SECTION 3-4

Driving ADC/DAC **Reference** Inputs

Walt Jung, Walt Kester

It might seem odd to include a section on voltage references in a book devoted primarily to op amp applications, but the relevance will shortly become obvious. Unfortunately, there is little standardization with respect to ADC/DAC voltage references. Some ADCs and DACs have internal references, while others do not. In some cases, the dc accuracy of a converter with an internal reference can often be improved by overriding the internal reference with a more accurate and stable external one.

Although the reference element itself can be either a bandgap, buried zener, or XFETTM (see Reference 1), practically all references have some type of output buffer op amp. The op amp isolates the reference element from the output and also provides drive capability. However, this op amp must obey the general laws relating to op amp stability, and that is what makes the topic of references relevant to the discussion. Figure 3-47 summarizes voltage reference considerations.

- · Data converter accuracy determined by the reference, whether internal or external
- · Bandgap, buried zener, XFET generally have on-chip output buffer op amp
- · Transient loading can cause instability and errors
- External decoupling capacitors may cause oscillation
- · Output may require external buffer to source and sink current
- · Reference voltage noise may limit system resolution

Figure 3-47: ADC/DAC voltage reference considerations

Note that a reference input to an ADC or DAC is similar to the analog input of an ADC, in that the internal conversion process can inject transient currents at that pin. This requires adequate decoupling to stabilize the reference voltage. Adding such decoupling might introduce instability in some reference types, depending on the output op amp design. Of course, a reference data sheet may not show any details of the output op amp, which leaves the designer in somewhat of a dilemma concerning whether or not it will be stable and free from transient errors. Fortunately, some simple lab tests can exercise a reference circuit for transient errors, and also determine stability for capacitive loading.

Chapter Three

Figure 3-48 shows the transients associated with the reference input of a typical successive approximation ADC. The ADC reference voltage input must be stabilized with a sufficiently large decoupling capacitor, in order to prevent conversion errors. The value of the capacitor required as C_B may range from below 1 μ F, to as high as 100 μ F. This capacitor must of course have a voltage rating greater than the reference voltage. Physically, it will be of minimum size when purchased in a surface-mount style.



Figure 3-48: Successive approximation ADCs can present a transient load to the reference

Note that in this case, a 1 μ F capacitor on the reference input is required to reduce the transients to an acceptable level. Note that the capacitor size can be electrically larger for further noise reduction—the tradeoff here is of course cost and PCB real estate. The AD780 will work with capacitors of up to 100 μ F.

A well-designed voltage reference is stable with heavy capacitive decoupling. Unfortunately, some are not, as shown in Figure 3-49, where the addition of C_L to the reference output (a 0.01 µF capacitor) actually increases the amount of transient ringing. Such references are practically useless in data converter applications, because some amount of local decoupling is almost always required at the converter.



Figure 3-49: Make sure reference is stable with large capacitive loads

A suitable op amp buffer might be added between the reference and the data converter. Many good references are available (such as the AD780) which are stable with an output capacitor. This type of reference

should be chosen for a data converter application, rather than incurring the further complication and expense of an op amp.

If very low noise levels are required from a reference, an additional low-pass filter followed by a low noise op amp can be used to achieve the desired performance. The reference circuit of Figure 3-50 is one such example (see References 2 and 3). This circuit uses external filtering and a precision low-noise op-amp to provide both very low noise and high dc accuracy. Reference U1 is a 2.5 V, 3.0 V, 5 V, or 10 V reference with a low noise buffered output. The output of U1 is applied to the R1–C1/C2 noise filter to produce a corner frequency of about 1.7 Hz.



Figure 3-50: Low-noise op amp with filtering yields reference noise performance (1.5V to 5 nV/ $\sqrt{\text{HZ}}$ @ 1 kHz

Electrolytic capacitors usually imply dc leakage errors, but the bootstrap connection of C1 causes its applied bias voltage to be only the relatively small drop across R2. This lowers the leakage current through R1 to acceptable levels. Since the filter attenuation is modest below a few Hertz, the reference noise still affects overall performance at low frequencies (i.e., <10 Hz).

A precision low noise unity-gain follower, such as the OP113, then buffers the output of the filter. With less than $\pm 150 \,\mu\text{V}$ of offset error and under $1 \,\mu\text{V}/^{\circ}\text{C}$ drift, the buffer amplifier's dc performance will not seriously affect the accuracy/drift of most references. For example, an ADR292E for U1 will have a typical drift of 3 ppm/°C, equivalent to 7.5 $\mu\text{V}/^{\circ}\text{C}$, higher than the buffer amplifier.

Almost any op amp will have a current limit higher than a typical IC reference, so this circuit allows greater current output. It also removes any load-related thermal errors that might occur when the reference IC is loaded directly.

Even lower noise op-amps are available, for 5–10 V use. The AD797 offers 1kHz noise performance less than $2nV/\sqrt{Hz}$ in this circuit, compared to about $5nV/\sqrt{Hz}$ for the OP113. With any buffer amplifier, Kelvin sensing can be used at the load point, a technique that eliminates I × R related output voltage errors.

References: Driving ADC/DAC Reference Inputs

- 1. Chapter 2, Walt Kester, Editor, **Practical Design Techniques for Power and Thermal Management**, Analog Devices, 1998, ISBN: 0-916550-19-2.
- Walt Jung, "Build an Ultra-Low-Noise Voltage Reference," Electronic Design Analog Applications Issue, June 24, 1993.
- 3. Walt Jung, "Getting the Most from IC Voltage References," **Analog Dialogue**, Vol. 28, No., 1994, pp. 13–21.

SECTION 3-5

Buffering DAC Outputs

Walt Kester, Paul Hendriks

General Considerations

Another important op amp application is buffering DAC outputs. Modern IC DACs provide either voltage or current outputs. Figure 3-51 shows three fundamental configurations, all with the objective of using an op amp for a buffered output voltage.



Figure 3-51: Buffering DAC outputs with op amps

Figure 3-51A shows a buffered voltage output DAC. In many cases, the DAC output can be used directly, without additional buffering. If an additional op amp buffer is needed, it is usually configured in a noninverting mode, with gain determined by R1 and R2.

There are two basic methods for dealing with a current output DAC. In Figure 3-51B, a voltage is simply developed across external load resistor, R_L . An external op amp can be used to buffer and/or amplify this voltage if required. Many DACs supply full-scale currents of 20 mA or more, thereby allowing reasonable voltages to be developed across fairly low value load resistors. For instance, fast settling video DACs typically supply nearly 30 mA full-scale current, allowing 1 V to be developed across a source and load terminated 75 Ω coaxial cable (representing a dc load of 37.5 Ω to the DAC output).

A direct method to convert the output current into a voltage is shown in Figure 3-51C, This circuit is usually called a current-to-voltage converter, or I/V. In this circuit, the DAC output drives the inverting input of an op amp, with the output voltage developed across the R2 feedback resistor. In this approach the DAC output always operates at virtual ground (which may give a linearity improvement vis-à-vis Figure 3-51B).

The general selection process for an op amp used as a DAC buffer is similar to that of an ADC buffer. The same basic specifications such as dc accuracy, noise, settling time, bandwidth, distortion, and so forth, apply to DACs as well as ADCs, and the discussion will not be repeated here. Rather, some specific application examples will be shown.

Chapter Three

Differential to Single-Ended Conversion Techniques

A general model of a modern current output DAC is shown in Figure 3-52. This model is typical of the AD976x and AD977x TxDACTM series (see Reference 1). Current output is more popular than voltage output, especially at audio frequencies and above. If the DAC is fabricated on a bipolar or BiCMOS process, it is likely that the output will sink current, and that the output impedance will be less than 500 Ω (due to the internal R/2R resistive ladder network). On the other hand, a CMOS DAC is more likely to source output current and have a high output impedance, typically greater than 100 k Ω .



Figure 3-52: Model of high speed DAC output

Another consideration is the output *compliance voltage*—the maximum voltage swing allowed at the output in order for the DAC to maintain its linearity. This voltage is typically 1 V to 1.5 V, but can vary depending upon the DAC. Best DAC linearity is generally achieved when driving a virtual ground, such as an op amp I/V converter.

Modern current output DACs usually have differential outputs, to achieve high CM rejection and reduce the even-order distortion products. Full-scale output currents in the range of 2 mA to 20 mA are common.

In most applications, it is desirable to convert the differential output of the DAC into a single-ended signal, suitable for driving a coax line. This can be readily achieved with an RF transformer, provided low frequency response is not required. Figure 3-53 shows a typical example of this approach. The high impedance current output of the DAC is terminated differentially with 50 Ω , which defines the source impedance to the transformer as 50 Ω .



Figure 3-53: Differential transformer coupling

The resulting differential voltage drives the primary of a 1:1 RF transformer, to develop a single-ended voltage at the output of the secondary winding. The output of the 50 Ω LC filter is matched with the 50 Ω load resistor R_L, and a final output voltage of 1 V p-p is developed.

The transformer not only serves to convert the differential output into a single-ended signal, but it also isolates the output of the DAC from the reactive load presented by the LC filter, thereby improving overall distortion performance.

An op amp connected as a differential to single-ended converter can be used to obtain a single-ended output when frequency response to dc is required. In Figure 3-54 the AD8055 op amp is used to achieve high bandwidth and low distortion (see Reference 2). The current output DAC drives balanced 25 Ω resistive loads, thereby developing an out-of-phase voltage of 0 V to 0.5 V at each output. The AD8055 is configured for a gain of 2, to develop a final single-ended ground-referenced output voltage of 2 V p-p. Note that because the output signal swings above and below ground, a dual-supply op amp is required.



Figure 3-54: Differential dc-coupled output using a dual supply op amp

The C_{FILTER} capacitor forms a differential filter with the equivalent 50 Ω differential output impedance. This filter reduces any slew induced distortion of the op amp, and the optimum cutoff frequency of the filter is determined empirically to give the best overall distortion performance.

Chapter Three

A modified form of the Figure 3-54 circuit can also be operated on a single supply, provided the CM voltage of the op amp is set to mid-supply (2.5 V). This is shown in Figure 3-55 below. The output voltage is 2 V p-p centered around a CM voltage of 2.5 V. This CM voltage can be either developed from the 5 V supply using a resistor divider, or directly from a 2.5 V voltage reference. If the 5 V supply is used as the CM voltage, it must be heavily decoupled to prevent supply noise from being amplified.



Figure 3-55: Differential dc-coupled output using a single-supply op amp

Single-Ended Current-to-Voltage Conversion

Single-ended current-to-voltage conversion is easily performed using a single op amp as an I/V converter, as shown in Figure 3-56. The 10 mA full-scale DAC current from the AD768 (see Reference 3) develops a 0 V to 2 V output voltage across the 200 Ω R_F.



Figure 3-56: Single-ended I/V op amp interface for precision 16-bit AD768 DAC

Driving the virtual ground of the AD8055 op amp minimizes any distortion due to nonlinearity in the DAC output impedance. In fact, most high resolution DACs of this type are factory trimmed using an I/V converter.

It should be recalled, however, that compared to a differential operating mode using the single-ended output of the DAC in this manner will cause degradation in the CM rejection and increased second-order distortion products.

The C_F feedback capacitor should be optimized for best pulse response in the circuit. The equations given in the diagram should only be used as guidelines. A more detailed analysis of this circuit is given in Reference 6.

Differential Current-to-Differential Voltage Conversion

If a buffered differential voltage output is required from a current output DAC, the AD813x-series of differential amplifiers can be used as shown in Figure 3-57.



Figure 3-57: Buffering high speed DACs using AD813x differential amplifier

The DAC output current is first converted into a voltage that is developed across the 25 Ω resistors. The voltage is amplified by a factor of 5 using the AD813x. This technique is used in lieu of a direct I/V conversion to prevent fast slewing DAC currents from overloading the amplifier and introducing distortion. Care must be taken so that the DAC output voltage is within its compliance rating.

The V_{OCM} input on the AD813x can be used to set a final output CM voltage within the range of the AD813x. If transmission lines are to be driven at the output, adding a pair of 75 Ω resistors will allow this.

An Active Low-Pass Filter for Audio DAC

Figure 3-58 shows an active low-pass filter which also serves as a current-to-voltage converter for the AD1853 sigma-delta audio DAC (see Reference 4). The filter is a 4-pole filter with a 3 dB cutoff frequency of approximately 75 kHz. Because of the high oversampling frequency (24.576 MSPS when operating the DAC at a 48 KSPS throughput rate), a simple filter is all that is required to remove aliased components above 12 MHz).



Figure 3-58: A 75 kHz 4-pole gaussian active filter for buffering the output of AD1853 stereo DAC

The diagram shows a single channel for the dual channel DAC output. U1A and U1B I/V stages form a 1-pole differential filter, while U2 forms a 2-pole multiple-feedback filter that also performs a differential-to-single-ended conversion.

A final fourth passive pole is formed by the 604 Ω resistor and the 2.2 nF capacitor across the output. The OP275 op amp was chosen for operation at U1 and U2, and for its quality audio characteristics (see Reference 5).

For further details of active filter designs, see Chapter 5 of this book.

References: Buffering DAC Outputs

- 1. Data sheet for AD9772A 14-Bit, 160 MSPS TxDAC+® with 2x Interpolation Filter, www.analog.com.
- 2. Data sheet for AD8055/AD8056 Low Cost, 300 MHz Voltage Feedback Amplifiers, www.analog.com.
- 3. Data sheet for AD768 16-Bit, 30 MSPS D/A Converter, www.analog.com.
- 4. Data sheet for AD1853 Stereo, 24-Bit, 192 kHz, Multibit Σ - Δ DAC, www.analog.com.
- 5. Data sheet for **OP275 Dual Bipolar/JFET, Audio Operational Amplifier**, www.analog.com.
- 6. Chapters 5, Walt Kester, Editor, **Practical Design Techniques for Sensor Signal Conditioning**, Analog Devices, 1999, ISBN: 0-916550-20-6.

This page intentionally left blank

CHAPTER 4 Sensor Signal Conditioning

- Section 4-1: Introduction
- Section 4-2: Bridge Circuits
- Section 4-3: Strain, Force, Pressure and Flow Measurements
- Section 4-4: High Impedance Sensors
- Section 4-5: Temperature Sensors

This page intentionally left blank

CHAPTER 4

Sensor Signal Conditioning

Walt Kester, James Bryant, Walt Jung Scott Wurcer, Chuck Kitchin

SECTION 4-1

Introduction

Walt Kester

This chapter of the book deals with various sensors and associated signal-conditioning circuitry involving the use of op amps and in amps. While the topic is generally very broad, the focus is to concentrate on circuit and signal processing applications of sensors rather than the details of the actual sensors themselves.

Strictly speaking, a *sensor* is a device that receives a signal or stimulus and responds with an electrical signal, while a *transducer* is a converter of one type of energy into another. In practice, however, the terms are often used interchangeably.

Sensors and their associated circuits are used to measure various physical properties such as temperature, force, pressure, flow, position, light intensity, and so forth. These properties act as the stimulus to the sensor, and the sensor output is conditioned and processed to provide the corresponding measurement of the physical property. We will not cover all possible types of sensors, only the most popular ones, and specifically, those that lend themselves to process control and data acquisition systems.

Sensors do not operate by themselves. They are generally part of a larger system consisting of signal conditioners and various analog or digital signal processing circuits. The *system* could be a measurement system, data acquisition system, or process control system, for example.

Sensors may be classified in a number of ways. From a signal-conditioning viewpoint it is useful to classify sensors as either *active* or *passive*. An *active* sensor requires an external source of excitation. Resistor-based sensors such as thermistors, resistance temperature detectors (RTDs), and strain gages are examples of active sensors, because a current must be passed through them and the corresponding voltage measured in order to determine the resistance value. An alternative would be to place the devices in a bridge circuit; however in either case, an external current or voltage is required.

On the other hand, *passive* (or *self-generating*) sensors generate their own electrical output signal without requiring external voltages or currents. Examples of passive sensors are thermocouples and photodiodes which generate thermoelectric voltages and photocurrents, respectively, which are independent of external circuits.

It should be noted that these definitions (*active* versus *passive*) refer to the need (or lack thereof) of external active circuitry to produce a sensor electrical output signal. It would seem equally logical to consider a thermocouple *active*, in the sense that it produces an output voltage without external circuitry; however the convention in the industry is to classify the sensor with respect to the external circuit requirement as defined above.

A logical way to classify sensors is with respect to the physical property the sensor is designed to measure. Thus we have temperature sensors, force sensors, pressure sensors, motion sensors, and so forth. However, sensors that measure different properties may have the same type of electrical output. For instance, Resistance Temperature Detector (RTD) is a variable resistance, as is a resistive strain gage. Both RTDs and strain gages are often placed in bridge circuits, and the conditioning circuits are therefore quite similar. In fact, bridges and their conditioning circuits deserve a detailed discussion. Figure 4-1 is an overview of basic sensor characteristics.

- Sensors: Convert a Signal or Stimulus (Representing a Physical Property) into an Electrical Output
- Transducers:
 Convert One Type of Energy into Another
- The Terms are often Interchanged
- Active Sensors Require an External Source of Excitation: RTDs, Strain-Gages
- Passive (Self-Generating) Sensors do not: Thermocouples, Photodiodes

Figure 4-1: An overview of sensor characteristics

The full-scale outputs of most sensors (passive or active) are relatively small voltages, currents, or resistance changes, and therefore their outputs must be properly conditioned before further analog or digital processing can occur. Because of this, an entire class of circuits has evolved, generally referred to as signal-conditioning circuits. Amplification, level translation, galvanic isolation, impedance transformation, linearization, and filtering are fundamental signal-conditioning functions that may be required. Figure 4-2 summarizes sensors and their outputs.

PROPERTY	SENSOR	ACTIVE/ PASSIVE	OUTPUT
Temperature	Thermocouple Silicon RTD Thermistor	Passive Active Active Active	Voltage Voltage/Curren Resistance Resistance
Force / Pressure	Strain Gage Piezoelectric	Active Passive	Resistance Voltage
Acceleration	Accelerometer	Active	Capacitance
Position	LVDT	Active	AC Voltage
Light Intensity	Photodiode	Active	Current

Figure 4-2: Typical sensors and their output formats

Whatever form the conditioning takes, the circuitry and performance will be governed by the electrical character of the sensor and its output. Accurate characterization of the sensor in terms of parameters appropriate to the application, e.g., sensitivity, voltage and current levels, linearity, impedances, gain, offset, drift, time constants, maximum electrical ratings, stray impedances, and other important considerations can spell the difference between substandard and successful application of the device, especially where high resolution, precision, or low level measurements are necessary.

Higher levels of integration now allow ICs to play a significant role in both analog and digital signal conditioning. ADCs specifically designed for measurement applications often contain on-chip programmable-gain amplifiers (PGAs) and other useful circuits, such as current sources for driving RTDs, thereby minimizing the external conditioning circuit requirements.

To some degree or another, most sensor outputs are nonlinear with respect to the applied stimulus and as a result, their outputs must often be linearized in order to yield correct measurements. In terms of the design approach choice towards linearization, the designer can take a route along either of two major paths.

Analog is one viable route, and such techniques may be used to perform an "analog domain" linearization function.

However, the recent introduction of high performance ADCs now allows linearization to be done much more efficiently and accurately in software. This "digital domain" approach to linearization eliminates the need for tedious manual calibration using multiple and sometimes interactive analog trim adjustments.

A quite common application of sensors is within process control systems. One example would be control of a physical property, such as temperature. A sample block diagram of how this might be implemented is illustrated in Figure 4-3.



Figure 4-3: A typical industrial process temperature control loop

In this system, an output from a temperature sensor is conditioned, transmitted over some distance, received, and then digitized by an ADC. The microcontroller or host computer determines if the temperature is above or below the desired value, and outputs a digital word to the digital-to-analog converter (DAC).

The DAC output is conditioned and drives the remotely located *actuator*, in this case a heater. Notice that the interface between the control center and the remote process is via the industry-standard 4-20 mA loop.

Digital techniques are becoming more and more popular in processing sensor outputs in data acquisition, process control, and measurement. 8-bit microcontrollers (8051-based, for example) generally have sufficient speed and processing capability for most applications.

By including the A/D conversion and the microcontroller programmability on the sensor itself, a "smart sensor" can be implemented with self-contained calibration and linearization features among others. However, such digital techniques aren't a major focus of this text, so the section references may be consulted for further information.

The remaining sections of the chapter deal with analog signal conditioning methods for a variety of sensor types.

References: Introduction

- 1. Walt Kester, Bill Chestnut, and Grayson King, *Smart Sensors*, Chapter 9 of **Practical Design Tech**niques for Sensor Signal Conditioning, Analog Devices, Inc., 1999, ISBN: 0-916550-20-6.
- Compatibility of Analog Signals for Electronic Industrial Process Instruments, ANSI/ISA-S50.1-1982 (Rev. 1992), www.isa.org.
- 3. Editors, "Fieldbuses: Look Before You Leap," EDN, November 5, 1998, p. 197.
- 4. "MicroConverter Technology Backgrounder," Whitepaper, Analog Devices, Inc., www.analog.com.
- 5. Scott MacKenzie, The 8051 Microcontroller, 3rd Ed., Prentice-Hall, 1999, ISBN: 0-13-780008-8.

SECTION 4-2 Bridge Circuits Walt Kester

An Introduction to Bridges

This section of the chapter, 4-2, discusses more fundamental bridge circuit concepts. To gain greatest appreciation of these ideas, it should be studied along with those sections discussing precision op amps within Chapters 1 and 2. The next section (4-3) focuses on the detailed application circuits relating to strain-gage-based sensors. These sections can be read sequentially if the reader already understands the design issues related to precision op amp applications.

Resistive elements are some of the most common sensors. They are inexpensive, and relatively easy to interface with signal-conditioning circuits. Resistive elements can be made sensitive to temperature, strain (by pressure or by flex), and light. Using these basic elements, many complex physical phenomena can be measured, such as: fluid or mass flow (by sensing the temperature difference between two calibrated resistances), dew-point humidity (by measuring two different temperature points), and so forth.

Sensor element resistance can range from less than 100 Ω to several hundred k Ω , depending on the sensor design and the physical environment to be measured. Figure 4-4 indicates the wide range of sensor resistance encountered. For example, RTDs are typically 100 Ω or 1000 Ω . Thermistors are typically 3500 Ω or higher.

Figure 1-1: Sonsor resistances used in				
Thermistors	$100\Omega - 10M\Omega$			
Resistance Temperature Devices (RTDs)	100Ω, 1000Ω			
Relative Humidity	$100k\Omega - 10M\Omega$			
Pressure Sensors	$350\Omega - 3500\Omega$			
Weigh-Scale Load Cells	$350\Omega - 3500\Omega$			
Strain Gages	120Ω, 350Ω, 3500Ω			

Figure 4-4: Sensor resistances used in bridge circuits span a wide dynamic range

Resistive sensors such as RTDs and strain gages produce relatively small percentage changes in resistance, in response to a change in a physical variable such as temperature or force. For example, platinum RTDs have a temperature coefficient of about $0.385\%/^{\circ}$ C. Thus, in order to accurately resolve temperature to 1°C, the overall measurement accuracy must be much better than 0.385 Ω when using a 100 Ω RTD.

Strain gages present a significant measurement challenge because the typical change in resistance over the entire operating range of a strain gage may be less than 1% of the nominal resistance value. Accurately measuring small resistance changes is therefore critical when applying resistive sensors.

A simple method for measuring resistance is to force a constant current through the resistive sensor, and measure the voltage output. This requires both an accurate current source and an accurate means of measuring the voltage. Any change in the current will be interpreted as a resistance change. In addition, the power dissipation in the resistive sensor must be small and in accordance with the manufacturer's recommendations, so that self-heating does not produce errors. As a result, the drive current must be small, which tends to limit the resolution of this approach.

A *resistance bridge*, shown in Figure 4-5, offers an attractive alternative for accurately measuring small resistance changes. This is a basic Wheatstone bridge (actually developed by S. H. Christie in 1833), and is a prime example. It consists of four resistors connected to form a quadrilateral, a source of excitation voltage V_B (or, alternately, a current) connected across one of the diagonals, and a voltage detector connected across the other diagonal. The detector measures the difference between the outputs of the two voltage dividers connected across the excitation voltage, V_B . The general form of the bridge output V_0 is noted in the figure.



Figure 4-5: The basic Wheatstone bridge produces an output null when the ratios of sidearm resistances match

There are two principal ways of operating a bridge such as this. One is by operating it as a null detector, where the bridge measures resistance indirectly by comparison with a similar standard resistance. On the other hand, it can be used as a device that reads a resistance difference directly, as a proportional voltage output.

When R1/R4 = R2/R3, the resistance bridge is said to be at a *null*, irrespective of the mode of excitation (current or voltage, ac or dc), the magnitude of excitation, the mode of readout (current or voltage), or the impedance of the detector. Therefore, if the ratio of R2/R3 is fixed at K, a null is achieved when R1 = K \cdot R4. If R1 is unknown and R4 is an accurately determined variable resistance, the magnitude of R1 can be found by adjusting R4 until an output null is achieved. Conversely, in sensor-type measurements, R4 may be a fixed reference, and a null occurs when the magnitude of the external variable (strain, temperature, and so forth.) is such that R1 = K \cdot R4.

Null measurements are principally used in feedback systems involving electromechanical and/or human elements. Such systems seek to force the active element (strain gage, RTD, thermistor, and so forth.) to balance the bridge by influencing the parameter being measured.

For the majority of sensor applications employing bridges, however, the deviation of one or more resistors in a bridge from an initial value is measured as an indication of the magnitude (or a change) in the measured variable. In these cases, the output voltage change is an indication of the resistance change. Because very small resistance changes are common, the output voltage change may be as small as tens of millivolts, even with the excitation voltage $V_B = 10 V$ (typical for a load cell application).

In many bridge applications, there may not just be a single variable element, but two, or even four elements, all of which may vary. Figure 4-6 shows a family of four voltage-driven bridges, those most commonly suited for sensor applications. In the four cases the corresponding equations for V_0 relate the bridge output voltage to the excitation voltage and the bridge resistance values. In all cases we assume a constant voltage drive, V_B . Note that since the bridge output is always directly proportional to V_B , the measurement accuracy can be no better than that of the accuracy of the excitation voltage.



Figure 4-6: The output voltage sensitivity and linearity of constant voltage drive bridge configurations differs according to the number of active elements

In each case, the value of the fixed bridge resistor "R" is chosen to be equal to the nominal value of the variable resistor(s). The deviation of the variable resistor(s) about the nominal value is assumed to be proportional to the quantity being measured, such as strain (in the case of a strain gage), or temperature (in the case of an RTD).

The *sensitivity* of a bridge is the ratio of the maximum expected change in the output voltage to the excitation voltage. For instance, if $V_B = 10$ V, and the full-scale bridge output is 10mV, then the sensitivity is 1mV/V. For the four cases of Figure 4-6, sensitivity can be said to increase going left-right, or as more elements are made variable.

The *single-element varying* bridge of Figure 4-6A is most suited for temperature sensing using RTDs or thermistors. This configuration is also used with a single resistive strain gage. All the resistances are nominally equal, but one of them (the sensor) is variable by an amount ΔR . As the equation indicates, the relationship between the bridge output and ΔR is not linear. For example, if $R = 100 \Omega$ and $\Delta R = 0.1 \Omega$ (0.1% change in resistance), the output of the bridge is 2.49875 mV for $V_B = 10$ V. The error is 2.50000 mV – 2.49875 mV, or 0.00125 mV. Converting this to a % of full-scale by dividing by 2.5 mV yields an end-point linearity error in percent of approximately 0.05%. (Bridge end-point linearity error is

calculated as the worst error in % FS from a straight line which connects the origin and the end point at FS, i.e., the FS gain error is not included). If $\Delta R = 1 \Omega$, (1% change in resistance), the output of the bridge is 24.8756 mV, representing an end-point linearity error of approximately 0.5%. The end-point linearity error of the single-element bridge can be expressed in equation form:

Single-Element Varying

Bridge End-Point Linearity Error \approx % Change in Resistance ÷ 2

It should be noted that the above nonlinearity refers to the nonlinearity of the bridge itself and not the sensor. In practice, most sensors themselves will exhibit a certain specified amount of nonlinearity, which must also be accounted for in the final measurement.

In some applications, the bridge nonlinearity noted above may be acceptable. If not, there are various methods available to linearize bridges. Since there is a fixed relationship between the bridge resistance change and its output (shown in the equations), software can be used to remove the linearity error in digital systems. Circuit techniques can also be used to linearize the bridge output directly, and these will be discussed shortly.

There are two cases to consider in the instance of a *two-element varying* bridge. In Case 1 (Figure 4-6B), both of the diagonally opposite elements change in the same direction. An example would be two identical strain gages mounted adjacent to each other, with their axes in parallel.

The nonlinearity for this case, 0.5%/%, the same as that of the single-element varying bridge of Figure 4-6A. However, it is interesting to note the sensitivity is now improved by a factor of 2, vis-à-vis the single-element varying setup. The two-element varying bridge is commonly found in pressure sensors and flow meter systems.

A second case of the two-element varying bridge, Case 2, is shown in Figure 4-6C. This bridge requires two identical elements that vary in *opposite* directions. This could correspond to two identical strain gages: one mounted on top of a flexing surface, and one on the bottom. Note that this configuration is now linear, and like two-element varying Case 1, it has twice the sensitivity of the Figure 4-6A configuration. Another way to view this configuration is to consider the terms $R + \Delta R$ and $R - \Delta R$ as comprising two sections of a linear potentiometer.

The *all-element varying* bridge of Figure 4-6D produces the most signal for a given resistance change, and is inherently linear. It is also an industry-standard configuration for load cells constructed from four identical strain gages. Understandably, it is also one of the most popular bridge configurations.

Bridges may also be driven from constant current sources, as shown in Figure 4-7, for the corresponding cases of single, dual, dual, and four active element(s). As with the voltage-driven bridges, the analogous output expressions are noted, along with the sensitivities.

Current drive, although not as popular as voltage drive, does have advantages when the bridge is located remotely from the source of excitation. One advantage is that the wiring resistance doesn't introduce errors in the measurement; another is simpler, less expensive cabling. Note also that with constant current excitation, all bridge configurations are linear except the single-element varying case of Figure 4-7A.

In summary, there are many design issues relating to bridge circuits, as denoted by Figure 4-8. After selecting the basic configuration, the excitation method must be determined. The value of the excitation voltage or current must first be determined, as this directly influences sensitivity. Recall that the full-scale bridge output is directly proportional to the excitation voltage (or current). Typical bridge sensitivities are 1mV/V to 10mV/V.



Figure 4-7: The output voltage sensitivity and linearity of constant current drive bridge configurations also differs according to the number of active elements

- Selecting Configuration (1-, 2-, 4-Element Varying)
- Selection of Voltage or Current Excitation
- Ratiometric Operation
- · Stability of Excitation Voltage or Current
- Bridge Sensitivity: FS Output / Excitation Voltage 1mV/V to 10mV/V Typical
- Full-scale Bridge Outputs: 10mV 100mV Typical
- Precision, Low Noise Amplification/Conditioning Techniques Required
- · Linearization Techniques May Be Required
- Remote Sensors Present Challenges

Figure 4-8: A number of bridge considerations impact design choices

Although large excitation voltages yield proportionally larger full-scale output voltages, they also result in higher bridge power dissipation, and thus raise the possibility of sensor resistor self-heating errors. On the other hand, low values of excitation voltage require more gain in the conditioning circuits, and also increase sensitivity to low level errors such as noise and offset voltages.

Regardless of the absolute level, the stability of the excitation voltage or current directly affects the overall accuracy of the bridge output, as is evident from the V_B and I_B terms in the output expressions. Therefore stable references and/or *ratiometric* drive techniques are required, to maintain highest accuracy.

Here, ratiometric simply refers to the use of the bridge drive voltage of a voltage-driven bridge (or a current-proportional voltage, for a current-driven bridge) as the reference input to the ADC that digitizes the

amplified bridge output voltage. In this manner the absolute accuracy and stability of the excitation voltage becomes a second-order error. Examples to follow further illustrate this point.

Amplifying and Linearizing Bridge Outputs

The output of a single-element varying bridge may be amplified by a single precision op amp connected as shown in Figure 4-9. Unfortunately this circuit, although attractive because of relative simplicity, has poor overall performance. Its gain predictability and accuracy are poor, and it unbalances the bridge due to loading from R_F and the op amp bias current. The R_F resistors must be carefully chosen and matched to maximize common mode rejection (CMR). Also, it is difficult to maximize the CMR while at the same time allowing different gain options. Gain is dependent upon the bridge resistances and R_F . In addition, the output is nonlinear, as the configuration does nothing to address the intrinsic bridge nonlinearity. In summary, the circuit isn't recommended for precision use.



Figure 4-9: Using a single op amp as a bridge amplifier

However, a redeeming feature of this circuit is that it is capable of single supply operation, with a solitary op amp. Note that the R_F resistor connected to the noninverting input is returned to $V_s/2$ (rather than ground) so that both positive and negative ΔR values can be accommodated, with the bipolar op amp output swing referenced to $V_s/2$.

A much better approach is to use an *instrumentation amplifier* (in amp) for the required gain, as shown in Figure 4-10. This efficient circuit provides better gain accuracy, with the in amp gain usually set with a single resistor, R_G. Since the amplifier provides dual, high-impedance loading to the bridge nodes, it does not unbalance or load the bridge. Using modern in amp devices with gains ranging from 10–1000, excellent common mode rejection and gain accuracy can be achieved with this circuit.

However, due to the intrinsic characteristics of the bridge, the output is still nonlinear (see expression). As noted earlier, this can be corrected in software (assuming that the in amp output is digitized using an analog-to-digital converter and followed by a microcontroller or microprocessor).

The in amp can be operated on either dual supplies as shown, or alternately, on a single positive supply. In the figure, this corresponds to $-V_s = 0$. This is a key advantage, due to the fact that all such bridge circuits bias the in amp inputs at $V_B/2$, a voltage range typically compatible with amplifier bias requirements. In amps such as the AD620 family, the AD623, and AD627 can be used in single (or dual) supply bridge applications, provided their restrictions on the gain and input and output voltage swings are observed.



Figure 4-10: A generally preferred method of bridge amplification employs an instrumentation amplifier for stable gain and high CMR

The bridge in this example is voltage driven, by the voltage V_B . This voltage can optionally be used for an ADC reference voltage, in which case it also is an additional output, V_{REF} .

Various techniques are available to linearize bridge outputs, but it is important to distinguish between the linearity of the bridge equation (discussed earlier), and the sensor response linearity to the phenomenon being sensed. For example, if the active sensor element is an RTD, the bridge used to implement the measurement might have perfectly adequate linearity; *yet the output could still be nonlinear*, due to the RTD device's intrinsic nonlinearity. Manufacturers of sensors employing bridges address the nonlinearity issue in a variety of ways, including keeping the resistive swings in the bridge small, shaping complementary non-linear response into the active elements of the bridge, using resistive trims for first-order corrections, and others. In the examples which follow, what is being addressed is the linearity error of the bridge configuration itself (as opposed to a sensor element within the bridge).

Figure 4-11 shows a single-element varying active bridge circuit, in which an op amp produces a forced bridge null condition. For this single-element varying case, only the op amp feedback resistance varies, with the remaining three resistances fixed.

As used here, the op amp output provides a buffered, ground referenced, low impedance output for the bridge measurement, effectively suppressing the $V_B/2$ CM bridge component at the op amp inputs.



Figure 4-11: Linearizing a single-element varying bridge (Method 1)

The circuit works by adding a voltage in series with the variable resistance arm. This voltage is equal in magnitude and opposite in polarity to the incremental voltage across the varying element, and is linear with ΔR . As can be noted, the three constant "R" valued resistances and the op amp operate to drive a constant current in the variable resistance. This is the basic mechanism that produces the linearized output.

This active bridge has a sensitivity gain of two over the standard single-element varying bridge (Figure 4-6A). The key point is that the bridge's incremental resistance/voltage output becomes linear, even for large values of ΔR . However, because of a still relatively small output signal, a second amplifier must usually follow this bridge. Note also that the op amp used in this circuit requires dual supplies, because its output must go negative for conditions where ΔR is positive.

Another circuit for linearizing a single-element varying bridge is shown in Figure 4-12. The top node of the bridge is excited by the voltage, V_B . The bottom of the bridge is driven in complementary fashion by the left op amp, which maintains a constant current of V_B/R in the varying resistance element, $R + \Delta R$.



Figure 4-12: Linearizing a single-element varying bridge (Method 2)

Like the circuit of Figure 4-11, the constant current drive for the single-element variable resistance provides the mechanism for linearity improvement. Also, because of the fact that the bridge left-side center node is ground-referenced by the op amp, this configuration effectively suppresses CM voltages. This has the virtue of making the op amp selection somewhat less critical. Of course, performance parameters of high gain, low offset/noise, and high stability are all still needed.

The output signal is taken from the right-hand leg of the bridge, and is amplified by a second op amp, connected as a noninverting gain stage. With the scaling freedom provided by the second op amp, the configuration is very flexible. The net output is linear, and has a bridge-output referred sensitivity comparable to the single-element varying circuit of Figure 4-11.

The Figure 4-12 circuit requires two op amps operating on dual supplies. In addition, paired resistors R1-R2 must be ratio matched and stable types, for overall accurate and stable gain. The circuit can be a practical one using a dual precision op amp, such as an AD708, the OP2177 or the OP213.

A closely related circuit for linearizing a voltage-driven, *two-element* varying bridge can be adapted directly from the basic circuit of Figure 4-11. This form of the circuit, shown in Figure 4-13, is identical to the previous single-element varying case, with the exception that the resistance between V_B and the op amp (+) input is now also variable (i.e., both diagonal R + ΔR resistances vary, in a like manner).



Figure 4-13: Linearizing a two-element varying voltage-driven bridge (Method 1)

For the same applied voltage V_B , this form of the circuit has twice the sensitivity, which is evident in the output expressions. A dual supply op amp is again required, and additional gain may also be necessary.

The two-element varying bridge circuit shown in Figure 4-14 uses an op amp, a sense resistor, and a voltage reference, set up in a feedback loop containing the sensing bridge. The net effect of the loop is to maintain a constant current through the bridge of $I_B = V_{REF}/R_{SENSE}$. The current through each leg of the bridge remains constant ($I_B/2$) as the resistances change, therefore the output is a linear function of ΔR . An in amp provides the additional gain.



Figure 4-14: Linearizing a two-element varying current-driven bridge (Method 2)

This circuit can be operated on a single supply with the proper choice of amplifiers and signal levels. If ratiometric operation of an ADC is desired, the V_{REF} voltage can be used to drive the ADC.

Driving Remote Bridges

Wiring resistance and noise pickup are the biggest problems associated with remotely located bridges. Figure 4-15 shows a 350 Ω strain gage, which is connected to the rest of the bridge circuit by 100 feet of 30-gage twisted-pair copper wire. The resistance of the wire at 25°C is 0.105 Ω /ft, or 10.5 Ω for 100 ft. The total lead resistance in series with the 350 Ω strain gage is therefore 21 Ω . The temperature coefficient of the copper wire is 0.385%/°C. Now we will calculate the gain and offset error in the bridge output due to a 10°C temperature rise in the cable. These calculations are easy to make, because the bridge output voltage is simply the difference between the output of two voltage dividers, each driven from a 10 V source.



Figure 4-15: Wiring resistance related errors with remote bridge sensor

The full-scale variation of the strain gage resistance (with flex) above its nominal 350 Ω value is +1% (+3.5 Ω), corresponding to a full-scale strain gage resistance of 353.5 Ω which causes a bridge output voltage of +23.45 mV. Notice that the addition of the 21 Ω R_{COMP} resistor compensates for the wiring resistance and balances the bridge when the strain gage resistance is 350 Ω . Without R_{COMP}, the bridge would have an output offset voltage of 145.63 mV for a nominal strain gage resistance of 350 Ω . This offset could be compensated for in software just as easily but, for this example, we chose to do it with R_{COMP}.

Assume that the cable temperature increases 10°C above nominal room temperature. This results in a total lead resistance increase of +0.404 Ω (10.5 $\Omega \times 0.00385$ /°C × 10°C) in each lead. *Note: The values in parentheses in the diagram indicate the values at 35°C*. The total additional lead resistance (of the two leads) is +0.808 Ω . With no strain, this additional lead resistance produces an offset of +5.44 mV in the bridge output. Full-scale strain produces a bridge output of +28.83 mV (a change of +23.39 mV from no strain). Thus the increase in temperature produces an offset voltage error of +5.44 mV (+23% full-scale) and a gain error of -0.06 mV (23.39 mV - 23.45 mV), or -0.26% full-scale. Note that these errors are produced solely by the 30-gage wire, and do not include any temperature coefficient errors in the strain gage itself.

The effects of wiring resistance on the bridge output can be minimized by the 3-wire connection shown in Figure 4-16. We assume that the bridge output voltage is measured by a high impedance device, therefore there is no current in the sense lead. Note that the sense lead measures the voltage output of a divider: the top half is the bridge resistor plus the lead resistance, and the bottom half is strain gage resistance plus the lead resistance. The nominal sense voltage is therefore independent of the lead resistance. When the strain gage resistance increases to full-scale (353.5Ω), the bridge output increases to +24.15 mV.

Increasing the temperature to 35°C increases the lead resistance by +0.404 Ω in each half of the divider. The full-scale bridge output voltage decreases to +24.13 mV because of the small loss in sensitivity, but



Figure 4-16: Remote bridge wiring resistance errors are reduced with 3-wire sensor connection

there is no offset error. The gain error due to the temperature increase of 10°C is therefore only -0.02 mV, or -0.08% of full-scale. Compare this to the +23% full-scale offset error and the -0.26% gain error for the 2-wire connection shown in Figure 4-14.

The 3-wire method works well for remotely located resistive elements which make up one leg of a singleelement varying bridge. However, all-element varying bridges are generally housed in a complete assembly, as in the case of a load cell. When these bridges are remotely located from the conditioning electronics, special techniques must be used to maintain accuracy.

Of particular concern is maintaining the accuracy and stability of the bridge excitation voltage. The bridge output is directly proportional to the excitation voltage, and any drift in the excitation voltage produces a corresponding drift in the output voltage.

For this reason, most all-element varying bridges (such as load cells) are six-lead assemblies: two leads for the bridge output, two leads for the bridge excitation, and two *sense* leads. To take full advantage of the additional accuracy that these extra leads allow, a method called Kelvin or 4-wire sensing is employed, as shown in Figure 4-17.



Figure 4-17: A Kelvin sensing system with a 6-wire voltage-driven bridge connection and precision op amps minimizes errors due to wire lead resistances

In this setup the drive voltage V_B is not applied directly to the bridge, but goes instead to the input of the upper precision op amp, which is connected in a feedback loop around the bridge (+) terminal. Although there may be a substantial voltage drop in the +FORCE lead resistance of the remote cable, the op amp will automatically correct for it, since it has a feedback path through the +SENSE lead. The net effect is that the upper node of the remote bridge is maintained at a precise level of V_B (within the capability of the op amp used, of course). A similar situation occurs with the bottom precision op amp, which drives the bridge (-) terminal to a ground level, as established by the op amp input ground reference. Again, the voltage drop in the -FORCE lead is relatively immaterial, because of the sensing at the -SENSE terminal.

In both cases, the sense lines go to high impedance op amp inputs, thus there is minimal error due to the bias current induced voltage drop across their lead resistance. The op amps maintain the required excitation voltage at the remote bridge, to make the voltage measured between the (+) and (–) sense leads always equal to $V_{\rm B}$.

Note—a subtle point is that the lower op amp will need to operate on dual supplies, since the drive to the –FORCE lead will cause the op amp output to go negative. Because of relatively high current in the bridge (~30 mA), current buffering stages at the op amp outputs are likely advisable for this circuit.

Although Kelvin sensing eliminates errors due to voltage drops in the bridge wiring resistance, the basic drive voltage V_B must still be highly stable since it directly affects the bridge output voltage. In addition, the op amps must have low offset, low drift, and low noise. Ratiometric operation can be optionally added, simply by using V_B to drive the ADC reference input.

The constant current excitation method shown in Figure 4-18 is another method for minimizing the effects of wiring resistance on the measurement accuracy. This system drives a precise current I through the bridge, proportioned as per the expression in the figure. An advantage of the circuit in Figure 4-18 is that it only uses one amplifier.



Figure 4-18: A 4-wire current-driven bridge scheme also minimizes errors due to wire lead resistances, plus allows simpler cabling

The accuracy of the reference, the sense resistor, and the op amp all influence the overall accuracy. While the precision required of the op amp should be obvious, one thing not necessarily obvious is that it may be required to deliver appreciable current, when I is more than a few mA (which it will be with standard 350 Ω bridges). In such cases, current buffering of the op amp is again in order.

Therefore for highest precision with this circuit, a buffer stage is recommended. This can be as simple as a small transistor, since the bridge drive is unidirectional.

System Offset Minimization

Maintaining an accuracy of 0.1% or better with a full-scale bridge output voltage of 20 mV requires that the sum of all offset errors be less than 20 μ V. Parasitic thermocouples are cases in point and, if not given due attention, can cause serious temperature drift errors. All dissimilar metal-metal connections generate voltages between a few and tens of microvolts for a 1°C temperature differential, are basic thermocouple facts of life.

Fortunately, however, within a bridge measurement system the signal connections are differential; therefore this factor can be used to minimize the impact of parasitic thermocouples.

Figure 4-19 shows some typical sources of offset error that are inevitable in a system. Within a differential signal path, only those thermocouple pairs whose junctions are actually at different temperatures will degrade the signal. The diagram shows a typical parasitic junction formed between the copper printed circuit board traces and the kovar pins of an IC amplifier.



Figure 4-19: Typical sources of offset voltage within bridge measurement systems

This thermocouple voltage is about 35 μ V/°C temperature differential. Note that this package-PC trace thermocouple voltage is significantly less when using a plastic package with a copper lead frame (recommended). Regardless of what package is used, all metal-metal connections along the signal path should be designed so that minimal temperature differences occur between the sides.

The amplifier offset voltage and bias currents are further sources of offset error. The amplifier bias current must flow through the source impedance. Any unbalance in either the source resistances or the bias currents produces offset errors. In addition, the offset voltage and bias currents are a function of temperature.

High performance low offset, low offset drift, low bias current, and low noise precision amplifiers such as the AD707, the OP177 or OP1177 are required. In some cases, chopper-stabilized amplifiers such as the AD8551/AD8552/AD8554 may be a solution. Ac bridge excitation such as that shown in Figure 4-20 can effectively remove offset voltage effects in series with a bridge output, V_o .



Figure 4-20: Ac bridge excitation minimizes system offset voltages

The concept is simple, and can be described as follows. The net bridge output voltage is measured under the two phased-sequence conditions, as shown. A first measurement (top) drives the bridge at the top node with excitation voltage V_B . This yields a first-phase measurement output V_A , where V_A is the sum of the desired bridge output voltage V_O and the net offset error voltage E_{OS} .

In the second measurement (bottom) the polarity of the bridge excitation is then reversed, and a second measurement, V_B , is made. Subtracting V_B from V_A yields 2 V_O , and the offset error term E_{OS} cancels as noted from the mathematical expression in the figure.

Obviously, a full implementation of this technique requires a highly accurate measurement ADC such as the AD7730 (see Reference 5) as well as a microcontroller to perform the subtraction.

Note that if a ratiometric reference is desired, the ADC must also accommodate the changing polarity of the reference voltage, as well as sense the magnitude. Again, the AD7730 includes this capability.

A very powerful combination of bridge circuit techniques is shown in Figure 4-21, an example of a high performance ADC. In Figure 4-21A is shown a basic dc operated ratiometric technique, combined with Kelvin sensing to minimize errors due to wiring resistance, which eliminates the need for an accurate excitation voltage.

The AD7730 measurement ADC can be driven from a single supply voltage of 5 V, which in this case is also used to excite the remote bridge. Both the analog input and the reference input to the ADC are high impedance and fully differential. By using the + and - SENSE outputs from the bridge as the differential reference voltage to the ADC, there is no loss in measurement accuracy if the actual bridge excitation voltage varies.

To implement ac bridge operation of the AD7730, an "H" bridge driver of P-Channel and N-Channel MOSFETs can be configured as shown in Figure 4-21B (note—dedicated bridge driver chips are available,



Figure 4-21: Ratiometric dc or ac operation with Kelvin sensing can be implemented using the AD7730 ADC

such as the Micrel MIC4427). This scheme, added to the basic functionality of the AD7730 configuration of 4-21A greatly increases the utility of the offset canceling circuit, as generally outlined in the preceding discussion of Figure 4-20.

Because of the on-resistance of the H-bridge MOSFETs, Kelvin sensing must also be used in these ac bridge applications. It is also important that the drive signals be nonoverlapping, as noted, to prevent excessive MOSFET switching currents. The AD7730

ADC has on-chip circuitry which generates the required nonoverlapping drive signals to implement this ac bridge excitation. All that needs adding is the switching bridge as noted in Figure 4-21B.

The AD7730 is one of a family of sigma-delta ADCs with high resolution (24 bits) and internal programmable gain amplifiers (PGAs) and is ideally suited for bridge applications. These ADCs have self- and system calibration features, which allow offset and gain errors due to the ADC to be minimized. For instance, the AD7730 has an offset drift of 5 nV/°C and a gain drift of 2 ppm/°C. Offset and gain errors can be reduced to a few microvolts using the system calibration feature.

References: Bridge Circuits

- 1. Ramon Pallas-Areny and John G. Webster, **Sensors and Signal Conditioning**, John Wiley, New York, 1991.
- 2. Dan Sheingold, Editor, **Transducer Interfacing Handbook**, Analog Devices, Inc., 1980, ISBN: 0-916550-05-2.
- 3. Sections 2, 3, Walt Kester, Editor, **1992 Amplifier Applications Guide**, Analog Devices, 1992, ISBN: 0-916550-10-9.
- 4. Sections 1, 6, Walt Kester, Editor, **System Applications Guide**, Analog Devices, 1993, ISBN: 0-916550-13-3.
- 5. Data sheet for AD7730 Bridge Transducer ADC, www.analog.com.

SECTION 4-3

Strain, Force, Pressure and Flow Measurements

Walt Kester

Strain Gages

The most popular electrical elements used in force measurements include the resistance strain gage, the semiconductor strain gage, and piezoelectric transducers. The strain gage measures force indirectly by measuring the deflection it produces in a calibrated carrier. Pressure can be converted into a force using an appropriate transducer, and strain gage techniques can then be used to measure pressure. Flow rates can be measured using differential pressure measurements, which also make use of strain gage technology. These principles are summarized in Figure 4-22.

- Strain: Strain Gage, Piezoelectric Transducers
- Force: Load Cell
- Pressure: Diaphragm to Force to Strain Gage
- Flow: Differential Pressure Techniques

Figure 4.22 Strain gages are directly or indirectly the basis for a variety of physical measurements

The resistance-based strain gage uses a resistive element that changes in length, hence resistance, as the force applied to the base on which it is mounted causes stretching or compression. It is perhaps the most well-known transducer for converting force into an electrical variable.

An *unbonded* strain gage consists of a wire stretched between two points. Force acting upon the wire (area = A, length = L, resistivity = ρ) will cause the wire to elongate or shorten, which will cause the resistance to increase or decrease proportionally according to:

$$\mathbf{R} = \rho \mathbf{L} / \mathbf{A}$$
 Eq. 4-1

and,

$$\Delta \mathbf{R}/\mathbf{R} = \mathbf{GF} \cdot \Delta \mathbf{L}/\mathbf{L}$$
 Eq. 4-2

where GF = Gage factor (2.0 to 4.5 for metals, and more than 150 for semiconductors).

In this expression, the dimensionless quantity $\Delta L/L$ is a measure of the force applied to the wire and is expressed in *microstrains* (1 $\mu\epsilon$ = 10⁻⁶ cm/cm) which is the same as parts-per-million (ppm).

From Eq. 4-2, note that larger gage factors result in proportionally larger resistance changes, hence this implies greater strain gage sensitivity. These concepts are summarized in the drawing of Figure 4-23.



Figure 4-23: Operating principles of a basic unbonded strain gage

A *bonded* strain gage consists of a thin wire or conducting film arranged in a coplanar pattern and cemented to a base or carrier. The basic form of this type of gage is shown in Figure 4-24.



Figure 4-24: A bonded wire strain gage

This strain gage is normally mounted so that as much as possible of the length of the conductor is aligned in the direction of the stress that is being measured, i.e., longitudinally. Lead wires are attached to the base and brought out for interconnection. Bonded devices are considerably more practical and are in much wider use than are the aforementioned unbonded devices.

Perhaps the most popular version is the *foil-type gage*, produced by photo-etching techniques, and using similar metals to the wire types. Typical alloys are of copper-nickel (Constantan), nickel-chromium (Ni-chrome), nickel-iron, platinum-tungsten, and so forth. This strain gage type is shown in Figure 4-25.


PHOTO ETCHING TECHNIQUE LARGE AREA STABLE OVER TEMPERATURE THIN CROSS SECTION GOOD HEAT DISSIPATION

Figure 4-25: A metal foil strain gage

Gages having *wire sensing elements* present a small surface area to the specimen; this reduces leakage currents at high temperatures and permits higher isolation potentials between the sensing element and the specimen. Foil sensing elements, on the other hand, have a large ratio of surface area to cross-sectional area and are more stable under extremes of temperature and prolonged loading. The large surface area and thin cross section also permit the device to follow the specimen temperature and facilitate the dissipation of self-induced heat.

Semiconductor Strain Gages

Semiconductor strain gages make use of the piezoresistive effect in certain semiconductor materials such as silicon and germanium in order to obtain greater sensitivity and higher level output.

Semiconductor gages can be produced to have either positive or negative changes when strained. They can be made physically small while still maintaining a high nominal resistance.

Semiconductor strain gage bridges may have 30 times the sensitivity of bridges employing metal films, but are temperature-sensitive and difficult to compensate. Their change in resistance with strain is also nonlinear. They are not in as widespread use as the more stable metal-film devices for precision work; however, where sensitivity is important and temperature variations are small, they may have some advantage.

Instrumentation is similar to that for metal-film bridges but is less critical because of the higher signal levels and decreased transducer accuracy. Figure 4-26 summarizes the relative performance of metal and semiconductor strain gages.

PARAMETER	METAL STRAIN GAGE	SEMICONDUCTOR STRAIN GAGE 0.001 to 3000 με 50 to 200	
Measurement Range	0.1 to 40,000 με		
Gage Factor	2.0 to 4.5		
Resistance, Ω	120, 350, 600,, 5000	1000 to 5000	
Resistance Tolerance	0.1% to 0.2%	1% to 2%	
Size, mm	0.4 to 150 Standard: 3 to 6	1 to 5	

Figure 4-26: A comparison of metal and semiconductor type strain gages

Piezoelectric force transducers are employed where the forces to be measured are dynamic (i.e., continually changing over the period of interest—usually of the order of milliseconds). These devices utilize the effect that changes in charge are produced in certain materials when they are subjected to physical stress. In fact, piezoelectric transducers are *displacement* transducers with quite large charge outputs for very small displacements, but they are invariably used as force transducers on the assumption that in an elastic material, displacement is proportional to force. Piezoelectric devices produce substantial output voltage in instruments such as accelerometers for vibration studies. Piezoelectric sensor output conditioning is discussed within Section 4-4 of this chapter.

Strain gages can be used to measure force, as shown in Figure 4-27, where a cantilever beam is slightly deflected by the applied force. Four strain gages are used to measure the flex of the beam, two on the top, and two on the bottom. The gages are connected in a four-element bridge configuration. Recall from Section 4-2 that this configuration gives maximum sensitivity and is inherently linear. This configuration also offers first-order correction for temperature drift in the individual strain gages.



Figure 4-27: A beam force sensor using a strain gage bridge

Strain gages are low-impedance devices, consequently they require significant excitation power to obtain reasonable levels of output voltage. A typical strain-gage-based load cell bridge will have a 350 Ω impedance and is specified as having a sensitivity in a range 3 mV-10 mV full scale, per volt of excitation.

The load cell is composed of four individual strain gages arranged as a bridge, as shown in Figure 4-28. For a 10 V bridge excitation voltage with a rating of 3 mV/V, 30 mV of signal will be available at full-scale loading.



Figure 4-28: A load cell comprising four strain gages is shown in physical (top) and electrical (bottom) representations

While increasing the drive to the bridge can increase the output, self-heating effects are a significant limitation to this approach—they can cause erroneous readings, or even device destruction. One technique for evading this limitation is to use a low duty cycle pulsed drive signal for the excitation.

Many load cells have the ±"SENSE" connections as shown, to allow the signal-conditioning electronics to compensate for dc drops in the wires (Kelvin sensing as discussed in Section 4-2). This brings the wires to a total of six for the fully instrumented bridge. Some load cells may also have additional internal resistors for temperature compensation purposes.

Pressures in liquids and gases are measured electrically by a variety of pressure transducers. A number of mechanical converters (including diaphragms, capsules, bellows, manometer tubes, and Bourdon tubes) are used to measure pressure by measuring an associated length, distance, or displacement, and to measure pressure changes by the motion produced, as shown by Figure 4-29.



Figure 4-29: Pressure sensors use strain gages for indirect pressure measurement

The output of this mechanical interface is then applied to an electrical converter such as a strain gage, or piezoelectric transducer. Unlike strain gages, piezoelectric pressure transducers are typically used for high frequency pressure measurements (such as sonar applications, or crystal microphones).

There are many ways of defining flow (mass flow, volume flow, laminar flow, turbulent flow). Usually the *amount* of a substance flowing (mass flow) is the most important, and if the fluid's density is constant, a volume flow measurement is a useful substitute that is generally easier to perform. One commonly used class of transducers, which measures flow rate indirectly, involves the measurement of pressure.

Flow can be derived by taking the differential pressure across two points in a flowing medium—one at a static point and one in the flow stream. *Pitot tubes* are one form of device used to perform this function, where flow rate is obtained by measuring the differential pressure with standard pressure transducers.

Differential pressure can also be used to measure flow rate using the *venturi* effect by placing a restriction in the flow. Although there are a wide variety of physical parameters being sensed, the electronics interface is very often strain gage based.

Bridge Signal Conditioning Circuits

The remaining discussions of this section deal with applications that apply the bridge and strain gage concepts discussed thus far in general terms.

An example of an all-element varying bridge circuit is a fatigue monitoring strain sensing circuit, as shown in Figure 4-30. The full bridge is an integrated unit, which can be attached to the surface on which the strain or flex is to be measured. In order to facilitate remote sensing, current mode bridge drive is used. The remotely located bridge is connected to the conditioning electronics through a 4-wire shielded cable. The OP177 precision op amp servos the bridge current to 10 mA, being driven from an AD589 reference voltage of 1.235 V. Current buffering of the op amp is employed in the form of the PNP transistor, for lowest op amp self-heating, and highest gain linearity.



The strain gage produces an output of 10.25 mV/1000 $\mu\epsilon$. The signal is amplified by the AD620 in amp, which is configured for a gain of 100 times, via an effective R_G of 500 Ω . Full-scale voltage calibration is set by adjusting the 100 Ω gain potentiometer such that, for a sensor strain of -3500 $\mu\epsilon$, the output reads -3.500 V; and for a strain of +5000 $\mu\epsilon$, the output registers +5.000 V. The measurement may then be digitized with an ADC which has a 10 V full-scale input range.

The 0.1 μ F capacitor across the AD620 input pins serves as an EMI/RFI filter in conjunction with the bridge resistance of 1 k Ω . The corner frequency of this filter is approximately 1.6 kHz.

Another example is a load cell amplifier circuit, shown in Figure 4-31. This circuit is more typical of a bridge workhorse application. It interfaces with a typical 350Ω load cell, and can be configured to accommodate typical bridge sensitivities over a range of 3 mV-10 mV/V.



A 10.000 V bridge excitation excitation is derived from an AD588 10 V reference, with an OP177 and 2N2219A used as a buffer. The 2N2219A is within the OP177 feedback loop and supplies the necessary bridge drive current (28.57 mA). This ensures that the op amp performance will not be compromised. The Kelvin sensing scheme used at the bridge provides for low errors due to wiring resistances, and a precision zener diode reference, the AD588, provides lowest excitation drift and scaling with temperature changes.

To ensure highest linearity is preserved, a low drift instrumentation amplifier is used as the gain stage. This design has a minimum number of critical resistors and amplifiers, making the entire implementation accurate, stable, and cost effective. In addition to low excitation voltage TC, another stability requirement is minimum in amp gain TC. Both factors are critical towards insuring stable circuit scaling over temperature.

With the use of the AD621B in amp as shown, the scaling is for a precise gain of 100 (as set by the Pin 1–8 jumper), for lowest in amp gain TC. The AD621B is specified for a very low gain TC, only 5 ppm/°C. The gain of 100 translates a 100 mV full-scale bridge output to a nominal 10 V output. Alternately, an AD620B could also be used, with the optional gain network consisting of the fixed 475 Ω resistor, and 100 Ω potentiometer for gain adjustment. This will provide a 50 ppm/°C gain TC for the in amp, plus the TC of the external parts (which should have low temperature coefficients).

While the lowest TC is provided by the fixed gain AD621 setup, it doesn't allow direct control of overall scaling. To retain the very lowest TC, scaling could be accomplished via a software autocalibration routine. Alternately, the AD588 and OP177 reference/op amp stage could be configured for a variable excitation voltage (as opposed to a fixed 10.000 V as shown). Variable gain in the reference voltage driver will effectively alter the excitation voltage as seen by the bridge, and thus provide flexible overall system scaling. Of course, it is imperative that such a scheme be implemented with low TC resistances.

As shown previously, a precision load cell is usually configured as a 350 Ω bridge. Figure 4-32 shows a precision load cell amplifier, within a circuit possessing the advantage of being powered from just a single power supply.



Figure 4-32: a single-supply load cell amplifier

As noted previously, the bridge excitation voltage must be both precise and stable, otherwise it can introduce measurement errors. In this circuit, a precision REF195 5 V reference is used as the bridge drive, allowing a TC as low as 5 ppm/°C. The REF195 reference can also supply more than 30 mA to a load, so it can drive a 350 Ω bridge (~14 mA) without need of a buffer. The dual OP213 is configured as a gain-of-100, two-op-amp in amp. The resistor network sets the gain according to the formula:

$$G = 1 + \frac{10 \,\text{k}\Omega}{1 \,\text{k}\Omega} + \frac{20 \,\text{k}\Omega}{196 \,\Omega + 28.7 \,\Omega} = 100$$
 Eq. 4-3

For optimum CMR, the 10 k Ω /1 k Ω resistor ratio matching should be precise. Close tolerance resistors (±0.5% or better) should be used, and all resistors should be of the same type.

For a zero volt bridge output signal, the amplifier will swing to within 2.5 mV of 0 V. This is the minimum output limit of the OP213. Therefore, if an offset adjustment is required, the adjustment should start from a positive voltage at V_{REF} and adjust V_{REF} downward until the output (V_{OUT}) stops changing. This is the point at which the amplifier limits the swing. Because of the single supply design, the amplifier cannot sense input signals that have negative polarity.

If linearity around or at zero volts input is required, or if negative polarity signals must be processed, the V_{REF} connection can be connected to a stable voltage that is mid-supply (i.e., 2.5 V) rather than ground. Note that when V_{REF} is not at ground, the output must be referenced to V_{REF} . An advantage of this type of referencing is that the output is now bipolar, with respect to V_{REF} .

The AD7730 24-bit sigma-delta ADC is ideal for direct conditioning of bridge outputs, and requires no interface circuitry (see Reference 10). A simplified connection diagram was shown in Figure 4.21A. The entire circuit operates on a single 5 V supply, which also serves as the bridge excitation voltage. Note that the measurement is ratiometric, because the sensed bridge excitation voltage is also used as the ADC reference. Variations in the 5 V supply do not affect the accuracy of the measurement.

The AD7730 has an internal programmable gain amplifier that allows a full-scale bridge output of $\pm 10 \text{ mV}$ to be digitized to 16-bit accuracy. The AD7730 has self- and system calibration features that allow offset and gain errors to be minimized with periodic recalibrations.

A "chop" or ac mode option minimizes the offset voltage and drift and operates similarly to a chopper-stabilized amplifier. The effective input voltage noise RTI is approximately 40nV rms, or 264 nV peak-to-peak. This corresponds to a resolution of 13 ppm, or approximately 16.5 bits. Gain linearity is also approximately 16 bits.

References: Strain, Force, Pressure And Flow Measurements

- 1. Ramon Pallas-Areny and John G. Webster, **Sensors and Signal Conditioning**, John Wiley, New York, 1991.
- 2. Dan Sheingold, Editor, **Transducer Interfacing Handbook**, Analog Devices, Inc., 1980, ISBN: 0-916550-05-2.
- 3. Sections 2, 3, Walt Kester, Editor, **1992 Amplifier Applications Guide**, Analog Devices, 1992, ISBN: 0-916550-10-9.
- 4. Sections 1, 6, Walt Kester, Editor, **System Applications Guide**, Analog Devices, 1993, ISBN: 0-916550-13-3.
- 5. Harry L. Trietley, Transducers in Mechanical and Electronic Design, Marcel Dekker, Inc., 1986.
- 6. Jacob Fraden, Handbook of Modern Sensors, 2nd Ed., Springer-Verlag, New York, NY, 1996.
- 7. The Pressure, Strain, and Force Handbook, Vol. 29, Omega Engineering, One Omega Drive, P.O. Box 4047, Stamford CT, 06907-0047, 1995. www.omega.com.
- 8. The Flow and Level Handbook, Vol. 29, Omega Engineering, One Omega Drive, P.O. Box 4047, Stamford CT, 06907-0047, 1995, www.omega.com.
- 9. Ernest O. Doebelin, Measurement Systems Applications and Design, 4th Ed., McGraw-Hill, 1990.
- 10. Data sheet for AD7730 Bridge Transducer ADC, www.analog.com.

SECTION 4-4

High Impedance Sensors Walt Kester, Scott Wurcer, Chuck Kitchin

Many popular sensors have output impedances greater than several megohms, and thus the associated signal-conditioning circuitry must be carefully designed to meet the challenges of low bias current, low noise, and high gain. Figure 4-33 lists a few examples of high impedance sensors.

A large portion of this section is devoted to a germane example, the analysis of a photodiode preamplifier. This application demonstrates many of the problems associated with high impedance sensor signal-conditioning circuits, and offers a host of practical solutions that can be applied to virtually all such sensors.

- Photodiode Preamplifiers
- Piezoelectric Sensors
- Humidity
- pH Monitors
- Chemical Sensors
- Smoke Detectors

Figure 4-33: High impedance sensors

Other examples of high impedance sensors to be discussed are piezoelectric sensors and charge-output sensors.

Photodiode Preamplifier Design

Photodiodes generate a small current that is proportional to the level of illumination. Their applications range from relatively low speed, wide dynamic range circuits to much higher speed circuits. Examples of the types of applications are precision light meters and high-speed fiber optic receivers.

One of the standard methods for specifying photodiode sensitivity is to state its short-circuit photocurrent (I_{sc}) for a given light level from a well-defined light source. The most commonly used source is an incandescent tungsten lamp running at a color temperature of 2850K.

At 100 fc (foot candles) of illumination (approximately the light level on an overcast day), the short-circuit current usually falls in a range of picoamps to hundreds of microamps for small area (less than 1 mm²) diodes.

The equivalent circuit for a photodiode is shown in Figure 4-34. The short-circuit current is very linear over 6 to 9 decades of light intensity, and is therefore often used as a measure of absolute light levels. The opencircuit forward voltage drop across the photodiode varies logarithmically with light level but, because of its large temperature coefficient, the diode voltage is seldom used as an accurate measure of light intensity.

The shunt resistance R_{SH} is usually on the order of 1000 M Ω at room temperature, and decreases by a factor of 2 for every 10°C rise in temperature. Diode capacitance C_J is a function of junction area and the diode bias voltage. A value of 50 pF at zero bias is typical for small-area diodes.



NOTE: R_{SH} HALVES EVERY 10°C TEMPERATURE RISE

Figure 4-34: A photodiode equivalent circuit

Photodiodes may be operated in either of two basic modes, as shown in Figure 4-35. These modes are with zero bias voltage (*photovoltaic* mode, left) or with a reverse-bias voltage (*photoconductive* mode, right).



Figure 4-35: Photodiode operating modes

The most precise linear operation is obtained in the photovoltaic mode, while higher switching speeds can be realized when the diode is operated in the photoconductive mode at the expense of linearity. Under these reverse-bias conditions, a small amount of current called *dark current* will flow—even when there is no illumination.

There is no dark current in the photovoltaic mode. In the photovoltaic mode, the diode noise is basically the thermal noise generated by the shunt resistance. In the photoconductive mode, shot noise due to conduction is an additional source of noise. Photodiodes are usually optimized during the design process for use in either the photovoltaic mode or the photoconductive mode, but not both.

Figure 4-36 shows the photosensitivity for a small photodiode (Silicon Detector Part Number SD-020-12-001). This diode has a basic sensitivity of 0.03 μ A/fc, and was chosen for the design example to follow. As this chart indicates, this photodiode's dynamic range covers six orders of magnitude.

ENVIRONMENT	ILLUMINATION (fc)	SHORT CIRCUIT CURRENT	
Direct Sunlight	1000	30µA	
Overcast Day	100	ЗμА	
Twilight	1	0.03µA	
Full Moonlit Night	0.1	3000pA	
Clear Night / No Moon	0.001	30pA	

Figure 4-36: Short circuit current versus light intensity for SD-020-12-001 photodiode (photovoltaic operating mode)

A convenient way to convert the photodiode current into a usable voltage is to use a low bias current op amp, configured as a current-to-voltage converter as shown in Figure 4-37. The diode bias is maintained at zero volts by the virtual ground of the op amp, and the short-circuit current is converted into a voltage. At maximum sensitivity the amplifier must be able to detect a diode current of 30 pA. This implies that the feedback resistor must be very large, and the amplifier bias current very small.



Figure 4-37: A simplified current-to-voltage converter uses a low bias current op amp and a high value feedback resistor

For example, 1000 M Ω will yield a corresponding voltage of 30 mV for this amount of current. Larger resistor values are impractical, so we will use 1000 M Ω for the most sensitive range. This will give an output voltage range of 10 mV for 10 pA of diode current and 10 V for 10 nA of diode current. This yields a range of 60 dB. For higher values of light intensity, the gain of the circuit must be reduced by using a smaller feedback resistor. For this range of maximum sensitivity, we should be able to easily distinguish between the light intensity on a clear, moonless night (0.001 fc), and that of a full moon (0.1 fc).

Notice that we have chosen to get as much gain as possible from one stage, rather than cascading two stages. This is in order to maximize the signal-to-noise ratio (SNR). If we halve the feedback resistor value, the signal level decreases by a factor of 2, while the noise due to the feedback resistor $(\sqrt{4kTR} Bandwidth)$ decreases by only $\sqrt{2}$. This reduces the SNR by 3 dB, assuming the closed loop bandwidth remains constant. Later in the analysis, we will find the resistors one of the largest overall output noise contributors.

To accurately measure photodiode currents in the tens of picoamps range, the bias current of the op amp should be no more than a few picoamps. This considerably narrows the choice. The industry-standard OP07 is an ultralow offset voltage ($10 \mu V$) bipolar op amp, but its bias current is 4 nA (4000 pA). Even superbeta bipolar op amps with bias current compensation (such as the OP97) have bias currents on the order of 100 pA at room temperature, but they might be suitable for very high temperature applications, as unlike FET amplifiers, the bias currents do not double for every 10°C increase.

A JFET input electrometer-grade op amp is chosen for our photodiode preamp, since it must operate only over a limited temperature range. Figure 4-38 summarizes the performance of several popular "electrometer grade" FET input op amps.

PART #	V _{OS} , MAX*	TC V _{OS} , TYP	I _B , MAX*	0.1Hz TO 10Hz NOISE, TYP	PACKAGE
AD549K	250µV	5µV/°C	100fA	4µV р-р	то-99
AD795JR	500µV	3µV/°C	ЗрА	1µV p-p	SOIC
AD820B	1000µV	2µV/°C	10pA	2µV р-р	SOIC, DIP

*25°C SPECIFICATION

Figure 4-38: Some JFET input electrometer grade op amps suitable for use in photodiode preamplifiers

As can be noted from this figure, the 25°C maximum bias current specification ranges from a few pA down to as low as 100 fA, and there are a number of packages types from which to choose. As will be seen shortly, the package finally chosen can and will affect the performance of the circuit in terms of the bias current realized within an application. This is due to relative ability to control the inevitable leakage currents in a design's production environment.

Of these devices, the AD549 and AD795 are fabricated on a BiFET process and use P-Channel JFETs as the input stage, as is shown in Figure 4-39. The rest of the op amp circuit is designed using bipolar devices. These BiFET op amps are laser trimmed at the wafer level, to minimize offset voltage and offset voltage drift. The offset voltage drift is minimized, by first trimming the input stage for equal currents in the JFET differential pair (drift trim resistors). A further trim of the JFET source resistors minimizes the input offset voltage trim resistors).

For these discussions, an AD795JR was selected for the photodiode preamplifier, with key specifications summarized in Figure 4-38. This allows high circuit performance in an SOIC packaged device.

Alternately, for even greater performance, the AD549 could be used. The AD549 uses the glass sealed TO-99 package, which allows the very highest performance in terms of low leakage. More on this follows.

Since the photodiode current is measured in terms of picoamperes, it should be understood that extremely close attention must be given to all potential leakage paths in the actual physical circuit. To put this in some



Figure 4-39: JFET input stage op amp with separate trims for offset voltage and drift

perspective, consider the a simple printed circuit card example with two parallel conductor traces on a high-quality, well-cleaned epoxy-glass PC board 0.05 inches apart and parallel for 1 inch. Such an insulator has a leakage resistance of approximately $10^{11} \Omega$ at +125°C. By simple application of Ohm's law, 15 V of bias between these runs produces a 150 pA current—sufficient to mask all signal levels below this current. Obviously then, low-level photodiode circuitry needs to employ all possible means of minimizing such parasitic currents. Unfortunately, they can arise from numerous sources, some of which can be quite subtle in origin.

Figure 4-40 illustrates the circuit elements subject to leakage for the photodiode circuit, as enclosed within the dotted lines. The feedback resistor is highly critical, and should be a close tolerance (1%), low TC (50 ppm/°C) unit. Typical units suitable for R2 will be manufactured with thin film or metal oxide construction on ceramic or glass, with glass insulation. It should be readily apparent that any shunt conductive paths



Figure 4-40: Critical leakage paths and components for a photodiode preamplifier circuit are those within the dotted line area

across this resistor's body can (and will) degrade or lower the net effective resistance, producing scaling errors. It is for this reason that such high value resistors are often glass enclosed, and can require special handling. Some sources of suitable high value resistors are listed in the section references. If used, compensation capacitor C2 should use the lowest loss dielectric possible. Typically this will mean a film type capacitor of Teflon, polypropylene, or polystyrene construction.

All connections to the op amp's summing junction should be kept short, clean, and free from manufacturing process chemicals and residues. In cases where an input cable is used to connect the photodiode to the preamp, it should be kept as short as possible, and should use Teflon or similar low loss dielectric insulation.

The above considerations deal mainly with the more obvious construction points towards optimizing accuracy and keeping leakage low. However, two of the more difficult leakage sources that can plague this circuit aren't quite as obvious. These are the op amp package-related parasitic leakages, which can occur from all op amp package pins adjacent to the input pin. Consider leakage as a high value resistance to Pin 2.

Although it doesn't show in this particular figure (since Pin 1 isn't actively used by the application), *any leakage from package Pin 1 can be very relevant*. Note also, that since Pin 3 is grounded, it prevents error from leakage between Pins 4 and 2. If however, Pin 1 has any significant voltage on it (which it does in the case of the offset null pin of the AD820BN DIP device) serious leakage will then occur between Pins 1 and 2. The AD795JR SOIC is immune to this leakage, as Pin 1 isn't connected internally. These comments serve to illustrate some of the subtleties of these leakage sources.

The situation just described for the AD820BN DIP packaged device is by no means unique, as Pin 1 is a standard offset trim pin on many op amps. This circumstance will always tend to leak current into any high impedance source seen at Pin 2. There are also cases for follower-connected stages where leakage is just as critical, if not more so. In such cases the leakage goes into Pin 3 as a high impedance, typically from Pin 4, which is $-V_s$. Fortunately however, there is a highly effective answer to controlling both of these leakage problems, and that is the use of circuit *guard* techniques.

Guarding is used to reduce parasitic leakage currents, by isolating a sensitive amplifier input from large voltage gradients across the PC board. It does this by interposing a conductive barrier or screen between a high voltage source and a sensitive input. The barrier intercepts the leakage which would otherwise flow into the sensitive node, and diverts it away. In physical terms, a guard is a low impedance conductor that completely surrounds an input line or node, and it is biased to a potential equal to the line's voltage.

Note that the low impedance nature of a guard conductor shunts leakage harmlessly away. The biasing of the guard to the same potential as the guarded pin reduces any possibility of leakage between the guard itself and the guarded node. The exact technique for guarding depends on the amplifier's mode of operation, i.e., whether the connection is inverting (like Figure 4-40), or a noninverting stage.

Figure 4-41 shows a PC board layout for guarding the inputs of the AD820 op amp, as operated within either an inverting (top) or a noninverting gain stage (bottom). This setup uses the DIP ("N") package, and would also be applicable to other devices where relatively high voltages occur at Pin 1 or 4. Using a standard 8-pin DIP outline, it can be noted that this package's 0.1" pin spacing allows a PC trace (the guard trace) to pass between adjacent pins. This is the key to implementing effective DIP package guarding—the complete surrounding of the guarded trace with a low impedance trace.

In the inverting mode (top), note that Pin 3 connected and grounded guard traces surround the op amp inverting input (Pin 2), and run parallel to the input trace. This guard would be continued out to and around the source device and feedback connection in the case of a photodiode (or around the input pad, in the case of a cable). In the follower mode (bottom), the guard voltage is the feedback divider tap voltage to Pin 2, i.e., the inverting input node of the amplifier. Although the feedback divider impedance isn't as low



Figure 4-41: Guard techniques for inverting and noninverting op amp stages using DIP package devices

in absolute terms as a direct ground, it is still quite effective. Even a 1 k Ω or so impedance here will still be many orders of magnitude lower than the Pin 3 impedance. In both the inverting and the noninverting modes, the guard traces should be located on both sides of the PC board, with top and bottom side traces connected with several vias. Things become slightly more complicated when using guarding techniques with the SOIC surface mount ("R") package, because the 0.05" pin spacing doesn't allow routing of PC board traces between the pins. But there still is an effective guarding answer, at least for the inverting case. Figure 4-42 shows the preferred method.



Figure 4-42: Guard techniques for inverting and noninverting op amp stages using SOIC package devices

In the AD795 SOIC "R" package, Pins 1, 5, and 8 are "no-connect" pins and can be used to route signal traces as shown. Thus in the case of the inverting stage (top), guarding is still completely effective, with dummy Pin 1 and Pin 3 acting as a grounded guard trace.

In the case of the follower stage (bottom), the guard trace must be routed around the $-V_s$ pin, and thus the Pin 4 to Pin 3 leakage is *not* fully guarded. For this reason, a very high impedance follower stage using an SOIC package op amp isn't generally recommended, as adequate guarding simply isn't possible. An exception to this caveat would apply the use of a *single-supply* op amp as a follower (for example, the AD820), in which case Pin 4 becomes grounded by default, and some degree of intrinsic guarding is established.

For extremely low bias current applications, such as for example with the AD549 with input bias current of 100 fA, the high impedance input signal connection of the op amp should be made to a virgin Teflon standoff insulator, as shown in Figure 4-43. Note—"virgin" Teflon is a solid piece of new Teflon material that has been machined to shape (as opposed to one welded together from powder or grains).



Figure 4-43: Input pin isolation technique using virgin Teflon standoff insulator

If mechanical and manufacturing considerations allow, the sensitive op amp input pin should be soldered directly to the Teflon standoff, rather than going through a PCB hole.

For TO-99 packaged devices, such as the AD549KH, two possible guarding choices present themselves. One method is to employ the device in a scheme like Figure 4-43, with the sensitive input pin going to the Teflon stand-off. Alternately, a round PCB layout scheme that is more amenable to the TO-99 package can be used, as shown in Figure 4-44.

This scheme uses a guard ring, which completely surrounds the input and feedback nodes, with the ring tied to the device's metal can through the Pin 8 connection. The guard ring is then also tied to either ground or the feedback divider, as suits the application. This setup can also be further modified, to use the more sensitive of the two inputs going to a *Teflon stand-off within the guard ring*, for the ultimate in performance.

Note that in all cases where control of leakage is critical, the PC board itself must be carefully cleaned and then sealed against humidity and dirt using a high quality conformal coating material. In addition to minimizing leakage currents, the entire circuit should be well shielded with a grounded metal shield to prevent stray signal pickup.



Figure 4-44: TO-99 package devices can use guard rings surrounding input pins 2 and 3 (PCB bottom view shown)

Preamplifier Offset Voltage and Drift Analysis

A photodiode preamp offset voltage and bias current model is shown in Figure 4-45. There are two important considerations in this circuit. First, the diode shunt resistance (R1) is a function of temperature—it halves every time the temperature increases by 10°C. At room temperature (25°C), R1 = 1000 M Ω , but at 70°C it decreases to 43 M Ω . This has a drastic impact on the circuit noise gain and hence the output offset voltage. In the example, at 25°C the dc noise gain is 2, but at 70°C it increases to 24.

The second circuit difficulty is that the input bias current doubles with every 10°C temperature rise. The bias current produces an output offset error equal to I_BR2 . At 70°C bias current increases to 72 pA, compared to 3 pA at room temperature. Normally, the addition of a resistor (R3) between the noninverting input of the op amp and ground, with a value of R1||R2 would yield a first-order cancellation of this effect. However, because R1 changes with temperature, this method simply isn't effective. In addition, if R3 is



Figure 4-45: AD795JR photodiode preamplifier offset error model

used, the bias current then develops a voltage across it, which in turn would be applied to the photodiode as a parasitic bias. Such a bias would cause the diode response to become nonlinear, thus the use of R3 is also undesirable from a linearity point-of-view.

The total referred-to-output (RTO) offset voltage preamp errors are summarized in Figure 4-46. Notice that at 70°C the total error is 87.2 mV. This error is acceptable for the design under consideration. The primary contributor to the error at high temperature is of course the bias current.

		1	1	1	1
		0°C	25°C	50°C	70°C
	V _{OS}	0.575mV	0.500mV	0.575mV	0.635mV
	Noise Gain	1.1	2	7	24
	, V _{OS} Error RTO	0.6mV	1.0mV	4.0mV	15.2mV
	I _B	0.6pA	3.0pA	18.0pA	72.0pA
	B Error RTO	0.6mV	3.0mV	18.0mV	72.0mV
1	Total Error RTO	1.2mV	4.0mV	22.0mV	87.2mV

Figure 4-46: AD795JR photodiode preamplifier offset error summary

Several steps can be taken to minimize amplifier temperature rise, and thus offset drift. Operating the amplifier at reduced supply voltages, minimizing the output drive requirements, and heat sinking are some ways to reduce this error. The addition of an external offset nulling circuit would minimize the initial input offset voltage error.

Thermoelectric Voltages as Sources of Input Offset Voltage

As discussed earlier in this chapter, thermoelectric potentials are generated by electrical connections are made between different metals. For example, the copper PC board electrical contacts to the kovar input pins of a TO-99 IC package can create an offset voltage of up to 40 μ V/°C, if the two bimetal junctions so formed are at different temperatures. Even ordinary solders, being composed of alloys different from PCB copper traces, can give rise to thermoelectric voltages. For example, common high tin content lead-tin solder alloys, when used with copper, create thermoelectric voltages on the order of 1 μ V/°C to 3 μ V/°C (see Reference 8). While some special cadmium-tin solders can reduce this voltage to 0.3 μ V/°C, cadmium solders aren't in general use for health reasons. Another possible low thermal EMF solder is a low tin alloy such as Sn10Pb90.

The best general solution to minimizing this spurious thermocouple problem is to ensure that the connections to the inverting and noninverting input pins of the IC are made with the same material, and that the PC board thermal layout is such that these two pins remain at the same temperature. Everything should be balanced from a thermal standpoint. In the case where a Teflon standoff is used as an insulated connection point for the inverting input (as in the case of this preamp), prudence dictates that connections to the noninverting inputs also be made in a similar manner to minimize possible thermoelectric effects, and in keeping with the principle of thermal symmetry.

Preamplifier AC Design, Bandwidth, and Stability

The key to the preamplifier ac design is an understanding of the circuit noise gain as a function of frequency. Plotting gain versus frequency on a log-log scale makes the analysis relatively simple (see Figure 4-47). This type of plot is also referred to as a Bode plot. The noise gain is the gain seen by a small voltage source in series with one of the op amp input terminals. It is also the same as the noninverting signal gain (the gain from "A" to the output). In the photodiode preamplifier, the signal current from the photodiode passes through the C2/R2 network. It is important to distinguish between the signal gain and the noise gain, because it is the noise gain characteristic that determines the net circuit stability, regardless of where the signal is actually applied.



Figure 4-47: A generalized noise gain Bode plot

Note that the net slope between the noise gain and the open loop gain curves, *at the point where they intersect*, determines system stability. For unconditional stability, the noise gain curve must intersect the open loop response with a net slope of less than 12 dB/octave (or 20 dB per decade). In the figure, the dotted (C2 = 0) line shows a noise gain that intersects the open loop gain at a net slope of 12 dB/octave, indicating an unstable condition. This is what would occur in the circuit, without a feedback capacitor.

The general equations for determining the break points and gain values in the Bode plot are also given in Figure 4-47. It is useful to examine these gain characteristics with increasing frequency. At low frequencies, the circuit noise gain is 1 + R2/R1, as indicated by the lowest frequency shelf (below 10 Hz). There are two key time constants in this circuit, τ_1 and τ_2 . The first comes into play as a zero in the noise gain transfer function, which occurs at a frequency of $f_1 = 1/2\pi\tau_1$, where $\tau_1 = R1 ||R2$ (C1 + C2). Stated simply, this frequency falls where the noise gain begins to increase to a new (higher) value from the low frequency gain of 1 + R2/R1 plateau. In the Figure 4-47 example f_1 occurs at 10 Hz.

Above f_1 , gain increases towards a high frequency gain plateau where the gain is 1 + C1/C2, which is indicated as the highest frequency shelf (above 100 Hz). The second time constant, τ_2 , comes into play as a pole of the transfer function, which occurs at a corner frequency, $f_2 = 1/2\pi\tau_2$, where $\tau_2 = R2C2$. It can also be noted that this is equal to the signal bandwidth, if the signal is applied at point "B."

Plotting the composite noise gain curve on the log-log graph is a simple matter of connecting the f_1 and f_2 breakpoints with a line having a 45° slope, after first sketching the flat low and high frequency gain plateaus. The point at which the high frequency noise gain intersects the op amp open loop gain is called the *closed loop bandwidth*. Notice that the *signal bandwidth* for a signal applied at point "B" is much less, and is $1/2 \pi R2C2$.

Figure 4-48 shows the noise gain plot for the photodiode preamplifier using actual circuit values. The choice of C2 determines the actual signal bandwidth and also the phase margin. In the example, a signal bandwidth of 16 Hz was chosen. Notice that a smaller value of C2 would result in a higher signal bandwidth and a corresponding reduction in phase margin. It is also interesting to note that although the signal bandwidth is only 16 Hz, the closed loop bandwidth is 167 kHz. This will have important implications with respect to the output noise voltage analysis to follow.



Figure 4-48: Noise gain of the AD795 photodiode preamplifier at 25°C

It is important to note that temperature changes do not significantly affect the stability of the circuit. Changes in R1 (the photodiode shunt resistance) affect only the low frequency noise gain and the frequency at which the zero in the noise gain response occurs. The high frequency noise gain is determined by the C1/C2 ratio.

Photodiode Preamplifier Noise Analysis

To begin a noise analysis, we first consider the AD795 input voltage and current noise spectral densities, as shown in Figure 4-49. The AD795 performance is truly impressive for a JFET input op amp: 1 μ V p-p typical 0.1 Hz to 10 Hz noise, and a 1/f corner frequency of 12 Hz, comparing favorably with all but the best bipolar op amps. As shown in the (right) figure, the current noise is much lower than for bipolar op amps, a key factor making it an ideal choice for high impedance applications.



Figure 4-49: AD795 Voltage and current noise density performance

The complete noise model for an op amp is shown in Figure 4-50. This model includes the reactive elements C1 and C2. Each individual output noise contributor is calculated by integrating the square of its spectral density over the appropriate frequency bandwidth and then taking the square root, as:

RMS OUTPUT NOISE DUE TO
$$V_1 = \sqrt{\int V_1(f)^2 df}$$
 Eq. 4-4

In most cases, this integration can be done by inspection of the graph of the individual spectral densities superimposed on a graph of the noise gain. The total output noise is then obtained by combining the individual components in a root-sum-squares manner. The table in the diagram shows how each individual source is reflected to the output, and the corresponding bandwidth for integration. The factor of 1.57 ($\pi/2$) is required to convert the single-pole bandwidth into its equivalent noise bandwidth.



Figure 4-50: A noise model of preamp

The resistor Johnson noise spectral density V_R is given by:

$$V_{R} = \sqrt{4kTR} \qquad \qquad Eq. 4-5$$

where *R* is the resistance in ohms, *k* is Boltzmann's constant $(1.38 \times 10^{-23} \text{ J/K})$, and *T* is the absolute temperature in Kelvins.

A simple way to compute this is to remember that the noise spectral density of a 1 k Ω resistor is 4 nV/ $\sqrt{\text{Hz}}$ at +25°C. The Johnson noise of another resistor value can be found by multiplying by the square root of the ratio of the resistor value to 1000 Ω . For example, a 4 k Ω resistor produces a noise density $\sqrt{4}$ times a 1 k Ω resistor, or 8 nV/ $\sqrt{\text{Hz}}$ (at +25°C).

Finally, note that Johnson noise is broadband, and its spectral density is constant with frequency.

Input Voltage Noise

In order to obtain the output voltage noise spectral density plot due to the input voltage noise, the input voltage noise spectral density plot is multiplied by the noise gain plot. This is easily accomplished using the Bode plot on a log-log scale. The total RMS output voltage noise due to the input voltage noise is then obtained by integrating the square of the output voltage noise spectral density plot, and then taking the square root. In most cases, this integration may be approximated. A lower frequency limit of 0.01 Hz in the 1/f region is normally used. If the bandwidth of integration for the input voltage noise is greater than a few hundred Hz, the input voltage noise spectral density may be assumed to be constant. Usually, the value of the input voltage noise spectral density at 1 kHz will provide sufficient accuracy.

It is important to note that the input voltage noise contribution must be integrated over the entire closedloop bandwidth of the circuit (the closed loop bandwidth, f_{cl} , is the frequency at which the noise gain intersects the op amp open loop response). This is also true of the other noise contributors that are reflected to the output by the noise gain (namely, the non inverting input current noise and the non inverting input resistor noise).

The inverting input noise current flows through the feedback network to produce a noise voltage contribution at the output. The input noise current is approximately constant with frequency, therefore, the integration is accomplished by multiplying the noise current spectral density (measured at 1 kHz) by the noise bandwidth which is 1.57 times the signal bandwidth ($1/2 \pi R2C2$). The factor of 1.57 ($\pi/2$) arises when single-pole 3 dB bandwidth is converted to equivalent noise bandwidth.

Johnson Noise Due to Feedforward Resistor R1

The noise current produced by the feedforward resistor R1 also flows through the feedback network to produce a contribution at the output. The noise bandwidth for integration is also 1.57 times the signal bandwidth.

Noninverting Input Current Noise

The noninverting input current noise, I_{N+} , develops a voltage noise across R3 that is reflected to the output by the noise gain of the circuit. The bandwidth for integration is therefore the closed-loop bandwidth of the circuit. However, there is no contribution at the output if R3 = 0 (or, if R3 is used, but it is bypassed with a large capacitor). This will usually be desirable when operating the op amp in the inverting mode.

Johnson Noise Due to Resistor in Noninverting Input

The Johnson voltage noise due to R3 is also reflected to the output by the noise gain of the circuit. Again, if R3 is bypassed sufficiently, it makes no significant contribution to the output noise.

Summary of Photodiode Circuit Noise Performance

Figure 4-51 shows the output noise spectral densities for each of the contributors at 25°C. As can be noted, there is no contribution due to I_{N+} or R3, since the noninverting input of the op amp is grounded.



Figure 4-51: Preamp output spectral noise densities (nV/\sqrt{Hz}) @25°C

Noise Reduction Using Output Filtering

From the above analysis, the largest contributor to the output noise voltage at 25°C is the input voltage noise of the op amp reflected to the output by the noise gain. This contributor is large primarily because the noise gain over which the integration is performed extends to a bandwidth of 167 kHz (the intersection of the noise gain curve with the open-loop response of the op amp). If the op amp output is filtered by a single-pole low-pass filter with a 20 Hz cutoff frequency ($\tau = 7.95$ ms), this contribution is reduced to less than 1 µVrms. The diagram for the final, filtered, optimized photodiode circuit design is shown in Figure 4-52.



Figure 4-52: AD795K preamp with output filter and offset null option

Notice that the same results would not be achieved simply by increasing the feedback capacitor, C2. Increasing C2 lowers the high frequency noise gain, but the integration bandwidth becomes proportionally higher. Larger values of C2 may also decrease the signal bandwidth to unacceptable levels.

The addition of the post-filter stage reduces output noise to 28.5 μ Vrms; approximately 75% of its former value, and the resistor noise and current noise are now the largest contributors to output noise. Practically, this filter can be either active or passive. Care will need to be taken, of course, that the filter circuit does not add any significant noise of its own to the signal. Filter design is discussed in greater detail in Chapter 5 of this book. The final circuit also includes an offset trim arrangement that is capable of nulling op amp offsets of up to ±1.5 mV.

Summary of Circuit Performance

Performance characteristics are summarized in Figure 4-53. The total output voltage drift over 0 to 70°C is 87.2 mV, corresponding to 87.2 pA of diode current. The offset nulling circuit shown on the noninverting input can be used to null out the room temperature offset. Note that this method is better than using the offset null pins because using the offset null pins will increase the offset voltage TC by about 3 μ V/°C for each millivolt nulled. In addition, the AD795 SOIC package does not have offset nulling pins.

The input sensitivity based on a total output voltage noise of 44 μ V is obtained by dividing the output voltage noise by the value of the feedback resistor R2. This yields a minimum detectable diode current of 44 fA. If a 12-bit ADC is used to digitize the 10 V full-scale output, the weight of the least significant bit (LSB) is 2.5 mV. The output noise level is much less than this.

- Output Offset Error (0°C to 70°C) : 87.2mV
- Output Sensitivity: 1mV/pA
- Output Photosensitivity: 30V / foot-candle
- Total Output Noise @ 25°C : 28.5µV RMS
- Total Noise RTI @ 25°C : 44fA RMS, or 26.4pA p-p
- Range with R2 = $1000M\Omega$: 0.001 to 0.33 foot-candles
- Bandwidth: 16Hz

Figure 4-53: AD795JR photodiode preamp performance summary

Photodiode Circuit Trade-off

Many trade-offs could be made in the basic photodiode circuit design we have described. More signal bandwidth can be achieved in exchange for a larger output noise level. Reducing the feedback capacitor C2 to 1 pF increases the signal bandwidth to approximately 160 Hz. Further reductions in C2 are not practical because the parasitic capacitance is probably in the order of 1 pF to 2 pF. Some small amount of feedback capacitance is also required to maintain stability.

If the circuit is to be operated at higher levels of illumination (greater than approximately 0.3 fc), the value of the feedback resistor can be reduced, thereby resulting in further increases in circuit bandwidth and less resistor noise.

If gain-ranging is to be used to measure higher light levels, extreme care must be taken in the design and layout of the additional switching networks to minimize leakage paths and other parasitic effects.

Compensation of a High Speed Photodiode I/V Converter

A classical I/V converter is shown in Figure 4-54. Note that it is the same as the previous photodiode preamplifier, if we assume that $R1 \gg R2$. The total input capacitance, C1, is the sum of the diode capacitance and the op amp input capacitance. Dynamically, this is a classical second-order system, and the following guidelines can be applied in order to determine the proper compensation.



Figure 4-54: Input capacitance compensation for an I/V converter

The net input capacitance, C1, forms a zero at a frequency f_1 in the noise gain transfer function as shown in the Bode plot.

$$f_1 = \frac{1}{2\pi R^2 C l}$$
 Eq. 4-6

Note that we are neglecting the effects of the compensation capacitor C2 and are assuming that it is small relative to C1 and will not significantly affect the zero frequency f_1 when it is added to the circuit. In most cases, this approximation yields results that are close enough, considering the other variables in the circuit.

If left uncompensated, the phase shift at the frequency of intersection, f_2 , will cause instability and oscillation. Introducing a pole at f_2 by adding the feedback capacitor C2 stabilizes the circuit and yields a phase margin of about 45 degrees.

$$f_2 = \frac{1}{2\pi R^2 C^2}$$
 Eq. 4-7

Since f_2 is the geometric mean of f_1 and the unity-gain bandwidth frequency of the op amp, f_u ,

These equations can be combined and solved for C2:

$$C2 = \sqrt{\frac{C1}{2\pi R2 \cdot f_u}}$$
 Eq. 4-9

This C2 value yields a phase margin of about 45 degrees; increasing it by a factor of 2 increases phase margin to about 65 degrees. In practice, an optimum C2 value should be determined experimentally, by varying it slightly to optimize the output pulse response.

Op Amp Selection for Wideband Photodiode I/V Converters

The op amp in the high speed photodiode I/V converter should be a wideband FET-input one in order to minimize the effects of input bias current and allow low values of photocurrents to be detected. In addition, if the equation for the 3 dB bandwidth, f_2 , is rearranged in terms of f_u , R2, and C1, then

$$f_2 = \sqrt{\frac{f_u}{2\pi R^2 C 1}}$$
 Eq. 4-10

where C1 is the sum of the diode capacitance, C_D , and the op amp input capacitance, C_{IN} . In a high speed application, the diode capacitance will be much smaller than that of the low frequency preamplifier design previously discussed—perhaps as low as a few pF.

By inspection of this equation, it is clear that in order to maximize f_2 , the FET-input op amp should have both a high unity gain bandwidth product, f_u , and a low input capacitance, C_{IN} . In fact, the ratio of f_u to C_{IN} is a good figure-of-merit when evaluating different op amps for this application.

Figure 4-55 compares a number of FET-input op amps suitable for photodiode preamps. By inspection, the AD823 op amp has the highest ratio of unity gain bandwidth product to input capacitance, in addition to relatively low input bias current.

For these reasons, the AD823 op amp was chosen for the wideband design.

	Unity GBW Product f _u (MHz)	Input Capacitance C _{IN} (pF)	f _u /C _{IN} (MHz/pF)	Input Bias Current I _B (pA)	Voltage Noise @ 10kHz (nV/√Hz)
AD823	16	1.8	8.9	3	16
AD843	34	6	5.7	600	19
AD744	13	5.5	2.4	100	16
AD845	16	8	2	500	18
OP42	10	6	1.6	100	12
AD745*	20	20	1	250	2.9
AD795	1	1	1	3	8
AD820	1.9	2.8	0.7	10	13
AD743	4.5	20	0.2	250	2.9

*Stable for Noise Gains \geq 5, Usually the Case, Since High Frequency Noise Gain = 1 + C1/C2, and C1 Usually \geq 4C2



High Speed Photodiode Preamp Design

The HP 5082-4204 PIN Photodiode will be used as an example for our discussion. Its characteristics are listed in Figure 4-56. It is typical of many PIN photodiodes.

- Sensitivity: 350µA @ 1mW, 900nm
- Maximum Linear Output Current: 100µA
- Area: 0.002cm² (0.2mm²)
- Capacitance: 4pF @ 10V Reverse Bias
- Shunt Resistance: $10^{11}\Omega$
- Risetime: 10ns
- Dark Current: 600pA @ 10V Reverse Bias

Figure 4-56: HP 5082-4204 photodiode characteristics

As in most high speed photodiode applications, the diode will be operated in the reverse-biased or *photo-conductive* mode. This greatly lowers the diode junction capacitance, but causes a small amount of *dark current* to flow even when the diode is not illuminated (we will show a circuit that compensates for the dark current error later in the section). This photodiode is linear with illumination up to approximately 50 μ A to 100 μ A of output current.

The available dynamic range is limited by the total circuit noise, and the diode dark current (assuming no dark current compensation).

 $C_{D} = 4pF, C_{IN} = 1.8pF$ $C_{1} = C_{D} + C_{IN} = 5.8pF$ $33.2k\Omega = 33.2k\Omega = 33.2k\Omega$ +15V $R2 = 100k\Omega$

C1

5.8pF

100kΩ

D2

0.1µF

LOW LEAKAGE POLYPROPYLENE AD823

_151/

D1. D2: HP-5082-4204

Using the circuit shown in Figure 4-57, assume that we wish to have a full scale output of 10 V for a diode current of 100 μ A. This determines the value of the feedback resistor R2 to be 10 V/100 μ A = 100 k Ω .



Using the diode capacitance, $C_D = 4$ pF, and the AD823 input capacitance, $C_{IN} = 1.8$ pF, the value of $C1 = C_D + C_{IN} = 5.8$ pF. Solving the above equations using C1=5.8pF, R2 = 100 k Ω , and $f_u = 16$ MHz, we find that:

$$f_1 = 274 \text{ kHz}$$

 $C2 = 0.76 \text{ pF}$
 $f_2 = 2.1 \text{ MHz}.$

In the final circuit shown, notice at the 100 k Ω resistor is replaced with three 33.2 k Ω film resistors to minimize stray capacitance. The feedback capacitor, C2, is a variable 1.5 pF ceramic and is adjusted in the final circuit for best bandwidth/pulse response. The overall circuit bandwidth is approximately 2 MHz.

The full-scale output voltage of the preamp for 100 μ A diode current is 10 V, and therefore the (uncompensated) error (RTO) due to the photodiode dark current of 600 pA is 60 μ V.

This dark current error can be effectively canceled using a second photodiode, D2, of the same type. This diode is biased with a voltage identical to D1 and, with nominally matched characteristics, will tend to conduct a similar dark current. In the circuit, this "dummy" dark current drives the 100 k Ω resistance in the noninverting input of the op amp. This produces a dark current proportional bias voltage which, due to the CM rejection of the op amp, has an end result of suppressing the dark current effects.

High Speed Photodiode Preamp Noise Analysis

As in most noise analyses, only the key contributors need be identified. Because the noise sources combine in an RSS manner, any single noise source that is at least three or four times as large as any of the others will dominate.

In the case of the wideband photodiode preamp, the dominant sources of output noise are the input voltage noise of the op amp, V_N , and the resistor noise due to R2, $V_{N,R2}$ (see Figure 4-58). The input current noise of the FET-input op amp is negligible. The shot noise of the photodiode (caused by the reverse bias) is negligible because of the filtering effect of the shunt capacitance C1.



Figure 4-58: Equivalent circuit of preamp for output noise analysis

The resistor noise is easily calculated by knowing that a 1 k Ω resistor generates about 4 nV/ \sqrt{z} , therefore, a 100 k Ω resistor generates 40 nV/ \sqrt{z} . The bandwidth for integration is the signal bandwidth, 2.1 MHz, yielding a total output rms noise of:

$$V_{N,R2}$$
RTO NOISE = $40\sqrt{1.57 \cdot 2.1 \cdot 10^6}$ = 73 µV rms

The factor of 1.57 converts the amplifier approximate single-pole bandwidth of 2.1 MHz into the *equivalent* noise bandwidth.

The output noise due to the input voltage noise is obtained by multiplying the noise gain by the voltage noise and integrating the entire function over frequency. This would be tedious if done rigorously, but a few reasonable approximations can be made which greatly simplify the math. Obviously, the low frequency 1/f noise can be neglected in the case of the wideband circuit. The primary source of output noise is due to the high frequency noise-gain peaking that occurs between f_1 and f_u . If we simply assume that the output noise is constant over the entire range of frequencies and use the maximum value for ac noise gain [1 + (C1/C2)], then

$$V_{\rm N}$$
 RTO NOISE $\approx V_{\rm N} \left(1 + \frac{\text{Cl}}{\text{C2}} \right) \sqrt{1.57 \text{f}_2} = 250 \ \mu\text{V} \,\text{rms}$

The total rms noise referred to the output is then the RSS value of the two components:

FOTAL RTO NOISE =
$$\sqrt{(73)^2 + (250)^2} = 260 \,\mu\text{V rms}$$

The total output dynamic range can be calculated by dividing the 10 V full-scale output by the total $260 \,\mu$ Vrms noise, and, converting to dB, yielding approximately 92 dB.

High Impedance Charge Output Sensors

High impedance transducers such as piezoelectric sensors, hydrophones, and some accelerometers require an amplifier that converts a transfer of charge into a voltage change. Due to the high dc output impedance of these devices, appropriate buffer amplifiers are required. The basic circuit for an inverting charge sensitive amplifier is shown in Figure 4-59.



Figure 4-59: Charge amplifier basic principles

There are basically two types of charge transducers: capacitive and charge-emitting. In a capacitive transducer, the voltage across the capacitor (V_c) is held constant. The change in capacitance, ΔC , produces a change in charge, $\Delta Q = \Delta CV_c$. This charge is transferred to the op amp output as a voltage, $\Delta V_{OUT} = -\Delta Q/C2 = -\Delta CV_c/C2$.

Charge-emitting transducers produce an output charge, ΔQ , and their output capacitance remains constant. This charge would normally produce an open-circuit output voltage at the transducer output equal to $\Delta Q/C$. However, since the voltage across the transducer is held constant by the virtual ground of the op amp (R1 is usually small), the charge is transferred to capacitor C₂ producing an output voltage $\Delta V_{OUT} = -\Delta Q/C2$. In an actual application, this charge amplifier only responds to ac inputs.

It should be noted that ac gain for this charge amplifier is determined by the *capacitance ratio*, not the resistances. This is unlike a conventional wideband amplifier, where the upper cutoff frequency is given by $f_2 = 1/2 \pi R2C2$, and the lower by $f_1 = 1/2 \pi R1C1$.

In the Figure 4-59 charge amplifier, with nominal transducer capacitance fixed, gain is set by C2. Typically, bias return resistor R2 will be a high value ($\geq 1 \mod \Omega$), and R1 a much lower value. The resulting frequency response will be relatively narrow and bandpass shaped, with gain and frequency manipulated by the relative values, as suitable to SPICE analysis.

Low Noise Charge Amplifier Circuit Configurations

Figure 4-60 shows two ways to buffer and amplify the output of a charge output transducer. Both require using an amplifier which has a very high input impedance, such as the AD745. The AD745 provides both low voltage noise and low current noise. This combination makes this device particularly suitable in applications requiring very high charge sensitivity, such as capacitive accelerometers and hydrophones.



Figure 4-60: Two basic charge amplifier configurations using the AD745 low noise FET op amp

The first (left) circuit in Figure 4-60 uses the op amp in the inverting mode. Amplification depends on the principle of conservation of charge at the inverting input of the amplifier. The charge on capacitor C_s is transferred to capacitor C_F , thus yielding an output voltage of $\Delta Q/C_F$. The amplifier's input voltage noise will appear at the output amplified by the ac noise gain of the circuit, $1 + C_s/C_F$.

The second (right) circuit shown in Figure 4-60 is simply a high impedance follower with gain. Here the noise gain (1 + R2/R1) is the same as the gain from the transducer to the output. Resistor R_B, in both circuits, is required as a dc bias current return.

To maximize dc performance over temperature, source resistances should be balanced at the amplifier inputs, as represented by the resistor R_B shown in Figure 4-60. For best noise performance, the source capacitance should also be balanced with the capacitor C_B .

In general, it is good practice to balance the source impedances (both resistive and reactive) as seen by the inputs of precision low noise BiFET amplifiers such as the AD743/AD745. Balancing the resistive components will optimize dc performance over temperature, as balancing mitigates the effects of any bias current errors. Balancing the input capacitance will minimize ac response errors due to the amplifier's nonlinear common-mode input capacitance and, as shown in Figure 4-60, noise performance will be optimized. In any FET input amplifier, the current noise of the internal bias circuitry can be coupled to the inputs via the gate-to-source capacitances (20 pF for the AD743 and AD745) and appears as excess input voltage noise. This noise component is correlated at the inputs, so source impedance matching will tend to cancel out its effect. Figure 4-60 shows the required external components for both inverting and noninverting configurations. For values of C_B greater than 300 pF, there is a diminishing impact on noise, and C_B can then be simply a polyester bypass capacitor of 0.01 µF or greater.

40dB Gain Piezoelectric Transducer Amplifier Operates on Reduced Supply Voltages for Lower Bias Current

Figure 4-61 shows a gain-of-100 piezoelectric transducer amplifier connected in the voltage-output mode. Reducing the op amp power supplies to ± 5 V reduces the effects of bias current in two ways: first, by lowering the total power dissipation and, second, by reducing the basic gate-to-junction leakage current. The addition of a clip-on heat sink such as the Aavid #5801 will further limit the internal junction temperature rise.

With ac coupling capacitor C1 shorted, the amplifier will operate over a range of 0° C to 85° C. If the optional ac coupling capacitor C1 is used, the circuit will operate over the entire -55° C to $+125^{\circ}$ C temperature range, but dc information is lost.



• \pm 5V Power Supplies Reduce I_B for 0°C to +85°C Operation, P_D = 80mW

• C1 Allows -55°C to +125°C Operation

Figure 4-61: A gain-of-100 piezoelectric transducer amplifier

Hydrophones

Interfacing the outputs of highly capacitive transducers such as hydrophones, some accelerometers, and condenser microphones to the outside world presents many design challenges. Previously designers had to use costly hybrid amplifiers consisting of discrete low-noise JFETs in front of conventional op amps to achieve the low levels of voltage and current noise required by these applications. Using AD743/AD745 monolithic ICs, designers can achieve almost the same level of performance of a hybrid approach.

In sonar applications, a piezo-ceramic cylinder is commonly used as the active element in the hydrophone. A typical cylinder has a nominal capacitance of around 6,000 pF with a series resistance of 10 Ω . The output impedance is typically 10⁸ Ω or 100 M Ω .

Since the hydrophone signals of interest are inherently ac with wide dynamic range, noise is an overriding concern for sonar system designs. The noise floor of the hydrophone and its preamplifier together limit the system sensitivity, and thus the overall hydrophone usefulness. Typical hydrophone bandwidths are in the 1 kHz to 10 kHz range. The AD743 and AD745 op amps, with their low noise voltages of $2.9 \text{nV}/\sqrt{\text{Hz}}$ and high input impedance of $10^{10} \Omega$ (or 10 G Ω) are ideal for use as hydrophone amplifiers.

Op Amp Performance: JFET versus Bipolar

The AD743 and AD745 op amps are the first monolithic JFET devices to offer the low input voltage noise comparable to a bipolar op amp, *without* the high input bias currents typically associated with bipolar op amps. Figure 4-62 shows input voltage noise versus input source resistance of the OP27 (bipolar-input) and the JFET-input AD745 op amps. Note: the noise levels of the AD743 and the AD745 are identical.

From this figure (left), it is clear that at high source impedances, the low current noise of the AD745 provides lower overall noise than a high performance bipolar op amp such as the OP27. It is also important to note that, with the AD745, this noise reduction extends down to low source impedances. At high source impedances, the lower dc current errors of the AD745 also reduce errors due to offset and drift, as shown in Figure 4-62 (right).



The AD743 and AD745 are related, companion amplifiers, which differ in their levels of internal compensation. The AD743 is internally compensated for unity gain stability. The AD745, stable for noise gains of 5 or higher, has a much higher bandwidth and slew rate. This makes the latter device especially useful as a high-gain preamplifier where it provides both high gain and wide bandwidth. The AD743 and AD745 also operate with very low levels of distortion; less than 0.0003% and 0.0002% (at 1 kHz), respectively.

A pH Probe Buffer Amplifier

A typical pH probe requires a buffer amplifier to isolate its 10^6 to $10^9 \Omega$ source resistance from external circuitry. Such an amplifier is shown in Figure 4-63. The low input current of the AD795JR minimizes the voltage error produced by the bias current and electrode resistance. The use of guarding, shielding, high insulation resistance standoffs, and other such standard picoamp methods previously discussed should be used to minimize leakage. Are all needed to maintain the accuracy of this circuit.





The slope of the pH probe transfer function, 50 mV per pH unit at room temperature, has an approximate +3500 ppm/°C temperature coefficient. The buffer amplifier shown as Figure 4-63 provides a gain of 20, and yields a final output voltage equal to 1 V/pH unit. Temperature compensation is provided by resistor RT, which is a special temperature compensation resistor, 1 k Ω , 1%, +3500 ppm/°C, #PT146 available from Precision Resistor Co., Inc. (see Reference 15).

References: High Impedance Sensors

- 1. Ramon Pallas-Areny and John G. Webster, **Sensors and Signal Conditioning**, John Wiley, New York, 1991.
- Dan Sheingold, Editor, Transducer Interfacing Handbook, Analog Devices, Inc., 1980, ISBN: 0-916550-05-2.
- Section 3, Walt Kester, Editor, 1992 Amplifier Applications Guide, Analog Devices, 1992, ISBN: 0-916550-10-9.
- 4. Walt Kester, Editor, System Applications Guide, Analog Devices, 1993, ISBN: 0-916550-13-3.
- 5. Walt Kester, Editor, Linear Design Seminar, Analog Devices, 1995, ISBN: 0-916550-15-X.
- Walt Kester, Editor, Practical Analog Design Techniques, Analog Devices, 1995, ISBN: 0-916550-16-8.
- 7. Walt Kester, Editor, High Speed Design Techniques, Analog Devices, 1996, ISBN: 0-916550-17-6.
- Jiri Dostal, Section 9.3.1, "Thermoelectric Voltages," Operational Amplifiers 2nd Ed., Butterworth-Heinemann, 1993, pp. 265–268.
- 9. Optoelectronics Data Book, EG&G Vactec, St. Louis, MO, 1990.
- 10. Silicon Detector Corporation, Camarillo, CA, Part Number SD-020-12-001 Data Sheet.
- 11. Photodiode 1991 Catalog, Hamamatsu Photonics, Bridgewater, NJ.
- Lewis Smith and Dan Sheingold, "Noise and Operational Amplifier Circuits," Analog Dialogue 25th Anniversary Issue, pp. 19–31, Analog Devices Inc., 1991.
- 13. James L. Melsa and Donald G. Schultz, Linear Control Systems, pp. 196–220, McGraw-Hill, 1969.
- 14. Jerald G. Graeme, Photodiode Amplifiers: Op Amp Solutions, McGraw-Hill, 1995.
- 15. Precision Resistor Co., Inc., 10601 75th St. N., Largo, FL, 33777-1427, 727 541-5771, www.precisionresistor.com.
- 16. **Ohmite Victoreen MAXI-MOX Resistors**, 3601 Howard Street, Skokie, IL 60076, 847 675-2600, www.ohmite.com/victoreen/.
- 17. Vishay/Dale RNX Resistors, 2300 Riverside Blvd., Norfolk, NE, 68701-2242, 402 371-0800, www.vishay.com.

This page intentionally left blank
SECTION 4-5

Temperature Sensors Walt Kester, James Bryant, Walt Jung

Temperature measurement is critical in many modern electronic devices, especially expensive laptop computers and other portable devices; their densely packed circuitry dissipates considerable power in the form of heat. Knowledge of system temperature can also be used to control battery charging, as well as to prevent damage to expensive microprocessors.

Compact high power portable equipment often has fan cooling to maintain junction temperatures at proper levels. In order to conserve battery life, the fan should only operate when necessary. Accurate control of the fan requires, in turn, knowledge of critical temperatures from an appropriate temperature sensor.

Accurate temperature measurements are required in many other measurement systems, for example within process control and instrumentation applications. Some popular types of temperature transducers and their characteristics are indicated in Figure 4-64. In most cases, because of low level and/or nonlinear outputs, the sensor output must be properly conditioned and amplified before further processing can occur.

THERMOCOUPLE	RTD	THERMISTOR	SEMICONDUCTOR	
Widest Range: -184°C to +2300°C	Range: –200°C to +850°C	Range: 0°C to +100°C	Range: -55°C to +150°C	
High Accuracy and Repeatability	Fair Linearity	Poor Linearity	Linearity: 1°C Accuracy: 1°C	
Needs Cold Junction Compensation	Requires Excitation	Requires Excitation	Requires Excitation	
Low-Voltage Output	Low Cost	High Sensitivity	10mV/K, 20mV/K, or 1μA/K Typical Output	

Figure 4-64: Some common types of temperature transducers

Except for the semiconductor sensors of the last column, all of the temperature sensors shown have nonlinear transfer functions. In the past, complex analog conditioning circuits were designed to correct for the sensor nonlinearity. These circuits often required manual calibration and precision resistors to achieve the desired accuracy. Today, however, sensor outputs may be digitized directly by high resolution ADCs. Linearization and calibration can then be performed digitally, substantially reducing cost and complexity.

Resistance Temperature Devices (RTDs) are accurate, but require excitation current and are generally used within bridge circuits such as those described earlier. *Thermistors* have the most sensitivity, but are also the most nonlinear. Nevertheless, they are popular in portable applications for measurement of battery and other critical system temperatures.

Modern *semiconductor temperature sensors* offer both high accuracy and linearity over about a -55° C to $+150^{\circ}$ C operating range. Internal amplifiers can scale output to convenient values, such as 10 mV/°C. They are also useful in cold-junction compensation circuits for wide temperature range thermocouples. Semiconductor temperature sensors are also integrated into ICs that perform other hardware monitoring functions—an example is the imbedded transistor sensors used within modern µP chips.

Thermocouple Principles and Cold-Junction Compensation

Thermocouples comprise a more common form of temperature measurement. In operation, thermocouples rely on the fact that two dissimilar metals joined together produce a voltage output roughly proportional to temperature. They are small, rugged, relatively inexpensive, and operate over the widest range of contact temperature sensors.

Thermocouples are especially useful for making measurements at extremely high temperatures (up to +2300°C), and within hostile environments. Characteristics of some common types are shown in Figure 4-65.

JUNCTION MATERIALS	TYPICAL USEFUL RANGE (°C)	NOMINAL SENSITIVITY (µV/°C)	ANSI DESIGNATION
Platinum (6%)/ Rhodium- Platinum (30%)/Rhodium	38 to 1800	7.7	В
Tungsten (5%)/Rhenium- Tungsten (26%)/Rhenium	0 to 2300	16	С
Chromel-Constantan	0 to 982	76	E
Iron-Constantan	0 to 760	55	E
Chromel-Alumel	-184 to 1260	39	к
Platinum (13%)/Rhodium- Platinum	0 to 1593	11.7	R
Platinum (10%)/Rhodium- Platinum	0 to 1538	10.4	S
Copper-Constantan	-184 to 400	45	т

Figure 4-65: Some common thermocouples and their characteristics

However, thermocouples produce only millivolts of output, and thus they typically require precision amplification for further processing. They also require *cold-junction-compensation* (CJC) techniques, to be discussed shortly. They are more linear than many other sensors, and their nonlinearity has been well characterized.

The most common metals used for thermocouples are iron, platinum, rhodium, rhenium, tungsten, copper, alumel (composed of nickel and aluminum), chromel (composed of nickel and chromium) and constantan (composed of copper and nickel).

Figure 4-66 shows the voltage-temperature curves for three commonly used thermocouples, Types J, K, and S, referred to a 0°C fixed-temperature reference junction. Of these, Type J thermocouples are the most sensitive, producing the highest output voltage for a given temperature change, but over a relatively narrow temperature span. On the other hand, Type S thermocouples are the least sensitive, but can operate over a much wider range.



These characteristics are very important considerations when designing thermocouple signal conditioning circuitry. A main consideration is the fact that virtually any thermocouple employed will have a relatively low output signal, thus they generally require the careful application of stable, high gain amplifiers.

To understand thermocouple behavior more fully, it is also necessary to consider the nonlinearities in their response to temperature differences. As noted, Figure 4-66 shows the relationships between sensing junction temperature and voltage output for a number of thermocouple types (in all cases, the reference *cold* junction is maintained at 0°C).

While close scrutiny of these data may reveal the fact that none of the responses are quite linear, the exact nature of the nonlinearity isn't so obvious. What is needed is another perspective on the relationships displayed by these curves, to gain better insight into how the various devices can be best utilized.

Figure 4-67 shows how the thermocouple *Seebeck* coefficient varies with sensor junction temperature. The Seebeck coefficient is the *change* of output voltage with *change* of sensor junction temperature (i.e., the first derivative of output with respect to temperature). Note that we are still considering the case where the reference junction is maintained at 0° C.



An ideal linear thermocouple would have a constant Seebeck coefficient with varying temperature, but in practice all thermocouples are nonlinear to some degree. In selecting a measurement thermocouple for a particular temperature range, we should therefore choose one whose Seebeck coefficient varies as little as possible over that range.

For example, a Type J thermocouple has a nominal Seebeck coefficient of 55 μ V/°C, which varies by less than 1 μ V/°C between 200°C and 500°C, making it ideal for measurements over this range (the flat region of the upper curve within Figure 4-67).

Presenting these data on thermocouples serves two purposes. First, Figure 4-66 illustrates the range and sensitivity of the three thermocouple types so that the system designer can, at a glance, determine that a Type S thermocouple has the widest useful temperature range, but a Type J thermocouple is the more sensitive of the three.

Second, the relative stability of the Seebeck coefficient over temperature provides a quick guide to a thermocouple's linearity. Using Figure 4-67, a system designer can choose a Type K thermocouple for its relatively linear Seebeck coefficient over the range of 400°C to 800°C, or a Type S over the range of 900°C to 1700°C. The behavior of a thermocouple's Seebeck coefficient is important in applications where variations of temperature rather than absolute magnitude are important. These data also indicate what performance is required of the associated signal-conditioning circuitry.

To successfully apply thermocouples we must also understand their basic operating principles. Consider the diagrams shown in Figure 4-68.

If we join two dissimilar metals, A and B, at any temperature above absolute zero, there will be a potential difference between them, i.e., their "thermoelectric EMF" or "contact potential," V1. This voltage is a function of the temperature of the measurement junction, T1, as is noted in Figure 4-68A. If we join two wires of metal A with metal B at two places, two measurement junctions are formed, T1 and T2 (Figure 4-68B). If the two junctions are at different temperatures, there will be a net EMF in the circuit, and a current I will flow, as determined by the EMF V1 – V2, and the total resistance R in the circuit (Figure 4-68B).



Figure 4-68: Thermocouple basics

If we open the circuit as in Figure 4-68C, the voltage across the break will be equal to the net thermoelectric EMF of the circuit; and if we measure this voltage, we can use it to calculate the temperature difference between the two junctions. We must remember that a thermocouple always measures the temperature *difference* between two junctions, not the absolute temperature at one junction. We can only measure the temperature at the measuring junction if we know the temperature of the other junction. This is the origin of the terms "reference" or "cold" junction.

But of course, it is not so easy to just measure the voltage generated by a thermocouple. *Any wire attached to a thermocouple is also a thermocouple itself*, and if care is not taken, errors can be introduced. Suppose that we attach a voltmeter to the circuit of Figure 4-68C, as shown in Figure 4-68D. The wires of the voltmeter will form further thermocouple junctions where they are attached (T3, T4). If both these additional junctions are at the same temperature (it does not matter exactly what temperature), then the "Law of Intermediate Metals" states that they will make no net contribution to the total EMF of the system. If they are at different temperatures, they will introduce errors.

Since *every pair of dissimilar metals in contact generates a thermoelectric EMF* (including copper/solder, kovar/copper [kovar is the alloy used for IC leadframes] and aluminum/kovar [at the bond inside the IC]), it is obvious that in practical circuits the problem is even more complex. In fact, it is necessary to take extreme care to ensure that both junctions of each junction pair in series with a thermocouple are at the same temperature, *except for the measurement and reference junctions themselves*.

Thermocouples generate a voltage, albeit a very small one, and don't require excitation for this most basic operation. As was shown in Figure 4-68D, however, two junctions are always involved (T1, the measurement junction temperature and T2, the reference junction temperature). If T2 = T1, then V2 = V1, and the output voltage V = 0. Thermocouple output voltages are often defined with respect to a reference junction temperature of 0°C (hence the term *cold* or *ice point* junction). In such a system the thermocouple provides a convenient output voltage of 0 V at 0°C. Obviously, to maintain system accuracy, the reference junction must remain at a well-defined temperature (but not necessarily 0°C).

A conceptually simple approach to this need is shown in Figure 4-69, the ice-point reference, where junction T2 is kept at 0°C by virtue of being immersed in an ice water slurry. Although this ice water bath is relatively easy to define conceptually, it is quite inconvenient to maintain.



Figure 4-69: A thermocouple cold junction reference system using an ice-point (0°C) T2 reference

Today the inconvenient ice/water bath reference is replaced by an electronic equivalent. A temperature sensor of another sort (often a semiconductor sensor, sometimes a thermistor) measures the cold junction temperature and is used to inject a voltage into the thermocouple circuit which compensates for the difference between the actual cold junction temperature and its ideal value (usually 0° C) as shown in Figure 4-70.



Figure 4-70: A semiconductor temperature sensor can be used to provide cold junction compensation

Ideally, the compensation voltage should be an exact match for the difference voltage required, which is why the diagram gives the voltage V(COMP) as f(T2) (a *function* of T2) rather than KT2, where K is a simple constant. In practice, since the cold junction is rarely more than a few tens of degrees from 0°C, and generally varies by little more than $\pm 10^{\circ}$ C, a linear approximation (V = KT2) to the more complex reality is usually sufficiently accurate and is often used. (Note—the expression for the output voltage of a thermocouple with its measuring junction at T°C and its reference at 0°C is a polynomial of the form V = K₁T + K₂T² + K₃T³ + ..., but the values of the coefficients K₂, K₃, and so forth. are very small for most common types of thermocouple). References 7 and 8 give the values of these coefficients for a wide range of thermocouples.

When electronic cold-junction compensation is used, it is common practice to eliminate the additional thermocouple wire, and terminate the thermocouple leads in the isothermal block, as shown in the arrangement of Figure 4-71.



In Figure 4-71, the Metal A-Copper and the Metal B-Copper junctions, if at temperature T2, are equivalent to the Metal A-Metal B thermocouple junction in Figure 4-70.

Type K Thermocouple Amplifier and Cold Junction Compensator

The circuit in Figure 4-72 conditions the output of a Type K thermocouple, while providing cold-junction compensation, operating over temperatures between 0°C and 250°C. The circuit operates from single +3.3 V to +5.5 V supplies with the AD8551, and has been designed to produce a basic output voltage transfer characteristic of 10 mV/°C. A Type K thermocouple exhibits a Seebeck coefficient of approximately $40 \,\mu$ V/°C; therefore, at the cold junction, the TMP35 voltage output sensor with a temperature coefficient of 10 mV/°C is used with divider R1 and R2, to introduce an opposing cold-junction temperature coefficient of $-40 \,\mu$ V/°C. This prevents the isothermal, cold-junction connection between the circuit's printed circuit board traces and the thermocouple's wires from introducing an error in the measured temperature. This compensation works extremely well for conditioning circuit ambient temperatures of 20°C to 50°C.



Figure 4-72: Using a TMP35 temperature sensor for cold junction compensation within a Type K thermocouple amplifier-conditioner

Over a 250°C measurement temperature range, the thermocouple produces an output voltage change of ~10 mV. Since the circuit's required full-scale output voltage change is 2.5 V, the required gain is ~250. Choosing R4 equal to 4.99 k Ω sets R5 ~1.24 M Ω . With a fixed 1% value for R5 of 1.21 M Ω , a 50 k Ω potentiometer is used with R5 for fine trim of the full-scale output voltage. The U1 amplifier should be a low drift, very high gain type. A chopper-stabilized AD8551 or an OP777 precision bipolar op amp is suitable for U1.

Both the AD8551 and the OP777 have rail-rail output stages. To extend low range linearity, bias resistor R3 is added to the circuit, supplying an output offset voltage of about 0.1 V (for a nominal supply voltage of 5 V). Note that this 10°C offset must be subtracted, when making final measurements referenced to the U1 output. Note also that R3 provides a useful open thermocouple detection function, forcing the U1 output to greater than 3 V should the thermocouple open. Resistor R7 balances the dc input impedance at the U1 (+) input, and the 0.1 μ F film capacitor reduces noise coupling.

Single-Chip Thermocouple Signal Conditioners

While the construction of thermocouple signal conditioners with op amps and other discrete circuit elements offers great flexibility, this does come at the expense of component count. To achieve a greater level of integration in the thermocouple conditioning function, dedicated thermocouple signal conditioners can be used.

One such solution lies with the AD594 and AD595 (see Reference 9), which are complete, single-chip instrumentation amplifiers and thermocouple cold junction compensators, as shown in Figure 4-73. Suitable for use with either Type J (AD594) or Type K (AD595) thermocouples, these two devices combine an ice point reference with a precalibrated scaling amplifier. They provide a high level (10 mV/°C) output working directly from a thermocouple input, without additional precision parts. Pin-strapping options allow the devices to be used as a linear amplifier-compensator, or as a switched output set-point controller with fixed or remote setpoint control.



Figure 4-73: Functional diagram of AD594 and AD595 thermocouple signal-conditioning amplifiers

The AD594 and AD595 can be used to amplify the cold-junction compensation voltage directly, thereby becoming a standalone, 10 mV/°C output Celsius transducer. In such applications it is very important that the IC chip be at the same temperature as the cold junction of the thermocouple; this is usually achieved by keeping the two in close proximity and isolated from any heat sources.

The AD594/AD595 structure includes a flexible thermocouple failure alarm output, which provides broken thermocouple indication. The devices can be powered from either dual or single power supplies (as low as 5 V), but the use of a negative supply also allows temperatures below 0°C to be measured. To minimize self-heating, an unloaded AD594/AD595 operates with 160 μ A of supply current, and can deliver ±5 mA to a load.

Although the AD594 is precalibrated by laser wafer trimming to match the characteristics of Type J thermocouples, and the AD595 for Type K, the temperature transducer voltages and gain control resistors are also made available at the package pins. So, if desired, the circuit can be recalibrated for other thermocouple types with the addition of external resistors. These terminals also allow more precise calibration for both thermocouple and thermometer applications. The AD594/AD595 are available in C and A performance grades, with calibration accuracies of $\pm 1^{\circ}$ C and $\pm 3^{\circ}$ C, respectively. Both are designed to be used with cold junctions between 0 to 50°C.

The 5 V powered, single-supply circuit shown of Figure 4-73 provides a scaled 10 mV/°C output capable of measuring a range of 0°C to 300°C. This can be from either a Type J thermocouple using the AD594, or a Type K with the AD595.

Resistance Temperature Detectors

The Resistance Temperature Detector (RTD), is a sensor whose resistance changes with temperature. Typically built of a platinum (Pt) wire wrapped around a ceramic bobbin, the RTD exhibits behavior which is more accurate and more linear than a thermocouple over wide temperature ranges.

Figure 4-74 illustrates the temperature coefficient of a 100 Ω RTD, and the Seebeck coefficient of a Type S thermocouple. Over the entire range (approximately –200°C to +850°C), the RTD is a more linear device. Hence, linearizing an RTD is less complex.



Figure 4-74: Resistance temperature detectors

Unlike a thermocouple, however, an RTD is a passive sensor, and a current excitation is required to produce an output voltage. The RTD's low temperature coefficient of 0.385%/°C requires high performance signalconditioning circuitry similar to that used by a thermocouple. However, the typical voltage drop seen across an RTD is much larger than a thermocouple's output voltage. A system designer may opt for large value RTDs with higher output, but large-valued RTDs exhibit slow response times. Furthermore, although the cost of RTDs is higher than that of thermocouples, they use copper leads, and thermoelectric effects from terminating junctions do not affect their accuracy. And finally, because their resistance is a function of the absolute temperature, RTDs do not require cold-junction compensation.

Caution must be exercised with the level of current excitation applied to an RTD, because excessive current can cause self-heating. Any self-heating changes the RTD temperature, and therefore results in a measurement error. Hence, careful attention must be paid to the design of the signal-conditioning circuitry so that

self-heating errors are kept below 0.5°C. Manufacturers specify self-heating errors for various RTD values and sizes, in both still and moving air. To reduce the error due to self-heating, the minimum current should be used to achieve the required system resolution, and the largest RTD value chosen that results in acceptable response time.

Another effect that can produce measurement error is voltage drop in RTD lead wires. This is especially critical with low value, 2-wire RTDs, because both the temperature coefficient and absolute value of the RTD resistance are small. If the RTD is located a long distance from the signal-conditioning circuitry, the connecting lead resistance can be ohms or tens of ohms. Even this small amount of lead resistance can contribute a significant error to the temperature measurement, as shown in Figure 4-75.



RESISTANCE TC OF COPPER = 0.40%/°C @ 20°C RESISTANCE TC OF Pt RTD = 0.385%/°C @ 20°C

Figure 4-75: A 100 Ω Pt RTD with 100-foot #30 AWG lead wires

To illustrate this point, assume that a 100 Ω platinum RTD with 30-gage copper leads is located about 100 feet from a controller's display console. The resistance of 30-gage copper wire is 0.105 Ω /ft, and the two leads of the RTD will contribute a total 21 Ω to the network. Uncorrected, this additional resistance will produce a 55°C measurement error. Obviously, the temperature coefficient of the connecting leads can contribute an additional, and possibly significant, error to the measurement.

To eliminate the effect of the lead resistance, a 4-wire technique is used. In Figure 4-76, a 4-wire (Kelvin) connection is made to the RTD. A constant current, I, is applied though the FORCE leads of the RTD, and the voltage across the RTD itself is measured remotely, via the SENSE leads. The measuring device can be a DVM or an in amp, and high accuracy can be achieved provided that the measuring device exhibits high



Figure 4-76: Use of Kelvin or 4-wire Pt RTD connections provides high accuracy

input impedance and/or low input bias current. Since the SENSE leads don't carry appreciable current, this technique is relatively insensitive to lead wire length. Some major sources of errors in this scheme are the stability of the constant current source, and the input impedance and/or bias currents in the amplifier or DVM, and the associated drift.

RTDs are generally configured in a four-resistor bridge circuit. The bridge output is amplified by an in amp for further processing. However, high resolution measurement ADCs allow the RTD output to be digitized directly. In this manner, linearization can be performed digitally, thereby easing the analog circuit requirements considerably.

For example, an RTD output can be digitized by one of the AD77xx series high resolution ADCs. Figure 4-77 shows a 100 Ω Pt RTD, driven with a 400 μ A excitation current source. Note that the 400 μ A RTD excitation current source also generates a 2.5 V reference voltage for the ADC, by virtue of flowing in a 6.25 k Ω resistor, R_{REF}, with the drop across this resistance being metered by the ADC's V_{REF} (+) and (–) input terminals.



Figure 4-77: A Pt RTD interfaced to the AD77xx series of high resolution ADCs

It should be noted that this simple scheme has great benefits (beyond the obvious one of simplicity). Variations in the magnitude of the 400 μ A excitation current do not affect circuit accuracy, since both the input voltage drop across the RTD as well as the reference voltage across R_{REF} vary ratiometrically with the excitation current. However, it should be noted that the 6.25 k Ω resistor must be a stable type with a low temperature coefficient, to avoid errors in the measurement. Either a wirewound resistor, or a very low TC metal film type is most suitable for R_{REF} within this application.

In this application, the ADC's high resolution and the gain of 1 to 128 input PGA eliminates the need for any additional conditioning. The high resolution ADC can in fact perform virtually all the conditioning necessary for an RTD, leaving any further processing such as linearization to be performed in the digital domain.

Thermistors

Similar in general function to RTDs, thermistors are low-cost temperature-sensitive resistors, constructed of solid semiconductor materials which exhibit a positive or negative temperature coefficient. Although positive temperature coefficient devices do exist, the most common thermistors are negative temperature coefficient (NTC) devices.



Figure 4-78: Resistance characteristics for a $10k\Omega$ NTC thermistor

Figure 4-78 shows the resistance-temperature characteristic of a commonly used NTC thermistor. Although the thermistor is the most nonlinear of the three temperature sensors discussed, it is also the most sensitive.

The thermistor's very high sensitivity (typically – 44,000 ppm/°C at 25°C) allows it to detect minute temperature variations not readily observable with an RTD or thermocouple. This high sensitivity is a distinct advantage over the RTD, in that 4-wire Kelvin connections to the thermistor aren't needed for lead wire error compensation. To illustrate this point, suppose a 10 k Ω NTC thermistor with a typical 25°C temperature coefficient of –44,000 ppm/°C were substituted for the 100 Ω Pt RTD in the example given earlier. The total lead wire resistance of 21 Ω would generate less than 0.05°C error in the measurement, using the thermistor in lieu of the RTD. This is roughly a factor of 500 improvement in error sensitivity over an RTD.

However, the thermistor's high sensitivity to temperature does not come without a price. As previously shown in Figure 4-78, the temperature coefficient of thermistors does not decrease linearly with increasing temperature as with RTDs, and as a result linearization is required for all but the most narrow temperature ranges. Thermistor applications are limited to a few hundred degrees at best, because thermistors are also more susceptible to damage at high temperatures.

Compared to thermocouples and RTDs, thermistors are fragile in construction and require careful mounting procedures to prevent crushing or bond separation. Although a thermistor's response time is short due to its small size, its small thermal mass also makes it very sensitive to self-heating errors.

Thermistors are very inexpensive, highly sensitive temperature sensors. However, we have noted that a thermistor's temperature coefficient can vary, from -44,000 ppm/°C at 25°C, to -29,000 ppm/°C at 100°C. Not only is this nonlinearity the largest source of error in a temperature measurement, it also limits useful applications to very narrow temperature ranges without linearization.

As shown in Figure 4-79, a parallel resistor combination exhibits a more linear variation with temperature compared to the thermistor itself. This approach to linearizing a thermistor simply shunts it with a fixed, temperature-stable resistor. Paralleling the thermistor with a fixed resistor increases the linearity significantly. Also, the sensitivity of the combination still is high compared to a thermocouple or RTD. The primary disadvantage of the technique is that linearization is only effective within a narrow range. However, it is possible to use a thermistor over a wide temperature range, if the system designer can tolerate a lower net sensitivity, in order to achieve improved linearity.



Figure 4-79: Linearization of NTC thermistor using a fixed shunt resistance

R, the value of the fixed shunt resistor, can be calculated from the following equation:

$$R = \frac{RT2 \cdot (RT1 + RT3) - 2 \cdot RT1 \cdot RT3}{RT1 + RT3 - 2 \cdot RT2}$$
Eq. 4-11

where RT1 is the thermistor resistance at T1, the lowest temperature in the measurement range, RT3 is the thermistor resistance at T3, the highest temperature in the range, and RT2 is the thermistor resistance at T2, the midpoint, T2 = (T1 + T3)/2.

For a typical 10 k Ω NTC thermistor, RT1 = 32,650 Ω at 0°C, RT2 = 6,532 Ω at 35°C, and RT3 = 1,752 Ω at 70°C. This results in a value of 5.17 k Ω for R. The accuracy needed in the associated signal-conditioning circuitry depends on the linearity of the network. For the example given above, the network shows a nonlinearity of -2.3°C/+2.0°C.

The output of the network can be applied to an ADC for digital conversion (with optional linearization) as shown in Figure 4-80. Note that the output of the thermistor network has a slope of approximately $-10 \text{ mV/}^{\circ}\text{C}$, which implies that an 8- or 10-bit ADC easily has more than sufficient resolution with a full scale range of 1 V or less. The further linearization can be applied to the data in the digital domain, if desired.



 $LINEARITY \approx \pm 2^{\circ}C, \ 0^{\circ}C \ TO \ +70^{\circ}C$

Figure 4-80: Linearized thermistor network with amplifier or ADC

Semiconductor Temperature Sensors

Modern semiconductor temperature sensors offer high accuracy and high linearity over an operating range of about -55° C to $+150^{\circ}$ C. Internal amplifiers can scale the output to convenient values, such as 10 mV/°C. They are also useful in cold-junction-compensation circuits for wide temperature range thermocouples.

All semiconductor temperature sensors make use of the relationship between a bipolar junction transistor's (BJT) base-emitter voltage to its collector current:

$$V_{BE} = \frac{kT}{q} \ln \left(\frac{I_{C}}{I_{S}} \right)$$
 Eq. 4-12

In this expression k is Boltzmann's constant, T is the absolute temperature, q is the charge of an electron, and I_s is a current related to the geometry and the temperature of the transistors. (The equation assumes a voltage of at least a few hundred mV on the collector, and ignores Early effects.)

If we take 'N' transistors identical to the first (see Figure 4-81) and allow the total current I_c to be shared equally among them, we find that the new base-emitter voltage applicable to this case is given by the equation

$$V_{N} = \frac{kT}{q} \ln\left(\frac{I_{C}}{N \cdot I_{S}}\right)$$
Eq. 4-13



Figure 4-81: The basic relationships for BJT-based semiconductor temperature sensors

Neither of these circuits is of much use by itself, because of the strong temperature dependence of I_s . However, if we have equal currents flowing in one BJT, as well as the N similar BJTs, then the expression for the *difference* between the respective base-emitter voltages (or ΔV_{BE}) is proportional to absolute temperature, and it does not contain I_s .

This then leads to a far more useful relationship, developed as follows:

$$\Delta \mathbf{V}_{\rm BE} = \mathbf{V}_{\rm BE} - \mathbf{V}_{\rm N} = \frac{kT}{q} \ln \left(\frac{\mathbf{I}_{\rm C}}{\mathbf{I}_{\rm S}} \right) - \frac{kT}{q} \ln \left(\frac{\mathbf{I}_{\rm C}}{\mathbf{N} \cdot \mathbf{I}_{\rm S}} \right)$$

$$\Delta V_{BE} = V_{BE} - V_{N} = \frac{kT}{q} \left[\ln \left(\frac{I_{C}}{I_{S}} \right) - \ln \left(\frac{I_{C}}{N \cdot I_{S}} \right) \right]$$
$$\Delta V_{BE} = V_{BE} - V_{N} = \frac{kT}{q} \ln \left[\left(\frac{I_{C}}{I_{S}} \right) / \left(\frac{I_{C}}{N \cdot I_{S}} \right) \right] = \frac{kT}{q} \ln \left(N \right)$$

This end result of this algebra is expressed in a single key equation, one worthy of restatement:

$$\Delta V_{BE} = \frac{kT}{q} \ln (N)$$
 Eq. 4-14

As one can note, Eq. 4-14 contains only the transistor emitter area ratio N and T as variables. Since N is fixed within a given design, it can be the basis of a transducer for T measurement. The circuit as shown in Figure 4-82 implements the above equation, and is popularly known as the "Brokaw Cell," after its inventor (see Reference 10).



Figure 4-82: The "Brokaw Cell" is both a silicon bandgap voltage-based reference as well as a temperature sensor

The voltage $\Delta V_{BE} = V_{BE} - V_N$ appears across resistor R2, as noted. The emitter current in Q2 is therefore $\Delta V_{BE}/R2$. The op amp's servo loop and the two resistors 'R' force an identical current to flow through Q1. The Q1 and Q2 currents are equal, and they are summed, flowing in resistor R1.

The corresponding voltage developed across R1 is V_{PTAT} , a voltage proportional to absolute temperature (PTAT). This is given by:

$$V_{PTAT} = \frac{2R1(V_{BE} - V_{N})}{R2} = 2\frac{R1}{R2}\frac{kT}{q}\ln(N)$$

Within the circuit, a voltage labeled as $V_{BANDGAP}$ appears at the base of Q1, and, as can be noted, is the sum of $V_{BE(Q1)}$ and V_{PTAT} . When this voltage is set by the design to be exactly equal to the bandgap voltage of silicon, it will then become independent of temperature. The voltage $V_{BE(Q1)}$ is complementary to absolute temperature (CTAT), and summing it with a properly proportioned V_{PTAT} from across R1 gives the desired end result; the bandgap voltage becomes constant with respect to temperature. Note that this assumes the proper choice of R1/R2 ratio and N, to make the summed voltage equal to $V_{BANDGAP}$, the silicon bandgap voltage (in this instance 1.205 V). This circuit has the virtues of dual application because of the above features. It is useful as a basic silicon bandgap temperature sensor, with either direct or scaled use of the voltage V_{PTAT} . It is also widely used a temperature stable reference voltage source, by suitable scaling of $V_{BANDGAP}$ to standard outputs of 2.500 V, 5.000 V, and so forth.

Current and Voltage Output Temperature Sensors

The concepts used in the bandgap voltage temperature sensor discussion above can also be used as the basis for a variety of IC temperature sensors, with linear, proportional-to-temperature outputs, of either current or voltage.

The AD592 device shown in Figure 4-83 is a two-terminal, current output sensor with a scale factor of 1 μ A/K. This device does not require external calibration, and is available in several accuracy grades. The AD592 is a TO92 packaged version of the original AD590 TO52 metal packaged temperature transducer device (see Reference 11).



AD592 Specified from –25°C to +105°C

Figure 4-83: Current output absolute temperature sensor

The simplest operating mode for current mode temperature sensors is to load them with a precision resistor of 1% or better tolerance, and read the output voltage developed with either an ADC or a scaling ampli-fier/buffer. Figure 4-84 shows this technique with an ADC, as applicable to the AD592. The resistor load R1 converts the basic scaling of the sensor $(1 \mu A/K)$ into a proportional voltage.

Choice of this resistor determines the overall sensitivity of the temperature sensor, in terms of V/K. For example, with a 1 k Ω precision resistor load as shown, the net circuit sensitivity becomes 1 mV/K. With a 5 V bias on the temperature sensor as shown, the AD592's full dynamic range is allowed with a 1 K Ω load. If a higher value R1 is used, higher bias voltage may be required, as the AD592 requires 4 V of operating headroom.



Figure 4-84: Current output temperature sensor driving a resistive load

The function just described is a Kelvin-scaled temperature sensor, so the ADC will be required to read the full dynamic range of voltage across R1. With an AD592, this span is from the range corresponding to -25° C (248 K) to 105° C (378 K), which is 0.248 V to 0.378 V. A 10-bit, 0.5 V scaled ADC can read this range directly with $\approx 0.5^{\circ}$ C resolution.

If a centigrade-scaled reading is desired, two options present themselves. For a traditional analog approach, the common terminal of the ADC input can easily be biased with a reference voltage corresponding to 0°C, or 0.273 V. Alternately, the 0°C reference can be inserted in the digital domain, with the advantage of no additional hardware requirement.

The AD592 is available in three accuracy grades. The highest grade version (AD592CN) has a maximum error @ 25°C of ± 0.5 °C and ± 1.0 °C error from -25°C to ± 105 °C, and a linearity error of ± 0.35 °C. The AD592 is available in a TO-92 package.

With regard to stand-alone digital output temperature sensors, it is worthy of note that such devices do exist, that is ADCs with built-in temperature sensing. The AD7816/AD7817/AD7818-series ADCs have on-board temperature sensors digitized by a 10-bit 9 µs conversion time switched capacitor SAR ADC. The device family offers a variety of input options for flexibility. The similar AD7416/AD7417/AD7418 have serial interfaces.

For a great many temperature sensing applications, a voltage mode output sensor is most appropriate. For this, there are a variety of standalone sensors that can be directly applied. In such devices the basic mode of operation is as a three-terminal device, using power input, common, and voltage output pins. In addition, some devices offer an optional shutdown pin.

The TMP35/TMP36 are low voltage (2.7 V to 5.5 V), SO-8 or TO-92 packaged voltage output temperature sensors with a 10 mV/°C scale factor, as shown in Figure 4-85. Supply current is below 50 μ A, providing very low self-heating (less than 0.1°C in still air).



output temperature sensors with shutdown capability

Output scaling of this device family differs in range and 25°C offset. The TMP35 provides a 250 mV output at 25°C, and reads temperature from 10°C to 125°C. The TMP36 is specified from -40°C to +125°C. and provides a 750 mV output at 25°C. Both the TMP35 and TMP36 have an output scale factor of +10 mV/°C.

An optional shutdown feature is provided for the SO8 package devices, which reduces the standby current to 0.5 μ A. This pin, when taken to a logic LOW, turns the device OFF, and the output becomes a high impedance state. If shutdown isn't used, the pin should be connected to +V_s.

The power supply pin of these voltage mode sensors should be bypassed to ground with a 0.1 μ F ceramic capacitor having very short leads (preferably surface mount) and located as close to the power supply pin as possible. Since these temperature sensors operate on very little supply current and could be exposed to very hostile electrical environments, it is also important to minimize the effects of EMI/RFI on these devices. The effect of RFI on these temperature sensors is manifested as abnormal dc shifts in the output voltage due to rectification of the high frequency noise by the internal IC junctions. In those cases where the devices are operated in the presence of high frequency radiated or conducted noise, a large value tantalum electrolytic capacitor (>2.2 μ F) placed across the 0.1 μ F ceramic may offer additional noise immunity.

Ratiometric Voltage Output Temperature Sensors

In some cases, it is desirable for the output of a temperature sensor to be *ratiometric* with respect its supply voltage. A series of ADI temperature sensors have been designed to fulfill this need, in the form of the AD2210x series (see references 12–14). Of this series, the AD22103 illustrated in Figure 4-86 has an output that is ratiometric with regard its nominal 3.3 V supply voltage, according to the equation:

$$V_{OUT} = \frac{V_s}{3.3V} \times \left(0.25V + \frac{28mV}{^{\circ}C} \times T_A \right)$$
 Eq. 4-15



Figure 4-86: Ratiometric voltage output temperature sensors

Note also that the Figure 4-86 circuit uses the 3.3 V AD22103 power supply voltage as the reference input to the ADC. This step eliminates the need for a separate precision voltage reference source. Within a system, this key point potentially can have a positive impact.

Operation of the AD22103 is accomplished with an on-chip temperature sensing resistance R_T , which operates similarly to the RTD types discussed earlier in this section. This resistance is fed from a resistance network comprised of R14, R16, and R32. R14 and R16 provide a current drive component for R_T that is proportional to the supply voltage, a factor that gives the AD22103 a basic sensitivity that is proportional to the supply. R_T , like a classic platinum RTD, exhibits a nonlinear resistance versus temperature behavior. This nonlinear characteristic of R_T is corrected by a positive feedback loop, composed of R32 along with the Thevenin equivalent of resistances R14 and R16.

Gain scaling for the changing R_T output voltage is provided by the op amp negative feedback loop, R18, and the Thevenin equivalent of resistances R24 and R22. This references the gain network of the op amp to the supply voltage, instead of ground. The various resistance networks around the op amp are actively trimmed at temperature, to calibrate the sensor for its rated offset and scaling.

The net combination of these factors allows the device to behave in accordance with the relationship of Eq. 4-15. The AD22103 is specified over a range of 0°C to 100°C, and has an accuracy better than ± 2.5 °C, along with a linearity better than ± 0.5 % of full scale, i.e., 0.5°C over 100°C.

Since the AD22103 is a single-supply part, the sensing of low temperatures necessarily involves a positive output offset. For example, for 3.3 V operation of this example, the output offset is simply the 0.25 V term of Eq. 4-15. Accordingly, the 0 to 100°C temperature span translates to an output swing of 0.25 V to 3.05 V.

Should it be desired, operation of the AD22103 device is also possible at higher supply voltages. Because of the ratiometric operation feature, this will necessarily involve a change to both the basic sensitivity and the offset. For example, in operating the AD22103 at 5 V, the output expression changes to:

$$V_{OUT} = \frac{V_s}{5V} \times \left(0.378V + \frac{42.42mV}{^{\circ}C} \times T_A \right)$$
Eq. 4-16

However, it should be noted that the fact that the AD22103 is ratiometric does *not* preclude operating the part from a fixed reference voltage. The nominal current drain of the AD22103 is 500 μ A, allowing a number of these sensors to be operated from a common reference IC without danger of overload (as well as other analog parts). For example, one such reference family is the REF19x series, which can supply output currents of up to 30 mA.

In addition to the above-described AD22103 3.3 V part, there is also a companion device, the AD22100. While basically quite similar to the AD22103, the AD22100 operates from a nominal 5 V power supply with reduced sensitivity, allowing operation over a range of -50 to $+150^{\circ}$ C. Over this range the rated accuracy of the AD22100 is 2% or better, and linearity error is 1% or less.

References: Temperature Sensors

- Ramon Pallas-Areny and John G. Webster, Sensors and Signal Conditioning, John Wiley, New York, 1991.
- Dan Sheingold, Editor, Transducer Interfacing Handbook, Analog Devices, Inc., 1980, ISBN: 0-916550-05-2.
- Sections 2, 3, Walt Kester, Editor, 1992 Amplifier Applications Guide, Analog Devices, 1992, ISBN: 0-916550-10-9.
- Sections 1, 6, Walt Kester, Editor, System Applications Guide, Analog Devices, 1993, ISBN: 0-916550-13-3.
- 5. Dan Sheingold, Editor, Nonlinear Circuits Handbook, Analog Devices, Inc., 1974.
- James Wong, "Temperature Measurements Gain from Advances in High-precision Op Amps," Electronic Design, May 15, 1986.
- 7. OMEGA Temperature Measurement Handbook, Omega Instruments, Inc.
- 8. Handbook of Chemistry and Physics, Chemical Rubber Co.
- 9. Joe Marcin, "Thermocouple Signal Conditioning Using the AD594/AD595," Analog Devices AN369.
- Paul Brokaw, "A Simple Three-Terminal IC Bandgap Voltage Reference," IEEE Journal of Solid-State Circuits, Vol. SC-9, No. 6, December, 1974, pp. 388–393.
- Mike Timko, "A Two-Terminal IC Temperature Transducer," IEEE Journal of Solid-State Circuits, Vol. SC-11, No. 6, December, 1976, pp. 784–788.
- 12. Data sheet for AD22103 3.3 V Supply, Voltage Output Temperature Sensor with Signal Conditioning, www.analog.com.
- 13. Data sheet for AD22100 Voltage Output Temperature Sensor with Signal Conditioning, www analog.com.
- Adrian P. Brokaw, "Monolithic Ratiometric Temperature Measurement Circuit," US Patent No. 5,030,849, filed June 30, 1989, issued July 9, 1991.



AD590 basic (left), and complete schematics (right)

Designed by Mike Timko, based on an original Paul Brokaw concept,¹ the AD590^{2,3,4} current mode IC temperature transducer was introduced in 1977. The AD590 established early an elegant method of accurate temperature measurement, based upon fundamental silicon transistor operating principles. It has been in ADI production since introduction, along with such related ICs as the AD592 discussed within this chapter.

The references below discuss operation in great detail, but suffice it to say that in the basic structure (left) a voltage proportional to absolute temperature, V_T , appears across resistor R. This makes the current I_T drawn from an external source proportional to absolute temperature. In the full circuit (right), trimmed thin film resistors implement a calibrated scaling for I_T of 1µA/K. Additional circuitry is added for startup and for increased accuracy, both with respect to applied voltage as well as against high temperature leakage.

A current-operated transducer such as this is quite convenient to operate, the output being impervious to long lead lengths, and also virtually noise-immune. The low scaling factor also makes the AD590 easy to operate on low voltage supplies without self-heating, yet high output impedance also holds calibration with higher applied voltages. Readout is simply accomplished with a single resistance, making a simple, two-component Kelvin-scaled thermometer possible.

Adrian P. Brokaw, "Digital-to-Analog Converter with Current Source Transistors Operated Accurately at Different Current Densities," US Patent No. 3,940,760, filed March 21, 1975, issued Feb. 24, 1976.

² Mike Timko, "A Two-Terminal IC Temperature Transducer," IEEE Journal of Solid-State Circuits, Vol. SC-11, No. 6, December, 1976, pp. 784–788.

³ Mike Timko, Goodloe Suttler, "1µA/K IC Temperature-to-Current Transducer," Analog Dialogue, Vol. 12, No. 1, 1978, pp. 3–5.

⁴ Michael P. Timko, Adrian P. Brokaw, "Integrated Circuit Two Terminal Temperature Transducer," US Patent No. 4,123,698, filed July 6, 1976, issued October 31, 1978.

This page intentionally left blank

CHAPTER 5 Analog Filters

- Section 5-1: Introduction
- Section 5-2: The Transfer Function
- Section 5-3: Time Domain Response
- Section 5-4: Standard Responses
- Section 5-5: Frequency Transformations
- Section 5-6: Filter Realizations
- Section 5-7: Practical Problems in Filter Implementation
- Section 5-8: Design Examples

This page intentionally left blank

CHAPTER 5 Analog Filters Hank Zumbahlen

SECTION 5-1

Introduction

Filters are networks that process signals in a frequency-dependent manner. The basic concept of a filter can be explained by examining the frequency-dependent nature of the impedance of capacitors and inductors. Consider a voltage divider where the shunt leg is a reactive impedance. As the frequency is changed, the value of the reactive impedance changes, and the voltage divider ratio changes. This mechanism yields the frequency dependent change in the input/output transfer function that is defined as the frequency-response.

Filters have many practical applications. A simple, single pole, low-pass filter (the integrator) is often used to stabilize amplifiers by rolling off the gain at higher frequencies where excessive phase shift may cause oscillations.

A simple, single pole, high pass filter can be used to block dc offset in high gain amplifiers or single supply circuits. Filters can be used to separate signals, passing those of interest, and attenuating the unwanted frequencies.

An example of this is a radio receiver, where the signal to be processed is passed through, typically with gain, while attenuating the rest of the signals. In data conversion, filters are also used to eliminate the effects of aliases in A/D systems. They are used in reconstruction of the signal at the output of a D/A as well, eliminating the higher frequency components, such as the sampling frequency and its harmonics, thus smoothing the waveform.

There are a large number of texts dedicated to filter theory. No attempt will be made to go heavily into much of the underlying math—Laplace transforms, complex conjugate poles and the like—although they will be mentioned.

While they are appropriate for describing the effects of filters and examining stability, in most cases examination of the function in the frequency domain is more illuminating.

An ideal filter will have an amplitude response that is unity (or at a fixed gain) for the frequencies of interest (called the *pass band*) and zero everywhere else (called the *stop band*). The frequency at which the response changes from pass band to stop band is referred to as the *cutoff frequency*.

Chapter Five

Figure 5-1(A) shows an idealized low-pass filter. In this filter the low frequencies are in the pass band and the higher frequencies are in the stop band.

The functional complement to the low-pass filter is the high-pass filter. Here, the low frequencies are in the stop band, and the high frequencies are in the pass band. Figure 5-1(B) shows the idealized high-pass filter.

If a high-pass filter and a low-pass filter are cascaded, a *band-pass* filter is created. The band-pass filter passes a band of frequencies between a lower cutoff frequency, f_1 , and an upper cutoff frequency, f_h . Frequencies below f_1 and above f_h are in the stop band. An idealized band-pass filter is shown in Figure 5-1(C).

A complement to the band-pass filter is the *bandreject*, or *notch* filter. Here, the pass bands include frequencies below f_1 and above f_h . The band from f_1 to f_h is in the stop band. Figure 5-1(D) shows a notch response.



Figure 5-1: Idealized filter responses

Unfortunately, the idealized filters defined above cannot be easily built. The transition from pass band to stop band will not be instantaneous, but instead there will be a transition region. Stop band attenuation will not be infinite.

The five parameters of a practical filter are defined in Figure 5-2.



The *cutoff frequency* (F_c) is the frequency at which the filter response leaves the error band (or the -3 dB point for a Butterworth response filter). The *stop band frequency* (F_s) is the frequency at which the minimum attenuation in the stop band is reached. The *pass band ripple* (A_{max}) is the variation (error band) in the pass band response. The *minimum pass band attenuation* (A_{min}) defines the minimum signal attenuation within the stop band. The steepness of the filter is defined as the *order* (M) of the filter. M is also the number of poles in the transfer function. A pole is a root of the denominator of the transfer function. Conversely, a zero is a root of the numerator of the transfer function. Each pole gives a -6 dB/octave or -20 dB/decade response. Each zero gives a +6 dB/octave, or +20 dB/decade response.

Note that not all filters will have all these features. For instance, all-pole configurations (i.e., no zeros in the transfer function) will not have ripple in the stop band. Butterworth and Bessel filters are examples of all-pole filters with no ripple in the pass band.

Typically, one or more of the above parameters will be variable. For instance, if you were to design an antialiasing filter for an ADC, you will know the cutoff frequency (the maximum frequency to be passed), the stop band frequency, (which will generally be the Nyquist frequency (= $\frac{1}{2}$ the sample rate)) and the minimum attenuation required (which will be set by the resolution or dynamic range of the system). Go to a chart or computer program to determine the other parameters, such as filter order, F₀, and Q, which determines the peaking of the section, for the various sections and/or component values.

It should also be pointed out that the filter will affect the phase of a signal, as well as the amplitude. For example, a single pole section will have a 90° phase shift at the crossover frequency. A pole pair will have a 180° phase shift at the crossover frequency. The Q of the filter will determine the rate of change of the phase. This will be covered more in depth in the next section.

This page intentionally left blank

SECTION 5-2 The Transfer Function

The S-Plane

Filters have a frequency dependent response because the impedance of a capacitor or an inductor changes with frequency. Therefore the complex impedances:

$$Z_L = sL$$
 Eq. 5-1

and

$$Z_{\rm C} = \frac{1}{\rm sC}$$
 Eq. 5-2

$$s = \sigma + j\omega$$
 Eq. 5-3

are used to describe the impedance of an inductor and a capacitor, respectively, where σ is the Neper frequency in nepers per second (NP/s) and ω is the angular frequency in radians per sec (rad/s).

By using standard circuit analysis techniques, the transfer equation of the filter can be developed. These techniques include Ohm's law, Kirchoff's voltage and current laws, and superposition, remembering that the impedances are complex. The transfer equation is then:

$$H(s) = \frac{a_{m}s^{m} + a_{m-1}s^{m-1} + \dots + a_{1}s + a_{0}}{b_{n}s^{n} + b_{n-1}s^{n-1} + \dots + b_{1}s + b_{0}}$$
Eq. 5-4

Therefore, H(s) is a rational function of *s* with real coefficients with the degree of *m* for the numerator and *n* for the denominator. The degree of the denominator is the order of the filter. Solving for the roots of the equation determines the poles (denominator) and zeros (numerator) of the circuit. Each pole will provide a -6 dB/octave or -20 dB/decade response. Each zero will provide a +6 dB/octave or +20 dB/decade response. These roots can be real or complex. When they are complex, they occur in conjugate pairs. These roots are plotted on the *s* plane (complex plane) where the horizontal axis is σ (real axis) and the vertical axis is ω (imaginary axis). How these roots are distributed on the *s* plane can tell us many things about the circuit. In order to have stability, all poles must be in the left side of the plane. If there is a zero at the origin, that is a zero in the numerator, the filter will have no response at dc (high pass or band pass).

Assume an RLC circuit, as in Figure 5-3. Using the voltage divider concept it can be shown that the voltage across the resistor is:

$$H(s) = \frac{Vo}{Vin} = \frac{RCs}{LCs^2 + RCs + 1}$$
Eq. 5-5

Substituting the component values into the equation yields:

$$H(s) = 10^3 \times \frac{s}{s^2 + 10^3 s + 10^7}$$

Factoring the equation and normalizing gives:

$$H(s) = 10^{3} \times \frac{s}{[s - (-0.5 + j3.122) \times 10^{3}] \times [s - (-0.5 - j3.122) \times 10^{3}]}$$

Figure 5-3: RLC circuit

This gives a zero at the origin and a pole pair at:

$$s = (-0.5 \pm j3.122) \times 10^3$$

Next, plot these points on the s plane as shown in Figure 5-4:



The above discussion has a definite mathematical flavor. In most cases there is more interest in the circuit's performance in real applications. While working in the s plane is completely valid, most of us don't think in terms of Nepers and imaginary frequencies.

F_o and Q

So if it is not convenient to work in the s plane, why go through the above discussion? The answer is that the groundwork has been set for two concepts that will be infinitely more useful in practice: F_0 and Q.

 F_o is the cutoff frequency of the filter. This is defined, in general, as the frequency where the response is down 3 dB from the passband. It can sometimes be defined as the frequency at which it will fall out of the pass band. For example, a 0.1 dB Chebyshev filter can have its F_o at the frequency at which the response is down > 0.1 dB.

The shape of the attenuation curve (as well as the phase and delay curves, which define the time domain response of the filter) will be the same if the ratio of the actual frequency to the cutoff frequency is examined, rather than just the actual frequency itself. Normalizing the filter to 1 rad/s, a simple system for designing and comparing filters can be developed. The filter is then scaled by the cutoff frequency to determine the component values for the actual filter.

Q is the "quality factor" of the filter. It is also sometimes given as α where:

$$\alpha = \frac{1}{Q}$$
 Eq. 5-6

This is commonly known as the *damping ratio*. ξ is sometimes used where:

$$\xi = 2\alpha$$
 Eq. 5-7

If Q is > 0.707, there will be some peaking in the filter response. If the Q is < 0.707, roll-off at F_0 will be greater; it will have a more gentle slope and will begin sooner. The amount of peaking for a 2-pole low-pass filter versus Q is shown in Figure 5-5.



Figure 5-5: Low-pass filter peaking versus Q

Rewriting the transfer function H(s) in terms of ω_0 and Q:

$$H(s) = \frac{H_0}{s^2 + \frac{\omega_0}{Q}s + {\omega_0}^2}$$
 Eq. 5-8

where H_0 is the pass band gain and $\omega_0 = 2\pi F_0$.

This is now the *low-pass prototype* that will be used to design the filters.

High-Pass Filter

Changing the numerator of the transfer equation, H(s), of the low-pass prototype to H_0s^2 transforms the low-pass filter into a high-pass filter. The response of the high-pass filter is similar in shape to a low-pass, just inverted in frequency.

The transfer function of a high-pass filter is then:

$$H(s) = \frac{H_0 s^2}{s^2 + \frac{\omega_0}{O} s + {\omega_0}^2}$$
Eq. 5-9

The response of a 2-pole high-pass filter is illustrated in Figure 5-6.



Figure 5-6: High-pass filter peaking versus Q

Band-Pass Filter

Changing the numerator of the low-pass prototype to $H_0\omega_0^2$ will convert the filter to a band-pass function. The transfer function of a band-pass filter is then:

$$H(s) = \frac{H_0 \omega_0^2}{s^2 + \frac{\omega_0}{Q} s + \omega_0^2}$$
 Eq. 5-10

 ω_0 here is the frequency (F₀ = 2 $\pi \omega_0$) at which the gain of the filter peaks.

 H_0 is the circuit gain and is defined:

$$H_0 = H/Q$$
 Eq. 5-11

Q has a particular meaning for the band-pass response. It is the selectivity of the filter. It is defined as:

$$Q = \frac{F_0}{F_H - F_L}$$
 Eq. 5-12

where F_L and F_H are the frequencies where the response is -3 dB from the maximum.

The bandwidth (BW) of the filter is described as:

$$BW = F_{H} - F_{L}$$
 Eq. 5-13

It can be shown that the resonant frequency (F_0) is the geometric mean of F_L and F_H , which means that F_0 will appear halfway between F_L and F_H on a logarithmic scale.

$$F_0 = \sqrt{F_H F_L} \qquad \qquad \text{Eq. 5-14}$$

Also, note that the skirts of the band-pass response will always be symmetrical around F_0 on a logarithmic scale.

The response of a band-pass filter to various values of Q are shown in Figure 5-7.



Figure 5-7: Band-pass filter peaking versus Q

A word of caution is appropriate here. Band-pass filters can be defined two different ways. The narrowband case is the classic definition that we have shown above.

In some cases, however, if the high and low cutoff frequencies are widely separated, the band-pass filter is constructed of separate high-pass and low-pass sections. Widely separated in this context means separated by at least two octaves (×4 in frequency). This is the wideband case.

Bandreject (Notch) Filter

By changing the numerator to $s^2 + \omega_z^2$, we convert the filter to a bandreject or notch filter. As in the bandpass case, if the corner frequencies of the bandreject filter are separated by more than an octave (the wideband case), it can be built out of separate low-pass and high-pass sections. The following convention will be adopted: A narrow band bandreject filter will be referred to as a *notch* filter and the wide band bandreject filter.

A notch (or bandreject) transfer function is:

$$H(s) = \frac{H_0(s^2 + \omega_z^2)}{s^2 + \frac{\omega_0}{Q}s + \omega_0^2}$$
Eq. 5-15

Chapter Five

There are three cases of the notch filter characteristics. These are illustrated in Figure 5-8. The relationship of the pole frequency, ω_0 , and the zero frequency, ω_z , determines if the filter is a standard notch, a low-pass notch or a high-pass notch.



If the zero frequency is equal to the pole frequency a standard notch exists. In this instance the zero lies on the $j\omega$ plane where the curve that defines the pole frequency intersects the axis.

A low-pass notch occurs when the zero frequency is greater than the pole frequency. In this case, ω_z lies outside the curve of the pole frequencies. What this means in a practical sense is that the filter's response below ω_z will be greater than the response above ω_z . This results in an elliptical low-pass filter.

A high-pass notch filter occurs when the zero frequency is less than the pole frequency. In this case, ω_z lies inside the curve of the pole frequencies. What this means in a practical sense is that the filter's response below ω_z will be less than the response above ω_z . This results in an elliptical high-pass filter.

The variation of the notch width with Q is shown in Figure 5-9.



All-Pass Filter

There is another type of filter that leaves the amplitude of the signal intact but introduces phase shift. This type of filter is called an *all-pass*. The purpose of this filter is to add phase shift (delay) to the response of

the circuit. The amplitude of an all-pass is unity for all frequencies. The phase response, however, changes from 0° to 360° as the frequency is swept from 0 to infinity. The purpose of an all-pass filter is to provide phase equalization, typically in pulse circuits. It also has application in single sideband, suppressed carrier (SSB-SC) modulation circuits.

~

The transfer function of an all-pass filter is:

$$H(s) = \frac{s^{2} - \frac{\omega_{0}}{Q}s + \omega_{0}^{2}}{s^{2} + \frac{\omega_{0}}{Q}s + \omega_{0}^{2}}$$
Eq. 5-16

Note that an allpass transfer function can be synthesized as:

$$H_{AP} = H_{LP} - H_{BP} + H_{HP} = 1 - 2H_{BP}$$
 Eq. 5-17

Figure 5-10 compares the various filter types.



Chapter Five

Phase Response

As mentioned earlier, a filter will change the phase of the signal as well as the amplitude. The question is, does this make a difference? Fourier analysis indicates a square wave is made up of a fundamental frequency and odd order harmonics. The magnitude and phase responses of the various harmonics are precisely defined. If the magnitude or phase relationships are changed, the summation of the harmonics will not add back together properly to give a square wave. It will instead be distorted, typically showing overshoot and ringing or a slow rise time. This would also hold for any complex waveform.

Each pole of a filter will add 45° of phase shift at the corner frequency. The phase will vary from 0° (well below the corner frequency) to 90° (well beyond the corner frequency). The start of the change can be more than a decade away. In multipole filters, each of the poles will add phase shift, so that the total phase shift will be multiplied by the number of poles $(180^{\circ} \text{ total shift for a two-pole system, } 270^{\circ} \text{ for a three-pole})$ system, and so forth).

The phase response of a single pole, low pass filter is:

$$\phi(\omega) = -\arctan\frac{\omega}{\omega_0} \qquad \text{Eq. 5-18}$$

The phase response of a low-pass pole pair is:

$$\phi(\omega) = -\arctan\left[\frac{1}{\alpha}\left(2\frac{\omega}{\omega_0} + \sqrt{4 - \alpha^2}\right)\right]$$

- $\arctan\left[\frac{1}{\alpha}\left(2\frac{\omega}{\omega_0} - \sqrt{4 - \alpha^2}\right)\right]$
Eq. 5-19

For a single pole high-pass filter the phase response is:

$$\phi(\omega) = \frac{\pi}{2} - \arctan\frac{\omega}{\omega_0} \qquad \text{Eq. 5-20}$$

The phase response of a high-pass pole pair is:

$$\phi(\omega) = \pi - \arctan\left[\frac{1}{\alpha}\left(2\frac{\omega}{\omega_0} + \sqrt{4 - \alpha^2}\right)\right]$$

- $\arctan\left[\frac{1}{\alpha}\left(2\frac{\omega}{\omega_0} - \sqrt{4 - \alpha^2}\right)\right]$
Eq. 5-21

The phase response of a band-pass filter is:

$$\phi(\omega) = \frac{\pi}{2} - \arctan\left(\frac{2Q\omega}{\omega_0} + \sqrt{4Q^2 - 1}\right)$$
Eq. 5-22
$$-\arctan\left(\frac{2Q\omega}{\omega_0} - \sqrt{4Q^2 - 1}\right)$$

The variation of the phase shift with frequency due to various values of Q is shown in Figure 5-11 (for lowpass, high-pass, band-pass, and all-pass) and in Figure 5-12 (for notch).


Figure 5-11: Phase response versus frequency



Figure 5-12: Notch filter phase response

It is also useful to look at the change of phase with frequency. This is the group delay of the filter. A flat (constant) group delay gives best phase response, but, unfortunately, it also gives the least amplitude discrimination. The group delay of a single low-pass pole is:

$$\tau(\omega) = -\frac{d\phi(\omega)}{d\omega} = \frac{\cos^2 \phi}{\omega_0}$$
 Eq. 5-23

For the low-pass pole pair it is:

$$\tau(\omega) = \frac{2\sin^2\phi}{\alpha\omega_0} - \frac{\sin 2\phi}{2\omega}$$
 Eq. 5-24

For the single high-pass pole it is:

$$t(\omega) = -\frac{d\phi(\omega)}{d\omega} = \frac{\sin^2 \phi}{\omega_0}$$
 Eq. 5-25

For the high-pass pole pair it is:

$$\tau(\omega) = \frac{2\sin^2\phi}{\alpha\omega_0} - \frac{\sin 2\phi}{2\omega}$$
 Eq. 5-26

And for the band-pass pole pair it is:

$$\tau(\omega) = \frac{2Q2\cos^2\phi}{\alpha\omega_0} + \frac{\sin 2\phi}{2\omega}$$
 Eq. 5-27

The Effect of Nonlinear Phase

A waveform can be represented by a series of frequencies of specific amplitude, frequency and phase relationships. For example, a square wave is:

$$F(t) = A\left(\frac{1}{2} + \frac{2}{\pi}\sin\omega t + \frac{2}{3\pi}\sin 3\omega t + \frac{2}{5\pi}\sin 5\omega t + \frac{2}{7\pi}\sin 7\omega t + ...\right)$$
 Eq. 5-28

If this waveform were passed through a filter, the amplitude and phase response of the filter to the various frequency components of the waveform could be different. If the phase delays were identical, the waveform would pass through the filter undistorted. If, however, the different components of the waveform were changed due to different amplitude and phase response of the filter to those frequencies, they would no longer add up in the same manner. This would change the shape of the waveform. These distortions would manifest themselves in what we typically call overshoot and ringing of the output.

Not all signals will be composed of harmonically related components. An amplitude modulated (AM) signal, for instance, will consist of a carrier and two sidebands at \pm the modulation frequency. If the filter does not have the same delay for the various waveform components, then "envelope delay" will occur and the output wave will be distorted.

Linear phase shift results in constant group delay since the derivative of a linear function is a constant.

SECTION 5-3

Until now the discussion has been primarily focused on the frequency domain response of filters. The time domain response can also be of concern, particularly under transient conditions. Moving between the time domain and the frequency domain is accomplished by the use of the Fourier and Laplace transforms. This yields a method of evaluating performance of the filter to a nonsinusoidal excitation.

The transfer function of a filter is the ratio of the output to input time functions. It can be shown that the impulse response of a filter defines its bandwidth. The time domain response is a practical consideration in many systems, particularly communications, where many modulation schemes use both amplitude and phase information.

Impulse Response

The impulse function is defined as an infinitely high, infinitely narrow pulse, with an area of unity. This is, of course, impossible to realize in a physical sense. If the impulse width is much less than the rise time of the filter, the resulting response of the filter will give a reasonable approximation actual impulse response of the filter response.

The impulse response of a filter, in the time domain, is proportional to the bandwidth of the filter in the frequency domain. The narrower the impulse, the wider the bandwidth of the filter. The pulse amplitude is equal to ω_c/π , which is also proportional to the filter bandwidth, the height being taller for wider bandwidths. The pulsewidth is equal to $2\pi/\omega_c$, which is inversely proportional to bandwidth. It turns out that the product of the amplitude and the bandwidth is a constant.

It would be a nontrivial task to calculate the response of a filter without the use of Laplace and Fourier transforms. The Laplace transform converts multiplication and division to addition and subtraction, respectively. This takes equations, which are typically loaded with integration and/or differentiation, and turns them into simple algebraic equations, which are much easier to deal with. The Fourier transform works in the opposite direction.

The details of these transforms will not be discussed here. However, some general observations on the relationship of the impulse response to the filter characteristics will be made.

It can be shown, as stated, that the impulse response is related to the bandwidth. Therefore, amplitude discrimination (the ability to distinguish between the desired signal from other, out-of-band signals and noise) and time response are inversely proportional. That is to say that the filters with the best amplitude response are the ones with the worst time response. For all-pole filters, the Chebyshev filter gives the best amplitude discrimination, followed by the Butterworth and then the Bessel.

If the time domain response were ranked, the Bessel would be best, followed by the Butterworth and then the Chebyshev. Details of the different filter responses will be discussed in the next section.

The impulse response also increases with increasing filter order. Higher filter order implies greater bandlimiting, therefore degraded time response. Each section of a multistage filter will have its own impulse response, and the total impulse response is the accumulation of the individual responses. The degradation in

the time response can also be related to the fact that as frequency discrimination is increased, the Q of the individual sections tends to increase. The increase in Q increases the overshoot and ringing of the individual sections, which implies longer time response.

Step Response

The step response of a filter is the integral of the impulse response. Many of the generalities that apply to the impulse response also apply to the step response. The slope of the rise time of the step response is equal to the peak response of the impulse. The product of the bandwidth of the filter and the rise time is a constant. Just as the impulse has a function equal to unity, the step response has a function equal to 1/s. Both of these expressions can be normalized, since they are dimensionless.

The step response of a filter is useful in determining the envelope distortion of a modulated signal. The two most important parameters of a filter's step response are the overshoot and ringing. Overshoot should be minimal for good pulse response. Ringing should decay as fast as possible, so as not to interfere with subsequent pulses.

Real life signals typically aren't made up of impulse pulses or steps, so the transient response curves don't give a completely accurate estimation of the output. They are, however, a convenient figure of merit so that the transient responses of the various filter types can be compared on an equal footing.

Since the calculations of the step and impulse response are mathematically intensive, they are most easily performed by computer. Many CAD (Computer Aided Design) software packages have the ability to calculate these responses. Several of these responses are also collected in the next section.

SECTION 5-4 Standard Responses

There are many transfer functions to satisfy the attenuation and/or phase requirements of a particular filter. The one chosen will depend on the particular system. The importance of the frequency domain response versus the time domain response must be determined. Also, both of these considerations might be traded off against filter complexity, and thereby cost.

Butterworth

The Butterworth filter is the best compromise between attenuation and phase response. It has no ripple in the pass band or the stop band, and because of this is sometimes called a maximally flat filter. The Butterworth filter achieves its flatness at the expense of a relatively wide transition region from pass band to stop band, with average transient characteristics.

The normalized poles of the Butterworth filter fall on the unit circle (in the *s* plane). The pole positions are given by:

$$-\sin\frac{(2K-1)\pi}{2n} + j\cos\frac{(2K-1)\pi}{2n} \quad K = 1, 2....n \qquad \text{Eq. 5-29}$$

where *K* is the pole pair number, and n is the number of poles.

The poles are spaced equidistant on the unit circle, which means the angles between the poles are equal.

Given the pole locations, ω_0 , and α (or Q) can be determined. These values can then be used to determine the component values of the filter. The design tables for passive filters use frequency and impedance normalized filters. They are normalized to a frequency of 1 rad/sec and impedance of 1 Ω . These filters can be denormalized to determine actual component values. This allows the comparison of the frequency domain and/or time domain responses of the various filters on equal footing. The Butterworth filter is normalized for a -3 dB response at $\omega_0 = 1$.

The values of the elements of the Butterworth filter are more practical and less critical than many other filter types. The frequency response, group delay, impulse response and step response are shown in Figure 5-15. The pole locations and corresponding ω_0 and α terms are tabulated in Figure 5-26.

Chebyshev

The Chebyshev (or Chevyshev, Tschebychev, Tschebyscheff or Tchevysheff, depending on how you translate from Russian) filter has a smaller transition region than the same-order Butterworth filter, at the expense of ripples in its pass band. This filter gets its name because the Chebyshev filter minimizes the height of the maximum ripple, which is the Chebyshev criterion.

Chebyshev filters have 0 dB relative attenuation at dc. Odd order filters have an attenuation band that extends from 0 dB to the ripple value. Even order filters have a gain equal to the pass band ripple. The number of cycles of ripple in the pass band is equal to the order of the filter.

The poles of the Chebyshev filter can be determined by moving the poles of the Butterworth filter to the right, forming an ellipse. This is accomplished by multiplying the real part of the pole by k_r and the imaginary part by k_I . The values k_r and k_I can be computed by:

$$K_r = \sinh A$$
 Eq. 5-30

$$K_{\rm I} = \cosh A$$
 Eq. 5-31

where:

$$A = \frac{1}{n} \sinh^{-1} \frac{1}{\epsilon}$$
 Eq. 5-32

where n is the filter order and:

$$\varepsilon = \sqrt{10^{R} - 1}$$
 Eq. 5-33

where:

$$R = \frac{R_{dB}}{10}$$
 Eq. 5-34

where:

$$R_{dB}$$
 = pass band ripple in dB Eq. 5-35

The Chebyshev filters are typically normalized so that the edge of the ripple band is at $\omega_0 = 1$. The 3 dB bandwidth is given by:

$$A_{3dB} = \frac{1}{n} \cosh^{-1}\left(\frac{1}{\varepsilon}\right)$$
 Eq. 5-36

This is shown in Table 1.

The frequency response, group delay, impulse response and step response are cataloged in Figures 5-16 to 5-20 on following pages, for various values of pass band ripple (0 .01 dB, 0.1 dB, 0.25 dB, 0.5 dB and 1 dB). The pole locations and corresponding ω_0 and α terms for these values of ripple are tabulated in Figures 5-27 to 5-31 on following pages.

ORDER	0.01dB	0.1dB	0.25dB	0.5dB	1dB
2	3.30362	1.93432	1.59814	1.38974	1.21763
3	1.87718	1.38899	1.25289	1.16749	1.09487
4	1.46690	1.21310	1.13977	1.09310	1.05300
5	1.29122	1.13472	1.08872	1.05926	1.03381
6	1.19941	1.09293	1.06134	1.04103	1.02344
7	1.14527	1.06800	1.04495	1.03009	1.01721
8	1.11061	1.05193	1.03435	1.02301	1.01316
9	1.08706	1.04095	1.02711	1.01817	1.01040
10	1.07033	1.03313	1.02194	1.01471	1.00842

Table 1: 3dB bandwidth to ripple bandwidth for Chebyshev filters

Bessel

Butterworth filters have fairly good amplitude and transient behavior. The Chebyshev filters improve on the amplitude response at the expense of transient behavior. The Bessel filter is optimized to obtain better transient response due to a linear phase (i.e., constant delay) in the pass band. This means that there will be relatively poorer frequency response (less amplitude discrimination).

The poles of the Bessel filter can be determined by locating all of the poles on a circle and separating their imaginary parts by:

2	Eq. 5-37
n	19.007

where *n* is the number of poles. Note that the top and bottom poles are distanced by where the circle crosses the j ω axis by:

1	Eq. 5-38
n	24.000

or half the distance between the other poles.

The frequency response, group delay, impulse response and step response for the Bessel filters are cataloged in Figure 5-21. The pole locations and corresponding ω_0 and α terms for the Bessel filter are tabulated in Figure 5-32.

Linear Phase with Equiripple Error

The linear phase filter offers linear phase response in the pass band, over a wider range than the Bessel, and superior attenuation far from cutoff. This is accomplished by letting the phase response have ripples, similar to the amplitude ripples of the Chebyshev. As the ripple is increased, the region of constant delay extends further into the stop band. This will also cause the group delay to develop ripples, since it is the derivative of the phase response. The step response will show slightly more overshoot than the Bessel and the impulse response will show a bit more ringing.

It is difficult to compute the pole locations of a linear phase filter. Pole locations are taken from the Williams book (see Reference 2), which, in turn, comes from the Zverev book (see Reference 1).

The frequency response, group delay, impulse response and step response for linear phase filters of 0.05° ripple and 0.5° ripple are given in Figures 5-22 and 5-23. The pole locations and corresponding ω_{0} and α terms are tabulated in Figures 5-33 and 5-34.

Transitional Filters

A transitional filter is a compromise between a Gaussian filter, which is similar to a Bessel, and the Chebyshev. A transitional filter has nearly linear phase shift and smooth, monotonic rolloff in the pass band. Above the pass band there is a break point beyond which the attenuation increases dramatically compared to the Bessel, and especially at higher values of n.

Two transition filters have been tabulated. These are the Gaussian to 6 dB and Gaussian to 12 dB.

The Gaussian to 6 dB filter has better transient response than the Butterworth in the pass band. Beyond the breakpoint, which occurs at $\omega = 1.5$, the roll-off is similar to the Butterworth.

The Gaussian to 12 dB filter's transient response is much better than Butterworth in the pass band. Beyond the 12 dB breakpoint, which occurs at $\omega = 2$, the attenuation is less than the Butterworth.

As is the case with the linear phase filters, pole locations for transitional filters do not have a closed form method for computation. Again, pole locations are taken from Williams's book (see Reference 2). These were derived from iterative techniques.

The frequency response, group delay, impulse response and step response for Gaussian to 12 dB and 6 dB are shown in Figures 5-24 and 5-25. The pole locations and corresponding ω_0 and α terms are tabulated in Figures 5-35 and 5-36.

Comparison of All-Pole Responses

The responses of several all-pole filters, namely the Bessel, Butterworth and Chebyshev (in this case of 0.5 dB ripple) will now be compared. An 8-pole filter is used as the basis for the comparison. The responses have been normalized for a cutoff of 1 Hz. Comparing Figures 5-13 and 5-14, it is easy to see the trade-offs in the response types. Moving from Bessel through Butterworth to Chebyshev, notice that the amplitude discrimination improves as the transient behavior gets progressively poorer.



of Bessel, Butterworth and Chebyshev filters



Figure 5-14: Comparison of Step and Impulse Responses of Bessel, Butterworth and Chebyshev

Elliptical

The previously mentioned filters are all-pole designs, which mean that the zeros of the transfer function (roots of the numerator) are at one of the two extremes of the frequency range (0 or ∞). For a low-pass filter, the zeros are at f = ∞ . If finite frequency transfer function zeros are added to poles an Elliptical filter (sometimes referred to as Cauer filters) is created. This filter has a shorter transition region than the Cheby-shev filter because it allows ripple in both the stop band and pass band. It is the addition of zeros in the stop

band that causes ripple in the stop band but gives a much higher rate of attenuation, the most possible for a given number of poles. There will be some "bounceback" of the stop band response between the zeros. This is the stop band ripple. The Elliptical filter also has degraded time domain response.

Since the poles of an elliptical filter are on an ellipse, the time response of the filter resembles that of the Chebyshev.

An elliptical filter is defined by the parameters shown in Figure 5-2, those being A_{max} , the maximum ripple in the pass band, A_{min} , the minimum attenuation in the stop band, F_c , the cutoff frequency, which is where the frequency response leaves the pass band ripple and F_s , the stop band frequency, where the value of A_{max} is reached.

An alternate approach is to define a filter order n, the modulation angle, θ , which defines the rate of attenuation in the transition band, where:

$$\theta = \sin^{-1} \frac{1}{F_s}$$
 Eq. 5-39

and ρ which determines the pass band ripple, where:

$$\rho = \sqrt{\frac{\epsilon^2}{1 + \epsilon^2}}$$
 Eq. 5-40

where ε is the ripple factor developed for the Chebyshev response, and the pass band ripple is:

$$R_{dB} = -10 \log(1 - \rho^2)$$
 Eq. 5-41

Some general observations can be made. For a given filter order n, and θ , A_{min} increases as the ripple is made larger. Also, as θ approaches 90°, F_s approaches F_c . This results in extremely short transition region, which means sharp roll-off. This comes at the expense of lower A_{min} .

As a side note, ρ determines the input resistance of a passive elliptical filter, which can then be related to the VSWR (Voltage Standing Wave Ratio).

Because of the number of variables in the design of an elliptic filter, it is difficult to provide the type of tables provided for the previous filter types. Several CAD (Computer Aided Design) packages can provide the design values. Alternatively several sources, such as Williams's (see Reference 2), provide tabulated filter values. These tables classify the filter by

$$C n \rho \theta$$

where the *C* denotes Cauer. Elliptical filters are sometime referred to as Cauer filters after the network theorist Wilhelm Cauer.

Maximally Flat Delay With Chebyshev Stop band

Bessel type (Bessel, linear phase with equiripple error and transitional) filters give excellent transient behavior, but less than ideal frequency discrimination. Elliptical filters give better frequency discrimination, but degraded transient response. A maximally flat delay with Chebyshev stop band filter takes a Bessel type function and adds transmission zeros. The constant delay properties of the Bessel type filter in the pass band are maintained, and the stop band attenuation is significantly improved. The step response exhibits no overshoot or ringing, and the impulse response is clean, with essentially no oscillatory behavior. Constant group delay properties extend well into the stop band for increasing n.

As with the elliptical filter, numeric evaluation is difficult. Williams's book (see Reference 2) tabulates passive prototypes normalized component values.

Inverse Chebyshev

The Chebyshev response has ripple in the pass band and a monotonic stop band. The inverse Chebyshev response can be defined that has a monotonic pass band and ripple in the stop band. The inverse Chebyshev has better pass band performance than even the Butterworth. It is also better than the Chebyshev, except very near the cutoff frequency. In the transition band, the inverse Chebyshev has the steepest roll-off. Therefore, the inverse Chebyshev will meet the A_{min} specification at the lowest frequency of the three. In the stop band there will, however, be response lobes which have a magnitude of:

$$\frac{\varepsilon^2}{(1-\varepsilon)}$$
 Eq. 5-42

where ε is the ripple factor defined for the Chebyshev case. This means that deep into the stop band, both the Butterworth and Chebyshev will have better attenuation, since they are monotonic in the stop band. In terms of transient performance, the inverse Chebyshev lies midway between the Butterworth and the Chebyshev.

The inverse Chebyshev response can be generated in three steps. First take a Chebyshev low-pass filter. Then subtract this response from 1. Finally, invert in frequency by replacing ω with 1/ ω .

These are by no means all the possible transfer functions, but they do represent the most common.

Using the Prototype Response Curves

In the following pages, the response curves and the design tables for several of the low-pass prototypes of the all-pole responses will be cataloged. All the curves are normalized to a -3 dB cutoff frequency of 1 Hz. This allows direct comparison of the various responses. In all cases the amplitude response for the 2- through 10-pole cases for the frequency range of 0.1 Hz. to 10 Hz. will be shown. Then a detail of the amplitude response in the 0.1 Hz to 2 Hz. pass band will be shown. The group delay from 0.1 Hz to 10 Hz and the impulse response and step response from 0 seconds to 5 seconds will also be shown.

To use these curves to determine the response of real life filters, they must be denormalized. In the case of the amplitude responses, this is accomplished simply by multiplying the frequency axis by the desired cutoff frequency F_c . To denormalize the group delay curves, divide the delay axis by $2\pi F_c$, and multiply the frequency axis by F_c , as before. Denormalize the step response by dividing the time axis by $2\pi F_c$. Denormalize the impulse response by dividing the time axis by $2\pi F_c$.

For a high-pass filter, simply invert the frequency axis for the amplitude response. In transforming a low-pass filter into a high-pass (or bandreject) filter, the transient behavior is not preserved. Zverev (see Reference 1) provides a computational method for calculating these responses.

In transforming a low-pass into a narrowband band-pass, the 0 Hz axis is moved to the center frequency F_0 . It stands to reason that the response of the band-pass case around the center frequency would then match the low-pass response around 0 Hz. The frequency response curve of a low-pass filter actually mirrors itself around 0 Hz, although negative frequency generally is not a concern.

To denormalize the group delay curve for a band-pass filter, divide the delay axis by π BW, where BW is the 3 dB bandwidth in Hz. Then multiply the frequency axis by BW/2. In general, the delay of the band-pass filter at F₀ will be twice the delay of the low-pass prototype with the same bandwidth at 0Hz. This is due to the fact that the low-pass to band-pass transformation results in a filter with order 2n, even though it is typically referred to it as having the same order as the low-pass filter from which it is derived. This approximation holds for narrowband filters. As the bandwidth of the filter is increased, some distortion of the curve occurs. The delay becomes less symmetrical, peaking below F₀.

The envelope of the response of a band-pass filter resembles the step response of the low-pass prototype. More exactly, it is almost identical to the step response of a low-pass filter having half the bandwidth. To determine the envelope response of the band-pass filter, divide the time axis of the step response of the low-pass prototype by πBW , where BW is the 3 dB bandwidth. The previous discussions of overshoot, ringing, and so forth can now be applied to the carrier envelope.

The envelope of the response of a narrowband band-pass filter to a short burst of carrier (that is where the burst width is much less than the rise time of the denormalized step response of the band-pass filter) can be determined by denormalizing the impulse response of the low pass prototype. To do this, multiply the amplitude axis and divide the time axis by π BW, where BW is the 3 dB bandwidth. It is assumed that the carrier frequency is high enough so that many cycles occur during the burst interval.

While the group delay, step, and impulse curves cannot be used directly to predict the distortion to the waveform caused by the filter, they are a useful figure of merit when used to compare filters.



Figure 5-15: Butterworth response



Figure 5-16: 0.01 dB Chebyshev response



Figure 5-17: 0.1 dB Chebyshev response



Figure 5-18: 0.25 dB Chebyshev response



Figure 5-19: 0.5 dB Chebyshev response



Figure 5-20: 1 dB Chebyshev response



Figure 5-21: Bessel response



Figure 5-22: Linear phase response with equiripple error of 0.05°



Figure 5-23: Linear phase response with equiripple error of 0.5°



Figure 5-24: Gaussian to 12 dB response



Figure 5-25: Gaussian to 6 dB response

Analog Filters

1000		REAL	IMAGINARY	S			-3 dB	PEAKING	PEAKING
ORDER	SECTION	PART	PART	Fo	α	Q	FREQUENCY	FREQUENCY	LEVEL
2	1	C.707*	0.7071	1.000	1.4142	0,7071	.5000		
з	1.1	0.5000	0.8650	1.0000	1.0000	1,3005	-	0.7071	1 2493
	2	1.0000		1.0000			1.0000		
4	1	0.9239	0.3827	1.0000	1.3473	0.5412	0.7195		
	2	0.3327	0.9239	1.0000	0.7654	1.3065		0 6409	3 01 32
5	1	0.8090	0.5878	1.0000	1.6180	0.5185	0.6686		
	2	1.3.90	0.3511	1.00.00	0.6180	1 5182		0.8906	4.6163
	.E.	1.0000		1.0000			0000.1		
6	ī	0.9359	0.2588	1.0000	1.9519	0.5176	0.6758		
	2	0.7071	0.7071	1.0000	1.4142	0.7071	1.0000		
	Э	U 2588	0.9650	1.00.000	0.5173	1, 13" 9	1000	0.5306	6 G210
7	1	0.9010	0.4335	1.0200	1.0019	0.5550	0.7449		
	2	0.6235	0,7818	1.0000	1.2470	0.3013	10001000	0.4717	0 2234
	a	0.2225	0.9749	1.0000	0.<453	2.2471		0.9492	7 2530
	4	1.0.00		1.0L00			1.000		
8	1	0.9008	0 1951	1.0000	1.9016	0.5095	0.6615		
	2	0.8315	0.5556	1.0000	1.5629	0.5015	0.6296		
	3	0.5556	0.8315	1.0000	1.1112	3006.0		0.6186	0 6876
	4	U1.J5*	8086.0	1.0000	0.3902	2,5625		0.9612	8 3429
9	1	0.9397	0 3420	1.000	1.0794	0.5321	0.7020		
	2	0.7560	0.8428	1.0000	1.5520	0.6627	0.9172		
	3	0.5000	0.6660	1.0000	1.0000	1.3000		0.7071	1 2498
	4	L1737	0.9848	1.0000	U.34/4	2.5785		U 5604	U 3*35
	E	1.0300		1.0000			0000		
10	T I	0.9377	0.1664	1.0000	1.9754	0.5062	0.6649		
	2	0.8310	0.4540	1.0000	1.7820	0.5612	0.7564		
	з	11/17	0.7021	1.0L00	1.4142	0,7071	1.0000	- Sector - 1	
	1	0.4540	0.8910	1.0000	0.9080	1.1013		0.7687	1.8407
	5	0.1564	0.3072	1.0000	0.3123	0.197E		0.5752	10 2020

Figure 5-26: Butterworth design table

1.000	1. State 1.	REAL	IMAGINARY	1.1		1	-3 dB	PEAKING	PEAKING
ORDER	SECTION	PART	PART	Fo	à	Q	FREQUENCY	FREQUENCY	LEVEL
2	1	1,6743	0 7075	0.9774	1.3738	0.7247		0.2172	0.010E
	1	7 4233	0.8565	0.9542	0.3750	1,1385		0.7558	2 0595
	2	1.8467		0.8467			0.8467		
1	1	1.6762	0 3325	0.7772	1.7425	0.5745	0.8069		
1.25	2	° 2001	fi 9241	0.95%	0.5051	1.7232	1.1.1.1.1.1.1.1	T BOUR	511H
E	1	0.5120	0 5979	0.779E	1.3126	0.7813		E.2689	0.0527
	2	0.1966	0.9512	0.9711	0.4028	2.4824	and the second se	0.9309	8.0772
	3	1,6328		0.6525			0,8328		
E	1	0.6335	0 2583	0.6531	1.7995	0.6567	0.4425	-	
	2	0 3906	0 7072	0.8575	0.9670	1 0342		0.5695	1.4482
	Э	2.1430	0 9563	0.9765	0.2929	3.4144		C.9554	10.7605
12	Ť	4346	0.4335	0.6175	1.4229	0.7025	11.5*-55		
1	2	5.3040	0 7813	0.8389	0.7247	1.3795		0.7204	3.4577
	3	C 1005	0.9750	0.9512	0.2212	4 5203		0.9009	15.1570
_	4	1.4876		0.4578		Deliver of	0.4876		0.00
6	1	_ 4268	U 1.51	0.4595	1.stau	U 549E	0.345*		
	2	0.3168	0 5553	0.6595	0.3957	1.0394		0.4561	1.3541
	3	- 2410	21000	0.0555	0.7576	1.790F		0.7956	/5 412F
	4	0.0849	0.9808	0.9545	0.1725	5 7975		0.9771	15.2977
5	I.	1,3686	0.3425	0.5.25	1,4621	0.6521	0.4844		
	2	0.3005	0.6425	0.7095	0.3470	1 1307		0.5682	2.3205
	3	1961	0.0561	ILB B	0.4417	2254		0.94.06	1 1145
	ă i	7.0681	0.9345	0.9572	0.1350	7 247=		0.9824	17:2049
	5	2,3923	0.0012	0 3525	0.1020		0,3923		11,1444.10
10	1	0.3522	0 1564	0.3554	1.5279	0.5471	0.2814		
	2	31/8	4-4	11:5542	1 14-9	1111/15		1039250	115412
	3	2622	0.7071	0.7507	0.2719	1.4584		0.6606	3.9742
	4	1 1619	0.89 ^r	0.9755	0.1576	2,7965		0 8762	9.0742
	5	0.0568	0.9377	0.9595	0.1128	8 8345		0.9861	13,9669

Figure 5-27: 0.01dB Chebyshev design table

ORDER	SECTION	REAL	IMAGINARY PART	For	α	a	-3 dB FREQUENCY	PEAKING	PEAKING
5		0.6104	0.7106	0.9568	*,5032	0.7973		0 3638	0 C909
s		3.3400	7.9354	r 9359	0.7458	13468		6 7952	\$ 1978
	8	0.6970		C 6970		1.0100	0 E9/C	0.000	
đ	×	0237	0.6252	0.9507	0.4580	2 * 52/		0.6997	7 (2137
1.1.1.1.1.1	2	0 6257	0.3833	0.6506	1,5160	0.6158	0.9590		
5	-	3 3842	0.5864	0.7027		0.5145		9 4457	3 7602
	21 - 33 33	0 1438 0 4749	0.952'	0.9839 0.4749	0.5040	8 2612	2 4745	9 \$407	3 4 2 26
3		0.3946	0.2590	0.4695	1 3682	0.5355	3 8875		
	2	0.2867	0.7077	C.7836	0.7509	1 2516		0.6470	5, 14, 8
	3	3 1045	0.9867	0.9794	Ú 2158	4 6348		a setc	13 3714
7	1	0.3176	6.151	0.5360	- 18:14	0.8464		0.9957	34357
	ž	0.2200	0.9623	C.8126	0.5414	1.3469		0.7607	5 6505
	3	3.0736	自由765	0.9787	0.1604	6 2325	100 Aug. 100	0.9722	.5 \$2.76
	- 2	T 3526		T.3528			0.3529		
ŝ	8	0.3058	0 °962	0.3628	1.6358	0.5032	0.295€		
	3	3.2529	T 6568	E 8168	0.8283	1 2073		0.1945	2 1537
	3	0 1752	0.6519	0.8497	0.4077	2.453*		0 \$137	7.6784
		3C80 C	B 9812	C 982 I	U 1237	8 0319		1 8743	S 1688
э		0.2622	0.342	T. 2340	12168	m 8219		0.\$197	3 2037
	2	0.2157	0.6430	C.677S	0,6306	1 5894		0 0004	4.4570
	2	0.1395	0.8363	0.8775	0.5130	3 *460		0.8550	10.0836
	2	0.0485	0.9852	0.9804	0.0982	10.1795		0.984C	20 1650
	0.	0.2790		0.2780			.0.548C		
1ť	-	0.2793	0.156%	C 2943	1 3942	0 = 902	0.2302		
	Ē	0.2249	0.454*	6.5067	0.8870	1 266		0.3945	1 5880
	7	0.1785	11/3/3	E. (285	0,4694	2.0437		0 F847	3 4750
		3 1:46	0.8913	0,8985	0,2551	3 5200		0.6339	1 5386
	0	0.0395	0.9520	0.9028	0.0799	12.5165		J 8872	21.9585

Figure 5-28: 0.1 dB Chebyshev design table

ORDER	SECTION	PART	PART	Fa	a	o	-3 dB	PEAKING	FEAKING LEVEL
2	-	0.5821	0.7154	0.9358	1.2355	0.9695		0.4425	0.2502
3	8	0.3062	0.8712	0.9234	0.3632	1.5079	-	0.8153	4 0734
	¥.	0.8124		0.6194			n 6.12=		
4		0.4601	0.38-1	1.5318	5215	0 3572	11 6470	Ante	SALA.
	2	0.1865	0.0272	0.9458	0.3944	2.5355		0.0082	8.2538
2	1.1	0.5247	0,5892	0.6727	C.0053	.0850		0.46*7	1 4 589
	2	0.1270	0.9533	0.9813	0.2520	3.576%	1	0.9452	11.8612
_	3	0.4016		0,4915			0 4013		
6		0.5284	0.2095	Ċ.4184		0.6371	0.2750		
	2	0.2404	0.7085	C.4430	0.3425	1,555/		0.3665	4/3121
	- 3	0.0680	0,9675	0.9715	6,4814	5.5905	-	0.9637	14.8753
7		0.2652	0.42±4	6.5350	1.3471	0.3598	1	0 34-1	1.5.173
	2	0.1835	0.7825	0.8340	0.4565	2,1905		0.7610	1 0443
	3	0.0655	0.2761	6.9783	0 132,9	7 467.9	1.00	0.9739	17 4835
	4	0.2944		C.2944			0 2944		
.6		0.2545	0,1955	0.3205	1.5062	0.5304	0 2822		
1000	2	0.2158	0.5661	C 5364	0 723.1	3827		0.5128	3 1258
	3	0,1441	0.8325	6.8447	0.6412	2,9309	1	0.8197	0.4683
		0.0608	6 98 °	0880	0 10/23	§ 7173		10.9201	13 7824
ų	~	0.2*73	0.5425	0.4088	2 0730	0.9820		0.2672	0.5854
	2	0.1774	0.5433	0.8673	6.53 7	5086.1		0.5184	5 8052
	3	0.1158	0.6667	C.8744	0.28/9	3,7755		0.5589	11 E133
	÷	0.0402	0.9853	0.9864	0.08 5	2 2659		0.9843	21 7812
	5	0.2315	and the second se	C.2575			0.2315		
10	*	0.2065	0.1565	U.2591	1.5940	0.3274	0.2257		
	2	0.1863	0.4543	0.4910	f, 7583	3173		0.41=3	3.0721
	3.	0.1475	0.7075	C.7228	0.4090	2,4461		0.5919	7.8545
	4	0.094.9	0.35 5	0.3965	0.2417	+ 7238		0 \$864	13 6344
	5	0.0327	0.9283	0.9698	0.0661	151230		0.8376	23.6957

Figure 5-29: 0.25 dB Chebyshev design table

Analog Filters

		REAL	IMAGINARY	-			-3 dB	PEAKING	PEAKING
ORDER	SECTION	PART	PART	Fo	α	Q	FREQUENCY	FREQUENCY	LEVEL
2	1	6.5129	0.7225	-,25 4	- 1377	0.9658		2.7072	0.5002
3	- 4	0.3693	0.5753	.0385	0.9861	1 706	1	0.0737	5 (501
	2	0.5365		0,6265			0.6285		
× .	t	1. 3872	0.2350	0.5.965	185	0.725*	II 535*		
1.000	2	ĉ.1605	0.0297	÷.65°3.	0.3402	2 939	the second second	1,5010	9.4915
5	1	0.2767	0.6502	0,6905	0 8490	1 1779		2,5522	2.2849
1.00%	2	0.1067	U 8550	0.17%	0.1200	4 545*		1,3054	13 2037
	5	0.3427		0.3623	<u> </u>		0.3823		
6	1	0.2784	0.2596	0.3063	* 4327	0.6856	0 3927		
100	, in 1	0.2037	0.7.391	0.7680	0.6522	1 8109		0.7071	5.6.126
	à	0.048	0.9387	1.01/4	0.1536	6.5119		1.0085	16 2893
*	1	0.2241	0,4548	0,2040	0.2161	1.0216		2.5839	* 7938
1	2	0,1550	0.7336	0.8228	0.3381	2 5767	11 21	0.7912	8 3380
	à	U.0553	0.8771	1.0081	0.1130	8 3487	a second second	1.0019	18,9515
	4	0.2487		0.2563	_		0.2562		
8	1	6.2444	0.1950	0.2968	4775	0 6767	0 2635		
1.000	2	0.1812	0 < 565	0 = 389	0.6208	1.6109		2 5581	i dSin
1	3	0.12:4	0.\$528	0.86 0	0.2885	3 4662		0.8430	10.8885
11 7	4	1. 2468	0.6834	* 1(06)2	10 10967	*1.5305	1.1	1 304 1	21 2452
9	1	0 1831	0.3425	0.2964	0.8428	10505		0.2947	602%
1 Cale	4	0.1493	0.0430	0.0727	0 4520	2 2120		0.9374	7 258
	ā	0.6974	0.8371	0.8884	0.9232	4 1779		5.8773	13.0759
	4	0.0338	0.9861	1.0048	0.0360	14.5829	and a second second	1,0034	23 2820
	5	0.1913		0 1984			0 ~ 384		
1.3	1	0.1786	0.1566	0.7338	4.851	0.6734	0.9721		
11.12	2	0.1568	0,4545	0.4807	0.62.5	1 5349		0.4267	4 2087
	.ä.	.0 1218	0.7.178	0.7485	0.8459	2.8807		0.8968	9.2520
1.0	4	0.0790	0.0915	0.0956	0.1782	5.6107		0.0805	15 0149
	5	6.6275	0.9387	0.9391	0.0556	7.3823	+	0.388.5	.25 - 0.08

Figure 5-30: 0.5 dB Chebyshev design table

ORDER	SECTION	REAL	IMAGINARY PART	Fo	a	Q	-3 dB FREQUENCY	PEAKING FREQUENCY	PEAKING
2	1	02565	0.7351	0.8525	1.0465	0.9564		0:2806	0.8889
Ş.	1	0.2257 0.45*3	1.8922	0.9005 0.45-3	C.4957	2.0(75	3 4513	7 8528	3 370 <u>3</u>
.4.	1	C 3193	C 3863	1.5.1.3	1.2746	0.7815		J 9174	0 1557
		0.1325	0.9339	0.8435	0.2003	3.5594		0.9245	1 1142
5	1	0.2265	C.59*6	C.6537	C7/149	1.3986		0.5467	5 5009
	1.2	0.0865 0.2800	0.9675	0.9614 0.2900	C 803	5 5553	9 2900	J \$536	14 8305
S	1	0.2268	C.2601	C.8451	1.3144	0.7605		9 1273	5 (813
1.00	3	6 550	f. 7403	f. 7273	0.4262	7.3462		3.6985	7 6030
	3	00605	0.9707	0.9725	0.1248	\$.0035		0.0688	15 0827
7	Ť.	6.4910	č.4354	C.4715	C.77*0	1,2071		9.3956	2 6576
	17	6. 253	C 7848	C 7318	C 3163	2 1655		0.774	10.027
-	3	0.0449	C.9785	0.9795	8100.0	10 8982		2 9775	20 7563
1000	1	6.30.3		6,2013			3 2319		
	1	6. 727	6, 1855	E 2813	1 3287	E 763.1		31.699	3.081
	2	6.1473	Ć.5574	0.5762	C.5112	1,9560		0.6373	\$ 1210
	8	1360.0	C 8237	C 8395	0.25(4	4 2657		0.8220	12 6599
-	4	0.0545	0.9830	C.9842	0.0702	14 2391	4	0 \$830	23 0750
Ð	t	0.1482	6.5427	0.3734	C.7038	1.2597		0 3090	2 7408
	9	0.1208	0.847.2	0.6554	C 3888	\$ (129		3 6398	8.6187
	3	0.0788	0.8070	C.8715	0.1809	5.5268		0.8643	4 8852
	1	10274	119869	0.8875	E 0555	18 13275		0.48885	95,1197
	5	Ç.+ 577		Č:1577	1.1.1.200	1.22	2 - 577		10. 1 Ch. 1
-10	1	0.1403	0.1567	0.2105	1.3341	0.7495		10308	0 0050
	2	C. 1263	C.45/5	C 5721	0.5563	1.8625		0.4368	5 7357
-	3	01:005	0.7084	0,7153	0.2809	3,5597		3 7012	11:47
	4	0.03/5	C.8823	E.8979	¢.34/1	E.9374		D 5805	× 8 54 BE
	- 4	0.0222	0.9895	C,9897	0.0442	22,2910		0 5893	29 5650

Figure 5-31: 1 dB Chebyshev design table

÷		REAL	IMAGINARY	1.000		1.1	-3 dB	PEAKING	PEAKING
ORDER	SECTION	PART	PART	Fo	a	Q	FREQUENCY	FREQUENCY	LEVEL
a	3	1 1050	0.6368	1 2754	1 72.28	0.5771	1020		
3	+	1.0509	1.0025	1 4524	1 4471	6.65.16	4 85		
	4	1,5270		1.3270			1.3270		
4		1.35.96	0.40/1	1.4192	1.9130	0.5215	0.9705		
	2	6.3877	1 2476	1.5512	1 24 14	6 8655		3 7822	0.2849
6	1	1.885	0.7261	1.5011	1 7745	0.5636	1876		
-	ź	0.9606	14753	1.7607	1.0211	0.9165	-	1.1204	15.4738
	3	1.5039		1.5065	_		1.5009		
Ŀ	1	1 5735	0 8215	1.6060	1 9596	0.5102	1.3638		
In setting 4	4	1.5856	0.972×	1.6913	1.6331	C.8112	* 4323		
-	3	0.0318	1 6640	1.9071	0.9773	1.0254		1 3785	1.5557
7	d.	1.9150	0.5856	1.7174	1.8784	0.5324	.2074		
	2	1.3797	7 1925	1.825	1.5132	E.BEDS	1927		
 A second s	3	0.2104	1.8375	2.0507	0.3879	1,1262		1 5961	1.2860
	1	1 3853		1.6853			1 5853		
đ	1	1 (627	0.272×	1 1832	1.9732	0.5060	1675		
1.11.1		0.8955	2 0044	2,4953	0.8155	1.2258		1 7932	3,5585
	2	1.3750	1.3978	1.9691	1/1037	1.7108		0.29/1	0.0006
	4	1,6419	0.8256	1.8378	1,7858	0.5907	1.3849		
ġ	- 11	1.8081	0.6126	1.8794	1.9342	0.5197	.0774		
1.1	S.	1 3532	0.0319	1.9188	1.8936	165631	5747		
	3	1.3688	1 5685	2.0815	1.3148	C.7606		0.7668	0.0807
	10. I. I.	1. 8782	2 1509	7 3736	0.7681	1.8220		1.9852	21920.5
	5	1.8675		1.8575	_		1.8575		
10	1.1	1.9335	0.2451	1.9490	1.9841	0.5040	.2665		
	.9	1 3137	0 7325	1.9876	1 8687	E 538E	4 77		
	2	1.36'5	1 2246	2.0678	1.6115	C.620E	1.7848		
	3	1.3648	1 7385	2 21 15	1 2846	6,8100	1	1.0785	6.263
S	5	0,5686	.2.2954	2.4580	0.7037	1.4150		2 1291	3.5944

Figure 5-32: Bessel design table

ORDER	SECTION	REAL	IMAGINARY PART	Fo	a	Q	-3 dB FREQUENCY	PEAKING FREQUENCY	PEAKING
2	1	1 0037	0 0580	1 2005	1,6676	0.5067	0 9909		
3	Ŷ.	3.854	1 0725	1 3710	1.2459	0.8026		.j 6497	0.2232
	2	1.0458		1.0458			1 0459		
a.	1	0.9549	U ≤ /49	1053	1.7945	03513	0.0056		
1.2	2	n 7048	1 =008	1 5865	0.8385	1.0350		18.64	- 5283
ê	1	ates c	R 8738	1.2480	1 287	C 6999	1:5357		
	2	0.0781	17085	1 8303	0.753	1.3541		1 5703	3.3334
	3	0.8430		9.6430					
Ê	1	0.8904	0 2112	9 2397	1.8158	0.8667	1.7228		_
	2	0.8253	1 2175	1 1701	1:201	0.0928		.).8575	0.6495
	- 5	U 6°52	1.8810	21743	U.6932	1,6655.9		1.5331	9.9365
2	1	0.9426	0.779	1 (475	1 # 304	0.6810	1 1036		
	7	9 7798	1 5351	17177	0.8976	1 113		1.3576	3162
	3	0.5727	2.2456	2 3175	0.4042	2.0233		2.1715	6.3048
	ন	0.881s		0.6816			:185 U		
2	1	0.2192	0.271*	9.5986	1.8218	0.5429	9.6690		
	2	0.7930	1:354	1 2604	1 058	0.8578	10000	0.7701	0.4705
	5	0.7213	1.2134	1.9516	0.7392	1.3528		1.3600	3.2627
i and	4	0.5341	2 170	2 5336	£ ±217	2 3713		24-78	7.6973
ε.	1	A 7852	11 7 126	1.6834	1 - 317	C 875	1.0107		
	2	0.7555	1 4127	1 6020	0.0432	1.0362		1.1937	1.5005
	3	0.6849	20,851	2 195C	0.6341	1 6020		1,96,97	5404
	4	0.5000	2 7 183	2 760/	0.8007	2,7274		.2 6657	8.8035
	5	0.7968	and the second	0 79BE			0.7962		
10	1	0 7592	U.9410	5 8324	1.824 *	0.5492	n ewae		
	2	0.7467	1 0 185	1 7657	1 318	0.8462		3 8941	0.41-5
	3	07158	1.6836	1.0295	0.7626	1.2778		1.5230	2,8507
	4	0.6476	2,8498	2 2086	0 5377	1.8558		2.2276	5 7162
	5	0.4777	2 9129	2 9517	0.3237	3.0005		2.6714	0516.9

Figure 5-33: Linear phase with equiripple error of $0.05^{\,\circ}$ design table

Analog Filters

and a second		REAL	IMAGINARY				-3 dB	PEAKING	PEAKING
ORDER	SECTION	PART	PART	Fo	a	Q	FREQUENCY	FREQUENCY	LEVEL
2	1	0.5286	ົມ ເຕັອກູ	1.1088	1,652*	C.6443	1 0000	- 7	
3	- 1	0 6 965	0.1518	1.3292	1.0406	0.9536		0.6210	0.9023
	2	0 8 257		0.8257		-	0.8257		
x.	0	0 7448	0.5132	0.5045	1.6468	0.8372	.) 7537		
_	2	0.6037	1 4 9 8 3	1 6154	0.7475	1.3579		1.5718	3.1817
6	1	0 0779	0.540	1 1588	1.1993	0.8552		0.0918	0.4570
	2	0 =/113	1 3256	1 2011	E 8 307	1.7582		1 7/36	5.2730
	â	0 7056	1	0 7056		0.010	3 7056		
Ê	- 1	0.6515	ñ 4374	A 7,850	1.6308	6.637	0.6527		
	2	Ú€167	1 2963	1.4355	0.8592	1,1539		1.1402	2.2642
	3	0.4898	7 0.382	2:545	0.4547	7 2016		2.9404	7.084.5
2	- 1	0.6180	0.8338	1.0385	1 325	E 8328		1 55.56	E 3793
	2	0.6816	1.0455	1 7453	0.6005	1.5964		1.5393	4,0355
	3	0.1592	2 2997	2 143*	1.276	26567		2 3649	8 3423
	4	0.6283		0.0283			0.0288		
8	4	0.579	0.3357	0.6958	1.6546	6.6007	0.0764		
-	9	0.5.665	1,1505	1.2824	1,8036	1.*\19		1,0014	2.3187
	3	0.5302	1 6.314	1 \$643	0.6399	1.857		1.8155	5 381 A
	4	0 4 1 4 8	2 5700	2 €112	0.3177	3.1475		2 5444	10.0703
	1	15685	0.7505	0.9489	1.1980	C 8541		0.5535	0.3581
	2	0.6546	1.5089	1.6076	E ESSE	1 4488		14038	8 77/8
	8	0 5176	2 2329	2 2922	0.4519	2.2130		2 1720	7,1270
	A	U.S.DUC	2.6028	2.6312	0.2702	3,5923		2.6740	11 1925
	2	0 9728		0.9728			3.9738		
10	1	0.0245	0.3487	0.6302	1.6355	0.6003	3 5215		
	2	U = 192	1 0428	1.1650	0.8916	1.1213		0.9044	1.9599
	5	1.505*	17261	1.7.988	6.5816	1 7308		1.8604	5 3681
	4	02711	2 3650	2431	0.3876	2,5902		2 5380	8.5984
	5	3073.0	2 \$ 340	8-0.189	0.5458	1.038	-	2,9709	12 2589

Figure 5-34: Linear phase with equiripple error of 0.5° design table

ORDER	SECTION	REAL PART	IMAGINARY PART	Fo	a	Q	-3 dB FREQUENCY	PEAKING FREQUENCY	PEAKING
2	0.9360	1	0.0360			0.0300			
	2	0.8270	- 6995	1 3.463	0.2553	1.0435		1 - 29,89	15005
	2	1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1	1:5560	0713	1.7112	0.58/9	0.8582		
-	-	1.5.176	F 0173	- 282.2	1 2586	7 702.7		E (5863	11.24
~	3	0.7158	(2153	11 745.7	1 6 7 24	3 5902	0.5005	0.0000	9. PH
	3	0.0151	CTT DO C	0.613	Desar	0.0412	0.8131		
3	ā	E 7015	E ZARS	0.823	1.7030	0.5879	0.8827		
	2	0.6337	733	4549	0.5.35	1.25.1		080	1.7305
	3	C.4475	2,1363	2,1827	0.4104	2.4365		2.0888	7.9227
Ť	~	C 2456	0.7703	n sten	1.3465	10011		T #332	0.0160
	0	C 54GP	1 8/187	1 31 15	R FR'DA	1.4F83		1 8 L/F	3.974-
	3	0.990e	3 2486	2 162	0.2880	3 : 218		1980	11 8 189
	÷	C.6291	1	0.325			0.825		
			I Personal II			1.00			
3	-	0.5441	0.3558	0.6354	1.7020	3,5876	0.5145		
	2	£,£17£	0,9962	1226	0.8220	1 1845		0.8512	1 7432
	3	117 328	* 610C	3672	0.6198	1.9260		15502	5 8952
	(<u>t</u>	81613	2 0703	2.0/87	0.1303	5,2571		÷ 0808	14.4545
3	(8)	0.4901	0.6192	0.7934	1.2505	0.7997		0.3700	0.2146
	2	0.4598	.2145	1.2076	0.7041	1.4208		.1253	3.6221
	3	C.3592	1,7428	* 7725	07057	2.4771		17955	2.0594
	10	0.7488	80C* 5	2.1056	11.1314	7.0704	and the second sec	2.0960	17 0107
	6	280e ()	1	0,5065			0.5065		
710.		0.4589	0.2794	0.5527	1.7028	3 5873	0.4283		
	2	6.4352	0.8289	0.9302	0.9297	1.0756	0000	0.7065	1.6004
	3	1.3996	1 3448	1,5928	0 6552	1.9011		2674	5,4591
	1	10.29008	1997	1.580/2	91,222,18	3:1024		1.7598	S 9318
	5	C.* 136	⊴ 0595	2.0620	0.*191	2080.6		2 0568	15:1751

Figure 5-35: Gaussian to 12 dB design table

ORDER	SECTION	REAL PART	IMAGINARY PART	Fo	a	Q	-3 dB FREQUENCY	PEAKING FREQUENCY	PEAKING LEVEL
6	1	0.9622	1.014	1.5548	1.25//	0.9630		0.75.25	0.2440
	2	6.9778	0 5629	1 0994	1.1785	0.5623	0.8888		
1	4	0.70-1	36236	3,63,36	1 6206	0.5640	0.26-25		
7	2	0.5304	1,5407	1 6647	0.7574	1.3203	0/1692	.46.56	3,3850
	6	0.3195	0.0254	1.031/	1 1929	0.6594		0.5460	0.5540
	9	1.3553	1.5658	1.605/	0.34/35	2.9600		1 5979	7.5004
	3	0.8650		0.665C			0.8650	P	, 19.000
							1000		
6	1	0.5435	0.8451	9.6426	1.6310	0.5914	0.5215	a desi	
	2	0.4672	0 9891	1 1028	C.8472	1.1234		0.6834	2.2892
	3	0.3304	1.503/	1522/	0.2885	3.45/15		14205	10 55.96
7	1	0.4580	3,5952	37431	1.2223	0.8132		0.3770	0,2874
	1	0.5640	1.1230	1 1861	0.6163	1.6255		1,0080	e,6906
	3	0.1622	1.4938	1.5015	C.2027	4,0328		4860	13,9067
	4	0.4828		0.4829			0.4825		
8	T.	6.4222	0.2640	0.4975	1.6358	0.5897	0.4628		
	2	0.3833	3.7716	J 8616	0.8858	1.1239		0.6697	.3722
	5	0.2678	1 20 60	1 2360	C.4388	2.5076		1.1765	7.4721
	4	0.1122	1.4795	1.4840	C.1512	6.6134		*.4755	16.4334
u.	1	0.5703	04704	0.5985	1.2365	O SUSE		0.0%05	0.2480
	Ľ	0.3230	S508.0	0.9626	0.6711	1.4901		0.8473	3,9831
	3	0.2303	1 2634	1 2843	0.3556	2.7811		.2421	9.0271
	4	0.0860	1.4740	1.4765	0.105	8.55.04		.4715	18,6540
	5	0.5942		0 3842			0.3842		
10	1	6 3389	0.2591	0.3952	1 6991	0.5485	0.3212		
	12	0.3164	0.6180	0.6943	0.9114	1.0872		0.5209	18184
	.ă i	0.2677	0.9852	1 0236	0.5244	1.9038		0.2481	5.2057
	4	0.1840	1 2745	1 2878	0.2371	3.4825			10.0284
	5	0.0674	1.4385	1 4405	0.0031	10.7401		1.437 <u>5</u>	20.3299

Figure 5-36: Gaussian to 6 dB design table

SECTION 5-5 Frequency Transformations

Until now, only filters using the low pass configuration have been examined. In this section, transforming the low-pass prototype into the other configurations: high pass, band pass, bandreject (notch) and all pass will be discussed.

Low Pass to High Pass

The low-pass prototype is converted to high-pass filter by scaling by 1/s in the transfer function. In practice, this amounts to capacitors becoming inductors with a value 1/C, and inductors becoming capacitors with a value of 1/L for passive designs. For active designs, resistors become capacitors with a value of 1/R, and capacitors become resistors with a value of 1/C. This applies only to frequency setting resistor, not those only used to set gain.

Another way to look at the transformation is to investigate the transformation in the s plane. The complex pole pairs of the low-pass prototype are made up of a real part, α , and an imaginary part, β . The normalized high-pass poles are then given by:

$$\alpha_{\rm HP} = \frac{\alpha}{\alpha^2 + \beta^2}$$
 Eq. 5-43

and:

A simple pole, α_0 , is transformed to:

$$\alpha_{\omega,\mathrm{HP}} = \frac{1}{\alpha_0}$$
 Eq. 5-45

Low-pass zeros, $\omega_{z,lp}$, are transformed by:

$$\omega_{\rm Z,HP} = \frac{1}{\omega_{\rm Z,LP}}$$
 Eq. 5-46

In addition, a number of zeros equal to the number of poles are added at the origin. After the normalized low pass prototype poles and zeros are converted to high pass, they are then denormalized in the same way as the low pass, that is, by frequency and impedance.

As an example a 3-pole 1dB Chebyshev low-pass filter will be converted to a high-pass filter.

From the design tables of the last section:

	$\alpha_{LP1} = 0.2257$
	$\beta_{LP1} = 0.8822$
	$\alpha_{LP2} = 0.4513$
This will transform to:	
	$\alpha_{\rm HP1} = 0.2722$
	$\beta_{\rm HP1} = 1.0639$
	$\alpha_{\rm HP2} = 2.2158$
Which then becomes:	
	$F_{01} = 1.0982$
	$\alpha = 0.4958$
	Q = 2.0173
	$F_{02} = 2.2158$

A worked out example of this transformation will appear in a later section.

A high-pass filter can be considered to be a low-pass filter turned on its side. Instead of a flat response at dc, there is a rising response of $n \times (20 \text{ dB/decade})$, due to the zeros at the origin, where n is the number of poles. At the corner frequency a response of $n \times (-20 \text{ dB/decade})$ due to the poles is added to the above rising response. This results in a flat response beyond the corner frequency.

.

Low Pass to Band Pass

Transformation to the band pass response is a little more complicated. Band-pass filters can be classified as either wideband or narrowband, depending on the separation of the poles. If the corner frequencies of the band pass are widely separated (by more than two octaves), the filter is wideband and is made up of separate low pass and high pass sections, which will be cascaded. The assumption made is that with the widely separated poles, interaction between them is minimal. This condition does not hold in the case of a narrowband band-pass filter, where the separation is less than two octaves. We will be covering the narrowband case in this discussion.

As in the high pass transformation, start with the complex pole pairs of the low-pass prototype, α and β . The pole pairs are known to be complex conjugates. This implies symmetry around dc (0 Hz.). The process of transformation to the band pass case is one of mirroring the response around dc of the low-pass prototype to the same response around the new center frequency F₀.

This clearly implies that the number of poles and zeros is doubled when the band pass transformation is done. As in the low pass case, the poles and zeros below the real axis are ignored. So an n^{th} order low-pass prototype transforms into an n^{th} order band pass, even though the filter order will be 2 n. An n^{th} order band-pass filter will consist of n sections, versus n/2 sections for the low-pass prototype. It may be convenient to think of the response as n poles up and n poles down.

The value of Q_{BP} is determined by:

$$Q_{BP} = \frac{F_0}{BW}$$
 Eq. 5-47

where BW is the bandwidth at some level, typically -3 dB.

A transformation algorithm was defined by Geffe (Reference 16) for converting low-pass poles into equivalent band-pass poles.

Given the pole locations of the low-pass prototype:

$$-\alpha \pm j\beta$$
 Eq. 5-48

and the values of F_0 and Q_{BP} , the following calculations will result in two sets of values for Q and frequencies, F_H and F_L , which define a pair of band-pass filter sections.

$$C = \alpha^2 + \beta^2 \qquad \text{Eq. 5-49}$$

$$D = \frac{2\alpha}{Q_{BP}}$$
 Eq. 5-50

$$E = \frac{C}{Q_{BP}^{2}} + 4$$
 Eq. 5-51

$$G = \sqrt{E^2 - 4D^2} \qquad \qquad Eq. 5-52$$

$$Q = \sqrt{\frac{E+G}{2D^2}}$$
 Eq. 5-53

Observe that the Q of each section will be the same.

The pole frequencies are determined by:

$$M = \frac{\alpha Q}{Q_{BP}}$$
 Eq. 5-54

$$W = M + \sqrt{M2 - 1}$$
 Eq. 5-55

$$F_{BP1} = \frac{F_0}{W}$$
 Eq. 5-56

$$\mathbf{F}_{\mathbf{BP2}} = \mathbf{W} \mathbf{F}_0$$
 Eq. 5-57

Each pole pair transformation will also result in two zeros that will be located at the origin.

A normalized low-pass real pole with a magnitude of α_0 is transformed into a band pass section where:

$$Q = \frac{Q_{BP}}{\alpha_0}$$
 Eq. 5-58

and the frequency is F_0 .

Each single pole transformation will also result in a zero at the origin.

Elliptical function low-pass prototypes contain zeros as well as poles. In transforming the filter the zeros must be transformed as well. Given the low-pass zeros at $\pm j\omega_z$, the band pass zeros are obtained as follows:

$$M = \frac{\alpha Q}{Q_{BP}}$$
 Eq. 5-59

$$W = M + \sqrt{M2 - 1}$$
 Eq. 5-60

$$F_{BP1} = \frac{F_0}{W}$$
 Eq. 5-61

$$F_{BP2} = W F_0$$
 Eq. 5-62

Since the gain of a band-pass filter peaks at F_{BP} instead of F_0 , an adjustment in the amplitude function is required to normalize the response of the aggregate filter. The gain of the individual filter section is given by:

$$A_{R} = A_{0}\sqrt{1 + Q^{2} \left(\frac{F_{0}}{F_{BP}} - \frac{F_{BP}}{F_{0}}\right)^{2}}$$
Eq. 5-63

Where:

 A_0 = gain a filter center frequency A_R = filter section gain at resonance F_0 = filter center frequency F_{BP} = filter section resonant frequency.

Again using a 3 pole 1dB Chebychev as an example:

$$\alpha_{LP1} = 0.2257$$

 $\beta_{LP1} = 0.8822$
 $\alpha_{LP2} = 0.4513$

A 3 dB bandwidth of 0.5 Hz. with a center frequency of 1 Hz. is arbitrarily assigned. Then:

$$Q_{BP} = 2$$

Going through the calculations for the pole pair the intermediate results are:

and:

$$F_{BP1} = 0.823391$$
 $F_{BP2} = 1.214489$
 $Q_{BP1} = Q_{BP2} = 9.029157$

And for the single pole:

 $F_{BP3} = 1$ $Q_{BP3} = 4.431642$

Again a full example will be worked out in a later section.

Low Pass to Bandreject (Notch)

As in the band pass case, a bandreject filter can be either wideband or narrowband, depending on whether or not the poles are separated by two octaves or more. To avoid confusion, the following convention will be adopted. If the filter is wideband, it will be referred to as a bandreject filter. A narrowband filter will be referred to as a notch filter.

One way to build a notch filter is to construct it as a band-pass filter whose output is subtracted from the input (1 - BP). Another way is with cascaded low pass and high pass sections, especially for the bandreject (wideband) case. In this case, the sections are in parallel, and the output is the difference.

Just as the band pass case is a direct transformation of the low-pass prototype, where dc is transformed to F_0 , the notch filter can be first transformed to the high-pass case, and then dc, which is now a zero, is transformed to F_0 .

A more general approach would be to convert the poles directly. A notch transformation results in two pairs of complex poles and a pair of second order imaginary zeros from each low-pass pole pair.

First, the value of Q_{BR} is determined by:

$$Q_{BR} = \frac{F_0}{BW}$$
 Eq. 5-64

where BW is the bandwidth at -3 dB.

Given the pole locations of the low-pass prototype

$$-\alpha \pm j\beta$$
 Eq. 5-65

and the values of F_0 and Q_{BR} , the following calculations will result in two sets of values for Q and frequencies, F_H and F_L , which define a pair of notch filter sections.

$$C = \alpha^2 + \beta^2$$
 Eq. 5-66

$$D = \frac{\alpha}{Q_{BR}C}$$
 Eq. 5-67

$$E = \frac{\beta}{Q_{BR}C}$$
 Eq. 5-68

$$F = E^2 - 4D^2 + 4$$
 Eq. 5-69

$$G = \sqrt{\frac{F}{2} + \sqrt{\frac{F2}{4}}} + D^2 E^2$$
 Eq. 5-70

$$H = \frac{D E}{G}$$
 Eq. 5-71

$$K = \frac{1}{2}\sqrt{(D+H)^2 + (E+G)^2}$$
 Eq. 5-72

$$Q = \frac{K}{D+H}$$
 Eq. 5-73

The pole frequencies are given by:

$$F_{BRI} = \frac{F_0}{K}$$
 Eq. 5-74

$$F_{BR2} = K F_0$$
 Eq. 5-75

$$F_z = F_0$$
 Eq. 5-76

where F_0 is the notch frequency and the geometric mean of F_{BR1} and F_{BR2} . A simple real pole, α_0 , transforms to a single section having a Q given by:

$$Q = Q_{BR} \alpha_0$$
 Eq. 5-78

with a frequency $F_{BR} = F_0$. There will also be transmission zero at F_0 .

In some instances, such as the elimination of the power line frequency (hum) from low level sensor measurements, a notch filter for a specific frequency may be designed.

Assuming that an attenuation of A dB is required over a bandwidth of B, the required Q is determined by:

$$Q = \frac{\omega_0}{B\sqrt{10^{.1 A} - 1}}$$
 Eq. 5-79

A 3-pole 1 dB Chebychev is again used as an example:

$$\alpha_{LP1} = 0.2257$$

 $\beta_{LP1} = 0.8822$
 $\alpha_{LP2} = 0.4513$

A 3 dB bandwidth of 0.1 Hz with a center frequency of 1 Hz is arbitrarily assigned. Then:

$$Q_{BR} = 10$$

Going through the calculations for the pole pair yields the intermediate results:

and:

$$F_{BR1} = 0.94061$$
 $F_{BR2} = 1.063139$
 $Q_{BR1} = Q_{BR2} = 37.10499$

And for the single pole:

 $F_{BP3} = 1$ $Q_{BP3} = 4.431642$

Once again a full example will be worked out in a later section.

Low Pass to All Pass

The transformation from low pass to all pass involves adding a zero in the right-hand side of the s plane corresponding to each pole in the left-hand side.

In general, however, the all-pass filter is usually not designed in this manner. The main purpose of the allpass filter is to equalize the delay of another filter. Many modulation schemes in communications use some form or another of quadrature modulation, which processes both the amplitude and phase of the signal.

All-pass filters add delay to flatten the delay curve without changing the amplitude. In most cases a closed form of the equalizer is not available. Instead the amplitude filter is designed and the delay calculated or measured. Then graphical means or computer programs are used to figure out the required sections of equalization.

Each section of the equalizer gives twice the delay of the low-pass prototype due to the interaction of the zeros. A rough estimate of the required number of sections is given by:

$n = 2 \Delta_{BW} \Delta_{T} + 1$

where Δ_{BW} is the bandwidth of interest in hertz and Δ_{T} is the delay distortion over Δ_{BW} in seconds.

This page intentionally left blank
Filter Realizations

Now that it has been decided *what* to build, it must be decided *how* to build it. That means it is necessary to decide which of the filter topologies to use. Filter design is a two-step process where it is determined what is to be built (the filter transfer function) and then how to build it (the topology used for the circuit).

In general, filters are built from one-pole sections for real poles, and two-pole sections for pole pairs. While a filter can be built from three-pole, or higher order sections, the interaction between the sections increases, and therefore, component sensitivities go up.

It is better to use buffers to isolate the various sections. In addition, it is assumed that all filter sections are driven from a low impedance source. Any source impedance can be modeled as being in series with the filter input.

In all of the design equation figures the following convention will be used:

- H = circuit gain in the passband or at resonance
- F_0 = cutoff or resonant frequency in Hertz
- ω_0 = cutoff or resonant frequency in radians/sec
- Q = circuit "quality factor." Indicates circuit peaking
- $\alpha = 1/Q =$ damping ratio

It is unfortunate that the symbol α is used for damping ratio. It is not the same as the α that is used to denote pole locations ($\alpha \pm j\beta$). The same issue occurs for Q. It is used for the *circuit* quality factor and also the *component* quality factor, which are not the same thing.

The circuit Q is the amount of peaking in the circuit. This is a function of the angle of the pole to the origin in the s plane. The component Q is the amount of loss in what should be lossless reactances. These losses are the parasitics of the components; dissipation factor, leakage resistance, ESR (equivalent series resistance), and so forth, in capacitors and series resistance and parasitic capacitances in inductors.

Single-Pole RC

The simplest filter building block is the passive RC section. The single pole can be either low pass or high pass. Odd order filters will have a single pole section.

The basic form of the low pass RC section is shown in Figure 5-37(A). It is assumed that the load impedance is high (> $\times 10$), so that there is no loading of the circuit. The load will be in parallel with the shunt arm of the filter. If this is not the case, the section will have to be buffered with an op amp. A low-pass filter can be transformed to a high-pass filter by exchanging the resistor and the capacitor. The basic form of the high-pass filter is shown in Figure 5-37(B). Again it is assumed that load impedance is high.



Figure 5-37: Single pole sections

The pole can also be incorporated into an amplifier circuit. Figure 5-38(A) shows an amplifier circuit with a capacitor in the feedback loop. This forms a low-pass filter since as frequency is increased, the effective feedback impedance decreases, which causes the gain to decrease.

Figure 5-38(B) shows a capacitor in series with the input resistor. This causes the signal to be blocked at dc. As the frequency is increased from dc, the impedance of the capacitor decreases and the gain of the circuit increases. This is a high-pass filter.

The design equations for single pole filters appear in Figure 5-66.



Figure 5-38: Single pole active filter blocks

Passive LC Section

While not strictly a function that uses op amps, passive filters form the basis of several active filters topologies and are included here for completeness.

As in active filters, passive filters are built up of individual subsections. Figure 5-39 shows low-pass filter sections. The full section is the basic two-pole section. Odd order filters use one half section which is a single pole section. The m-derived sections, shown in Figure 5-40, are used in designs requiring transmission zeros as well as poles.





Figure 5-40: Passive filter blocks (low pass m-derived)

A low-pass filter can be transformed into a high-pass (see Figures 5-41 and 5-42) by simply replacing capacitors with inductors with reciprocal values and vise versa so:

 $C_{HP} = \frac{1}{L_{LP}}$ Eq. 5-81 $C_{HP} = \frac{1}{L_{LP}}$ Eq. 5-81



Figure 5-42: Passive filter blocks (high pass m-derived)

Transmission zeros are also reciprocated in the transformation so:

$$\omega_{Z,HP} = \frac{1}{\omega_{Z,HP}}$$
 Eq. 5-82

The low-pass prototype is transformed to band pass and bandreject filters as well by using the table in Figure 5-43.

For a passive filter to operate, the source and load impedances must be specified. One issue with designing passive filters is that in multipole filters each section is the load for the preceding sections and also the source impedance for subsequent sections, so component interaction is a major concern. Because of this, designers typically make use of tables, such as in Williams's book (Reference 2).

$$L_{\rm HP} = \frac{1}{C_{\rm LP}}$$
 Eq. 5-80

and



Figure 5-43: Low pass \rightarrow band pass and high pass \rightarrow bandreject transformation

Integrator

Any time a frequency-dependent impedance is put in a feedback network the inverse frequency response is obtained. For example, if a capacitor, which has a frequency dependent impedance that decreases with increasing frequency, is put in the feedback network of an op amp, an integrator is formed, as in Figure 5-44.

The integrator has high gain (i.e., the open-loop gain of the op amp) at dc. An integrator can also be thought of as a low-pass filter with a cutoff frequency of 0 Hz.



Figure 5-44: Integrator

General Impedance Converter

Figure 5-45 is the block diagram of a general impedance converter. The impedance of this circuit is:

$$Z = \frac{Z1 \ Z3 \ Z5}{Z2 \ Z4}$$
 Eq. 5-83

By substituting one or two capacitors into appropriate locations (the other locations being resistors), several impedances can be synthesized (see Reference 25).

One limitation of this configuration is that the lower end of the structure must be grounded.

Active Inductor

Substituting a capacitor for Z4 and resistors for Z1, Z2, Z3, and Z5 in the GIC results in an impedance given by:

$$Z_{11} = \frac{\text{sC R1 R3 R5}}{\text{R2}}$$
 Eq. 5-84

By inspection it can be shown that this is an inductor with a value of:

$$L = \frac{C R1 R3 R5}{R2}$$
Eq. 5-85

This is just one way to simulate an inductor as shown in Figure 5-46.



Figure 5-45: General impedance converter

Figure 5-46: Active inductor

Frequency-Dependent Negative Resistor (FDNR)

By substituting capacitors for two of the Z1, Z3, or Z5 elements, a structure known as a frequencydependant negative resistance (FDNR) is generated. The impedance of this structure is

$$Z_{11} = \frac{sC^2 R2 R4}{R5}$$
 Eq. 5-86

This impedance, which is called a D element, has the value:

$$D = C^2 R4$$
 Eq. 5-87

assuming

$$C1 = C2 \text{ and } R2 = R5$$
 Eq. 5-88

The three possible versions of the FDNR are shown in Figure 5-47.



Figure 5-47: Frequency-dependent negative resistor blocks

There is theoretically no difference in these three blocks, and so they should be interchangeable. In practice, though, there may be some differences. Circuit (a) is sometimes preferred because it is the only block to provide a return path for the amplifier bias currents.

For the FDNR filter (see Reference 24), the passive realization of the filter is used as the basis of the design. As in the passive filter, the FDNR filter must then be denormalized for frequency and impedance. This is typically done before the conversion by 1/s. First take the denormalized passive prototype filter and transform the elements by 1/s. This means that inductors, whose impedance is equal to sL, transform into a resistor with an impedance of L. A resistor of value R becomes a capacitor with an impedance of R/s; and a capacitor of impedance 1/sC transforms into a frequency dependent resistor, D, with an impedance of $1/s^2C$. The transformations involved with the FDNR configuration and the GIC implementation of the D element

are shown in Figure 5-48. We can apply this transformation to low-pass, high-pass, band-pass, or notch filters, remembering that the FDNR block must be restricted to shunt arms.



Figure 5-48: 1/s transformation

A worked out example of the FDNR filter is included in the next section.

A perceived advantage of the FDNR filter in some circles is that there are no op amps in the direct signal path, which can add noise and/or distortion, however small, to the signal. It is also relatively insensitive to component variation. These advantages of the FDNR come at the expense of an increase in the number of components required.

Sallen-Key

The Sallen-Key configuration, also known as a voltage control voltage source (VCVS), was first introduced in 1955 by R. P. Sallen and E. L. Key of MIT's Lincoln Labs (see Reference 14). It is one of the most widely used filter topologies and is shown in Figure 5-49. One reason for this popularity is that this configuration shows the least dependence of filter performance on the performance of the op amp. This is



due to the fact that the op amp is configured as an amplifier, as opposed to an integrator, which minimizes the gain-bandwidth requirements of the op amp. This infers that for a given op amp, it is possible to design a higher frequency filter than with other topologies since the op amp gain bandwidth product will not limit the performance of the filter as it would if it were configured as an integrator. The signal phase through the filter is maintained (noninverting configuration).

Another advantage of this configuration is that the ratio of the largest resistor value to the smallest resistor value and the ratio of the largest capacitor value to the smallest capacitor value (component spread) are low, which is good for manufacturability. The frequency and Q terms are somewhat independent, but they are very sensitive to the gain parameter. The Sallen-Key is very Q-sensitive to element values, especially for high Q sections. The design equations for the Sallen-Key low pass are shown in Figure 5-67.

There is a special case of the Sallen-Key low-pass filter. If the gain is set to 2, the capacitor values, as well as the resistor values, will be the same.

While the Sallen-Key filter is widely used, a serious drawback is that the filter is not easily tuned, due to interaction of the component values on F_0 and Q.

To transform the low pass into the high pass we simply exchange the capacitors and the resistors in the frequency determining network (i.e., *not* the op amp gain resistors). This is shown in Figure 5-50. The comments regarding sensitivity of the filter given above for the low pass case apply to the high pass case as well. The design equations for the Sallen-Key high pass are shown in Figure 5-68.

The band-pass case of the Sallen–Key filter has a limitation (see Figure 5-51). The value of Q will determine the gain of the filter, i.e., it can not be set independently, as in the low pass or high pass cases. The design equations for the Sallen-Key band-pass are shown in Figure 5-69.

A Sallen-Key notch filter may also be constructed, but it has a large number of undesirable characteristics. The resonant frequency, or the notch frequency, cannot be adjusted easily due to component interaction. As in the band pass case, the section gain is fixed by the other design parameters, and there is a wide spread in component values, especially capacitors. Because of this, and the availability of easier to use circuits, it is not covered here.



Figure 5-50: Sallen-Key high-pass filter

Figure 5-51: Sallen-Key band-pass filter

Multiple Feedback

The multiple feedback filter uses an op amp as an integrator as shown in Figure 5-52. Therefore, the dependence of the transfer function on the op amp parameters is greater than in the Sallen-Key realization. It is hard to generate high Q, high frequency sections due to the limitations of the open-loop gain of the op amp. A rule of thumb is that the open-loop gain of the op amp should be at least 20 dB (×10) above the amplitude response at the resonant (or cutoff) frequency, including the peaking caused by the Q of the filter. The peaking due to Q will cause an amplitude, A_0 :

$$A_0 = H Q \qquad \qquad \text{Eq. 5-89}$$

where *H* is the gain of the circuit. The multiple feedback filter will invert the phase of the signal. This is equivalent to adding the resulting 180° phase shift to the phase shift of the filter itself.

The maximum to minimum component value ratios is higher in the multiple feedback case than in the Sallen-Key realization. The design equations for the multiple feedback low pass are given in Figure 5-70.



Figure 5-52: Multiple feedback low pass

Comments made about the multiple feedback low pass case apply to the high-pass case as well (see Figure 5-53). Note that resistors and capacitors are again swapped to convert the low pass case to the high pass case. The design equations for the multiple feedback high pass are given in Figure 5-71.

The design equations for the multiple feedback band pass case (see Figure 5-54) are given in Figure 5-72.



Figure 5-53: Multiple feedback high pass

Figure 5-54: Multiple feedback band pass

This circuit is widely used in low Q (< 20) applications. It allows some tuning of the resonant frequency, F_0 , by making R2 variable. Q can be adjusted (with R5) as well, but this will also change F_0 .

Tuning of F_0 can be accomplished by monitoring the output of the filter with the horizontal channel of an oscilloscope, with the input to the filter connected to the vertical channel. The display will be a Lissajous pattern. This pattern will be an ellipse that will collapse to a straight line at resonance, since the phase shift will be 180°. Output could also be adjusted for maximum output, which will also occur at resonance, but this is usually not as precise, especially at lower values of Q where there is a less pronounced peak.

State Variable

The state-variable realization (see Reference 11) is shown in Figure 5-55, along with the design equations in Figure 5-73. This configuration offers the most precise implementation, at the expense of many more circuit elements. All three major parameters (gain, Q and ω_0) can be adjusted independently, and low pass, high pass, and band pass outputs are available simultaneously. Note that the low pass and high pass outputs are inverted in phase while the band pass output maintains the phase. The gain of each of the outputs of the filter is also independently variable. With an added amplifier section summing the low pass and high pass sections the notch function can also be synthesized. By changing the ratio of the summed sections, low pass notch, standard notch and high pass notch functions can be realized. A standard notch may also be realized by subtracting the band pass output from the input with the added op amp section. An all-pass filter may also be built with the four amplifier configuration by subtracting the band pass output from the input. In this instance, the band pass gain must equal 2.



Figure 5-55: State-variable filter

Since all parameters of the state-variable filter can be adjusted independently, component spread can be minimized. Also, variations due to temperature and component tolerances are minimized. The op amps used in the integrator sections will have the same limitations on op amp gain-bandwidth as described in the multiple feedback section.

Tuning the resonant frequency of a state-variable filter is accomplished by varying R4 and R5. While it is not necessary to tune both, if varying over a wide range it is generally preferable. Holding R1 constant, tuning R2 sets the low pass gain and tuning R3 sets the high-pass gain. Band pass gain and Q are set by the ratio of R6 and R7.

Since the parameters of a state variable filter are independent and tunable, it is easy to add electronic control of frequency, Q and ω_0 . This adjustment is accomplished by using an analog multiplier, multiplying DACs (MDACs) or digital pots, as shown in one of the examples in a later section. For the integrator sections adding the analog multiplier or MDAC effectively increases the time constant by dividing the voltage driving the resistor, which, in turn, provides the charging current for the integrator capacitor. This in effect raises the resistance and, in turn, the time constant. The Q and gain can be varied by changing the ratio of the various feedback paths. A digital pot will accomplish the same feat in a more direct manner, by directly changing the resistance value. The resultant tunable filter offers a great deal of utility in measurement and control circuitry. A worked out example is given in Section 8 of this chapter.

Biquadratic (Biquad)

A close cousin of the state variable filter is the biquad as shown in Figure 5-56. The name of this circuit was first used by J. Tow in 1968 (Reference 11) and later by L. C. Thomas in 1971 (see Reference 12). The name derives from the fact that the transfer function is a quadratic function in both the numerator and the denominator. Hence the transfer function is a biquadratic function. This circuit is a slight rearrangement of the state variable circuit. One significant difference is that there is not a separate high pass output. The band pass output inverts the phase. There are two low pass outputs, one in phase and one out of phase. With the addition of a fourth amplifier section, high pass, notch (low-pass, standard and high-pass) and all-pass filters can be realized. The design equations for the biquad are given in Figure 5-74.

Referring to Figure 5-74, the all-pass case of the biquad, R8 = R9/2 and R7 = R9. This is required to make the terms in the transfer function line up correctly. For the high pass output, the input, band pass and second low pass outputs are summed. In this case the constraints are that R1 = R2 = R3 and R7 = R8 = R9.

Like the state variable, the biquad filter is tunable. Adjusting R3 will adjust the Q. Adjusting R4 will set the resonant frequency. Adjusting R1 will set the gain. Frequency would generally be adjusted first followed by Q and then gain. Setting the parameters in this manner minimizes the effects of component value interaction.



Figure 5-56: Biquad filter

Dual Amplifier Band Pass (DAPB)

The Dual Amplifier band-pass filter structure is useful in designs requiring high Qs and high frequencies. Its component sensitivity is small, and the element spread is low. A useful feature of this circuit is that the Q and resonant frequency can be adjusted more or less independently.

Referring to Figure 5-57, the resonant frequency can be adjusted by R2. R1 can then be adjusted for Q. In this topology it is useful to use dual op amps. The match of the two op amps will lower the sensitivity of Q to the amplifier parameters.

It should be noted that the DABP has a gain of 2 at resonance. If lower gain is required, resistor R1 may be split to form a voltage divider. This is reflected in the addendum to the design equations of the DABP, Figure 5-75.



Figure 5-57: Dual amplifier band-pass filter

Twin-T Notch

The Twin T is widely used as a general-purpose notch circuit as shown in Figure 5-58. The passive implementation of the twin T (i.e., with no feedback) has a major shortcoming of having a Q that is fixed at 0.25. This issue can be rectified with the application of positive feedback to the reference node. The amount of the signal feedback, set by the R4/R5 ratio, will determine the value of Q of the circuit, which, in turn, determines the notch depth. For maximum notch depth, the resistors R4 and R5 and the associated op amp can be eliminated. In this case, the junction of C3 and R3 will be directly connected to the output.

Tuning is not easily accomplished. Using standard 1% components a 60 dB notch is as good as can be expected, with 40 dB–50 dB being more typical.

The design equations for the Twin T are given in Figure 5-76.



Figure 5-58: Twin-T notch filter

Bainter Notch

A simple notch filter is the Bainter circuit (see Reference 21). It is composed of simple circuit blocks with two feedback loops as shown in Figure 5-59. Also, the component sensitivity is very low.

This circuit has several interesting properties. The Q of the notch is not based on component matching as it is in every other implementation, but is instead only dependent on the gain of the amplifiers. Therefore, the



Figure 5-59: Bainter notch filter

notch depth will not drift with temperature, aging, and other environmental factors. The notch frequency may shift, but not the depth.

Amplifier open loop gain of 10⁴ will yield a Q_z of > 200. It is capable of orthogonal tuning with minimal interaction. R6 tunes Q and R1 tunes ω_z . Varying R3 sets the ratio of ω_0/ω_z produces low pass notch (R4 > R3), notch (R4 = R3) or high pass notch (R4 < R3).

The design equations of the Bainter circuit are given in Figure 5-77.

Boctor Notch

The Boctor circuits (see References 22, 23), while moderately complicated, uses only one op amp. Due to the number of components, there is a great deal of latitude in component selection. These circuits also offer low sensitivity and the ability to tune the various parameters more or less independently.

There are two forms, a low pass notch (Figure 5-60) and a high-pass notch (Figure 5-61). For the low pass case, the preferred order of adjustment is to tune ω_0 with R4, then Q_0 with R2, next Q_z with R3 and finally ω_z with R1.



In order for the components to be realizable we must define a variable, k1, such that:

$$\frac{\omega_0^2}{\omega_z^2} < k1 < 1$$
 Eq. 5-90

The design equations are given in Figure 5-78 for the low pass case and in Figure 5-79 for the high pass case. In the high-pass case circuit gain is require and it applies only when

$$Q < \frac{1}{1 - \frac{\omega_z^2}{\omega_0^2}}$$
 Eq. 5-91

but a high pass notch can be realized with one amplifier and only two capacitors, which can be the same value. The pole and zero frequencies are completely independent of the amplifier gain. The resistors can be trimmed so that even 5% capacitors can be used.

"1 – Band Pass" Notch

As mentioned in the state variable and biquad sections, a notch filter can be built as 1 - BP. The band pass section can be any of the all-pole band pass realizations discussed above, or any others. Keep in mind whether the band pass section is inverting as shown in Figure 5-62 (such as the multiple feedback circuit) or noninverting as shown in Figure 5-63 (such as the Sallen-Key), since we want to subtract, not add, the band pass output from the input.

It should be noted that the gain of the band pass amplifier must be taken into account in determining the resistor values. Unity gain band pass would yield equal values.



Figure 5-62: 1 - BP filter for inverting band pass configurations



Figure 5-63: 1 - BP filter for noninverting band pass configurations

First Order All-Pass

The general form of a first order all-pass filter is shown in Figure 5-64. If the function is a simple RC highpass (Figure 5-64A), the circuit will have a have a phase shift that goes from -180° at 0 Hz. and 0° at high frequency. It will be -90° at $\omega = 1/RC$. The resistor may be made variable to allow adjustment of the delay at a particular frequency.



Figure 5-64: First order all-pass filters

If the function is changed to a low-pass function (Figure 5-64B), the filter is still a first order all-pass and the delay equations still hold, but the signal is inverted, changing from 0° at fv to -180° at high frequency.

Second Order All-Pass

A second order all-pass circuit shown in Figure 5-65 was first described by Delyiannis (see Reference 17). The main attraction of this circuit is that it requires only one op amp. Remember also that an all-pass filter can also be realized as 1 - 2BP.



Figure 5-65: Second order all-pass filter

Any of the all-pole realizations discussed above may be used to build the filter, but be aware of whether or not the BP inverts the phase. Be aware also that the gain of the BP section must be 2. To this end, the DABP structure is particularly useful, since its gain is fixed at 2.

Figures 5-66 through 5-81 following summarize design equations for various active filter realizations. In all cases, H, ω_0 , Q, and α are given, being taken from the design tables.



Figure 5-66: Single-pole filter design equations



SALLEN-KEY LOW PASS

Figure 5-67: Sallen-Key low pass design equations

SALLEN-KEY HIGH PASS C1 R1 ┨┠ О оит IN \bigcirc \sim C2 + R2 ≶ \sim R3 $\frac{+Hs^2}{s^2 + \alpha \omega_0 s + \omega_0^2}$ ≶ R4

$$\frac{V_{o}}{V_{IN}} = \frac{H s^{2}}{s^{2} + s \left[\frac{\frac{C2}{R2} + \frac{C1}{R2} + (1 - H)\frac{C2}{R1}}{C1 C2}\right] + \frac{1}{R1 R2 C1 C2}}$$

CHOOSE:	C1	R3
THEN:	$k = 2 \pi F_0 C1$	$R4 = \frac{R3}{(H-1)}$
	C2 = C1	
	$R1 = \frac{\alpha + \sqrt{\alpha^2 + (H-1)}}{4k}$	
	$R2 = \frac{4}{\alpha + \sqrt{\alpha^2 + (H-1)}} * \frac{1}{k}$	

Figure 5-68: Sallen-Key high pass design equations

О оит R2 ≩ C1 R1 IN C ⊥ ⊤ R3 ≩ C2 = R4 $\frac{+H \omega_0 s}{s^2 + \alpha \omega_0 s + \omega_0^2}$ ₹ R5 $\frac{V_{o}}{V_{IN}} = \frac{Hs\frac{1}{R1C2}}{s^{2} + s\left[\frac{C1}{R3} + \frac{(C1 + C2)}{R1} + \frac{C2}{R2} + \frac{C1}{R2}}{C1 C2}\right] + \frac{1}{R3 C1 C2}\left(\frac{R1 + R2}{R1 R2}\right)}$ CHOOSE: C1 R4 $R5 = \frac{R4}{(H-1)}$ $k = 2\pi F_0 C1$ THEN: $C2 = \frac{1}{2}C1$ $R1 = \frac{2}{k}$ $R2 = \frac{2}{3k}$ $R3 = \frac{4}{k}$ $H = \frac{1}{3} \left(6.5 - \frac{1}{Q} \right)$

SALLEN-KEY BAND PASS

Figure 5-69: Sallen-Key band pass design equations

MULTIPLE FEEDBACK LOW PASS



Figure 5-70: Multiple feedback low pass design equations

MULTIPLE FEEDBACK HIGH PASS



Figure 5-71: Multiple feedback high pass design equations

MULTIPLE FEEDBACK BAND PASS



Figure 5-72: Multiple feedback band-pass design equations



Figure 5-73A: State-variable design equations









Figure 7-73C: State-variable design equations



BIQUADRATIC (A)



Figure 5-74A: Biquad design equations



Figure 5-74B: Biquad design equations

DUAL AMPLIFIER BANDPASS ₩~-R4 OUT О Ş R5 $\frac{+H\,\omega_{_{0}}\,s}{s^{^{2}}+\alpha\,\omega_{_{0}}\,s+\omega_{_{0}}^{^{2}}}$ С R3 W/-R2 ≤ R1 IN ~~~ ($\frac{2}{R1 C}$ С $\frac{V_0}{V_{IN}}$ $\frac{1}{R1C} + \frac{1}{R2 R3 C^2}$ $s^{2} + s$ С CHOOSE: R4 $R = \frac{1}{2\pi F_0 C}$ R5 = R4THEN: R1 = QRR2 = R3 = RR2 ≷ FOR GAINS LESS THAN 2 (GAIN = A_v): R1A $R1A = \frac{2R1}{A_v}$ IN () ~~~~ $R1B = \frac{R1A A_{v}}{2 - A_{v}}$ ≶ R1B C

Figure 5-75: Dual amplifier band pass design equations



Figure 5-76: Twin-T notch design equations



BAINTER NOTCH

Figure 5-77: Bainter notch



GIVEN ω_0 , ω_z , Q_0 CHOOSE R6 R5 C1 R4 = $\frac{1}{1}$

$$R4 = \frac{1}{\omega_0 C1 \ 2Q_0} \qquad R3 = \left(\frac{R6}{R1} + 2\frac{C1}{C2}\right)R5$$
$$R2 = \frac{R4 \ R6}{R4 = R6} \qquad C2 = 4Q_0^2 \frac{R4}{R6}C1$$
$$R1 = \frac{1}{2} \left(\frac{R6 \ \omega_2^2}{R4 \ \omega_0^2} - 1\right)$$

Figure 5-78: Boctor notch, low pass



WHERE:
$$R_{EQ1} = R1 || R3 || R6$$

 $R_{EQ2} = R2 + (R4 || R5)$

GIVEN: F_z F₀ H

 $F_z \ Q \ H$



$$F_0 = F_Z \sqrt{\frac{1}{1 - \frac{1}{2Q^2}}}$$

$$Y = \frac{1}{Q \left(1 - \frac{F_z^2}{F_0^2} \right)}$$

Figure 5-79A: Boctor high pass design equations

or



Figure 5-79B: Boctor high-pass design equations (continued)

FIRST ORDER ALL PASS



PHASE SHIFT (
$$\phi$$
) = -2 TAN⁻¹ $\left(\frac{\text{RC}}{2 \pi \text{F}}\right)$

GROUP DELAY = $\frac{2 \text{ R C}}{(2 \pi \text{ F R C})^2 + 1}$

GIVEN A PHASE SHIFT OF ϕ AT A FREQUENCY = F

 $R C = 2 \pi F TAN\left(-\frac{\phi}{2}\right)$



Figure 5-80: First order all pass design equations

SECOND ORDER ALL PASS



CHOOSE: C

 $k = 2 \pi F_0 C$ $R2 = \frac{2Q}{k}$ $R1 = \frac{1}{2kQ}$ R3 = R1 $R4 = \frac{Q}{2}$

Figure 5-81: Second order all pass
SECTION 5-7

Practical Problems in Filter Implementation

In the previous sections filters were dealt with as mathematical functions. The filter designs were assumed to have been implemented with "perfect" components. When the filter is built with real-world components, design trade-offs must typically be made.

In building a filter with an order greater the two, multiple second and/or first order sections are used. The frequencies and Qs of these sections must align precisely or the overall response of the filter will be affected. For example, the antialiasing filter design example in the next section is a 5th order Butterworth filter, made up of a second order section with a frequency $(F_o) = 1$ and a Q = 1.618, a second order section with a frequency $(F_o) = 1$ and a Q = 1.618, a second order section with a frequency $(F_o) = 1$ and a Q = 0.618, and a first order section with a frequency $(F_o) = 1$ (for a filter normalized to 1 rad/sec). If the Q or frequency response of any of the sections is off slightly, the overall response will deviate from the desired response. It may be close, but it won't be exact. As is typically the case with engineering, a decision must be made as to which trade-offs should be made. For instance, do we really need a particular response exactly? Is there a problem if there is a little more ripple in the pass band? Or if the cutoff frequency is at a slightly different frequency? These are the types of questions that face a designer, and will vary from design to design.

Passive Components (Resistors, Capacitors, Inductors)

Passive components are the first problem. When designing filters, the calculated values of components will most likely not be available commercially. Resistors, capacitors, and inductors come in standard values. While custom values can be ordered, the practical tolerance will probably still be $\pm 1\%$ at best. An alternative is to build the required value out of a series and/or parallel combination of standard values. This increases the cost and size of the filter. Not only is the cost of components increased, so are the manufacturing costs, both for loading and tuning the filter. Furthermore, success will be still limited by the number of parts that are used, their tolerance, and their tracking, both over temperature and time.

A more practical way is to use a circuit analysis program to determine the response using standard values. The program can also evaluate the effects of component drift over temperature. The values of the sensitive components are adjusted using parallel combinations where needed, until the response is within the desired limits. Many of the higher end filter CAD programs include this feature.

The resonant frequency and Q of a filter are typically determined by the component values. Obviously, if the component value is drifting, the frequency and the Q of the filter will drift. This, in turn, will cause the frequency response to vary. This is especially true in higher order filters.

Higher order implies higher Q sections. Higher Q sections means that component values are more critical, since the Q is typically set by the ratio of two or more components, typically capacitors.

In addition to the initial tolerance of the components, the effects of temperature/time drift must be evaluated. The temperature coefficients of the various components may be different in both magnitude and sign. Capacitors, especially, are difficult in that not only do they drift, but the temperature coefficient (TC) is also a function of temperature, as shown in Figure 5-82. This represents the temperature coefficient of a (relatively) poor film capacitor, which might be typical for a Polyester or Polycarbonate type. *Linear TC* in film capacitors can be found in the polystyrene, polypropylene, and Teflon dielectrics. In these types TC is on the order of 100 ppm/°C–200 ppm/°C, and if necessary, this can be compensated with a complementary TC elsewhere in the circuit.



Figure 5-82: A poor film capacitor temperature coefficient

The lowest TC dielectrics are NPO (or COG) ceramic ($\pm 30 \text{ ppm/°C}$), and polystyrene (-120 ppm/°C). Some capacitors, mainly the plastic film types, such as polystyrene and polypropylene, also have a limited temperature range.

While there is infinite choice of the values of the passive components for building filters, in practice there are physical limits. Capacitor values below 10 pF and above 10 μ F are not practical. Electrolytic capacitors should be avoided. Electrolytic capacitors are typically very leaky. A further potential problem is if they are operated without a polarizing voltage, they become nonlinear when the ac voltage reverse biases them. Even with a dc polarizing voltage, the ac signal can reduce the instantaneous voltage to 0 V or below. Large values of film capacitors are physically very large.

Resistor values of less than 100 Ω should be avoided, as should values over 1 M Ω . Very low resistance values (under 100 Ω) can require a great deal of drive current and dissipate a great deal of power. Both of these should be avoided. And low values and very large values of resistors may not be as readily available. Very large values tend to be more prone to parasitics since smaller capacitances will couple more easily into larger impedance levels. Noise also increases with the square root of the resistor value. Larger value resistors also will cause larger offsets due to the effects of the amplifier bias currents.

Parasitic capacitances due to circuit layout and other sources affect the performance of the circuit. They can form between two traces on a PC board (on the same side or opposite side of the board), between leads of adjacent components, and just about everything else you can (and in most cases can't) think of. These capacitances are usually small, so their effect is greater at high impedance nodes. Thus, they can be controlled most of the time by keeping the impedance of the circuits down. Remember that the effects of stray capacitance are frequency-dependent, being worse at high frequencies because the impedance drops with increasing frequency.

Parasitics are not just associated with outside sources. They are also present in the components themselves.

A capacitor is more than just a capacitor in most instances. A real capacitor has inductance (from the leads and other sources) and resistance as shown in Figure 5-83. This resistance shows up in the specifications



Figure 5-83: Capacitor equivalent circuit

as leakage and poor power factor. Obviously, users would like capacitors with very low leakage and good power factor (see Figure 5-84).

CAPACITOR COMPARISON CHART

TYPE Polystyrene	TYPICAL DA 0.001% to 0.02%	ADVANTAGES Inexpensive Low DA Good Stability (~120ppm/°C)	DISADVANTAGES Damaged by Temperature > 85°C Large High Inductance Vendors Limited
Polypropylene	0.001% to 0.02%	Inexpensive Low DA Stable (~200ppm/°C) Wide Range of Values	Damaged by Temperature > 105°C Large High Inductance
Teflon	0.003% to 0.02%	Low DA Available Good Stability Operational above 125 °C Wide Range of Values	Expensive Large High Inductance
Polycarbonate	0.1%	Good Stability Low Cost Wide Temperature Range Wide Range of Values	Large DA Limits to 8-Bit Applications High Inductance
Polyester	0.3% to 0.5%	Moderate Stability Low Cost Wide Temperature Range Low Inductance (Stacked Film)	Large DA Limits to 8-Bit Applications High Inductance (Conventional)
NP0 Ceramic	<0.1%	Small Case Size Inexpensive, Many Vendors Good Stability (30ppm/°C) 1% Values Available Low Inductance (chip)	DA Generally Low (May not be Specified) Low Maximum Values (≤ 10nF)
Monolithic Ceramic (High K)	>0.2%	Low Inductance (chip) Wide Range of Values	Poor Stability Poor DA High Voltage Coefficient
Mica	>0.003%	Low Loss at HF Low Inductance Good Stability 1% Values Available	Quite Large Low Maximum Values (≤ 10nF) Expensive
Aluminum Electrolytic	Very High	Large Values High Currents High Voltages Small Size	High Leakage Usually Polarized Poor Stability, Accuracy Inductive
Tantalum Electrolytic	Very High	Small Size Large Values Medium Inductance	High Leakage Usually Polarized Expensive Poor Stability, Accuracy

Figure 5-84: Capacitor comparison chart

In general, it is best to use plastic film (preferably Teflon or polystyrene) or mica capacitors and metal film resistors, both of moderate to low values in our filters.

One way to reduce component parasitics is to use surface-mounted devices. Not having leads means that the lead inductance is reduced. Also, being physically smaller allows more optimal placement. A disadvantage is that not all types of capacitors are available in surface mount. Ceramic capacitors are popular surface-mount types, and of these, the NPO family has the best characteristics for filtering. Ceramic capacitors may also be prone to microphonics. Microphonics occurs when the capacitor turns into a motion sensor, similar to a strain gage, and turns vibration into an electrical signal, which is a form of noise.

Resistors also have parasitic inductances due to leads and parasitic capacitance. The various qualities of resistors are compared in Figure 5-85.

DISCRETE	TYPE Carbon Composition	ADVANTAGES Lowest Cost High Power/Small Case Size Wide Range of Values	DISADVANTAGES Poor Tolerance (5%) Poor Temperature Coefficient (1500 ppm/°C)
	Wirewound	Excellent Tolerance (0.01%) Excellent TC (1ppm/°C) High Power	Reactance is a Problem Large Case Size Most Expensive
	Metal Film	Good Tolerance (0.1%) Good TC (<1 to 100ppm/°C) Moderate Cost Wide Range of Values Low Voltage Coefficient	Must be Stabilized with Burn-In Low Power
	Bulk Metal or Metal Foil	Excellent Tolerance (to 0.005%) Excellent TC (to <1ppm/°C) Low Reactance Low Voltage Coefficient	Low Power Very Expensive
	High Megohm	Very High Values (10 ⁸ to $10_{14}\Omega$) Only Choice for Some Circuits	High Voltage Coefficient (200ppm/V) Fragile Glass Case (Needs Special Handling) Expensive
NETWORKS	Thick Film	Low Cost High Power Laser-Trimmable Readily Available	Fair Matching (0.1%) Poor TC (>100ppm/°C) Poor Tracking TC (10ppm/°C)
	Thin Film	Good Matching (<0.01%) Good TC (<100ppm/°C) Good Tracking TC (2ppm/°C) Moderate Cost Laser-Trimmable Low Capacitance Suitable for Hybrid IC Substrate	Often Large Geometry Limited Values and Configurations

RESISTOR COMPARISON CHART

Figure 5-85: Resistor comparison chart

Limitations of Active Elements (Op Amps) in Filters

The active element of the filter will also have a pronounced effect on the response. In developing the various topologies (Multiple Feedback, Sallen-Key, State-Variable, and so forth), the active element was always modeled as a "perfect" operational amplifier. That is, it has:

- 1) infinite gain
- 2) infinite input impedance
- 3) zero output impedance

none of which vary with frequency. While amplifiers have improved a great deal over the years, this model has not yet been realized.

The most important limitation of the amplifier has to do with its gain variation with frequency. All amplifiers are band limited. This is due mainly to the physical limitations of the devices with which the amplifier is constructed. Negative feedback theory tells us that the response of an amplifier must be first order (-6 dB per octave) when the gain falls to unity in order to be stable. To accomplish this, a real pole is usually introduced in the amplifier so the gain rolls off to <1 by the time the phase shift reaches 180° (plus some phase margin, hopefully). This roll-off is equivalent to that of a single-pole filter. So in simplistic terms, the transfer function of the amplifier is added to the transfer function of the filter to give a composite function. How much the frequency-dependent nature of the op amp affects the filter is dependent on which topology is used as well as the ratio of the filter frequency to the amplifier bandwidth.

The Sallen-Key configuration, for instance, is the least dependent on the frequency response of the amplifier. All that is required is for the amplifier response to be flat to just past the frequency where the attenuation of the filter is below the minimum attenuation required. This is because the amplifier is used as a gain block. Beyond cutoff, the attenuation of the filter is reduced by the roll-off of the gain of the op amp. This is because the output of the amplifier is phase-shifted, which results in incomplete nulling when fed back to the input. There is also an issue with the output impedance of the amplifier rising with frequency as the open loop gain rolls off. This causes the filter to lose attenuation.

The state-variable configuration uses the op amps in two modes, as amplifiers and as integrators. As amplifiers, the constraint on frequency response is basically the same as for the Sallen-Key, which is flat out to the minimum attenuation frequency. As an integrator, however, more is required. A good rule of thumb is that the open-loop gain of the amplifier must be greater than 10 times the closed-loop gain (including peaking from the Q of the circuit). This should be taken as the absolute minimum requirement. What this means is that there must be 20 dB loop gain, minimum. Therefore, an op amp with 10 MHz unity gain bandwidth is the minimum required to make a 1 MHz integrator. What happens is that the effective Q of the circuit increases as loop gain decreases. This phenomenon is called *Q enhancement*. The mechanism for Q enhancement is similar to that of slew rate limitation. Without sufficient loop gain, the op amp virtual ground is no longer at ground. In other words, the op amp is no longer behaving as an op amp. Because of this, the integrator no longer behaves like an integrator.

The multiple feedback configuration also places heavy constraints on the active element. Q enhancement is a problem in this topology as well. As the loop gain falls, the Q of the circuit increases, and the parameters of the filter change. The same rule of thumb as used for the integrator also applies to the multiple feedback topology (loop gain should be at least 20 dB). The filter gain must also be factored into this equation.

In the FDNR realization, the requirements for the op amps are not as clear. To make the circuit work, we assume that the op amps will be able to force the input terminals to be the same voltage. This implies that the loop gain be a minimum of 20 dB at the resonant frequency. Also it is generally considered to be advantageous to have the two op amps in each leg matched. This is easily accomplished using dual op amps. It is also a good idea to have low bias current devices for the op amps so, all other things being equal, FET input op amps should be used.

In addition to the frequency-dependent limitations of the op amp, other of its parameters may be important to the filter designer.

One is input impedance. We assume in the "perfect" model that the input impedance is infinite. This is required so that the input of the op amp does not load the network around it. This means that we probably want to use FET amplifiers with high impedance circuits.

There is also a small frequency-dependent term to the input impedance, since the effective impedance is the real input impedance multiplied by the loop gain. This is not usually a major source of error, since the network impedance of a high frequency filter should be low.

Distortion resulting from Input Capacitance Modulation

Another subtle effect can be noticed with FET input amps. The input capacitance of a FET changes with the applied voltage. When the amplifier is used in the inverting configuration, such as with the multiple feedback configuration, the applied voltage is held to 0 V. Therefore there is no capacitance modulation. However, when the amplifier is used in the noninverting configuration, such as in the Sallen-Key circuit, this form of distortion can exist.

There are two ways to address this issue. The first is to keep the equivalent impedance low. The second is to balance the impedance seen by the inputs. This is accomplished by adding a network into the feedback leg of the amplifier which is equal to the equivalent input impedance. Note that this will only work for a unity gain application.

As an example, which is taken from the OP176 data sheet, a 1 kHz high-pass Sallen-Key filter is shown (Figure 5-86). Figure 5-87 shows the distortion for the uncompensated version (curve A1) as well as with the compensation (curve A2). Also shown is the same circuit with the impedances scaled up by a factor of 10 (B1 uncompensated, B2 compensated). Note that the compensation improves the distortion, but not as much as having low impedance to start with.



Figure 5-86: Compensation for input capacitance voltage modulation



Figure 5-87: Distortion due to input capacitance modulation

Similarly, the op amp output impedance affects the response of the filter. The output impedance of the amplifier is divided by the loop gain, therefore the output impedance will rise with increasing frequency. This may have an effect with high frequency filters if the output impedance of the stage driving the filter becomes a significant portion of the network impedance.

The fall of loop gain with frequency can also affect the distortion of the op amp, since there is less loop gain available for correction. In the multiple feedback configuration the feedback loop is also frequency-dependent, which may further reduce the feedback correction, resulting in increased distortion. This effect is counteracted somewhat by the reduction of distortion components in the filter network (assuming a low-pass or band-pass filter).

All of the discussion so far is based on using classical voltage feedback op amps. Current feedback, or transimpedance, op amps offer improved high frequency response, but are unusable in any topologies discussed except the Sallen-Key. The problem is that capacitance in the feedback loop of a current feedback amplifier usually causes it to become unstable. Also, most current feedback amplifiers will drive only a small capacitive load. Therefore, it is difficult to build classical integrators using current feedback amplifiers. Some current feedback op amps have an external pin that may be used to configure them as a very good integrator, but this configuration does not lend itself to classical active filter designs.

Current feedback integrators tend to be noninverting, which is not acceptable in the state variable configuration. Also, the bandwidth of a current feedback amplifier is set by its feedback resistor, which would make the Multiple Feedback topology difficult to implement. Another limitation of the current feedback amplifier in the Multiple Feedback configuration is the low input impedance of the inverting terminal. This would result in loading of the filter network. Sallen-Key filters are possible with current feedback amplifiers, since the amplifier is used as a noninverting gain block. New topologies that capitalize on the current feedback amplifiers' superior high frequency performance, and compensate for its limitations, will have to be developed.

The last thing to be aware of is exceeding the dynamic range of the amplifier. Qs over 0.707 will cause peaking in the response of the filter (see Figures 5-5 through 5-7). For high Qs, this could cause overload of the input or output stages of the amplifier with a large input. Note that relatively small values of Q can cause significant peaking. The Q times the gain of the circuit must stay under the loop gain (plus some margin; again, 20 dB is a good starting point). This holds for multiple amplifier topologies as well. Be aware of internal node levels, as well as input and output levels. As an amplifier overloads, its effective Q decreases, so the transfer function will appear to change even if the output appears undistorted. This shows up as the transfer function changing with increasing input level.

We have been dealing mostly with low-pass filters in these discussions, but the same principles are valid for high pass, band pass, and bandreject as well. In general, things like Q enhancement and limited gain/

bandwidth will not affect high-pass filters, since the resonant frequency will hopefully be low in relation to the cutoff frequency of the op amp. Remember, though, that the highpass filter will have a low pass section, by default, at the cutoff frequency of the amplifier. Band-pass and bandreject (notch) filters will be affected, especially since both tend to have high values of Q.

The general effect of the op amp's frequency response on the filter Q is shown in Figure 5-88.

As an example of the Q enhancement phenomenon, consider the Spice simulation of a 10 kHz band pass Multiple Feedback



Figure 5-88: Q enhancement

filter with Q = 10 and gain = 1, using a good high frequency amplifier (the AD847) as the active device. The circuit diagram is shown in Figure 5-89. The open loop gain of the AD847 is greater than 70 dB at 10 kHz as shown in Figure 5-91(A). This is well over the 20 dB minimum, so the filter works as designed as shown in Figure 5-90.



Figure 5-91: AD847 and OP-90 Bode plots

We now replace the AD847 with an OP-90. The OP-90 is a dc precision amplifier and so has a limited bandwidth. In fact, its open loop gain is less than 10 dB at 10 kHz (see Figure 5-91(B)). This is not to imply that the AD847 is in all cases better than the OP-90. It is a case of misapplying the OP-90.

From the output for the OP-90, also shown in Figure 5-90, it can be seen that the magnitude of the output has been reduced, and the center frequency has shifted downward.

SECTION 5-8

Design Examples

Several examples will now be worked out to demonstrate the concepts previously discussed

Antialias Filter

As an example, passive and active antialiasing filters will now be designed based upon a common set of specifications. The active filter will be designed in four ways: Sallen-Key, Multiple Feedback, State Variable, and Frequency-Dependent Negative Resistance (FDNR).

The specifications for the filter are given as follows:

- 1) The cutoff frequency will be 8 kHz.
- 2) The stopband attenuation will be 72 dB. This corresponds to a 12-bit system.
- 3) Nyquist frequency of 50 kSPS.
- 4) The Butterworth filter response is chosen in order to give the best compromise between attenuation and phase response.

Consulting the Butterworth response curves (Figure 5-14, reproduced in Figure 5-92), it can be seen that for a frequency ratio of 6.25 (50 kSPS/8 kSPS), a filter order of 5 is required.





Now consulting the Butterworth design table (Figure 5-25), the normalized poles of a 5th order Butterworth filter are:

STAGE	F _o	α
1	1.000	1.618
2	1.000	0.618
3	1.000	

403

The last stage is a real (single) pole, thus the lack of an alpha value. It should be noted that this is not necessarily the order of implementation in hardware. In general, you would typically put the real pole last and put the second order sections in order of decreasing alpha (increasing Q) as we have done here. This will avoid peaking due to high Q sections possibly overloading internal nodes. Another feature of putting the single pole at the end is to bandlimit the noise of the op amps. This is especially true if the single pole is implemented as a passive filter.

For the passive design, we will choose the zero input impedance configuration. While "classic" passive filters are typically double terminated, that is with termination on both source and load ends, the concern here is with voltage transfer not power transfer so the source termination will not be used. From the design table (see Reference 2, p. 313), we find the normalized values for the filter (see Figure 5-93).



These values are normalized for a 1 rad/s filter with a 1 Ω termination. To scale the filter we divide all reactive elements by the desired cutoff frequency, 8 kHz (= 50265 rad/sec, = $2\pi 8 \times 10^3$). This is commonly referred to as the frequency scale factor (FSF). We also need to scale the impedance.

For this example, an arbitrary value of 1000Ω is chosen. To scale the impedance, we multiply all resistor and inductor values and divide all capacitor values by this magnitude, which is commonly referred to as the impedance scaling factor (Z).

After scaling, the circuit looks like Figure 5-94.



For the Sallen-Key active filter, use the design equations shown in Figure 5-49. The values for C1 in each section are arbitrarily chosen to give reasonable resistor values. The implementation is shown in Figure 5-95.



Figure 5-95: Sallen-Key implementation

The exact values have been rounded to the nearest standard value. For most active realization to work correctly, it is required to have a zero-impedance driver, and a return path for dc due to the bias current of the op amp. Both of these criteria are approximately met when you use an op amp to drive the filter.

In the above example the single pole has been built as an active circuit. It would have been just as correct to configure it as a passive RC filter. The advantage to the active section is lower output impedance, which may be an advantage in some applications, notably driving an ADC input that uses a switched capacitor structure.

This type of input is common on sigma-delta ADCs as well as many other CMOS type of converters. It also eliminates the loading effects of the input impedance of the following stage on the passive section.

Figure 5-96 shows a multiple feedback realization of the filter. It was designed using the equations in Figure 5-52. In this case, the last section is a passive RC circuit.



Figure 5-96: Multiple feedback implementation

An optional buffer could be added after the passive section, if desired. This would give many of the advantages outlined above, except for bandlimiting the noise of the output amp. By using one of the above two filter realizations, we have both an inverting and a noninverting design.

The state-variable filter, shown in Figure 5-97, was designed with the equations in Figure 5-55. Again, we have rounded the resistor values to the nearest standard 1% value.



Figure 5-97: State variable implementation

Obviously this filter implementation has many more parts than either the Sallen-Key or the multiple feedback. The rational for using this circuit is that stability is improved and the individual parameters are independently adjustable.

The frequency dependent negative resistance (FDNR) realization of this filter is shown in Figure 5-98.



Figure 5-98: FDNR Implementation

In the conversion process from passive to FDNR, the D element is normalized for a capacitance of 1 F. The filter is then scaled to a more reasonable value (0.01 μ F in this case).

In all of the above implementations standard values were used instead of the calculated values. Any variation from the ideal values will cause a shift in the filter response characteristic, but often the effects are minimal. The computer can be used to evaluate these variations on the overall performance and determine if they are acceptable.

To examine the effect of using standard values, take the Sallen-Key implementation. Figure 5-99 shows the response of each of the three sections of the filter. While the Sallen-Key was the filter used, the results from any of the other implementations will give similar results.



Figure 5-99: Individual section response

Figure 5-100 then shows the effect of using standard values instead of calculated values. Notice that the general shape of the filter remains the same, just slightly shifted in frequency. This investigation was done only for the standard value of the resistors. To understand the total effect of component tolerance the same type of calculations would have to be done for the tolerance of all the components and also for their temperature and aging effects.



Figure 5-100: Effect of using standard value resistors

In active filter applications using op amps, the dc accuracy of the amplifier is often critical to optimal filter performance. The amplifier's offset voltage will be passed by the low-pass filter and may be amplified to produce excessive output offset. For low frequency applications requiring large value resistors, bias currents flowing through these resistors will also generate an output offset voltage.

In addition, at higher frequencies, an op amp's dynamics must be carefully considered. Here, slew rate, bandwidth, and open-loop gain play a major role in op amp selection. The slew rate must be fast as well as symmetrical to minimize distortion.

Transformations

In the next example the transformation process will be investigated.

As mentioned earlier, filter theory is based on a low pass prototype, which is then manipulated into the other forms. In these examples the prototype that will be used is a 1 kHz, 3-pole, 0.5 dB Chebyshev filter. A Chebyshev was chosen because it would show more clearly if the responses were not correct, a Butterworth would probably be too forgiving in this instance. A 3-pole filter was chosen so that a pole pair and a single pole would be transformed.

The pole locations for the LP prototype were taken from Figure 5-30. They are:

STAGE	α	β	Fo	α
1	0.2683	0.8753	1.0688	0.5861
2	0.5366		0.6265	

The first stage is the pole pair and the second stage is the single pole. Note the unfortunate convention of using α for 2 entirely separate parameters. The α and β on the left are the pole locations in the s-plane. These are the values that are used in the transformation algorithms. The α on the right is 1/Q, which is what the design equations for the physical filters want to see.

The Sallen-Key topology will be used to build the filter. The design equations in Figure 5-67 (pole pair) and Figure 5-66 (single pole) were then used to design the filter. The schematic is shown in Figure 5-101.

Figure 5-101: Low pass prototype



Using the equation string described in Section 5, the filter is now transformed into a high-pass filter. The results of the transformation are:

STAGE	α	β	Fo	α
1	0.3201	1.0443	0.9356	0.5861
2	1.8636		1.596	

A word of caution is warranted here. Since the convention of describing a Chebyshev filter is to quote the end of the error band instead of the 3 dB frequency, the F_0 must be divided (for high pass) by the ratio of ripple band to 3 dB bandwidth (Table 1, Section 4).

The Sallen-Key topology will again be used to build the filter. The design equations in Figure 5-68 (pole pair) and Figure 5-66 (single pole) where then used to design the filter. The schematic is shown in Figure 5-102.

Figure 5-102: High pass transformation



Figure 5-103 shows the response of the low pass prototype and the high pass transformation. Note that they are symmetric around the cutoff frequency of 1 kHz. Also note that the errorband is at 1 kHz, not the -3 dB point, which is characteristic of Chebyshev filters.

The low pass prototype is now converted to a band-pass filter. The equation string outlined in Section 5-5 is used for the transformation. Each pole of the prototype filter will transform into a pole pair. Therefore the 3-pole prototype, when transformed, will have six poles (3-pole pairs). In addition, there will be six zeros at the origin.





Part of the transformation process is to specify the 3 dB bandwidth of the resultant filter. In this case this bandwidth will be set to 500 Hz. The results of the transformation yield:

STAGE	F_0	Q	A_0
1	804.5	7.63	3.49
2	1243	7.63	3.49
3	1000	3.73	1

The reason for the gain requirement for the first two stages is that their center frequencies will be attenuated relative to the center frequency of the total filter. Since the resultant Q_s are moderate (less than 20) the Multiple Feedback topology will be chosen. Figure 5-72 was then used to design the filter sections.

Figure 5-104 is the schematic of the filter and Figure 5-105 shows the filter response.



Note that again there is symmetry around the center frequency. Also the 500 Hz bandwidth is not 250 Hz either side of the center frequency (arithmetic symmetry). Instead the symmetry is geometric, which means that for any two frequencies (F_1 and F_2) of equal amplitude are related by:

$$\mathbf{F}_0 = \sqrt{\mathbf{F}_1 * \mathbf{F}_2}$$

Lastly the prototype will be transformed into a bandreject filter. For this the equation string in Section 5-5 is used. Again, each pole of the prototype filter will transform into a pole pair. Therefore, the 3-pole prototype, when transformed, will have six poles (3-pole pairs).

As in the band pass case, part of the transformation process is to specify the 3 dB bandwidth of the resultant filter. Again in this case this bandwidth will be set to 500 Hz. The results of the transformation yield:

STAGE	F_0	Q	F _{0Z}
1	763.7	6.54	1000
2	1309	6.54	1000
3	1000	1.07	1000

Note that there are three cases of notch filters required. There is a standard notch ($F_0 = F_z$, section 3), a low pass notch ($F_0 < F_z$, section 1) and a high pass notch ($F_0 > F_z$, section 2). Since there is a requirement for all three types of notches, the Bainter Notch is used to build the filter. The filter is designed using Figure 5-77. The gain factors K1 and K2 are arbitrarily set to 1. Figure 5-106 is the schematic of the filter.



Figure 5-106: Bandreject transformation

The response of the filter is shown in Figure 5-107 and in detail in Figure 5-108. Again, note the symmetry around the center frequency. Again the frequencies have geometric symmetry.

CD Reconstruction Filter

This design was done for a magazine article describing a high quality outboard D/A converter for use with digital audio sources (see Reference 26).

A reconstruction filter is required on the output of a D/A converter because, despite the name, the output of a D/A converter is not really an analog voltage but, instead, a series of steps. The converter will put out a discrete voltage, which it will then hold until the next sample is asserted. The filter's job is to remove the



high frequency components, smoothing out the waveform. This is why the filter is sometimes referred to as a smoothing filter. This also serves to eliminate the aliases of the conversion process. The "standard" in the audio industry is to use a third order Bessel function as the reconstruction filter. The reason to use a Bessel filter is that it has the best phase response. This helps to preserve the phase relationship of the individual tones in the music. The price for this phase "goodness" is that the amplitude discrimination is not as good as some other filter types. If we assume that we are using 8× oversampling of the 48 kSPS data stream in the D/A converter the aliases will appear at 364 kHz (8 × 48 k – 20 k). The digital filter that is used in the interpolation process will eliminate the frequencies between 20 kHz and 364 kHz. If we assume that the bandedge is 30 kHz, we have a frequency ratio of approximately 12 (364 \div 30). We use 30 kHz as the band edge, rather than 20 kHz to minimize the roll-off due to the filter in the pass band. In fact, the complete design for this filter includes a shelving filter to compensate for the pass band roll-off. Extrapolating from Figure 5-20, a third order Bessel will only provide on the order of 55 dB attenuation at 12 × Fo. This is only about 9-bit accuracy.

By designing the filter as 7^{th} order, and by designing it as a linear phase with equiripple error of 0.05°, the stopband attenuation can be increased to about 120 dB at 12 × Fo. This is close to the 20-bit system.

The filter will be designed as an FDNR type. This is an arbitrary decision. Reasons to choose this topology are its low sensitivities-to-component tolerances and the fact that the op amps are in the shunt arms rather than in the direct signal path.

The first step is to find the passive prototype. To do this, use the charts in Williams's book. Then get the circuit shown in Figure 5-109A. Next perform a translation in the s-plane. This gives the circuit shown in Figure 5-109B. This filter is scaled for a frequency of 1 Hz and an impedance level of 1 Ω . The D structure of the converted filter is replaced by a GIC structure that can be physically realized. The filter is then denormalized by frequency (30 kHz) and impedance (arbitrarily chosen to be 1 k Ω). This gives a frequency-scaling factor (FS) of 1.884 ×10⁵ (= 2π (3 ×10⁴)). Next, arbitrarily choose a value of 1 nF for the capacitor. This gives an impedance-scaling factor (Z) of 5305 (= (C_{OLD}/C_{NEW})/FSF).

Then multiply the resistor values by Z. This results in the resistors that had the normalized value of 1 Ω will now have a value of 5.305 k Ω . For the sake of simplicity adopt the standard value of 5.36 k Ω . Working backwards, this will cause the cutoff frequency to change to 29.693 kHz. This slight shift of the cutoff frequency will be acceptable.

The frequency scaling factor is then recalculated with the new center frequency and this value is used to denormalize the rest of the resistors. The design flow is illustrated in Figure 5-109. The final schematic is shown it Figure 5-109D.



Analog Filters



The performance of the filter is shown in Figure 5-110(A–D).



Figure 5-110: CD filter performance

Digitally Programmable State Variable Filter

One of the attractive features of the state variable filter is that the parameters (gain, cutoff frequency, and "Q") can be individually adjusted. This attribute can be exploited to allow control of these parameters.

To start, the filter is slightly reconfigured. The resistor divider that determines Q (R6 & R7 of Figure 5-54) is changed to an inverting configuration. The new filter schematic is shown in Figure 5-111. The resistors R1, R2, R3, and R4 (of Figure 5-111) are then replaced by CMOS multiplying DACs. Note that R5 is implemented as the feedback resistor implemented in the DAC. The schematic of this circuit is shown in Figure 5-112.







Figure 5-112: Digitally controlled state variable filter

The AD7528 is an 8-bit dual MDAC. The AD825 is a high speed FET input op amp. Using these components the frequency range can be varied from around 550 Hz to around 150 kHz (Figure 5-113). The Q can be varied from approximately 0.5 to over 12.5 (Figure 5-114). The gain of the circuit can be varied from 0 dB to -48 dB (Figure 5-115).

Analog Filters





The operation of the DACs in controlling the parameters can best be thought of as the DACs changing the effective resistance of the resistors. This relationship is

DAC EQUIVALENT RESISTANCE
$$\frac{256 * \text{DAC RESISTANCE}}{\text{DAC CODE (DECIMAL)}}$$

This, in effect, varies the resistance from 11 k Ω to 2.8 M Ω for the AD7528.

One limitation of this design is that the frequency is dependent on the ladder resistance of the DAC. This particular parameter is not controlled. DACs are trimmed so that the ratios of the resistors, not their absolute values, are controlled. In the case of the AD7528, the typical value is 11 k Ω . It is specified as 8 k Ω min. and 15 k Ω max. A simple modification of the circuit can eliminate this issue. The cost is two more op amps (Figure 5-116). In this case, the effective resistor value is set by the fixed resistors rather than the DAC's resistance. Since there are two integrators the extra inversions caused by the added op amps cancel.



Figure 5-116: Improved digitally variable integrator

As a side note, the multiplying DACs could be replaced by analog multipliers. In this case the control would obviously be an analog rather than a digital signal. We also could just as easily have used a digital pot in place of the MDACs. The difference is that instead of increasing the effective resistance, the value of the pot would be the maximum.

60 Hz Notch Filter

A very common problem in instrumentation is that of interference of the telemetry that is to be measured. One of the primary sources of this interference is the power line. This is particularly true of high impedance circuits. Another path for this noise is ground loops. One possible solution is to use a notch filter to remove the 60 Hz. component. Since this is a single frequency interference, the Twin-T circuit will be used.

Since the maximum attenuation is desired and the minimum notch width is desired, the maximum Q of the circuit is desired. This means the maximum amount of positive feedback is used (R5 open and R4 shorted). Due to the high impedance of the network, a FET input op amp is used.

The filter is designed using Figure 5-78. The schematic is shown in Figure 5-117 and the response in Figure 5-118.



Figure 5-117: 60 Hz Twin-T notch filter



Figure 5-118: 60 Hz notch response

References: Analog Filters

- 1. A. I. Zverev, Handbook of Filter Synthesis, John Wiley, 1967.
- 2. A. B. Williams, Electronic Filter Design Handbook, McGraw-Hill, 1981, ISBN: 0-07-070430-9.
- 3. M. E. Van Valkenburg, Analog Filter Design, Holt, Rinehart & Winston, 1982.
- 4. M. E. Van Valkenburg, Introduction to Modern Network Synthesis, John Wiley and Sons, 1960.
- 5. Zverev and H. J. Blinchikoff, Filtering in the Time and Frequency Domain, John Wiley and Sons, 1976.
- 6. S. Franco, **Design with Operational Amplifiers and Analog Integrated Circuits**, McGraw-Hill 1988, ISBN: 0-07-021799-8.
- 7. W. Cauer, Synthesis of Linear Communications Networks, McGraw-Hill, New York, 1958.
- 8. Budak, Passive and Active Network Analysis and Synthesis, Houghton Mifflin Company, Boston, 1974.
- 9. L. P. Huelsman and P. E. Allen, **Introduction to the Theory and Design of Active Filters**, McGraw Hill, 1980, ISBN: 0-07-030854-3.
- 10. R. W. Daniels, Approximation Methods for Electronic Filter Design, McGraw-Hill, New York, 1974.
- 11. J. Tow, "Active RC Filters a State-space Realization," Proc. IEEE, 1968, vol.56, pp. 1137–1139.
- L.C. Thomas, "The Biquad: Part I Some Practical Design Considerations," IEEE Trans. Circuits and Systems, Vol. CAS-18, 1971, pp. 350–357.
- L.C. Thomas, "The Biquad: Part I A Multipurpose Active Filtering System," IEEE Trans. Circuits and Systems, Vol. CAS-18, 1971, pp. 358–361.
- R. P. Sallen and E. L. Key, "A Practical Method of Designing RC Active Filters," IRE Trans. Circuit Theory, Vol. CT-2, 1955, pp. 74–85.
- 15. P. R. Geffe, "How to Build High-Quality Filters out of Low-Quality Parts," **Electronics**, Nov. 1976, pp. 111–113.
- 16. P. R. Geffe, "Designers Guide to Active Band-pass Filters," EDN, Apr. 5 1974, pp. 46-52.
- 17. T. Delyiannis, "High-Q Factor Circuit with Reduced Sensitivity," **Electronics Letters**, 4, Dec. 1968, p. 577.
- J. J. Friend, "A Single Operational-Amplifier Biquadratic Filter Section," IEEE ISCT Digest Technical Papers, 1970 p. 189.
- L. Storch, "Synthesis of Constant-Time-Delay Ladder Networks Using Bessel Polynomials," Proceedings of IRE, Vol. 42, 1954, pp. 1666–1675.
- K. W. Henderson and W. H. Kautz, "Transient Response of Conventional Filters," IRE Trans. Circuit Theory, Vol. CT-5, 1958, pp. 333–347.
- J. R. Bainter, "Active Filter Has Stable Notch and Response Can be Regulated," Electronics, Oct. 2 1975, pp.115-117.
- S. A. Boctor, "Single Amplifier Functionally Tunable Low-Pass Notch Filter," IEEE Trans. Circuits and Systems, Vol. CAS-22, 1975, pp. 875–881.
- S. A. Boctor, "A Novel Second-Order Canonical RC-Active Realization of High-Pass-Notch Filter," Proc. 1974 IEEE Int. Symp. Circuits and Systems, pp. 640–644.

- L. T. Burton, "Network Transfer Function Using the Concept of Frequency Dependant Negative Resistance," IEEE Trans. Circuit Theory, Vol. CT-16, 1969, pp. 406–408.
- L. T. Burton and D. Trefleaven, "Active Filter Design Using General Impedance Converters," EDN, Feb. 1973, pp.68–75.
- 26. H. Zumbahlen, "A New Outboard DAC, Part 2," Audio Electronics, Jan. 1997, pp. 26–32, 42.
- 27. M. Williamsen, "Notch-Filter Design," Audio Electronics, Jan. 2000, pp. 10–17.
- 28. W. Jung, "Bootstrapped IC Substrate Lowers Distortion in JFET Op Amps," Analog Devices AN232.
- 29. H. Zumbahlen, "Passive and Active Filtering," Analog Devices AN281.
- 30. P. Toomey & W. Hunt, "AD7528 Dual 8-Bit CMOS DAC," Analog Devices AN318.
- 31. W. Slattery, "8th Order Programmable Low-pass Analog Filter Using 12-Bit DACs," Analog Devices AN209.
- 32. CMOS DAC Application Guide, Analog Devices.

This page intentionally left blank

CHAPTER 6 Signal Amplifiers

- Section 6-1: Audio Amplifiers
- Section 6-2: Buffer Amplifiers/Driving Cap Loads
- Section 6-3: Video Amplifiers
- Section 6-4: Communications Amplifiers
- Section 6-5: Amplifier Ideas
- Section 6-6: Composite Amplifiers

This page intentionally left blank

CHAPTER 6 Signal Amplifiers Walt Jung, Walt Kester

SECTION 6-1 Audio Amplifiers Walt Jung

Audio Preamplifiers

Audio signal preamplifiers (preamps) represent the low-level end of the dynamic range of practical audio circuits using modern IC devices. In general, amplifying stages with input signal levels of 10 mV or less fall into the preamp category. This section discusses some basic types of audio preamps, which are:

Microphone—including preamps for dynamic, electret, and phantom-powered microphones, using transformer input circuits, operating from dual and single supplies.

Phonograph—including preamps for moving magnet and moving coil phono cartridges in various topologies, with detailed response analysis and discussion.

In general, when working signals drop to a level of ≈ 1 mV, the input noise generated by the first system amplifying stage becomes critical for wide dynamic range and good signal-to-noise ratio. For example, if internally generated noise of an input stage is 1 μ V and the input signal voltage 1 mV, the best signal-to-noise ratio possible is just 60 dB.

In a given application, both the input voltage level and impedance of a source are usually fixed. Thus, for best signal-to-noise ratio, the input noise generated by the first amplifying stage must be minimized when operated from the intended source. This factor has definite implications to the preamp designer, as a "low noise" circuit for low impedances is quite different from one with low noise operating from a high impedance.

Successfully minimizing the input noise of an amplifier requires a full understanding of all the various factors that contribute to total noise. This includes the amplifier itself as well as the external circuit in which it is used; in fact, *the total circuit environment must be considered* both to minimize noise and maximize dynamic range and signal fidelity.

A further design complication is the fact that not only is a basic gain or signal scaling function to be accomplished, but *signal frequency response* may also need to be altered in a predictable manner. Microphone preamps are an example of wideband, flat frequency response, low noise amplifiers. In contrast to this, phonograph preamp circuits not only scale the signal, they also impart a specific frequency response characteristic to it. A major part of the design for the RIAA phono preamps of this section is a systematic analysis process, which can be used to predictably select components for optimum performance in frequency response terms. This leads to very precise functioning, and excellent correlation between a computer-based design and measured lab operation.

Chapter Six

Microphone Preamplifiers

The microphone preamplifier (mic preamp) is a basic low level audio amplification requirement. Mic preamps can assume a variety of forms, considering the wide range of possible signal levels, the microphone types, and their impedances. These factors influence the optimum circuit for a specific application. Discussed in this section are mic preamps that work with both high and low impedance microphones, both with and without phantom power, and with transformer input stages.

Single-Ended, Single-Supply High-Impedance Mic Preamp

A very simple form of mic preamp is shown in Figure 6-1. This is a noninverting stage with a single-ended input, most useful with high-impedance microphones such as dynamic and piezoelectric types. As shown, it has adjustable gain of 20 dB–40 dB via R_{GAIN} , and is useful with audio sources with 600 Ω or greater source impedances.



Figure 6-1: A single-ended, single-supply mic preamp

The U1 op amp can greatly affect the overall performance, not only in general amplification terms but also in suitability for single supply operation (as shown here). In terms of noise performance, the U1 device should have a low input noise with \geq 500 Ω sources, with the external circuit values adjusted so that the source impedance (microphone) dominates the overall source resistance.

For very low noise on 5 V supplies, very few devices are suitable. Among these the dual SSM2135 or the OP213, and AD822/AD823 stand out, and are recommended as first choices. For very low power, minimal quiescient current parts like the AD8541 can be considered. Many other low noise devices can also work well in this circuit for total supply voltages of 10 V or more, for example the OP275, and OP270/OP470 types. The circuit is also easily adapted for dual supply use, as noted below.

In this circuit, gain-determining resistors R1||R2 (where R2a + R2b = R_{GAIN}) are scaled such that their total resistance is less than the expected source impedance, that is 1 k Ω or less. This minimizes the contribution of the gain resistors to input noise, at high gain. As noted, gain of the circuit is adjusted in the feedback path via resistor R_{GAIN} . In a system sense, control of a microphone or other low level channel signal level is preferably done *after* it has undergone some gain, as the case here. R_{GAIN} can of course be a fixed value.

Because of the single supply operation, input/output coupling is via polar capacitors, namely C1, C2, and C3. C4 is a noise filter, and C5 a bypass. For lowest noise in the circuit, the amplifier biasing must also be

noiseless; that is, free from noise added directly or indirectly by the biasing (see Reference 1). Resistors with dc across them should have low excess noise (film types), or be ac-bypassed. Thus R1, R2, R3, R4, R7, and R8 are preferably metal films, with R7–R8 bypassed. A 2.2 V bias provided from R7–R8 biases the output of U1 to near midsupply. If higher supply voltage is used, R7–R8 can be adjusted for maximum output with a particular amplifier. For example, with low bias current, rail-rail output op amps, R7 and R8 should be high, equal values ($\geq 100 \text{ k}\Omega$).

While the OP213 or SSM2135 for U1 is optimum when operating from lower impedance sources, FET input types such as the AD82x families (or a select CMOS part) is preferable for high impedance sources, such as crystal or ceramic mics. To adapt the circuit for this, R3 and R4 should be $1M\Omega$ or more, and C1 a 0.1 µF film capacitor.

Bandwidth using the OP213 or SSM2135 is about 30 kHz at maximum gain, or about 20 kHz for similar conditions with the AD822 (or AD820). Distortion and noise performance will reflect the U1 device and source impedance. With a shorted input, an SSM2135 measures output noise of about 110 μ V rms at a gain of 100, with a 1 kHz THD + N of 0.022% at 1 V rms into a 2 k Ω load. The AD820 measures about 200 μ V rms with 0.05% THD + N for similar conditions. For both, the figures improve at lower gains.

The circuit of Figure 6-1 is a good one if modest performance and simplicity are required, but requires attention to details. The input cable to the microphone must be shielded, and no longer than required. Similar comments apply to a cable for R_{GAIN} (if remote).

To adapt this circuit for dual supply use, R3 is returned to ground as noted, plus the bias network of R7, R8, and C4 is eliminated. U1 is operated on symmetric supplies (± 5 V, ± 15 V, and so forth), with the $-V_s$ rail bypassed similar to $+V_s$. Coupling caps C1, C2, and C3 are retained, but must be polarized to matched the amplifier used (or nonpolar types). Although microphones with output impedances of less than 600 Ω can be used with this circuit, the noise performance will not be optimum. Also, many of these typically require a balanced input interface. Subsequent circuits show methods of optimizing noise with low impedance, balanced output microphones, as suited for professional applications.

Chapter Six

Electret Mic Preamp Interface

A popular mic type for speech recording and other noncritical applications is the electret type. This is a permanently polarized condenser mic, typically with a built-in common-source FET amplifier. The amplified output signal is taken from the same single ended lead which supplies the microphone with dc power, typically from a 3-10 V dc source.

Figure 6-2 illustrates a basic interface circuit that is useful in powering and scaling the output signal of an electret mic for further use. In this case the scaled output signal from this interface is fed into the LEFT and RIGHT inputs of a 5 V supply powered CODEC for digitization and processing. Dc phantom power is fed to the mic capsules by the R_A - C_A - R_B decoupling network from the 5 V supply, and the ac output signal is tapped off by C_{IN} -R2, and fed to U1. The R_B resistors will vary with different mics and supply voltages, and the values shown are typical. For a quiet mic supply voltage, a filtered/scaled V_R can be generated by the optional U2 connection shown.



Figure 6-2: An electret mic interface for 5 V powered CODECs

The U1 dual scaling amplifier is an SSM2135 or AD822, and is used to normalize the mic signal to either a 1 V rms line level or 100 mV rms mic level typically required by CODEC inputs, and also to low-pass filter it prior to digitization. With a wide variety of electret mics and operating parameters, some signal level scaling is often required.

The scaling gain is simply R1/R2, and R2 is selected to provide a gain "G," to yield 0.1 V rms at the mic inputs of the CODEC, with the rated output from the mic. The U1 stages are inverting, so G can be greater or less than unity, i.e., other than 4 as is shown here, to normalize any practical input signal to an optimum CODEC level. The amplifier's low-pass corner frequency is set by the time constant R1-C, which results in a –3 dB point of 36 kHz. Bias for the U1 stages is provided from the CODEC, via the reference or CMOUT pins, typically a 2.25 V–2.5 V reference voltage. The low frequency time constants C_{IN} -R_B/R2 and C_0 -20 k Ω are wideband to minimize LF phase shift. These (nonpolar) capacitors can be reduced to 1 µF or less, for narrowband uses.

Transformer-Coupled Low-Impedance Microphone Preamps

For any op amp, the best noise performance is attained when the characteristic *noise resistance* of the amplifier, R_n , is equal to the source resistance, R_s . Examples of microphone preamps that make use of this factor are discussed in this section. They utilize an input matching transformer to more closely optimize an amplifier to a source impedance which is unequal to the amplifier R_n . A basic circuit operating on this principle is shown in Figure 6-3. In order to select an optimum transformer turns ratio for a given source resistance (R_s), calculate the characteristic R_n of the op amp in use.



Figure 6-3: Transformer input mic preamplifier with 28 dB to 50 dB gain

 R_n must first be calculated from the op amp's e_n and i_n data as:

$$R_n = \frac{e_n}{i_n}$$
 Eq. 6-1

where e_n is in $\,V\!/\sqrt{Hz}\,$ and i_n is in $\,V\!/\sqrt{Hz}\,$. A turns ratio for T1 may be calculated as:

$$\frac{N_s}{N_p} = \sqrt{\frac{R_n}{R_s}}$$
 Eq. 6-2

where N_s/N_p is the transformer secondary/primary turns ratio. For the OP275 op amp, the values of e_n and i_n are 7 nV/ \sqrt{Hz} and 1.5 pA/ \sqrt{Hz} , respectively; thus,

$$R_n = \frac{e_n}{i_n} = \frac{7 \times 10^{-9}}{1.5 \times 10^{-12}} = 4.7 k\Omega$$

Since both e_n and i_n vary with frequency, R_n will also vary with frequency. Therefore, a value calculated for R_n from the data sheet (such as above) is most accurate at the specified frequency. If the amplifier is to be optimized for a specific frequency, the e_n and i_n values should be for that frequency. However, audio amplifiers are wideband circuits, so latitude is due here. When available, a minimum noise-figure plot for the amplifier will allow graphical determination of the optimum source resistance for noise.

For this case, an optimum transformer turns ratio can be calculated to provide the optimum R_n to the op amp, working from a given R_s . For example, if R_s is 150 Ω , an optimum turns ratio for an OP275 (or other amplifier) with an R_n of 4.7 k Ω will be:

$$\frac{N_{s}}{N_{p}} = \sqrt{\frac{R_{n}}{R_{s}}} = \sqrt{\frac{4.7 \times 10^{3}}{1.5 \times 10^{2}}} \approx 5.6$$

Other examples matching these criteria would include OP27 family types.

Transformers are catalogued in fairly narrow and specific impedance ranges, so a unit with a rated secondary impedance in the range of 5 k Ω to 10 k Ω will be useful (the amplifier minimum noise impedance is reasonably broad). A suitable unit for this purpose is the Jensen JT-110K-HPC. Note that T1 must be adequately shielded and otherwise suitable for operation in low level environments. The use of the matching transformer allows the circuit to achieve an equivalent input noise (referred to the transformer input) that is only a few decibels above the theoretical limit, or very close to the thermal noise of the source resistance. For example, the thermal noise of a 150 Ω resistor in a 20 kHz noise bandwidth at room temperature is 219 nV. A real circuit has a higher input referred noise, due to the transformer plus op amp noise.

An additional advantage of the transformer lies in the *effective voltage gain* that it provides, due to the step up turns ratio. For a given circuit total numeric gain, G_{total} , this reduces the gain required from the op amp U1, $G_{(U1)}$, to:

$$G_{(U1)} = \frac{G_{total}}{N_s / N_p}$$
Eq. 6-3

Thus, in the composite circuit of Figure 6-3 gain G_{total} is the product of the transformer step up, N_s/N_p , and (R1 + R2)/R1, which is $G_{(UI)}$. This has advantages of allowing more amplifier loop gain, thus greater bandwidth and accuracy, lower distortion, and so forth.

The transformer input example mic preamp stage of Figure 6-3 uses the JT-110K-HPC transformer for T1 with a primary/secondary ratio of about 1/8 (150 Ω /10 k Ω). The op amp section has a variable gain of about 3.3-41 times, which, in combination with the 17.8 dB transformer gain, yields a composite gain of 28 dB to 50 dB (26 to 300 times). Transient response of the transformer plus U1 amplifier is excellent. U1 here is one-half an OP275, operating on ±18 V power. Supplies should be well regulated and decoupled close to U1, particularly with low impedance loads. *Care should be used to operate U1 below maximum voltage rating*. The OP275 is rated for maximum supplies of ±22 V.

For best results, passive components should be high quality, such as 1% metal film resistors, a reverse log taper film pot for R2a, and low ESR capacitors for C1 and C3. Microphone phantom powering (see References 2 and 3) can be used, simply by adding the $\pm 0.1\%$ matched 6.81 k Ω resistors and a 48 V dc source, as shown. Close matching of the dc feed resistors is recommended by the transformer manufacturer whenever phantom power is used, to optimize CMR and to minimize the transformer's primary dc current flow (see Reference 4). Note that use of phantom powering has little or no effect on the preamp, since the transformer decouples the CM dc variations at the primary. CMR in an input transformer such as the JT-110K-HPC is typically 85 dB or more at 1 kHz, and substantially better at lower frequencies.

THD + N performance versus frequency of this OP275 mic preamp is shown in the family of curves in Figure 6-4. The test conditions are 35 dB gain, and successive input sweeps resulting in outputs of 0.5, 1, 2, and 5 V rms into 600 Ω . For these distortion tests as well as most of those following throughout these sections, THD + N frequency sweeps at various levels are used for sensitivity to slewing related distortions (see References 5–7), and output loaded tests are used for sensitivity to load related nonlinearities.


Figure 6-4: Transformer-coupled mic preamplifier THD + N (%) versus frequency (Hz) for 35 dB gain, outputs of 0.5, 1, 2, and 5 V rms into 600 Ω

For the OP275 data shown in Figure 6-4, there are three interest regions, a sub-100 Hz region where distortion is largely transformer-related, a 100 Hz–3 kHz region where distortion is lowest, and a greater-than-3 kHz region where it again rises. For most of the spectrum THD + N is $\leq 0.01\%$ for medium outputs, and slightly higher at high frequencies.

The -3 dB bandwidth of this circuit is about 100 kHz, and is dominated by the JT-110K-HPC transformer and its termination network, assuming a 150 Ω source impedance. Conversely, for higher or lower source impedances, the bandwidth will lower or rise in proportion, so application of this circuit should take this into account. For example, capacitor microphone capsules with emitter follower outputs appear as a $\approx 15 \Omega$ source.

Very Low Noise Transformer Coupled Mic Preamp

A high performance low noise mic preamp is shown in Figure 6-5, using a lower ratio transformer, the Jensen JT-16A. This transformer has a lower nominal step up ratio of about 2/1, and is optimized for use with



Figure 6-5: Low noise transformer input 20 dB to 50 dB gain mic preamp

lower noise resistance amplifiers such as the AD797. As can be noted from the figure, the general topology is similar to the previous transformer coupled preamp, but some details allow premium levels of performance.

This preamp has a selectable gain feature, using GAIN switch S1 to alter R2 of the feedback network. This varies U1's gain (and thus overall gain) over a range of 20 dB–50dB, making the preamp suitable for a wide range of uses. With the R2 step values shown, gain is selectable in 5 dB increments. This ranges from 50 dB with R2 (total) = 15 Ω , down to 20 dB with R2 (total) = 588.5 Ω . The transformer provides a fixed gain of about 5.6 dB.

Inasmuch as the AD797 has high precision as well as low distortion audio characteristics, this circuit can be dc-coupled quite effectively. This has the worthwhile advantage of eliminating large electrolytic coupling caps in the gain network and in the output coupling between U1 and V_{OUT} . This is accomplished as follows:

The initial device offset of the AD797 is 80 μ V(max), a factor that allows a relatively simple trim by OFF-SET trimmer R7 to null offset. R7 has a range of ±150 μ V at the AD797 input, with noise well decoupled by C3. With the preamp warmed up well, and working at a midrange gain setting of 35 dB, the offset can be trimmed out. This is best done with the servo temporarily defeated, by grounding test point TP1. Under this condition the V_{OUT} dc level is then trimmed to <1 mV, via R7. This nulls out the residual offset of the AD797, and also ensures that the gain-range network sees minimal dc, which minimizes "pops" with gain changes. The offset shift thereafter with gain is only a few mV, and is of little concern, since the servo circuit of U2A and U2B holds the longer-term dc offset to 100 μ V or less, with little gain interaction. Note that for the gain-change scheme to work properly, S1 *must* be a shorting (make-before-break) type.

THD + N performance versus frequency of this mic preamp is shown in Figure 6-6, for conditions of 35 dB gain, and successive input sweeps resulting in outputs of 0.5, 1, 2, and 5 V rms into 600 Ω . From these data it is essentially clear that the only distortion in the circuit is due to the transformer, which is small and occurs only at the low frequencies. Above 100 Hz, the apparent distortion is noise limited, to the highest frequencies.



Figure 6-6: Low noise transformer input mic preamp THD + N (%) versus frequency (Hz) for 35 dB gain, outputs of 0.5, 1, 2, and 5 V rms into 600 Ω

The -3 dB bandwidth of this circuit is just under 150 kHz, and while this is essentially dominated by the JT-16A transformer and termination, bandwidth does reduce slightly at the highest gain (50 dB). Like the previous transformer coupled circuit, this circuit also assumes a 150 Ω source impedance, and similar application caveats apply.

The basic circuit as shown is single-ended with V_{OUT} taken from R8. However, a transformer can be simply added, as an option for driving balanced lines. When this is done, a nickel core type is suggested, for lowest distortion. One type suitable would be a Jensen JT-11-DM (or similar). It is coupled to the U1 output via a 10 Ω resistor.

Just as shown the circuit is suited for local, higher impedance loads of $1 \text{ k}\Omega$ and more. For very high levels of output drive or to drive long lines, a dedicated high current output driver should be used with U1, as generally described in the "Line Drivers" section. This can be most simply implemented by making U1 a composite amplifier, using a AD797 input section plus a follower-type output stage. A good choice for this would be a BUF04 IC, connected between Pin 6 of the AD797 and the remaining circuitry. The buffer will isolate the U1 stage, allowing it to operate with highest linearity with difficult loads. Note also that $\pm 17 \text{ V}$ supplies won't be necessary with the AD797 unless extreme voltage swings are required. More conventional ($\pm 15 \text{ V}$) supplies will minimize the U1 heating.

References: Microphone Preamplifiers

- C. D. Motchenbacher, F. C. Fitchen, Low-Noise Electronic Design, Wiley, New York, 1973, ISBN 0-471-61950-7.
- 2. G. Bore, "Powering Condenser Microphones," db, June 1970.
- 3. "ANSI Standard 268-15 (Revision 1987, amendments 1989, 1990, 1991)." American National Standards Institute, 11 W. 42nd St., New York, NY, 10036.
- 4. Steve Hogan, "Standard Mic Input Application," Jensen Transformers Application note JT99-0003, November, 1992.
- W. Jung, M. Stephens, C. Todd, "Slewing Induced Distortion & Its Effect on Audio Amplifier Performance—With Correlated Measurement/Listening Results," presented at 57th AES convention, May 1977, AES preprint # 1252.
- W. Jung, M. Stephens, C. Todd, "An Overview of SID and TIM," Parts 1–3, Audio, June, July, August, 1979.
- Walter G. Jung, Audio IC Op Amp Applications, 3rd Ed., Howard W. Sams & Co., 1987, ISBN 0-672-22452-6.
- 8. W. Jung, A. Garcia, "A Low Noise Microphone Preamp with a Phantom Power Option," Analog Devices AN242, November 1992.
- 9. Walt Jung, "Audio Preamplifiers, Line Drivers, and Line Receivers," within Chapter 8 of Walt Kester, **System Application Guide**, Analog Devices, Inc., 1993, ISBN 0-916550-13-3, pp. 8-1 to 8-100.
- 10. Walt Jung, "Microphone Preamplifiers for Audio," Analog Dialogue, Vol. 28, No. 2, 1994, pp. 12–18.

RIAA Phono Preamplifiers

An example of an audio range preamplifier application requiring equalized frequency response is the RIAA phono preamp. While LP record sales have faded with the establishment of new digital media, for completeness equipment is still designed to include phono playback stages. RIAA preamp stages, as amplifiers with predictable, nonflat frequency response, have more general application connotations. The design techniques within this section are specific to RIAA as an example, but they are also applicable to other frequency dependent amplitude designs in general. The techniques are also useful as a study tool, considering the various approaches advanced to optimize the function of high performance gain with predictable equalization (EQ). These last two points make these discussions useful in a much broader sense.

Some RIAA Basics

The RIAA equalization curve (see Reference 1) is shown in Figure 6-7, expressed as it is relative to dc. This curve indicates maximum gain below 50 Hz (f1), with two high frequency inflection points. Above f1, the gain rolls off at 6 dB/octave until a first high frequency breakpoint is reached at 500 Hz (f2). Gain then remains relatively constant until a second high frequency breakpoint is reached at about 2.1 kHz (f3), where it again rolls off at 6 dB/octave through the remainder of the audio region and above.



Use of a low frequency roll-off (f0, not shown) is at the option of the designer. Frequency response can be extended towards dc, or, alternately, rolled off at a low frequency below 50 Hz. When applied, this roll-off is popularly called a "rumble" filter, as it reduces turntable/record related low frequency disturbances, lessening low frequency driver overload. This roll-off may or may not coincide with a fourth time constant (below).

However, gain at the frequencies f1, f2, and f3 describes the basic RIAA curve. In the standard, this is described in terms of three corresponding time constants, T1, T2, and T3, defined as 3180 µs, 318 µs, and 75 µs, respectively (Reference 1, again). The T1–T3 are here described as they correspond to *ascend*-*ing frequency*, the reverse of the terminology in Reference 1 (however, the time constants themselves are identical). In some literature one may occasionally find the frequencies corresponding to T1, T2, and T3 referenced. These exact frequencies can be found simply by the basic relationship of:

$$f = 1/(2 \bullet \pi \bullet T)$$
Eq. 6-4

So, for the three time constants specified, the frequencies are:

 $fl = 1/T1 = 1/(2 \bullet \pi \bullet 3180E - 6) = 50 \text{ Hz}$ $f2 = 1/T2 = 1/(2 \bullet \pi \bullet 318E - 6) = 500 \text{ Hz}$ $f3 = 1/T3 = 1/(2 \bullet \pi \bullet 75E - 6) = 2122 \text{ Hz}$

An IEC amendment to the basic RIAA response adds a fourth time constant of 7950 μ s, corresponding to an f0 of 20 Hz when used (see Reference 2). Use of this roll-off has never been standardized in the US, and isn't treated in detail here.

The characteristic gain in dB for an RIAA preamp is generally specified relative to a 1kHz reference frequency. For convenience in evaluating the RIAA curve numerically, Figure 6-8 is a complete 10 kHz–100 kHz relative decibel table for the three basic RIAA time constants. From these data several key points

Signal Amplifiers

FREQ	VDB(6) ⁽¹⁾	VDB(5) ⁽²⁾
1.000E+01	1.974E+01	-1.684E-01
1.259E+01	1.965E+01	-2.639E-01
1.585E+01	1.950E+01	-4.109E-01
1.995E+01	1.928E+01	-6.341E-01
2.512E+01	1.895E+01	-9.654E-01
3.162E+01	1.847E+01	-1.443E+00
3.981E+01	1.781E+01	-2.103E+00
5.012E+01	1.694E+01	-2.975E+00
6.310E+01	1.584E+01	-4.067E+00
7.943E+01	1.455E+01	-5.362E+00
1.000E+02	1.309E+01	-6.823E+00
1.259E+02	1.151E+01	-8.398E+00
1.585E+02	9.877E+00	-1.003E+01
1.995E+02	8.236E+00	-1.167E+01
2.512E+02	6.645E+00	-1.327E+01
3.162E+02	5.155E+00	-1.476E+01
3.981E+02	3.810E+00	-1.610E+01
5.012E+02	2.636E+00	-1.727E+01
6.310E+02	1.636E+00	-1.828E+01
7.943E+02	7.763E–01	-1.913E+01
1.000E+03	8.338E-07	-1.991E+01
1.259E+03	-7.682E-01	-2.068E+01
1.585E+03	-1.606E+00	-2.152E+01
1.995E+03	-2.578E+00	-2.249E+01
2.512E+03	-3.726E+00	-2.364E+01
3.162E+03	-5.062E+00	-2.497E+01
3.981E+03	-6.572E+00	-2.648E+01
5.012E+03	-8.227E+00	-2.814E+01
6.310E+03	-9.992E+00	-2.990E+01
7.943E+03	-1.184E+01	-3.175E+01
1.000E+04	-1.373E+01	-3.365E+01
1.259E+04	-1.567E+01	-3.558E+01
1.585E+04	-1.763E+01	-3.754E+01
1.995E+04	-1.960E+01	-3.951E+01
2.512E+04	-2.158E+01	-4.149E+01
3.162E+04	-2.357E+01	-4.348E+01
3.981E+04	-2.557E+01	-4.548E+01
5.012E+04	-2.756E+01	-4.747E+01
6.310E+04	-2.956E+01	-4.947E+01
7.943E+04	-3.156E+01	-5.147E+01
1.000E+05	-3.356E+01	-5.347E+01

Notes: ⁽¹⁾ Denotes 1 kHz 0 dB reference

(2) Denotes dc 0 dB reference

Figure 6-8: Idealized RIAA frequency response referred to 1 kHz and to dc

can be observed: If the 1 kHz gain is taken as the zero dB reference, frequencies below or above show higher or lower dB levels, respectively (Note 1, column 2). With a dc 0 dB reference, it can be noted that the 1 kHz gain is 19.91 dB below the dc gain (Note 2, column 3).

Expressed in terms of a gain ratio, this means that in an ideal RIAA preamp the 1 kHz gain is always 0.101 times the dc gain. The constant 0.101 is unique to all RIAA preamp designs following the above curve, therefore it can be designated as " K_{RIAA} ", or:

$$K_{RIAA} = 0.101$$
 Eq. 6-5

This constant logically shows up in the various gain expressions of the RIAA preamp designs following. In all examples discussed here (and virtually all RIAA preamps in general), the shape of the standard RIAA curve is fixed, so specifying gain for a given frequency (1 kHz) also defines the gain for all other frequencies.

It can also be noted from the RIAA curve of Figure 6-7 that the gain characteristic continues to fall at higher frequencies. This implies that an amplifier with unity-gain stability for 100% feedback is ultimately required, which can indeed be true, when a standard feedback configuration is used. Many circuit approaches can be used to accomplish RIAA phono-playback equalization; however, all must satisfy the general frequency response characteristic of Figure 6-7.

Equalization Networks for RIAA Equalizers

Two equalization networks well suited in practice to RIAA phono reproduction are illustrated in Figure 6-9a and 6-9b, networks N1 and N2. Both networks with values as listed can yield with high accuracy the three standard RIAA time constants of 3180 µs, 318 µs, and 75 µs as outlined by network theory (see References 3-6). For convenience, both theoretical values for the ideal individual time constants are shown at the left, as well as closest fit standard "no trim" values to the right. Designers can, of course, parallel and/or series RC values as may be deemed appropriate, adhering to network theory.

There are of course an infinite set of possible RC combinations from which to choose network values, but practicality should rule any final selection. A theoretical starting point for a network value selection can begin with *any* component, but in practice the much smaller range of available capacitors suggests their



Figure 6-9: Two RIAA EQ networks (T1 = 3180 µs, T2 = 318 µs, T3 = 75 µs)

selection first, then resistors, since they have a much broader span of (stock) values. Note that precision film resistors can in fact be obtained (on special order) in virtually *any* value, up to several megohms. The values listed here are those taken as standard from the E96 series.

Very high standards of EQ accuracy are possible, to tolerances of noticeably better than ± 0.1 dB (see for example data from Reference 8, also quoted in 6). In the design process, there are several distinct general aspects of EQ component selection which can impact the ultimate accuracy. These are worth placing in perspective before starting a design.

The *selection tolerance* of the component defines how far an ideal (zero manufacturing tolerance) component deviates from the theoretical value. A good design will seek to minimize this error by using either carefully selected standard values, or series and/or shunt combinations, so as to achieve selection tolerance of less than 1%, preferably zero.

The *manufacturing tolerance* of the component defines how far an otherwise ideal component deviates from its stated catalog value, such as $\pm 1\%$, $\pm 2\%$, and so forth. This can obviously be controlled by tighter specifications, but usually at some premium, particularly with capacitors of $\pm 1\%$ or less. Note that a "hidden" premium here can be long delivery times for certain values. Care should be taken to use standard stock values with capacitors—even to the extent that multiple standard values may be preferable (three times 0.01 µF for 0.03 µF, as an example).

Topology-related parasitics must also be given attention, as they can also potentially wreck accuracy. Amplifier gain-bandwidth is one possible source of parasitic EQ error. However, a more likely error source is the parasitic zero associated with active feedback equalizers. If left uncompensated below 100 kHz, this alone can be a serious error.

In any event, for high equalization accuracy to be "real," once a basic solid topology is selected, the designer must provide for the qualification of components used, by precise measurement and screening, or tight purchase tolerances. An alternative is iterative trimming against a reference standard such as that of Reference 9, but this isn't suited for production. An example is the data of Reference 8, derived with the network of Reference 9. If used, the utility of such a trim technique lies in the reduction of the equipment accuracy burden. While the comparator used needs to have high *resolution*, the accuracy is transferred to the network comparison standard used.

It should be understood that an appropriately selected high quality network will allow excellent accuracy, for example either N1 or N2 with the "closest fit" (single component) values of exact value yield a broadband error of about ± 0.15 dB. Accuracy about three times better than this is achieved with the use of N1 and the composite C2, as noted. The composite C2 is strongly suggested, as without it there is a selection error of about 3%.

It is also strongly recommended that only the highest quality components be employed for use in these networks, for obvious reasons. Regardless of the quality of the remainder of the circuit, it is surely true that the equalization accuracy and fidelity can be no better than the quality of those components used to define the transfer function. Thus only the best available components are used in the N1 (or N2) RC network, selected as follows:

Capacitors—should have close initial tolerance (1%-2%), a low dissipation factor and low dielectric absorption, be noninductive in construction, and have stably terminated low-loss leads. These criteria in general are best met by capacitors of the Teflon, polypropylene and polystyrene film families, with 1%-2% polypropylene types being preferred as the most practical. Types to definitely avoid are the "high K" ceramics. In contrast, "low K" ceramic types, such as "NPO" or "COG" dielectrics, have excellent dissipation factors. See the passive component discussions of Chapter 7 on capacitors, as well as the component-specific references at the end of this section.

Resistors—should also be close tolerance ($\leq 1\%$), have low nonlinearity (low voltage coefficient), be temperature stable, with solid stable terminations and low-loss noninductive leads. Types that best meet these criteria are the bulk metal foil types and selected thick films, or selected military grade RN55 or RN60 style metal film resistor types. See the passive component discussions of Chapter 7 on resistors, as well as the component-specific references at the end of this section.

It should be noted also that the specific component values suggested might not be totally optimum from a low impedance, low noise standpoint. But, practicalities will likely deter using appreciably lower ones. For example, one could reduce the input resistance of either network down to say 1 k Ω , and thus lower the input referred noise contribution of the network. But, this in turn would necessitate greater drive capability from the amplifier stage, and raise the C values up to 1 μ F-3 μ F, where they are large, expensive, and most difficult to obtain. This may be justified for some uses, where performance is the guiding criterion rather than cost effectiveness, or the amplifiers used are sufficiently low in noise to justify such a step. Regardless of the absolute level of impedance used, in any case the components should be adequately shielded against noise pickup, with the outside foils of C1 or C2 connected either to common or a low impedance point.

These very same N1/N2 networks can suffice for both active and passive type equalization. Active (feedback) equalizers use the network simply by returning the input resistor R1 to common, that is jumpering points 1–3, and employing the network as a two-terminal impedance between points 1+3, and 2. Passive equalizers use the same network in a three-terminal mode, placed between two wideband gain blocks.

RIAA Equalizer Topologies

Many different circuit topologies can be used to realize an RIAA equalizer. Dependent upon the output level of the phono cartridge to be used, the 1 kHz gain of the preamp can range from 30 dB to more than 50 dB.

Magnetic phono cartridges in popular use consist of two basic types: moving magnet (MM) and moving coil (MC). The moving magnet types, which are the most familiar, are suitable for the first two circuits described. The moving coil cartridge types are higher performance devices; they are less commonplace but still highly popular.

Functionally, both types of magnetic cartridges perform similarly, and both must be equalized for flat response in accordance with the RIAA characteristic. A big difference in application, however, is the fact that moving magnet types have typical sensitivities of about 1 mV of output for each cm/s of recorded velocity. In moving coil types, sensitivity on the order of 0.1 mV is more common (for a similar velocity). In application then, a moving coil RIAA preamp must have more gain than one for moving magnets. Typically, 1 kHz gains are 40 dB–50 dB for moving coils, but only 30 dB–40 dB for moving magnets. Noise performance of a moving coil preamp can become a critical performance factor however, because of low-output voltage and low impedance involved—typically this is in the range of just 3 Ω –40 Ω . The following circuit examples illustrate techniques that are useful to these requirements.

Actively Equalized RIAA Preamp Topologies

The most familiar RIAA topology is shown in general form in Figure 6-10, and is called an active feedback equalizer, as the network N used to accomplish the EQ is part of an active feedback path (see References 10, 11). In these and all of the following discussions it is assumed that the input from the pickup is appropriately terminated by R_t – C_t , which are selected for flat *cartridge* frequency response driving U1. The following discussions deal with the *amplification* frequency response, given this ideal input signal.

Assuming an adequately high gain amplifier for U1, the gain/frequency characteristics of this circuit are determined largely by the network. The gain of the stage is set by the values of the network N and R_3 , and



Figure 6-10: Active feedback RIAA equalizer

the U1 output is a low impedance, V_{OUT} . The 1 kHz gain of this stage is defined by the RIAA curve and resistors R_1 and R_3 , and is:

$$G = 0.101 \bullet [1 + (R_1/R_3)]$$
 Eq. 6-6

where 0.101 is the constant K_{RIAA} . R_1 is within N; R_4 and R_5 are discussed momentarily.

As noted previously, an ideal RIAA response continues to fall with increasing frequency, and can in fact be less than unity at some high frequency (Figure 6-7, again). But, the basic U1 topology of Figure 6-10 can't achieve this, as the minimum gain seen at the output of U1 approaches unity at some (high) parasitic zero frequency, where the network equivalent series capacitive impedance of C_1 and C_2 is equal to R_3 . At this zero frequency, the response from U1 simply levels off and ceases to track the RIAA curve.

However, in terms of practical consequence the error created by this zero may or may not be of significance, dependent upon where the zero falls (as determined by gain). If well above audibility (i.e., ≥ 100 kHz), it will introduce a small equalization error at the upper end of the audio range. For example, if it falls at 100 kHz, the 20 kHz error is only about 0.3 dB. Fortunately, this error is easily compensated by a simple low-pass filter after the amplifier, R_5-C_3 . The filter time constant is set to match the zero T_4 , which is:

$$T4 = R_3 \bullet C_{EQUIV} \qquad Eq. 6-7$$

where R_3 is the value required for a specific gain in the design.

 C_{EOUIV} is the series equivalent capacitance of network capacitors C_1 and C_2 , or:

$$C_{EQUIV} = (C_1 \bullet C_2) / (C_1 + C_2)$$
Eq. 6-8

Here the C_{EQUIV} is 7.6 nF and $R_3 200 \Omega$, so T4 = 1.5µs The product of R5 and C4 are set equal to T4, so picking a R5 value solves for C4 as:

$$C4 = T4/R5$$
 Eq. 6-9

The 1.5 μ s R₅-C₄ time constant is realized with R₅ = 499 Ω and C4 = 3 nF. This design step increases the output impedance, making it more load susceptible. This should be weighed against the added parts and loading. In general, R₅ should be low, i.e., $\leq 1 \ k\Omega$.

In some designs, a resistor R_4 (dotted in Figure 6-10) may be used with N (for example, for purposes of amplifier stability at a gain higher than unity). With R_4 , T4 is calculated as:

$$T4 = (R_3 + R_4) \bullet C_{EOUIV}$$
 Eq. 6-10

The R_5 - C_3 product is again chosen to be equal to this T4 (more on this below).

The next two schematics illustrate variations of the most popular approach to achieving a simple RIAA phono preamp, using active feedback, as just described. Figure 6-11 is a high performance, dc-coupled version using precision 1% metal film resistors and 1% or 2% capacitors of polystyrene or polypropylene type. Amplifier U1 provides the gain, and equalization components R1-R2-C1-C2 form the RIAA network, providing accurate realization with standard component values. N1 is the network, with 1 and 3 common.



Figure 6-11: A dc-coupled active feedback RIAA moving magnet preamp

As mentioned, input RC components R_t - C_t terminate the moving magnet cartridge with recommended values (shown as typical). In terms of desired amplifier parameters for optimum performance, they are considerably demanding. For lowest noise from a cartridge's inductive source, the amplifier should have an input voltage noise density of 5 nV/ \sqrt{Hz} or less (favoring a bipolar), and an input current noise density of 1 pA/ \sqrt{Hz} or less (favoring a FET). In either case, the 1/F noise corner should be as low as possible.

For bipolar-input amplifiers, dc input-bias current can be a potential problem when direct coupling to the cartridge, so in this circuit only a very low input bias current type is suggested. If a bipolar input amplifier is used for U1, it should have an input current of <<100 nA for minimum dc offset problems (assuming a typical phono cartridge of $\approx 1 \text{ k}\Omega$ resistance). Examples are the OP27, OP270 families. FET-input amplifiers generally have negligible bias currents but also tend typically to have higher voltage noise. FET-input types useful for U1 are the AD845 and OP42, even though their voltage noise is not as low as the best of the bipolar devices mentioned. On the plus side, they both have a high output current and slew rate, for low distortion driving the feedback network load (approximately the R3 value at high frequencies). Of the two, the OP42 has lower noise, the AD845 higher output current and slew rate.

For high gain accuracy at high stage gains, the amplifier should have a high gain-bandwidth product; preferably >5 MHz at audio frequencies. Because of the 100% feedback through the network at high frequencies, the U1 amplifier must be unity-gain-stable. To minimize noise from sources other than the amplifier, gain resistor R_3 is set to a relatively low value, which generates a low voltage noise in relation to the amplifier.

RIAA accuracy is quite good using the stock equalizer values. A PSpice simulation run is shown in Figure 6-12 for the suggested gain of 34 dB. In this expanded scale plot over the 20 kHz–20 kHz range, the error relative to the 1 kHz gain is less than ± 0.1 dB.

As can be noted from Figure 6-12, the relative amplitude is expanded, to easily show response errors. A perfect response would be a straight line at 0 dB, meaning that the circuit under test had exactly the same gain as an ideal RIAA amplifier of the same 1 kHz gain. This high sensitivity in the simulation is done via the use of a feature in PSpice allowing the direct entry of Laplace statements (see Reference 10). With this evaluation tool, the ideal transfer function of an RIAA equalizer can be readily generated. The key parameters are the three time constants described above, and the ideal dc gain.



Figure 6-12: Relative error (B) versus frequency for dc-coupled active feedback RIAA moving magnet preamp, gain of 34 dB

The syntax to enable this mode of comparison is contained in the listing of Figure 6-13, which is the PSpice CIR file for the circuit of Figure 6-11. The Laplace details are all contained within the dotted box, and need only the editing of one value, "ENORM," for gain normalization from one circuit to another (see boldface). In this case ENORM is set to 490.7, to match the ideal R1 and R3 values of Figure 6-13. When the analysis is run, a difference display of the circuit-under-test and the ideal outputs (i.e., VdB(56)–VdB(5)) shows the relative response (Figure 6-12). Vertical axis scaling is easily adjusted for sensitivity, and is ±300 mB as displayed in Figure 6-12.

```
RIAA34LP: 34 dB gain RIAA preamp with AD845
.OPT ACCT LIST NODE OPTS NOPAGE LIBRARY
.AC DEC 10 10 100KHZ
.LIB D:\PS\ADLIB\AD_RELL.LIB
.PRINT AC VDB(5) VDB(56)
. PROBE
VIN 1 0 AC 1E-3
VCC
    52 0 +15V
VEE 53 0 -15V
* ----- V(5) = idealized RIAA frequency response ------
* Uses Laplace feature of PSpice Analog Behavioral option
* for frequency response reference.
* ENORM = ideal U1 DC gain = 1+(R1/R3) Use ideal values for R1, R3
\star T1 - T3 are time constants desired (in \mu s)\,.
* Input = node 1, Laplace Output = node 5
.PARAM ENORM = \{490.7\}
.PARAM T1 = {3180} ; Reference RIAA constants, do not alter!
.PARAM T2 = {318} ; Reference RIAA constants, do not alter!
.PARAM T3 = {75} ; Reference RIAA constants, do not alter!
                     ; Reference RIAA constants, do not alter!
ERIAA 5 0 LAPLACE {ENORM*V(1)}={(1+(T2*1E-6)*S)/((1+(T1*1E-6)*S)*(1+(T3*1E-6)*S))}
RDUMMY5 5 0 1E9
*
 _____
*
    (+) (-) V+ V- OUT
XU3 1 21 52 53 55 AD845
* Active values Theoretical values
R1 55 21 97.6K
                       ; 97.9k
                       ; 7.8931563k
R2 21 8 7.87K
C1 55 8 30NF
                       ; 30nF
C2 21 8 10.3NF
                       ; 10.2881nF R3 21 0 200
                                                         ; 199.9148
C3 55 100 10E-6
R6 100 0 100K R5 100 56 499
C4 56 0
         3.0000E-9
.END
```

Figure 6-13: An example PSpice circuit file that uses the Laplace feature for ideal RIAA response comparison

The 1 kHz gain of this circuit can be calculated from Eq. 6-6. For the values shown, the gain is just under 50 times (\approx 34 dB). Higher gains are possible by decreasing R₃, but gains >40 dB may show increasing equalization errors, dependent upon amplifier bandwidth. For example, R₃ can be 100 Ω for a gain of

about 100 times (\approx 40 dB). Note that if R₃ is changed to 100 Ω , C₄ should also be changed to 1.5 nF, to satisfy Eq. 6-9.

Dependent upon the amplifier in use, this circuit is capable of very low distortion over its entire range, generally below 0.01% at levels up to 7 V rms, assuming ± 15 V supplies. Higher output with ± 17 V supplies is possible, but will require a heat sink for the AD845. U2 is an optional unity-gain buffer useful with some op amps, particularly at higher gains or with a low-Z network. But this isn't likely to be necessary with U1 an AD845.

For extended low-frequency response, C_3 and R_6 are the large values, with C_3 preferably a polypropylene film type. If applied, the alternate values form a simple 6 dB per octave rumble filter with a 20 Hz corner. As can be noted from the figure's simplicity, C_3 is the only dc blocking capacitor in the circuit. Since the dc circuit gain is on the order of 54 dB, the amplifier used must be a low offset-voltage device, with an offset voltage that is insensitive to the source. Since these preamps are high gain, low level circuits (\geq 50 dB of gain at 50 Hz/60 Hz), supply voltages should be well regulated and noise-free, and reasonable care should be taken with the shielding and conductor routing in their layout.

Alternately, an inexpensive ac-coupled form of this circuit can be built with higher bias current, low noise bipolar op amps, for example the OP275, $I_B = 350 \text{ nA}(\text{max})$, which would tend to make direct coupling to a cartridge difficult. This form of the circuit is shown in Figure 6-14, and can be used with many unity gain stable bipolar op amps.



Figure 6-14: An ac-coupled active feedback moving magnet RIAA preamp

Here input ac coupling to U1 is added with C_5 , and the cartridge termination resistance R_t is made up of the R_6 - R_7 parallel equivalent. R_3 of the feedback network is ac-grounded via C_4 , a large value electrolytic. These measures reduce the dc offset at the output of U1 to a few mV. Nearest 5% values are also used for the network components, making it easily reproducible and inexpensive. C_3 is a nonpolar electrolytic type, and the R_3 - C_4 time constant as shown provides a corner frequency of ≤ 1 Hz at the 34 dB gain.

Frequency response of this version (not shown) isn't quite as good as that of Figure 6-11, but is still within ± 0.2 dB over 20 Hz–20 kHz (neglecting the effects of the low frequency roll-off). If a tighter frequency response is desired, the N1 network values can be adjusted. With a higher rated maximum supply voltage for the OP275, the power supplies of this version can be ± 21 V if desired, for outputs up to 10 V rms.

There is another, very useful variation on the actively equalized RIAA topology. This is one that operates at appreciably higher gain and with lower noise, making it suitable for operation with higher output moving coil (MC) cartridges. In this design example, shown in Figure 6-15, the basic circuit is used is quite similar to that of Figure 6-11. The lower R_t and C_t values shown are typical for moving coil cartridges. They are of course chosen per the manufacturer's recommendations (in particular the resistance).



Figure 6-15: A low noise dc-coupled active feedback RIAA moving coil preamp with 45 dB of gain

To make it suitable for a high-output MC cartridge, a very low-noise FET op amp is used for U1, the AD745. The AD745 is stable at a minimum gain of five times, as opposed to the unity-gain stable op amps of the prior examples. This factor requires a modification to gain resistors R1-R3. This is the inclusion of an extra resistor, R4. With the ratio shown, R3 and R4 form a 5/1 voltage divider for the voltage seen at the bottom of network N (the R1-R2-C2 node). This satisfies U1's gain-of-five stability requirement.

In this gain setup, R3 is still used for the gain adjustment, and R1-R2-C1-C2 still form the basic N1 RIAA network. With R4 used, Eq. 6-9 is used to calculate the T4 time constant. With C4 chosen as a standard value, R5 is then calculated. With these N1 network values and a 45 dB 1 kHz gain, R3 is 56.2 Ω , which is still suitable as a low noise value operating with either an AD745 or an OP37 used for U1.

Some subtle points of circuit operation are worth noting. The dc gain of this circuit is close to 1800, which can result in saturation of U1 if offset isn't sufficiently low. Fortunately, the AD745 has a maximum offset of 1.5 mV over temperature, making the output referred offset always less than 3 V. While this may limit the maximum output swing some due to asymmetrical clipping, 5 V rms or more of swing should be available operating from ± 15 V supplies. Coupling capacitor C3 decouples the dc output offset at U2, so any negative consequences of dc-coupling the U1 gain path are minimal.

For minimal loading of U1 and maximum linearity at high gains, the unity-gain buffer amplifier U2 is used, a BUF04. The BUF04 is internally configured for unity-gain operation, and needs no additional components. Note that this buffer is optional, and is not absolutely required. Other buffer amplifiers are discussed later in this chapter.

This Figure 6-15 circuit was analyzed with PSpice using the Laplace comparison technique earlier described, and the results are displayed in Figure 6-16. As was true previously, the vertical scaling of this



Figure 6-16: Relative error (B) versus frequency for dc-coupled active feedback RIAA moving coil preamp, gain of 45 dB (simulation)

display is very sensitive; $\pm 300 \text{ mB}$ (or $\pm 0.3 \text{ dB}$). Thus, placed in context, gain errors relative to 1 kHz over 20 kHz–20 kHz are extremely small, $\approx 0.1 \text{ dB}$. Lab measurements of the circuit were also consistent with the simulation. Of course in terms of audible effects, errors of $\pm 0.1 \text{ dB}$ or less aren't likely to be apparent.

Distortion/noise measurements of the circuit are essentially dominated by noise (as opposed to actual distortion) measuring $\sim 0.01\%$ THD + N or less, over output levels ranging from 0.5 to 5 V rms, from 20 Hz–20 kHz. Of course, as with any high gain circuit, layout and lead dress into the circuit are extremely critical to noise, and must be arranged for minimum susceptibility. Supply voltages must be low in noise, and well regulated.

This exercise has illustrated both the basic design process of the active RIAA equalizer, as well as a convenient SPICE analysis method to optimize the design for best frequency response. It is not suggested that the exact network values shown of the examples are the only ones suitable. To the contrary, great many sets of values can be used with success comparable to that shown above.

This final active equalizer circuit example is the best of the bunch, and has a virtue of being easily adapted for other operating conditions; i.e., higher gain, other networks, and so forth. For example, note that even lower noise MC operation is possible, by using the $\leq 1 \text{ nV}/\sqrt{\text{Hz}}$ AD797 for U1, and scaling the N1 RC components further downward. This will have the desirable effect of making R3 lower than 50 Ω , which minimizes the R1–R4 network's noise. Note that gains of 50 dB or more are also possible, suitable for very low output moving coil cartridges (given suitable attention to worst-case U1 offsets).

Passively Equalized RIAA Preamp Topologies

Another RIAA design approach is the so-called *passively equalized* preamp (see Reference 11). This topology consists of two high quality, wideband gain blocks, separated by a three terminal passive network, N (N can be either network N1 or N2). The gain blocks are assumed very wide in bandwidth, so in essence the preamp's entire frequency response is defined by the passive network, thus the name passively equalized.