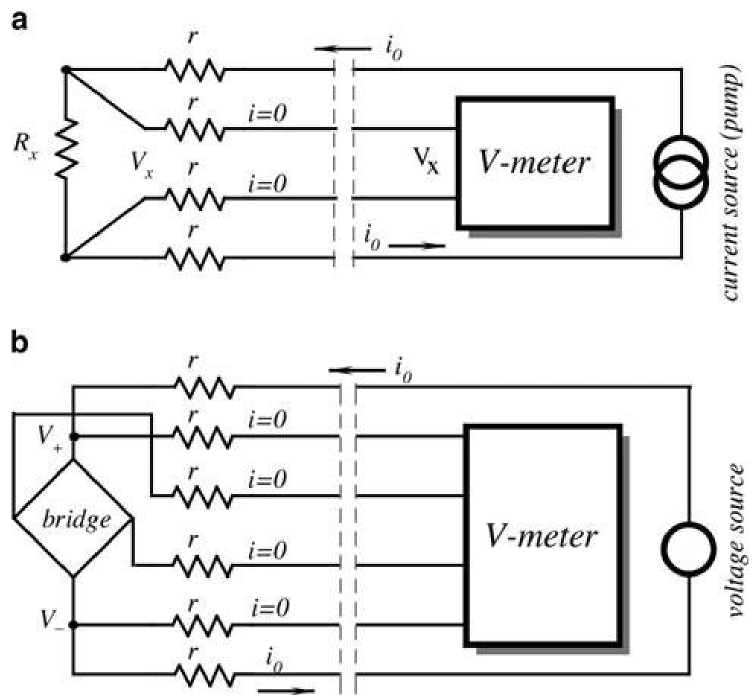


Fig. 5.42 Remote measurements of resistances four-wire method (**a**); six-wire measurement of a bridge (**b**)

which is independent of any resistances *r* of the connecting wires. The 4-wire method is a very powerful means of measuring resistances of remote detectors and is used in industry and science quite extensively.

5.12.3 Six-Wire Sensing

When a Wheatstone bridge circuit is remotely located, voltage across the bridge plays an important role in the bridge temperature stability. That voltage often should be either measured or controlled. Long transmitting wires may introduce unacceptably high resistance in series with the bridge excitation voltage, which interferes with the temperature compensation. The problem may be solved by providing two additional wires to feed the bridge with voltage and to dedicate two wires to measuring the voltage across the bridge (Fig. 5.42b). The actual excitation voltage across the bridge and the bridge differential output voltage are measured by a high-input impedance voltmeter with negligibly small input currents. Thus, the accurate bridge voltages are available at the data processing site without being affected by long transmission lines.

5.13 Noise in Sensors and Circuits

Noise in sensors and circuits may present a substantial source of errors and should be seriously considered. “Like diseases, noise is never eliminated, just prevented, cured, or endured, depending on its nature, seriousness, and the cost/difficulty of treating” [11]. There are two basic classifications of noise for a given circuit: they are inherent noise, which is noise arising within the circuit, and interference (transmitted) noise, which is noise picked up from outside the circuit.

Any sensor, no matter how well it was designed, never produces an electric signal that is an ideal representation of the input stimulus. Often, it is a matter of judgment to define the goodness of the signal. The criteria for this are based on the specific requirements to accuracy and reliability. Distortions of the output signal can be either systematic or stochastic. The former are related to the sensor’s transfer function, its linearity, dynamic characteristics, etc. They all are the result of the sensor’s design, manufacturing tolerances, material quality, and calibration. During a reasonably short time, these factors either do not change or drift relatively slowly. They can be well-defined, characterized, and specified (see Chap. 2). In many applications, such a determination may be used as a factor in the error budget and can be accounted for. Stochastic disturbances, on the other hand, often are irregular, unpredictable to some degree and may change rapidly. Generally, they are termed noise, regardless of their nature and statistical properties. It should be noted that word noise, in association with audio equipment noise, is often mistaken for an irregular, somewhat fast changing signal. We use this word in a much broader sense for all disturbances, either in stimuli, environment, or in the components of sensors and circuits from dc to the upper operating frequencies.

5.13.1 Inherent Noise

A signal, which is amplified and converted from a sensor into a digital form, should be regarded not just by its magnitude and spectral characteristics, but also in terms of a digital resolution. When a conversion system employs an increased digital resolution, the value of the least-significant bit (LSB) decreases. For example, the LSB of a 10-bit system with a 5 V full scale is about 5 mV, the LSB of 16 bits is 77 μV. This by itself poses a significant problem. It makes no sense to employ, say a 16-bit resolution system, if a combined noise is, for example, 300 μV. In a real world, the situation is usually much worse. There are almost no sensors that are capable of producing a 5 V full-scale output signals. Most of them require an amplification. For instance, if a sensor produces a full-scale output of 5 mV, at a 16-bit conversion it would correspond to a LSB of 77 nV, an extremely small signal which makes amplification an enormous task by itself. Whenever a high resolution of a conversion is required, all sources of noise must be seriously considered. In the circuits, noise can be produced by the monolithic amplifiers and other components, which are required for the feedback, biasing, bandwidth limiting, etc.

Input offset voltages and bias currents may drift. In dc circuits, they are indistinguishable from low magnitude signals produced by a sensor. These drifts are usually slow (within a bandwidth of tenths and hundredths of a Hz), therefore they are often called ultralow frequency noise. They are equivalent to randomly (or predictable, say with temperature) changing voltage and current offsets and biases. To distinguish them from the higher frequency noise, the equivalent circuit (Fig. 5.3) contains two additional generators. One is a voltage offset generator e_0 and the other is a current bias generator i_0. The noise signals (voltage and current) result from physical mechanisms within the resistors and semiconductors that are used to fabricate the circuits. There are several sources of noise whose combined effect is represented by the noise voltage and current generators.

One cause for noise is a discrete nature of electric current because current flow is made up of moving charges, and each charge carrier transports a definite value of charge (charge of an electron is 1.6×10^{-19} C). At the atomic level, current flow is very erratic. The motion of the current carriers resembles popcorn popping. This was chosen as a good analogy for current flow and has nothing to do with the "popcorn noise," which we will discuss below. As popcorn, the electron movement may be described in statistical terms. Therefore, one never can be sure about very minute details of current flow. The movement of carriers is temperature related and noise power, which in turn, is also temperature related. In a resistor, these thermal motions cause Johnson noise to result [12]. The mean-square value of noise voltage (which is representative of noise power) can be calculated from

$$\bar{e}_n^2 = 4kTR\Delta f \ [\mathrm{V}^2/\mathrm{Hz}], \tag{5.50}$$

where $k = 1.38 \times 10^{-23}$ J/K (Boltzmann constant), T is temperature in K, R is the resistance in Ω, and Δf is the bandwidth over which the measurement is made, in Hz.

For practical purposes, noise density per $\sqrt{Hz}$ generated by a resistor at room temperature may be estimated from a simplified formula $\overline{e_n} \approx 0.13\sqrt{\mathrm{Hz}}$ in $\mathrm{nV}\sqrt{\mathrm{Hz}}$. For example, if noise bandwidth is 100 Hz and the resistance of concern is 10 MΩ (10^7 Ω), the average noise voltage is estimated as $\overline{e_n} \approx 0.13\sqrt{10^7}\sqrt{100} = 4,111\ \mathrm{nV} \approx 4\ \mu\mathrm{V}$.

Even a simple resistor is a source of noise. It behaves as a perpetual generator of electric signal. Naturally, relatively small resistors generate extremely small noise, however, in some sensors Johnson noise must be taken into account. For instance, a pyroelectric detector uses a bias resistor on the order of 50 GΩ. If a sensor is used at room temperature within a bandwidth of 100 Hz, one may expect the average noise voltage across the resistor to be on the order of 0.3 mV, a pretty high value. To keep noise at bay, bandwidths of the interface circuits must be maintained small, just wide enough to pass the minimum required signal. It should be noted that noise voltage is proportional to square root of the bandwidth. It implies that if we reduce the bandwidth 100 times, noise voltage will be reduced by a factor of 10. Johnson noise magnitude is constant over a broad range of frequencies. Hence, it is often called white noise because of the similarity to white light, which is composed of all the frequencies in visible spectrum.

Another type of noise results because of dc current flow in semiconductors. It is called shot noise, the name suggested by Schottky not in association with his own name but rather because this noise sounded like "a hail of shot striking the target" (nevertheless, shot noise is often called Schottky noise). Shot noise is also white noise. Its value becomes higher with the increase in the bias current. This is the reason why in FET and CMOS semiconductors the current noise is quite small. For a bias current of 50 pA, it is equal to about $4\,\mathrm{fA}/\sqrt{Hz}$, an extremely small current that is equivalent to movement of about 6,000 electrons per second. A convenient equation for shot noise is

$$i_{\mathrm{sn}} = 5.7 \times 10^{-4}\sqrt{I\Delta f}, \tag{5.51}$$

where I is a semiconductor junction current in picoamperes and Δf is a bandwidth of interest in Hz.

An additional ac noise mechanism exists at low frequencies (Fig. 5.43). Both the noise voltage and noise current sources have a spectral density roughly proportional to $1/f$, which is called the pink noise, because of the higher noise contents at lower frequencies (lower frequencies are also at red side of the visible spectrum). This $1/f$ noise occurs in all conductive materials, therefore it is also associated with resistors. At extremely low frequencies it is impossible to separate the $1/f$ noise from dc drift effects. The $1/f$ noise is sometimes called a flicker noise. Mostly it is pronounced at frequencies below 100 Hz, where many sensors operate. It may dominate Johnson and Schottky noise and becomes a chief source of errors at these frequencies. The magnitude of pink noise depends on current passing through the resistive or semiconductive material. Nowadays progress in semiconductor technology resulted in significant reduction of $1/f$ noise in semiconductors, however, when designing a circuit, it is a good engineering practice to use metal film or wirewound resistors in sensors and the front stages of interface circuits wherever significant currents flow through the resistor and low noise at low frequencies is a definite requirement.

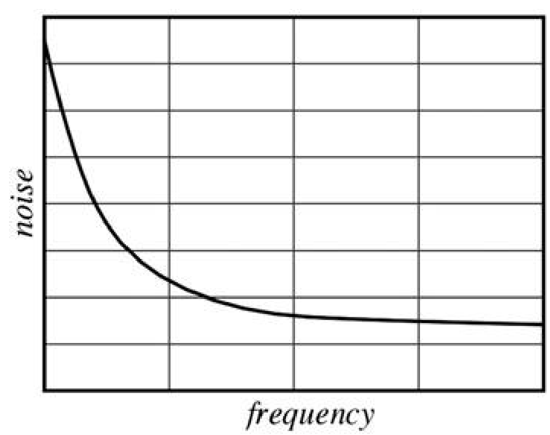


Fig. 5.43 Spectral distribution of $1/f$ "pink" noise

A peculiar ac noise mechanism is sometimes seen on the screen of an oscilloscope when observing the output of an operational amplifier, a principal building block of many sensor interface circuits. It looks like a digital signals transmitted from outer space; noise has a shape of square pulses having variable duration of many milliseconds. This abrupt type of noise is called popcorn noise because of the sound it makes coming from a loudspeaker. Popcorn noise is caused by defects that are dependent on the integrated circuits manufacturing techniques. Thanks to advances fabricating technologies, this type of noise is drastically reduced in modern semiconductor devices.

A combined noise from all voltage and current sources is given by sum of squares of individual noise voltages:

$$e_E = \sqrt{e_{n1}^2 + e_{n2}^2 + \cdots + (R_1 i_{n1})^2 + (R_1 i_{n2})^2 + \cdots}. \tag{5.52}$$

A combined random noise may be presented by its root mean square (r.m.s.) value, that is

$$E_{rms} = \sqrt{\frac{1}{T}\int_0^T e^2 dt}, \tag{5.53}$$

where T is time of observation, e is noise voltage and t is time.

Also, noise may be characterized in terms of the peak values, which are the differences between the largest positive and negative peak excursions observed during an arbitrary interval. For some applications, in which peak-to-peak (*p-p*) noise may limit the overall performance (in a threshold-type devices), *p-p* measurement may be essential. Yet, due to a generally Gaussian distribution of noise signal, *p-p* magnitude is very difficult to measure in practice. Because *r.m.s.* values are so much easier to measure repeatedly, and they are the most usual form for presenting noise data noncontroversially, the Table 5.3 should be useful for estimating the probabilities of exceeding various peak values given by the *r.m.s.* values. The casually observed *p-p* noise varies between 3 × r.m.s. and 8 × *r.m.s.*, depending on the patience of observer and amount of data available.

Table 5.3 Peak-to-peak value vs. *r.m.s.* (for Gaussian distribution)

Nominal p-p voltage	% of time that noise will exceed nominal p-p value
2 × r.m.s.	32.0%
3 × r.m.s.	13.0%
4 × r.m.s.	4.6%
5 × r.m.s.	1.2%
6 × r.m.s.	0.27%
7 × r.m.s.	0.046%
8 × r.m.s.	0.006%

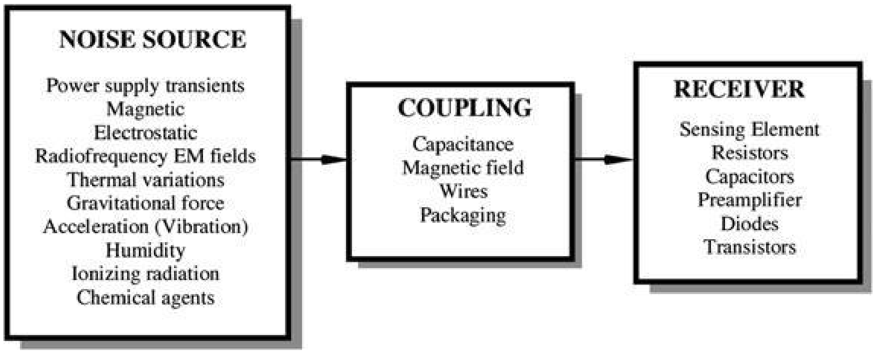


Fig. 5.44 Sources and coupling of transmitted noise

5.13.2 Transmitted Noise

A large portion of environmental stability is attributed to immunity of a sensor and an interface circuit to noise, which is originated in external sources. Figure 5.44 shows a diagram of the transmitted noise propagation. Noise comes from a source which often cannot be identified. Examples of the sources are: voltage surges in power lines, lightnings, change in ambient temperature, sun activity, etc. These interferences propagate toward the sensor and the interface circuit, and to present a problem eventually must appear at the output. However, before that, they somehow must affect the sensing element inside the sensor, its output terminals or the electronic components in the circuit. Both the sensor and circuit act as receivers of the interferences.

There can be several classifications of transmitted noise, depending on how it affects the output signal, how it enters the sensor or circuit, etc. With respect to its relation to the output signals, noise can be either additive or multiplicative.

Additive noise e_n is added to the useful signal V_s and mixed with it as a fully independent voltage (or current)

$$V_{out} = V_s + e_n. \tag{5.54}$$

An example of such a disturbance is depicted in Fig. 5.45b. It can be seen that the noise magnitude does not change when the actual (useful) signal changes. As long as the sensor and interface electronics can be considered linear, the additive noise magnitude is totally independent of the signal magnitude and, if the signal is equal to zero, the output noise still will be present.

Multiplicative noise affects the sensor's transfer function or the circuit's nonlinear components in such a manner as V_s signal's value becomes altered or *modulated* by the noise:

$$V_{out} = [1 + N(t)]V_s, \tag{5.55}$$

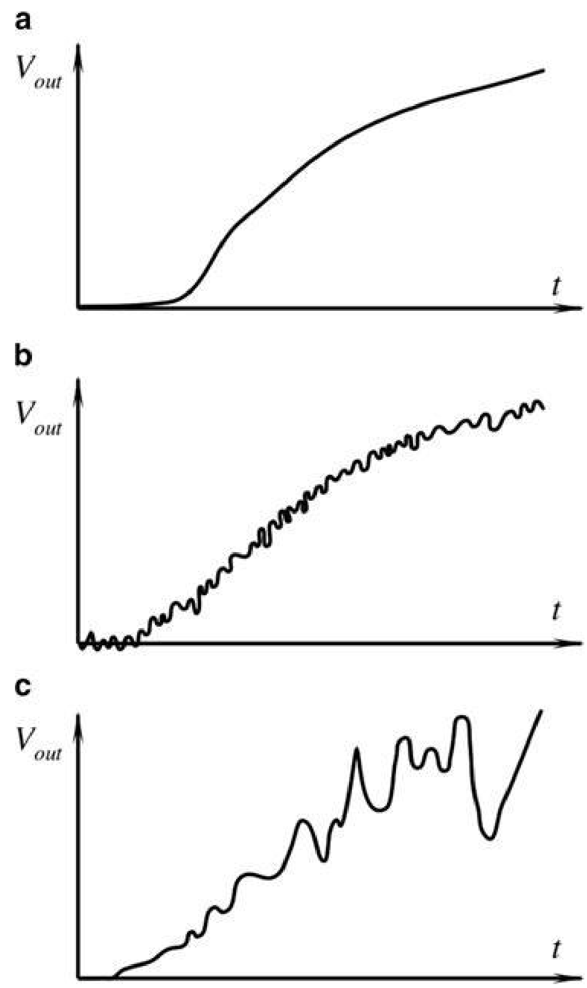


Fig. 5.45 Types of noise noise-free signal (**a**); additive noise (**b**); multiplicative noise (**c**)

where $N(t)$ is a function of noise. An example of such noise is shown in Fig. 5.45c. Multiplicative noise at the output disappears or becomes small (it also becomes additive) when the signal magnitude nears zero. Multiplicative noise grows together with the signal's V_s magnitude. As its name implies, multiplicative noise is a result of multiplication (which is essentially a nonlinear operation) of two values where one is a useful signal and the other is a noise dependent spurious signal.

To improve noise stability against transmitted additive noise, quite often sensors are combined in pairs, that is, they are fabricated in a dual form whose output signals are subtracted from one another (Fig. 5.46). This method is called a differential technique (see also Sect. 5.10). One sensor of the pair (it is called the main sensor) is subjected to a stimulus of interest s_1, while the other (reference) is shielded from stimulus perception. Since additive noise is specific for the linear or quasilinear sensors and circuits, the reference sensor does not have to be subjected to any particular stimulus. Often, it may be equal to zero. It is anticipated that both sensors are subjected to identical transmitted noise (nose generated inside the sensor cannot be cancelled by a differential technique), which it is said, is

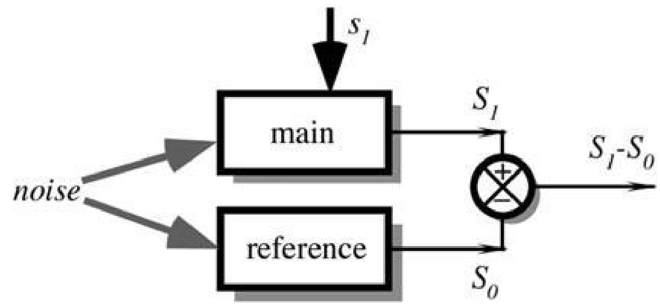


Fig. 5.46 Differential technique

a common-mode noise. This means that noisy effects at both sensors are in-phase and have the same magnitude. If both sensors are identically influenced by common-mode spurious stimuli, the subtraction removes the noise component. Such a sensor is often called either a dual or a differential sensor. The quality of noise rejection is described by a number which is called the common-mode rejection ratio (CMRR):

$$\text{CMMR} = 0.5\frac{S_1 + S_0}{S_1 - S_0}, \tag{5.56}$$

where S_1 and S_0 are output signals from the main and reference sensors, respectively. CMRR may depend on magnitude of stimuli and usually becomes smaller at greater input signals. The ratio shows how many times stronger the actual stimulus will be represented at the output, with respect to a common mode noise having the same magnitude. The value of the CMRR is a measure of the sensor's symmetry. To be an effective means of noise reduction, both sensors must be positioned as close as possible to each other, they must be very identical and subjected to the same environmental conditions. Also, it is very important that the reference sensor is reliably shielded from the actual stimulus, otherwise the combined differential response will be diminished.

To reduce transmitted multiplicative noise, a ratiometric technique is quite powerful (see Sect. 5.9 for circuit description). Its principle is quite simple. The sensor is fabricated in a dual form where one part is subjected to the stimulus of interest and both parts are subjected to the same environmental conditions, which may cause transmitted multiplicative noise. The second sensor is called reference because a constant environmentally stable reference stimulus s_o is applied to its input. For example, the output voltage of a sensor in a narrow temperature range may be approximated by equation

$$V_1 \approx [1 + \alpha(T - T_0)]f(s_1), \tag{5.57}$$

where α is the temperature coefficient of the sensor's transfer function, T is the temperature, and T_0 is the temperature at calibration. The reference sensor whose reference input is s_0 generates voltage

$$V_0 \approx [1 + \alpha(T - T_0)]f(s_0). \tag{5.58}$$

We consider ambient temperature as a transmitted multiplicative noise which affects both sensors in the same way. Taking ratio of the above equations we arrive at

$$\frac{V_1}{V_0} = \frac{1}{f(s_0)} f(s_1). \tag{5.59}$$

Since $f(s_0)$ is constant, the ratio is not temperature-dependent. It should be emphasized however that the ratiometric technique is useful only when the anticipated noise has a multiplicative nature, while a differential technique works only for additive transmitted noise. Neither technique is useful for inherent noise, which is generated internally in sensors and circuits.

While inherent noise is mostly Gaussian, the transmitted noise is usually less suitable for conventional statistical description. Transmitted noise may be periodic, irregularly recurring, or essentially random, and it ordinarily may be reduced substantially by taking precautions to minimize electrostatic and electromagnetic pickup from power sources at line frequencies and their harmonics, radio broadcast stations, arcing of mechanical switches, and current and voltage spikes resulting from switching in reactive (having inductance and capacitance) circuits. Such precautions may include filtering, decoupling, shielding of leads and components, use of guarding potentials, elimination of ground loops, physical reorientation of leads, components and wires, use of damping diodes across relay coils and electric motors, choice of low impedances where possible, and choice of power supply and references having low noise. Transmitted noise from vibration may be reduced by proper mechanical design. A list outlining some of the sources of transmitted noise, their typical magnitudes, and some ways of dealing with them is shown in Table 5.4.

Table 5.4 Typical sources of transmitted noise (adapted from [13])

External source	Typical magnitude	Typical cure
60/50 Hz power	100 pA	Shielding; attention to ground loops; isolated power supply
120/100 Hz supply ripple	3 μV	Supply filtering
180/150 Hz magnetic pickup from saturated 60/50 Hz transformers	0.5 μV	Reorientation of components
Radio broadcast stations	1 mV	Shielding
Switch-arcing	1 mV	Filtering of 5 to 100 MHz components; attention to ground loops and shielding
Vibration	10 pA (10–100 Hz)	Proper attention to mechanical coupling; elimination of leads with large voltages near input terminals and sensors
Cable vibration	100 pA	Use a low noise (carbon coated dielectric) cable
Circuit boards	$0.01 - 10\,\mathrm{pA}/\sqrt{\mathrm{Hz}}$ below 10 Hz	Clean board thoroughly; use Teflon insulation where needed and guard well

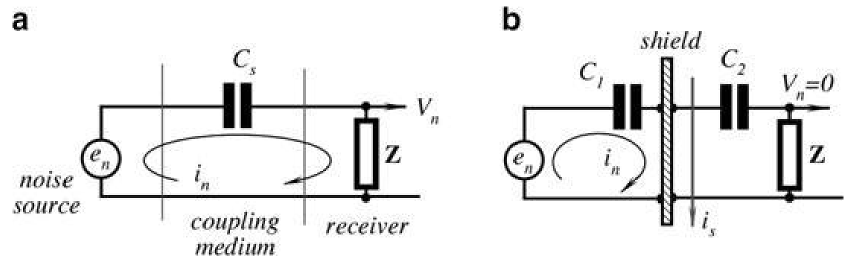


Fig. 5.47 Capacitive coupling (**a**) and electric shield (**b**)

The most frequent channel for the coupling of electrical noise is a "parasitic" capacitance. Such a coupling exists everywhere. Any object is capacitively coupled to another object. For instance, a human standing on isolated earth develops a capacitance to ground on the order of 700 pF, electrical connectors have a pin-to-pin capacitance of about 2 pF, an optoisolator has an emitter-detector capacitance of about 2 pF. Figure 5.47a shows that an electrical noise source is connected to the sensor's internal impedance **Z** through a coupling capacitance C_s. That impedance may be a simple resistance or a combination of resistors, capacitors, inductors, and nonlinear elements, like diodes. Voltage across the impedance is a direct result of the change rate in the noise signal, the value of coupling capacitance C_s and impedance **Z**. For instance, a pyroelectric detector may have an internal impedance, which is equivalent to a parallel connection of a 30 pf capacitor and a 50 GΩ resistor. The sensor may be coupled through just 1 pf to a moving person who has the surface electrostatic charge on the body resulting in static voltage of 1,000 V. If we assume that the main frequency of human movement is 1 Hz, the sensor would pickup the electrostatic interference of about 30 V. This is 3 to 5 orders of magnitude higher than the sensor would normally produce in response to thermal radiation received from the human body.

Since some sensors and virtually all electronic circuits have nonlinearities, high-frequency interference signals, generally called RFI (radiofrequency interference) or EMI (electromagnetic interferences), may be rectified and appear at the output as a dc or slow changing voltage.

5.13.3 Electric Shielding

Interferences attributed to electric fields can be significantly reduced by appropriate shielding of the sensor and circuit, especially of high impedance and nonlinear components. Each shielding problem must be analyzed separately and carefully. It is very important to identify the noise source and how it is coupled to the circuit. Improper shielding and guarding may only make matters worse or create a new problem.

A shielding serves two purposes [14]. First, it confines noise to a small region. This will prevent noise from getting into nearby circuits. However, the problem with such shields is that the noise captured by the shield can still cause problems if the return path that the noise takes is not carefully planned and implemented by an understanding of the ground system and making the connections correctly.

Second, if noise is present in the circuit, shields can be placed around critical parts to prevent the noise from getting into sensitive portions of the detectors and circuits. These shields may consist of metal boxes around circuit regions or cables with shields around the center conductors.

As it was shown in Sect. 3.1, the noise that resulted from the electric fields can be well controlled by metal enclosures because charge q cannot exist on the interior of a closed conductive surface. Coupling by a mutual, or stray, capacitance can be modeled by a circuit shown in Fig. 5.47a. Here e_n is a noise source. It may be some kind of a part or component whose electric potential varies. C_s is the stray capacitance (having impedance Z_s at a particular frequency) between the noise source and the circuit impedance Z, which acts as a receiver of the noise. Voltage V_n is a result of the capacitive coupling. A noise current is defined as

$$i_n = \frac{V_n}{Z + Z_s}, \tag{5.60}$$

and actually produces noise voltage

$$V_n = \frac{e_n}{\left(1 + \frac{Z_c}{Z}\right)}. \tag{5.61}$$

For example, if $C_s = 2.5$pf, Z= 10 kΩ (resistor) and $e_n = 100$ mV, at 1.3 MHz, the output noise will be 20 mV.

One might think that 1.3 MHz noise is relatively easy to filter out from low-frequency signals produced by a sensor. In reality, it cannot be done, because many sensors and, especially the front stages of the amplifiers, contain nonlinear components (*pn*-semiconductor junctions), which act as rectifiers. As a result, the spectrum of high-frequency noise shifts into a low-frequency region, making the noise signal similar to voltage produced by a sensor.

When a shield is added, the change to the situation is shown in Fig. 5.47b. With the assumption that the shield has zero impedance, the noise current at the left side will be $i_n = e_n/Z_c$. On the other side of the shield, noise current will be essentially zero since there is no driving source at the right side of the circuit. Subsequently, the noise voltage over the receiving impedance will also be zero and the sensitive circuit becomes effectively shielded from the noise source. One must be careful, however, that there is no significant currents i_s flow over the shield. Coupled with the shield resistance, these may generate additional noise. There are several practical rules that must be observed when applying electrostatic shields.

- An electrostatic shield, to be effective, should be connected to the reference potential of any circuitry contained within the shield. If the signal is connected to a ground (chassis of the frame or to earth), the shield must be connected to that ground. Grounding of shield is useless if the signal is not returned to the ground.

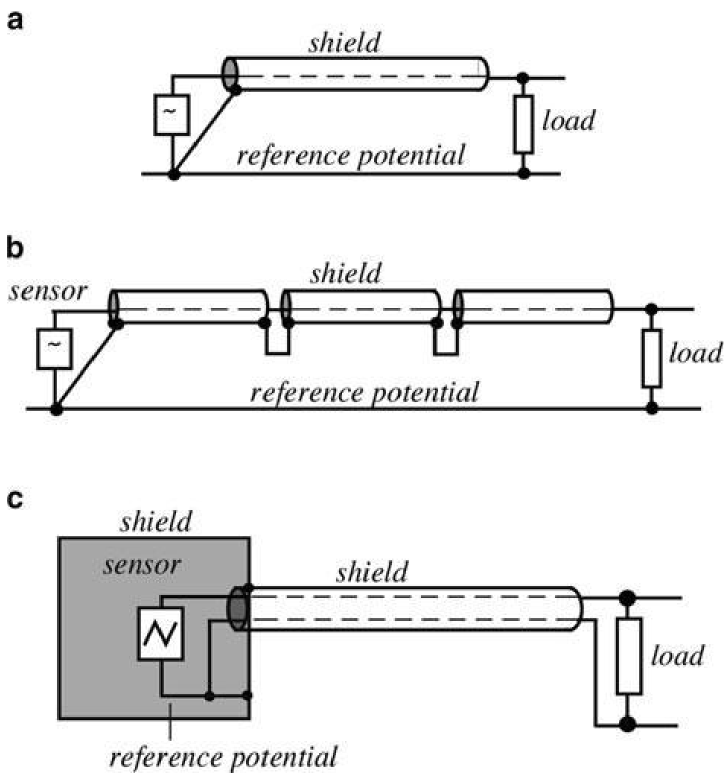
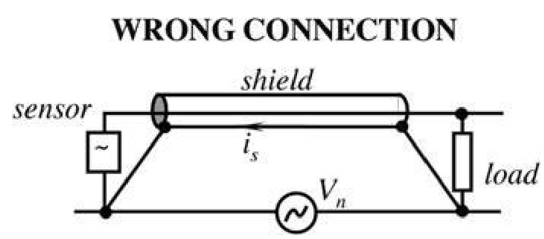


Fig. 5.48 Connections of an input cable to a reference potential

Fig. 5.49 Cable shield is erroneously grounded at both ends



- If a shielding cable is used, its shield must be connected to the signal referenced node at the signal source side (Fig. 5.48a).
- If the shield is split into sections, as might occur if connectors are used, the shield for each segment must be tied to those for the adjoining segments, and ultimately connected only to the signal referenced node (Fig. 5.49b).
- The number of separate shields required in a data acquisition system is equal to the number of independent signals that are being measured. Each signal should have its own shield, with no connection to other shields in the system, unless they share a common reference potential (signal "ground"). In that case all connections must be made by a separate jumping wire connected to each shield at a single point.

- A shield must be grounded only at one point, preferably next to the sensor. A shielded cable must *never* be grounded at both ends (Fig. 5.49). The potential difference (V_n) between two "grounds" will cause shield current i_s to flow, which may induce a noise voltage into the center conductor via magnetic coupling.
- If a sensor is enclosed into a shield box and data are transmitted via a shielded cable (Fig. 5.48c), the cable shield must be connected to the box. It is a good practice to use a separate conductor for the reference potential ("ground") inside the shield, and not to use the shield for any other purposes except shielding: do not allow shield current to exist.
- Never allow the shield to be at any potential with respect to the reference potential (except in case of driven shields as shown in Fig. 5.4b). The shield voltage couples to the center conductor (or conductors) via a cable capacitance.
- Connect shields to a ground via short wires to minimize inductance. This is especially important when both analog and digital signals are transmitted.

5.13.4 Bypass Capacitors

The bypass capacitors are used to maintain low power supply impedance at the point of a load. Parasitic resistance and inductance in supply lines mean that the power supply impedance can be quite high. As the frequency goes up, the inductive parasitic becomes troublesome and may result in the circuit oscillation or ringing effects. Even if the circuit operates at lower frequencies, the bypass capacitors are still important as high-frequency noise may be transmitted to the circuit and power supply conductors from external sources, for instance radio stations. At high frequencies, no power supply or regulator has zero output impedance. What type of capacitor to use is determined by the application, frequency range of the circuit, cost, board space, and some other considerations. To select a bypass capacitor, one must remember that a practical capacitor at high frequencies may be far away from the idealized capacitor, which is described in textbooks.

A generalized equivalent circuit of a capacitor is shown in Fig. 5.50. It is comprised of a nominal capacitance C, leakage resistance r_l, lead inductances L, and resistances R. Further, it includes dielectric absorption terms r and c_a, which are manifested in capacitor's "memory." In many interface circuits, especially amplifiers, analog integrators and current (charge)-to-voltage converters, dielectric

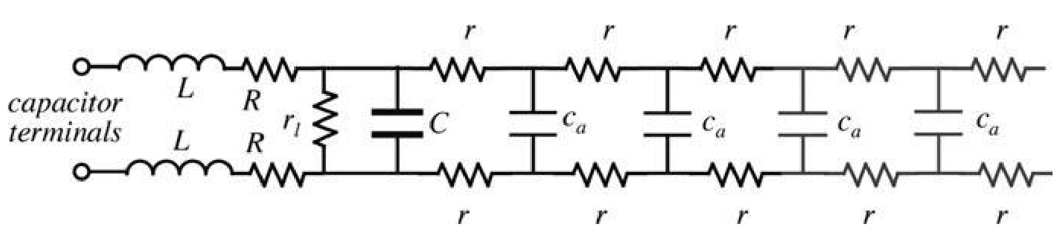


Fig. 5.50 Equivalent circuit of a capacitor

absorption is a major cause for errors. In such circuits, film capacitors should be used whenever possible.

In bypass applications, r_l and dielectric absorption are second order terms but series R and L are of importance. They limit the capacitor's ability to damp transients and maintain a low power supply output impedance. Often, bypass capacitors must be of large values (10 μF or more) so they can absorb longer transients, thus electrolytic capacitors are often employed. Unfortunately, these capacitors have large series R and L. Usually, tantalum capacitors offer better results, however, a combination of aluminum electrolytic with nonpolarized (ceramic or film) capacitors may offer even further improvement. Nowadays, high-volume ceramic capacitors are available for low price. A combination of wrong types of bypass capacitors may lead to ringing, oscillation and crosstalk between data communication channels. The best way to specify a correct combination of bypass capacitors is to first try them on a breadboard.

5.13.5 Magnetic Shielding

Proper shielding may dramatically reduce noise resulting from electrostatic and electrical fields. Unfortunately, it is much more difficult to shield against magnetic fields because it penetrates conducting materials. A typical shield placed around a conductor and grounded at one end has little if any effect on the magnetically induced voltage in that conductor. As magnetic field B_o penetrates the shield, its amplitude drops exponentially (Fig. 5.51b). The skin depth δ of the shield is the depth required for the field attenuation by 37% of that in the air. Table 5.5 lists typical values of δ for several materials at different frequencies. At high frequencies, any material from the list may be used for effective shielding, however at a lower range steel yields a much better performance.

For improving low-frequency magnetic field shielding, a shield consisting of a high-permeability magnetic material (e.g., mumetal) should be considered. However, the mumetal effectiveness drops at higher frequencies and strong magnetic fields. An effective magnetic shielding can be accomplished with thick steel shields

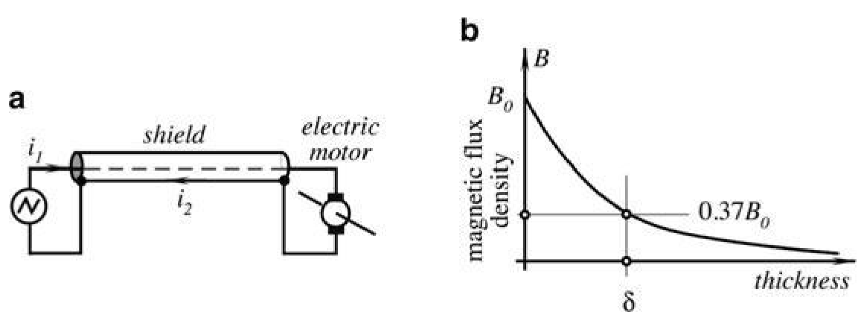


Fig. 5.51 Reduction of a transmitted magnetic noise by powering a load device through a coaxial cable (**a**); Magnetic shielding improves with the thickness of the shield (**b**)

Table 5.5 Skin depth, δ, in mm versus frequency (adapted from [15])

Frequency	Copper	Aluminum	Steel
60 Hz	8.5	10.9	0.86
100 Hz	6.6	8.5	0.66
1 kHz	2.1	2.7	0.20
10 kHz	0.66	0.84	0.08
100 kHz	0.2	0.3	0.02
1 MHz	0.08	0.08	0.008

at higher frequencies. Since magnetic shielding is very difficult, the most effective approach at low frequencies is to minimize the strength of magnetic fields, minimize the magnetic loop area at the receiving end, and selecting the optimal geometry of conductors. Some useful practical guidelines are as follows:

- Locate the receiving circuit as far as possible from the source of the magnetic field.
- Avoid running wires parallel to the magnetic field; instead, cross the magnetic field at right angles.
- Shield the magnetic field with an appropriate material for the frequency and strength.
- Use a twisted pair of wires for conductors carrying the high-level current that is the source of the magnetic field. If the currents in the two wires are equal and opposite, the net field in any direction over each cycle of twist will be zero. For this arrangement to work, none of the current can be shared with another conductor, for example, a ground plane, which may result in ground loops.
- Use a shielded cable with the high-level source circuit's return current carried by the shield (Fig. 5.51a). If the shield current i_2 is equal and opposite to that of the center conductor i_1, the center conductor field and the shield field will cancel, producing a zero net field. This case seems a violation of a rule "no shield currents" for the receiver's circuit, however, the shielded cable here is not used to electrostatically shield the center conductor. Instead, the geometry produces a cancellation of the magnetic field which is generated by a current supplied to a "current-hungry" device (an electric motor in this example).
- Since magnetically induced noise depends on the area of the receiver loop, the induced voltage due to magnetic coupling can be reduced by making the loop's area smaller.

What is the receiver's loop? Figure 5.52 shows a sensor, which is connected to the load circuit via two conductors having length L and separated by distance D. The rectangular circuit forms a loop area $a = L{\cdot}D$. The voltage induced in series with the loop is proportional to the area and cosine of its angle to the field. Thus, to minimize noise, the loop should be oriented at right angles to the field, and its area should be minimized.

The area can be decreased by reducing the length of the conductors and/or decreasing the distance between the conductors. This is easily accomplished with a twisted pair, or at least with a tightly cabled pair of conductors. It is a good practice to pair the conductors so that the circuit wire and its return path will always

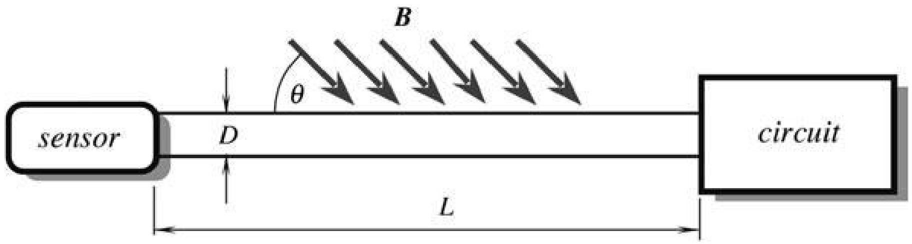


Fig. 5.52 Receiver's loop is formed by long conductors

be together. This requirement shall not be overlooked. For instance, if wires are correctly positioned by a designer, a service technician may reposition them during the repair work. A new wire location may create a disastrous noise level. Hence, a general rule is to know the area and orientation of the wires and permanently secure the wiring.

5.13.6 Mechanical Noise

Vibration and acceleration effects are also sources of transmitted noise in sensors, which otherwise should beimmune to them. These effects may alter transfer characteristics (multiplicative noise) or the sensor may generate spurious signals (additive noise). If a sensor incorporates certain mechanical elements, vibration along some axes with a given frequency and amplitude may cause resonant effects. For some sensors, an acceleration is a source of noise. For instance, pyroelectric detectors possess piezoelectric properties. The main function of a pyroelectric detector is to respond to thermal gradients. However, such environmental mechanical factors as a fast changing air pressure, strong wind or structural vibrations cause the sensor to respond with output signals, which often are indistinguishable from responses to normal stimuli. If this is the case, a differential noise cancellation may be quite efficient (see Sect. 5.8).

5.13.7 Ground Planes

For many years, ground planes have been known to electronic engineers and printed circuit designers as a "mystical and ill-defined" cure for spurious circuit operation [16]. Ground planes are primarily useful for minimizing circuit inductance. They do this by utilizing the basic magnetic theory. Current flowing in a wire produces an associated magnetic field (Sect. 3.3.1). The field's strength is proportional to the current i and inversely related to the distance r from the conductor:

$$B = \frac{\mu_0 i}{2\pi r}. \tag{5.62}$$

Thus, we can imagine a current-carrying wire surrounded by a magnetic field. Wire inductance is defined as energy stored in the field setup by the wire's current. To compute the wire's inductance requires integrating the field over the wire's length and the total area of the field. This implies integrating on the radius from the wire surface to infinity. However, if two wires carrying the same current in opposite directions are in close proximity, their magnetic fields are canceled. In this case, the virtual wire inductance is much smaller. An opposite flowing current is called *return current*. This is the underlying reason for ground planes. A ground plane provides a return path directly under the signal carrying conductor through which return current can flow. Return current has a direct path to ground, regardless of the number of branches associated with the conductor. Currents will always flow through the return path of the lowest impedance. In a properly designed ground plane, this path is directly under the signal conductor.

In practical circuits, a ground plane is one side of the board and the signal conductors are on the other. In the multilayer boards, a ground plane is usually sandwiched between two or more conductor planes. Aside from minimizing parasitic inductance, ground planes have additional benefits. Their flat surface minimizes resistive losses due to "skin effect" (ac current travel along a conductor's surface). Additionally, they aid the circuit's high-frequency stability by referring stray capacitance to the ground. Even though ground planes are very beneficial for digital circuits using them for current return of analog sensor signals are dangerous likely digital currents in a ground will create strong interferences in the analog part of the circuit.

Some practical suggestions:

- Make ground planes of as much area as possible on the components side (or inside for the multilayer boards). Maximize the area especially under traces that operate with high frequency or digital signals.
- Mount components that conduct fast transient currents (terminal resistors, ICs, transistors, decoupling capacitors, etc.) as close to the board as possible.
- Wherever a common ground reference potential is required, use separate conductors for the reference potential and connect them all to the ground plane at a common point to avoid voltage drops due to ground currents.
- Use separate nonoverlapping ground planes for digital and analog sections of the circuit board and connect them at one point only at the power supply terminals.
- Keep the trace length short. Inductance varies directly with length and no ground plane will achieve perfect cancellation.

5.13.8 Ground Loops and Ground Isolation

When a circuit is used for low-level input signals, a circuit itself may generate enough noise and interferences to present a substantial problem for accuracy. Sometimes, when a circuit is correctly designed on paper, a bench breadboard shows quite a satisfactory performance, however, when a production prototype with the printed circuit board is tested, the accuracy requirement is not met. A difference between a bread board and PC board prototypes may be in the physical layout of conductors. Usually, conductors between electronic components are quite specific; they may connect a capacitor to a resistor, a gate of a JFET transistor to the output of an operational amplifier, etc. However, there are at least two conductors which, in most cases, are common for the majority of the electronic circuit. These are the power supply bus and the ground bus. Both of them may carry undesirable signals from one part of the circuit to another, specifically, they may couple strong output signals to the sensors and input stages.

A power supply bus carries supply currents to all stages. A ground bus also carries supply currents, but, in addition, it is often used to establish a reference base for an electrical signal. For any measurement circuit cleanliness of a reference base is essential. Interaction of the two functions (power supply and reference) may lead to a problem which is known as *ground loop*. We illustrate it in Fig. 5.53a where a sensor is connected to a positive input of an amplifier which may have a substantial gain. The amplifier is connected to the power supply and draws current i which is returned to the ground bus as i'. A sensor generates voltage V_s, which is fed to the positive input of the amplifier. A ground wire is connected to the circuit at point a, right next to the sensor's terminal. A circuit has no visible error sources, nevertheless, the output voltage contain substantial error. A noise source is developed in a wrong connection of ground wires. Figure 5.53b shows that the ground conductor is not ideal. It may have some finite resistance R_g and inductance L_g. In this example, supply current while returning to the battery from the amplifier passes through the ground bus between points b and a resulting in voltage drop V_g. This drop, however small, may be comparable with the signal produced by the sensor. It should be noted that voltage V_g is serially connected with the sensor and is directly applied to the amplifier's input. In other words, the sensor is not referenced to a clean ground. Ground currents may also contain high-frequency components, then the bus inductance will produce quite strong spurious high-frequency signals, which not only add

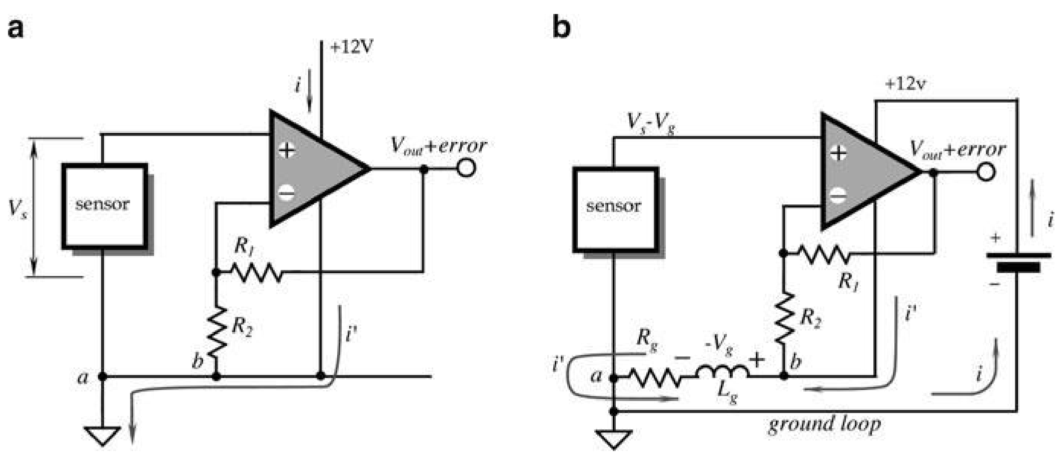


Fig. 5.53 Wrong connection of a ground terminal to a circuit (**a**); Path of a supply current through the ground conductors (**b**)

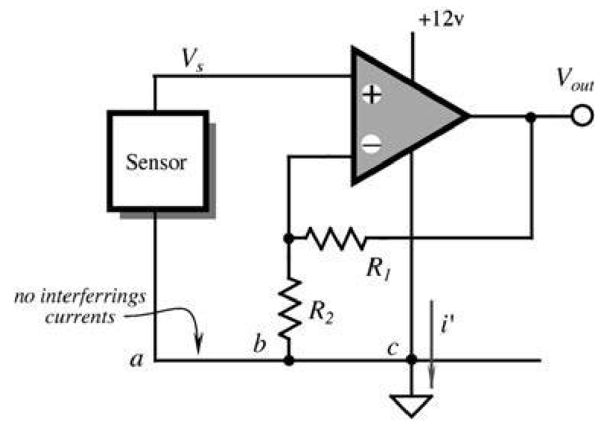


Fig. 5.54 Correct grounding of a sensor and interface circuit

noise to the sensor, but may cause circuit instability as well. For example, let us consider a thermopile sensor, which produces voltage corresponding to 100 μV/°C of the object's temperature. A low-noise amplifier has quiescent current, $i = 2$ mA, which passes through the ground loop having resistance $R_g = 0.2\ \Omega$. The ground loop voltage $V_g = iR_g = 0.4$ mV corresponds to an error of -4°C!

The cure is usually quite simple: ground loops must be broken. The most critical rule of the circuit board design: never use the same conductor for the reference potential and power supply currents. A circuit designer should always separate a reference ground from the current carrying grounds, especially serving digital devices. Thus, it is advisable to have at least three grounds: reference, analog, and digital.

The reference ground shall be used only for connecting the sensor components that produce low level input signals, all analog amplifier input components that need be referenced to a ground potential, and the reference input of an A/D converter. The analog ground shall be used exclusively for return currents from the analog interface circuits. And the digital ground shall be used only for binary signals, like microprocessors, digital gates, etc. There may be a need for additional "grounds," for example those that carry relatively strong currents, especially containing high-frequency signals (LEDs, relays, motors, heaters, etc.). Figure 5.54 shows that moving the ground connection from the sensor's point *a* to the power terminal point *c* prevents formation of spurious voltages across the ground conductor connected to the sensor and a feedback resistor R_2.

A rule of thumb is to join all "grounds" on a circuit board only at one point, preferably at the power source. Grounding at two or more spots may form ground loops, which often is very difficult to diagnose.

5.13.9 Seebeck Noise

This noise is a result of the Seebeck effect (Sect. 3.9), which is manifested as the generation of an electromotive force (e.m.f.) when two dissimilar metals are joined

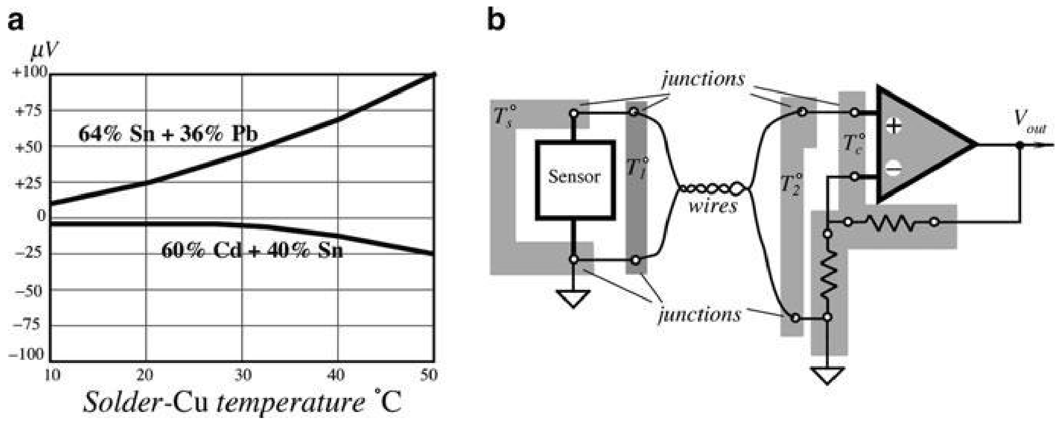


Fig. 5.55 Seebeck *e.m.f.* developed by solder-copper joints (**a**) (adapted from [17]); Maintaining joints at the same temperature reduces Seebeck noise (**b**)

together. The Seebeck e.m.f. is small and for many sensors may be simply ignored. However, when absolute accuracy on the order of 10–100 µV is required, that noise must be taken into account. The connection of two dissimilar metals produces a temperature sensor. However, when temperature sensing is not a desired function, a thermally induced *e.m.f.* is a spurious signal. In electronic circuits, connection of dissimilar metals can be found everywhere: connectors, switches, relay contacts, sockets, wires, etc. For instance, the copper PC board cladding connected to kovar™[5] input pins of an integrated circuit creates an offset voltage of 40 $\mu V \cdot \Delta T$ where ΔT is the temperature gradient in °C between two dissimilar metal contacts on the board. The common lead-tin solder, when used with the copper cladding creates a thermoelectric voltage between 1 and 3 µV/°C. There are special cadmium-tin solders available to reduce these spurious signals down to 0.3 µV/°C. Figure 5.55a shows Seebeck *e.m.f.* for two types of solder. Connection of two identical wires fabricated by different manufacturers may result in voltage having slope on the order of 200 nV/°C.

In many cases, Seebeck e.m.f. may be eliminated by a proper circuit layout and thermal balancing. It is a good practice to limit the number of junctions between the sensor and the front stage of the interface circuit. Avoid connectors, sockets, switches and other potential sources of *e.m.f.* to the extent possible. In some cases this will not be possible. In these instances, attempt to balance the number and type of junctions in the circuit's front stage so that differential cancellations occur. Doing this may involve deliberately creating and introducing junctions to offset necessary junctions. Junctions, which intent to produce cancellations, must be maintained at the same temperature. Figure 5.55b shows a remote sensor connection to an amplifier where the sensor junctions, input terminal junctions, and amplifier components junctions are all maintained while at different but properly arranged temperatures. Such thermally balanced junctions must be maintained at a

[5]Trademark of Westinghouse Electric Corp.

close physical proximity and preferably on common heat sinks. Air drafts and temperature gradients in the circuit boards and sensor enclosures must be avoided.

5.14 Calibration

Many sensors have natural manufacturer's tolerances that move their transfer functions beyond the acceptable accuracy limits. We discussed the calibrating principles in detail in Chap. 2. Now, let us briefly examine some electrical techniques that can be employed in the interface circuits to trim the overall transfer functions including both the sensor and the interface circuit. The non-exhaustive list of techniques is as follows:

1. A "trimpot" or "trimcap," a variable resistor or capacitor to adjust the sensor's internal resistance or capacitance. This is the oldest and most traditional method. Figure 5.56a illustrates use of the trimpot R_p in the resistive Wheatstone bridge. Presently, the trimming resistors may have digital rather than mechanical adjustments. In digital potentiometers (for example, from www.maxim-ic.com) an 8-bit code can set the resistance value and be changed at any time. A more permanent nonmechanical trimming of a resistor can be done by using methods developed by Microbridge Technologies, Inc. (www.mbridgetech.com). The trimming resistor (given by the manufacturer a strange name "rejustor") is a semiconductor whose ohmic resistance can be modified by heating pulses applied through separate terminals (Fig. 5.56c) from a special trimming equipment. Resistance can be reduced from the nominal as-manufactured value by at least 30% and then can be adjusted up or down within the active range. The programmed resistance value is stored in the physical properties of the semiconductor and retained until it is reprogrammed. No power is required to maintain the resistance value; it is strictly a passive circuit element.

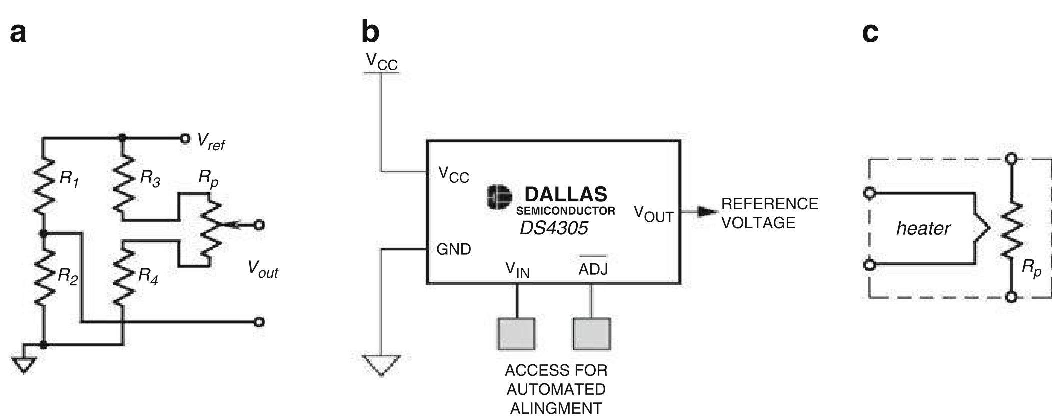


Fig. 5.56 Trimming circuits resistive bridge circuit with a trimming resistor (**a**); digitally adjustable voltage reference (**b**); and "rejustor," thermally trimmed resistor (**c**)

2. A matching, that is, selectable resistor or capacitor to match the corresponding resistance or capacitance of the sensor.
3. Adjust the sensor's reference signal. This can be done by modifying one of the excitation parameters of the active sensor's excitation signal, for example, amplitude or frequency. Alternatively, a reference voltage can be adjusted at the A/D converter. Figure 5.56b illustrates a digitally programmable voltage reference.
4. Digital code that is stored in a nonvolatile memory of the microprocessor and during every measurement is used to correct the sensor's response. The code is created during the calibration procedure. It may include numerous sensors characteristics, for example, coefficients for a polynomial approximation.

5.15 Batteries for Low-Power Sensors

Modern development of integrated sensors and need for long-term remote monitoring and data acquisition demand use of reliable and high-energy density power sources. History of battery development goes back to Volta and shows a remarkable progress during last decades. Well-known old electrochemical power sources improve dramatically. Examples are C-Zn, alkaline, Zn-air, NiCd, and lead-acid batteries. Nowadays, newer systems such as secondary Zn-air, Ni-metal-hydride, and especially lithium batteries are growing in use as new devices are designed around their higher voltage and superior shelf life. The Li-MnO_2 system dominates the commercial market where they range from miniature flat cells to "D" size cells.

All batteries can be divided into two groups: primary – single use devices and secondary (rechargeable) – multiple use devices.

Often, batteries are characterized by energy per unit weight, however, for miniature sensor applications energy per unit volume often becomes more critical (see in Appendix Table A-20).

In general, the energy delivered by a battery depends upon the rate at which power is withdrawn. Typically, as the current is increased the amount of energy delivered is decreased. Battery energy and power are also affected by construction of battery, the size, and the duty cycle of current delivery. Manufacturers usually specify batteries as ampere-hours or watt-hours when discharged at a specific rate to a specific voltage cut-off.

If the battery capacitance is C (in mA·hour) and the average current drain is I (mA), the time of a battery discharge (lifetime for a primary cell) is defined as

$$t = \frac{C}{In} \tag{5.63}$$

where n is a duty cycle. For instance, if the battery is rated as having capacity of 100 mA h, the circuit operating current consumption is about 5 mA and the circuit works only 3 min every hour (duty cycle is 3/60), the battery will last approximately for

$$t = \frac{C}{In} = \frac{100}{5\frac{3}{60}} = 400\,\mathrm{h}$$

Yet, the manufacturer's specification shall be used with a grain of salt and only as a guideline, because the specified discharge rate is rarely coincides with the actual power consumption. It is highly recommended to determine battery life experimentally, rather than rely on the calculation. When designing the electronic circuit, its power consumption shall be determined during various operating modes and over the operating temperature range. Then, these values of power consumption shall be used in simulation of the battery load to determine the useful life with a circuit-specific cut-off voltage in mind. Sometimes, a circuit draws high currents during short times (pulse mode) and the battery ability to deliver such pulse current should be evaluated. If a battery cannot deliver high pulse current, a parallel electrolytic capacitor serving as a storage tank may be considered.

It should be noted that the accelerated life tests of a battery shall be used with caution, since as it was noted above, the useful capacity of a battery greatly depends on the load, operational current profile, and a duty cycle.

5.15.1 *Primary Cells*

The construction of a battery cell determines its performance and cost. Most primary cells (disposable batteries) employ single thick electrodes arranges in parallel or concentric configuration and aqueous electrolytes. Most small secondary cells (rechargeable batteries) are designed differently; they use "wound" or "jelly roll" construction, in which long thin electrodes are wound into a cylinder and placed into a metal container. This results in a higher power density, but with decreased energy density and higher cost. Due to low conductivity of electrolytes, many lithium primary cells also use "wound" construction [18].

5.15.1.1 Leclanche (Carbon-Zinc) Batteries

These batteries use zinc as anode. There are two types of them. One uses natural manganese dioxide as the cathode with ammonium chloride electrolyte. A "premium" version uses electrolytic manganese dioxide as the cathode and a zinc chloride electrolyte. These batteries are still the most popular world-wide, especially in the Orient, being produced by over 200 manufacturers. Their use is about equal to that of the alkaline in Europe but is only near 25% of alkaline in the U.S.A. These batteries are preferred when high power density is not required, shelf life is not critical, but the low cost is a dominating factor.

5.15.1.2 Alkaline Manganese Batteries

Demand for these batteries grew significantly, especially after a major improvement—elimination of mercury from the zinc anode. The alkaline batteries are capable of delivering high currents, have improved power/density ratio and at least 5 years of shelf life (Table A.20)

5.15.1.3 Primary Lithium Batteries

Most of these batteries are being produced in Japan and China. The popularity of lithium-manganese dioxide cells grows rapidly thanks to their higher operating voltage, wide range of sizes and capacities, and excellent shelf life (Table A.21). Lithium iodine cells have very high energy density and allow up to 10 years of operation in a pacemaker (implantable heart rate controller). However, these batteries are designed with a low conductivity solid-state electrolyte and allow operation with very low current drain (in the order of microamperes), which in cases when passive sensors are employed often is quite sufficient.

Amount of lithium is batteries is quite small, because just 1 g is sufficient for producing capacity of 3.86Ah. Lithium cells are exempt from environmental regulations, but still are considered hazardous because of their flammability.

5.15.2 Secondary Cells

Secondary cells (Tables A.22 and A.23) are rechargeable batteries.

Sealed lead acid batteries offer small size at large capacities and allow about 200 cycles of life at discharge times as short as 1 h. The main advantages of these cells are low initial cost, low self-discharge, an ability to support heavy loads, to withstand harsh environments. Besides, these batteries have long life. The disadvantages include relatively large size and weight as well as potential environmental hazard due to presence of lead and sulfuric acid.

Sealed nickel-cadmium (NiCd) and nickel-metal hydrate (Ni-MH) are the most widely used secondary cells, being produced at volumes over 1 billion cells per year. Typical capacity for a "AA" cell is about 800 mAh and even higher from some manufacturers. This is possible thanks to use a high-porosity nickel foam or felt instead of traditional sintered nickel as carrier for the active materials. The NiCd cells are quite tolerant of overcharge and overdischarge. An interesting property of NiCd is that charging is endothermic process, which is the battery absorbs heat, while other batteries warm up when charging. Cadmium, however, presents potential environmental problem. Bi-MH and modern NiCd do not exhibit "memory" effect, that is, partial discharge does not influence their ability to fully recharge. The nickel-metal hydrate cells is nearly direct replacement for NiCd, yet they yield better capacity, but have somewhat poorer self-discharge.

A lithium polymer battery contain a nonliquid electrolyte, which makes it a solid-state battery. This allows to fabricate it in any size and shape, however, these batteries are most expensive.

Rechargeable alkaline batteries have low cost and good power density. However, their life cycles are quite low.

References

1. Widlar RJ (1980) Working with high impedance Op Amps, AN24, *Linear Application Handbook*. National Semiconductor
2. Pease RA (1983) Improve circuit performance with a 1-op-amp current pump. *EDN*, 85–90, Jan. 20
3. Sheingold DH (ed) (1986) *Analog-Digital Conversion Handbook*. 3rd ed., Prentice-Hall, Englewood Cliffs, NJ
4. Williams J (1990) Some techniques for direct digitization of transducer outputs, AN7, *Linear Technology Application Handbook*
5. Park YE, Wise KD (1983) An MOS switched-capacitor readout amplifier for capacitive pressure sensors. *IEEE Custom IC Conf* 380–384
6. Stafford KR, Gray PR, Blanchard RA (1974) A complete monolithic sample/hold amplifier. *IEEE J Solid-State Circuits* 9:381–387
7. Cho ST, Wise KD (1991) A self-testing ultrasensitive silicon microflow sensor. *Sensor Expo Proceedings*, 208B-1
8. Weatherwax S (1991) Understanding constant voltage and constant current excitation for pressure sensors. *SenSym Solid-State Sensor Handbook*. ©Sensym, Inc.
9. Coats MR (1991) New technology two-wire transmitters. *Sensors* 8(1)
10. Sheingold DH (ed) (1974) *Nonlinear Circuits Handbook*. Analog Devices, Inc. Northwood, MA
11. Johnson JB (1928) Thermal agitation of electricity in conductors. *Phys Rev*
12. The Best of Analog Dialogue. © Analog Devices, Inc. 1991
13. Rich A (1991) Shielding and guarding. *Best of Analog Dialogue*, ©Analog Devices, Inc.
14. Ott HW (1976) *Noise Reduction Techniques in Electronic Systems*. Wiley, New York
15. Williams J (1990) High speed comparator techniques, AN13. *Linear Applications Handbook*. ©Linear Technology Corp.
16. Pascoe G (1977) The choice of solders for high-gain devices. *New Electronics* (U.K.), Feb. 6
17. Powers RA (1995) Batteries for low power electronics. *Proc IEEE* 83(4):687–693
18. AVR121 (2005) Enhancing ADC resolution by oversampling. Atmel Application Note 8003A-AVR-09/05
19. Bell DA (1981) *Solid State Pulse Circuits*. 2nd ed. Reston Publishing Company, Inc., Reston, VA