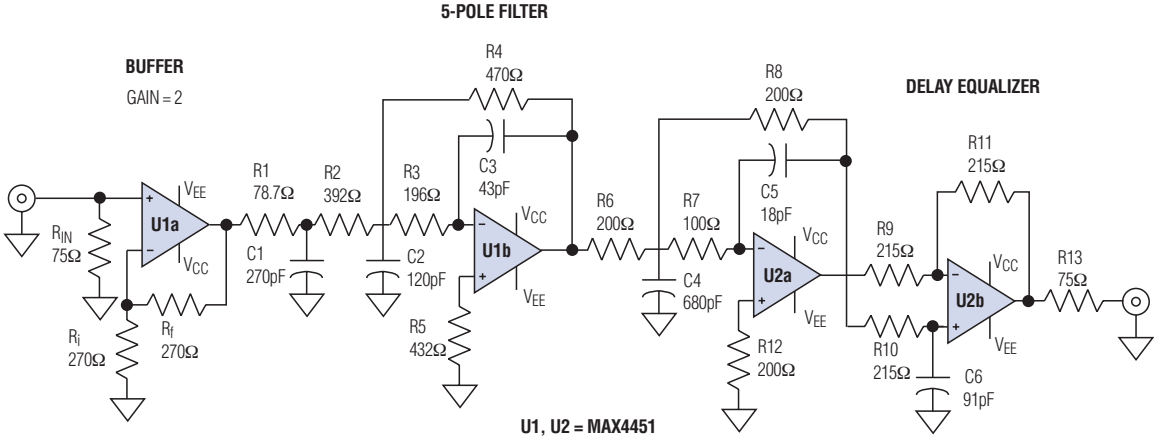
DALLAS SEMICONDUCTOR MAXIM

Engineering Journal

Volume Forty-Nine



This schematic represents a 5-pole, 5.75MHz Butterworth filter for ITU-601 antialiasing, using a Rauch circuit with a delay equalizer. (See article inside, page 3.)

News Brief

MAXIM REPORTS REVENUES AND EARNINGS FOR THE FOURTH QUARTER AND FISCAL YEAR 2003 AND DOUBLES QUARTERLY DIVIDEND FROM $0.04 PER SHARE TO $0.08 PER SHARE

Maxim Integrated Products, Inc., (MXIM) reported net revenues of $295.0 million for its fiscal fourth quarter ending June 28, 2003, an $8.8 million increase over the $286.2 million reported for the third quarter of fiscal 2003. Diluted earnings per share were $0.24 for the fourth quarter on net income of $81.7 million compared to diluted earnings per share of $0.23 on net income of $77.6 million reported for the third quarter of fiscal 2003. Dallas Semiconductor contributed 18.6% of the Corporation's fourth quarter earnings per share.

For the 2003 fiscal year, Maxim reported net revenues of $1.153 billion compared to $1.025 billion for last year, a 12.5% increase. Net income for the 2003 fiscal year was $309.6 million compared to $259.2 million reported for fiscal 2002. Dallas Semiconductor contributed 16.8% of the Company's fiscal 2003 net income. Diluted earnings per share for the Company grew 24.7% from the $0.73 per share reported for fiscal 2002 to $0.91 per share in fiscal 2003.

During the fourth quarter of fiscal 2003, the Company repurchased 1.3 million shares of its common stock for $46.7 million and acquired $28.4 million of property and equipment. Year-end cash, cash equivalents, and short-term investments increased $107.3 million during the quarter to $1.2 billion and increased $398.5 million or 52.1% from the end of fiscal 2002. Accounts receivable increased by $2.9 million in the fourth quarter to $126.8 million on increased revenues, and inventories decreased $2.6 million to $121.2 million.

Research and development expense increased to $67.2 million or 22.8% of net revenues in the fourth quarter, compared to $66.8 million or 23.3% of net revenues in the third quarter. The increase in research and development spending was the result of hiring additional engineers and increased spending to support new product development efforts. Selling, general and administrative expenses remained relatively unchanged during the quarter.

Fourth quarter bookings were approximately $313 million. Turns orders received in the quarter were $161 million (turns orders are customer orders that are for delivery within the same quarter and may result in revenue within the same quarter if the Company has available inventory that matches those orders). Fourth quarter ending backlog shippable within the next 12 months was approximately $227 million, including approximately $199 million requested for shipment in the first quarter of fiscal 2004. The Company's third quarter ending backlog shippable within the next 12 months was approximately $219 million, including approximately $196 million that was requested for shipment in the fourth quarter.

Jack Gifford, Chairman, President, and Chief Executive Officer, commented on the fiscal year: "Maxim's net revenues for the 2003 fiscal year increased 12% from fiscal 2002 net revenues, and our diluted earnings per share for the year were up 25% for the same period. Revenues for Dallas Semiconductor grew 4% year over year, and their earnings per share grew 42%. As a comparison, Maxim without Dallas Semiconductor saw revenue growth of 15% year over year and earnings per share growth of 21%. The Company's gross margins and operating margins improved year over year as we controlled costs and improved efficiencies. Unlike many semiconductor companies, Maxim has been highly profitable during the industry downturn, in part because of the success of our proprietary products and in part because of our effective expense management."

Mr. Gifford continued: "Maxim remained committed to investing in new product and process development this year, with an annual R&D expense of 23.6% of net revenues. Our continued focus on R&D resulted in the introduction of 1959 engineering man-months of new products this past product announcement year, and we have had important design wins in the next generation of cellular phones, digital cameras, 802.11 wireless LAN, high-speed data networks, flat panel displays, and other new electronics. We believe that Maxim continues to be well positioned to gain market share as corporate and consumer spending levels improve and our customers become more confident."

Mr. Gifford concluded: "With consideration for our stockholders and the Company's profitability, strong cash position, and business outlook, the Company's Board of Directors has increased this quarter's dividend from $0.04 per share to $0.08 per share. Payment will be made on September 5, 2003 to stockholders of record on August 22, 2003."

Active filters for video

Originally, video filters were passive LC circuits surrounded by amplifiers. Smaller, more efficient designs can currently be achieved by combining the amplifier with an RC filter. Sensitivity analysis and predistortion methods developed in the 1960s have, moreover, overcome the poor performance that gave early video filters a bad reputation.

High-performance op amps and specialized software for the PC enable the design of wide-bandwidth active filters, but those advantages do not address the requirements of any specific application. For video filters, the particular application and signal format add nuance to each circuit design. The two major video applications follow.

Antialiasing filters: These devices are placed before an analog-to-digital converter (ADC) to attenuate signals above the Nyquist frequency, which is one half the sample rate of the ADC. These filters are usually designed with the steepest possible response to reject everything above the cutoff frequency[1]. For ITU-601 applications and others, such performance is achieved using analog filters combined with digital filters and an oversampling ADC. For applications such as PC graphics, very little filtering is required.

Reconstruction filters: Also called (sinx)/x or zero-order-hold correctors, these filters are placed after a digital-to-analog converter (DAC) to remove multiple images created by sampling, though not to remove the DAC clock. Reconstruction filters are seldom as selective as antialiasing filters, because the DAC's hold function also acts as a filter—an action that lowers the required selectivity, but introduces loss in the response. The available video formats are RGB, component video, composite video, and RGB PC graphics.

All applications and formats require a video filter to be "phase linear," a condition specified by the parameter called group delay (delay versus frequency). The degree of phase linearity required depends on the application and the video format. For example, antialiasing filters and component formats are more tightly specified than are reconstruction applications and composite video. Requirements for the various applications and formats are specified by NTSC, PAL/DVB, ITU, SMPTE, and VESA.

This article compares different filters to determine the optimum design for a given application or format. Rauch and Sallen-Key realizations are compared for their GBW-to-cutoff ratios, using predistortion and element sensitivity techniques to achieve accuracy in the design. Those filters to be considered are:

- An ITU-601 antialiasing filter
- A 20MHz antialiasing and reconstruction filter
- An HDTV reconstruction filter

Filters and their characteristics

Whether used for antialiasing or reconstruction, the filter must have a lowpass characteristic to pass the video frame rate. One should, therefore, be wary of AC-coupling. Lowpass filters are categorized by their amplitude characteristic or by the name of the polynomial that describes it (Bessel, Butterworth, Chebyshev, or Cauer). **Figure 1** shows these characteristics normalized to a 1-rad bandwidth. Typically, a filter with the best selectivity and the minimum number of poles (to minimize cost) would be chosen, but the additional need for phase linearity limits the available choices.

Phase linearity and group delay

A filter's phase linearity is specified as envelope delay or group delay (GD) versus frequency. A flat group delay indicates all frequencies are delayed by the same amount, which preserves the shape of the waveform in the time domain. Thus, absolute group delay is not as important as the variation in group delay. A separate specification called channel-to-channel variation, which is specified as "time coincidence," should not be confused with group delay.

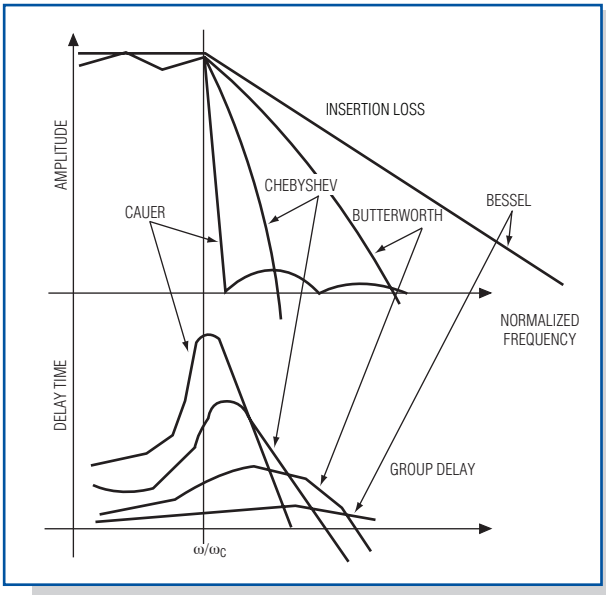


Figure 1. Amplitude and group delay vs. frequency for various filter types are normalized to a 1-rad bandwidth.

Though not desirable for video, how much group-delay variation is acceptable, and why? The answer depends on the application and the video format. For example, ITU-470 specifies group delay very loosely for composite video. However, ITU-601 specifies it tightly to ensure generational stability, both for MPEG-2 compression and to control phase jitter before serialization. So, what filter characteristics are considered necessary to ensure phase linearity?

The group-delay curves in Figure 1 show a peak near the cutoff frequency ($\omega/\omega_c = 1$). That is a problem caused by the steep phase change near the cutoff frequency. To get an idea of scale, a 3-pole, 6MHz Butterworth filter has a group-delay variation of 20ns–25ns over its bandwidth. Increasing the number of poles or the filter's selectivity increases that variation. Other, more exotic filters[2] used to minimize group-delay variation include Bessel, phase approximation, Thompson-Butterworth, and LeGendre. Nevertheless, the Butterworth characteristic is most often used for video.

Group-delay problems with component video

All formats and applications are sensitive to group-delay variation. The degree of sensitivity depends on the number of signals and their bandwidths. Composite NTSC/PAL has only one signal, with group delay specified in ITU-470. Those requirements are easily met. RGB and component video each have multiple signals. The RGB signals have equal bandwidths while component-video signals do not, making group-delay matching easy with RGB, but difficult with component video.

Because Pb and Pr signals have half the bandwidth of the luma (Y) signal, their group delay is double that of the Y signal[3]. One solution is to slow down the Y signal by adding delay stages. Another solution is to equalize bandwidths by doubling the sample rates of Pb and Pr, which raises the 4:2:2 sampling rate to 4:4:4[4], allowing the signal to be treated as RGB. The additional Pb and Pr samples are discarded during antialiasing or averaged in reconstruction applications.

The other component-video format, S-VHS, can be somewhat confusing. The Y channel is the same as in YPbPr, but the chroma signal (C) looks like it should be bandpass filtered rather than lowpass filtered. As for YPbPr signals, bandpass filtering causes group-delay and timing problems and, therefore, should not be implemented. Unless analog encoding is done, Y and C can be lowpass filtered with the same filter. S-VHS is more forgiving of bandwidth than of problems caused by trying to equalize the delay. S-VHS is typically seen in reconstruction applications, for which the main concern is correct timing between Y and C.

Choosing an op amp

After choosing a filter characteristic, the next step is to implement it with an actual circuit. The most commonly used, single-op-amp circuits are the Sallen-Key configuration in noninverting form and the Rauch configuration in inverting form. An important consideration for op amps operating in the wide bandwidths of video applications is the minimum gain-bandwidth (GBW). Video signals are large, typically $2V_{P-P}$, so the large-signal GBW is referenced. This parameter is not to be confused with the $2V_{P-P}$ 0.1dB GBW, which is much lower.

For filter circuits, how much larger than the filter's cutoff frequency does the op-amp GBW have to be? For a Rauch (inverting) filter, the phase argument of the characteristic is:

$$\text{Arg}[K(j\omega)]_{inv} = -(\omega_c/GBW_{rad})(1+Rf/Ri) \quad \text{(Eq 1)}$$

For a Sallen Key (noninverting) filter:

$$\text{Arg}[K(j\omega)]_{noninv} = \pi -(\omega_c/GBW_{rad})(1+Rf/Ri) \quad \text{(Eq 2)}$$

where R_f and R_i are the gain-set resistors in ohms, GBW_{rad} is the op amp's gain-bandwidth product, and ω_c is the filter's cutoff frequency in radians per second. Set the gain by introducing values for R_f and R_i[5], and solve for (ω_c/GBW_{rad}). A unity-gain Rauch circuit has $R_f/R_i = 1$, and a Sallen-Key circuit has $R_f/R_i = 0$. Thus, for the same phase error, a Sallen-Key requires half the GBW of a Rauch circuit. As the required gain increases, they converge and leave little advantage for the Sallen-Key in terms of GBW, but other issues must be considered as well.

Predistortion, bandwidth, Q, and element sensitivity

Anything less than an infinite GBW_{rad}/ω_c ratio causes the closed-loop poles of a filter to move. That is why an actual filter often exhibits a lower bandwidth (ω_c) than does the paper design[6]. This can be compensated for by increasing the design bandwidth, which is known as predistortion. Formulas for the Sallen-Key and Rauch circuits (listed in **Tables 1** and **2**) allow us to calculate a design bandwidth that provides the actual bandwidth needed. Component tolerance must then be taken into account.

To determine component tolerance, a sensitivity function[7] is needed: S_X^Y gives the ratio between a change in the value of part X and the consequent change in parameter Y. For example, Table 1 shows that the Q in a Sallen-Key circuit (vs. a Rauch circuit) has a large sensitivity to variations in C1 and C2. That means a Sallen-Key is less tolerant of parasitics than is a Rauch. The point is that S_X^Y lets the effect be predicted, and then the design can be created accordingly. Next, some typical designs are considered.

Table 1. Component sensitivities including BW and Q predistortion formulas (Sallen-Key realization, ω_0 = 1rad/sec)

Sensitivity Function S_X^Y	Gain K = 3 - 1/Q (R1 = R2 = C1 = C2 = 1)	Gain K = 1 (R1 = R2 = 1)	Gain K = 2 (R1 = C1 = 1)
S_x^ω (x = R1, R2, C1, C2)	-1/2	-1/2	-1/2
S_K^Q	14	50	10
S_{R1}^Q	4.5	0	4.5
S_{R2}^Q	-4.5	0	-4.5
S_{C1}^Q	9.5	1/2	5.5
S_{C2}^Q	-9.5	-1/2	-5.5
S_{Ra}^K	-9/14	N/A	-1/2
S_{Rb}^K	9/14	N/A	1/2
ω_C (actual)	ω_C (design)[1 - 1/2(3 - 1/Q)$^2\omega_C$ / GBW]	ω_C (design)[1 - ω_CQ / GBW]	—
Q (actual)	Q (design)[1 + 1/2(3 - 1/Q)$^2\omega_C$ / GBW]	Q (design)[1 + ω_CQ / GBW]	—

Table 2. Component sensitivities including BW and Q predistortion formulas (Rauch realization, ω_0 = 1rad/sec)

Sensitivity Function S_X^Y	Gain K = 1 (R1 = R2 = R3 = 1)	Gain K = 2 (R1 = 1, R3 = H_O, R2 = (H_O/ 1 + H_O))	Gain K = 2 (C1 = 1, C2 = C1 / 100)
S_x^ω (x = R2, R3, C1, C2, S_{R1}^ω = 0)	-1/2	-1/2	-1/2
S_{R1}^Q	1/3	1/3	1/3
S_{R2}^Q	-1/6	0	0
S_{C1}^Q	1/2	1/2	1/2
S_{C2}^Q	-1/2	-1/2	-1/2
S_{R3}^K	1	1	1
S_{R1}^K	-1	-1	-1
S_{R3}^Q	1/6	0	0
ω_C (actual)	ω_C (design)[1 - 3ω_CQ / 2GBW]	—	—
Q (actual)	Q (design)[1 + 3ω_CQ / 2GBW]	—	—

Design of antialiasing filters

For antialiasing filters, selectivity is determined by a template for ITU-601 like the one in **Figure 2**. The specified bandwidth is 5.75MHz ±0.1dB, with an insertion loss of 12dB at 6.75MHz and 40dB at 8MHz, and with a group-delay variation of ±3ns over the 0.1dB bandwidth. Such performance is too difficult for an analog filter alone, but 4x oversampling modifies the requirements to 12dB at 27MHz and 40dB at 32MHz.

Using software or normalized curves[8], one can find that a 5-pole Butterworth filter with -3dB bandwidth of 8.45MHz satisfies the requirement for selectivity, though not for group delay. For the latter, a delay stage is needed, for which the important op-amp parameter is the 0.1dB, $2V_{P-P}$ bandwidth[9]. That number should be used in equations 1 and 2 to get an accurate design. A schematic for this application, with curves showing its gain and group-delay characteristics, is based on 4x oversampling (**Figures 3a** and **3b**).

PC video is considered next. VESA does not specify templates for antialiasing or reconstruction filters. The XGA resolution (1024 x 768 at 85Hz) has a sampling rate of 94.5MHz and a Nyquist frequency of 47.25MHz. For >35dB attenuation at the Nyquist frequency, a Rauch realization of a 20MHz, 4-pole Butterworth filter (**Figures 4a** and **4b**) is used. Again, the MAX4450/4451 are chosen for their excellent transient response and large-signal bandwidth (175MHz at $2V_{P-P}$).

Reconstruction filters

Reconstruction filtering after a DAC is among the more poorly understood applications. Some designers think reconstruction filters are introduced to remove the sample

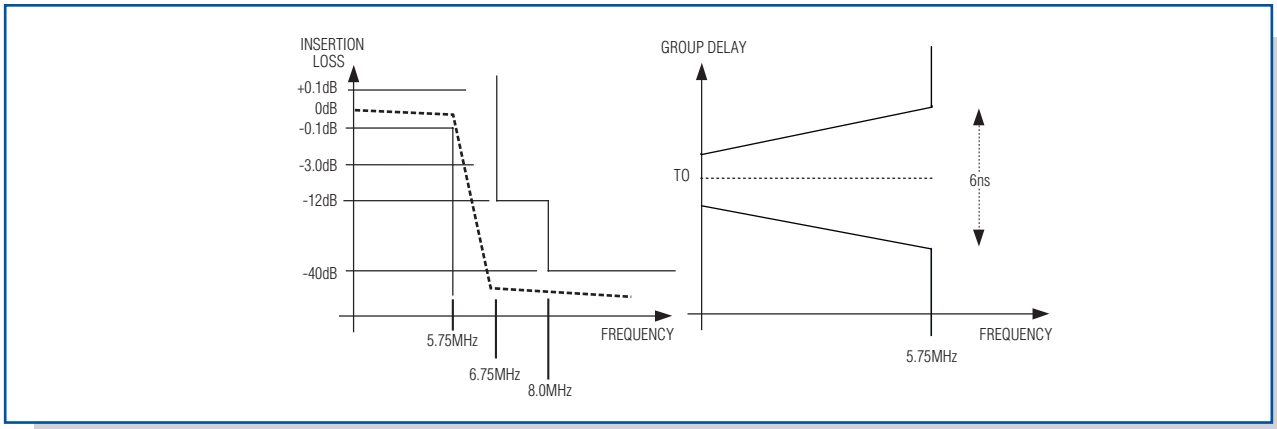


Figure 2. This filter template illustrates antialiasing requirements in accordance with the ITU-R BT.601-5 standard.

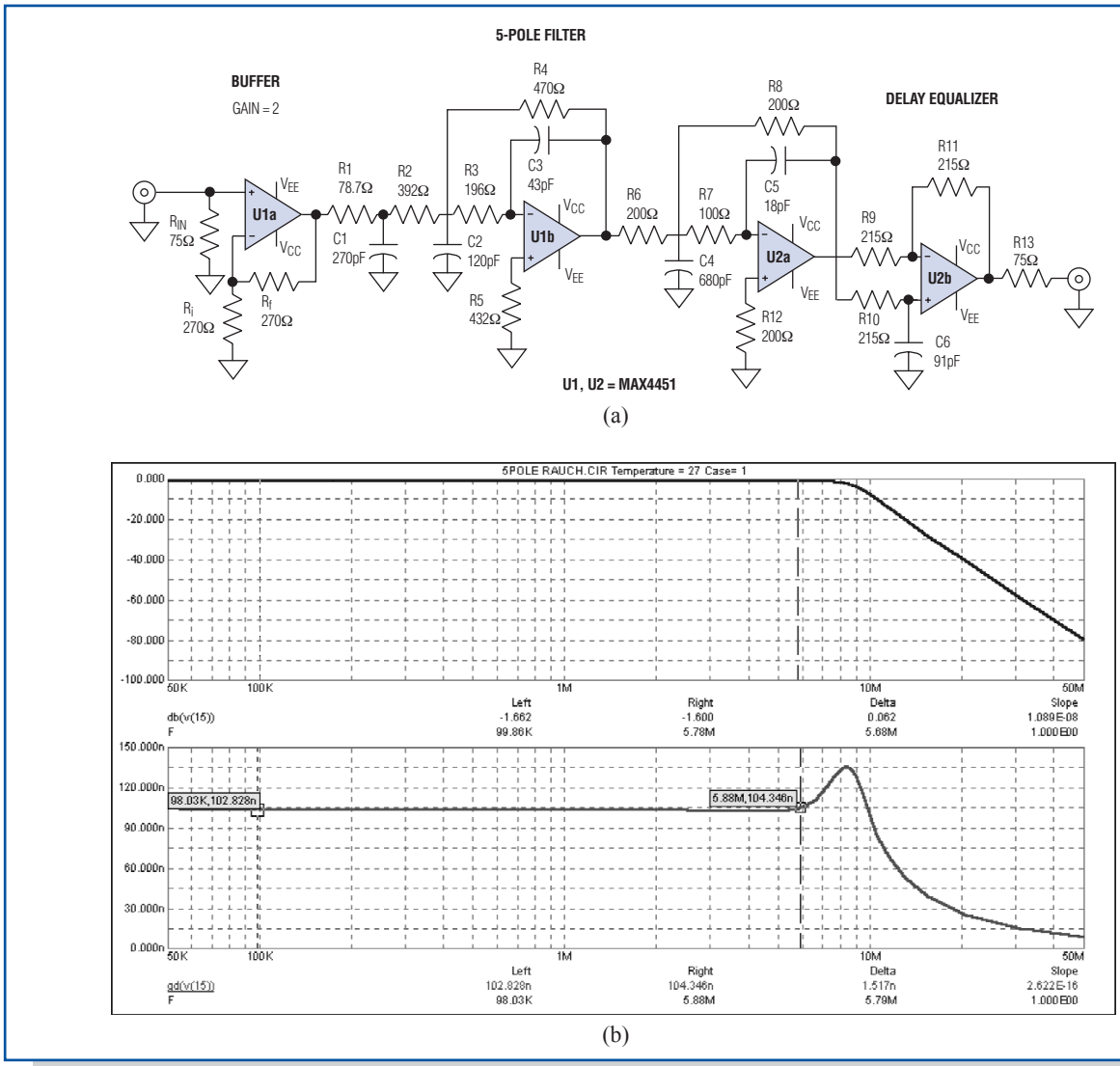


Figure 3. This schematic (a) and output response (b) represent a 5-pole, 5.75MHz Butterworth filter for ITU-601 antialiasing, using a Rauch circuit with a delay equalizer.

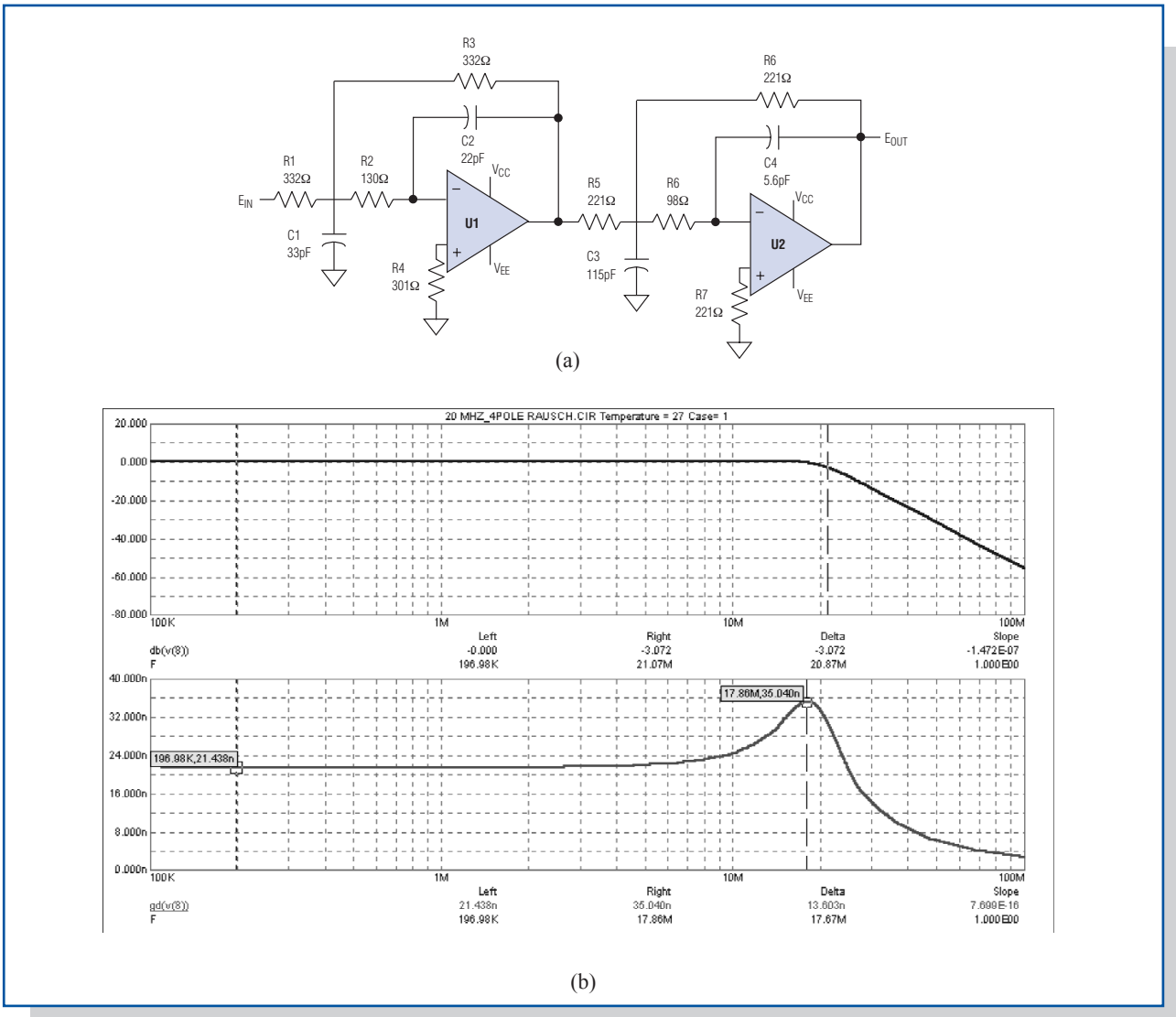


Figure 4. This schematic (a) and output response (b) depict a 4-pole, 20MHz Butterworth filter for XGA graphics antialiasing, using a Rauch circuit.

clock, but nothing is further from the truth. When a signal is sampled, the samples are composed of multiple recurring signal images centered on harmonics of the sample clock. A reconstruction filter removes all but the baseband sample. If the antialiasing filter has served its purpose, the DAC output looks like image A in **Figure 5**, and then all samples to the right of it should be removed. Thus, reconstruction is similar to antialiasing except that, because each sample exists only for an instant, the DAC holds each for one clock period, thereby creating the familiar staircase approximation to a sloping line.

The hold function corresponds to a digital filter whose characteristic[10] is similar to that of a Butterworth or Bessel filter (**Figure 6**). Notice that the response is decreased by 4dB at half the sample frequency. The second objective of a reconstruction filter is to restore that loss, which requires an amplitude equalizer like the circuit shown in **Figure 7a**. The equalizer is based on a delay stage and has a response like a Bessel filter. It can be designed from the DAC sample rate (F_S). **Figure 7b** shows the DAC's frequency response with and without an amplitude equalizer. Like the delay stage, it can be included in any reconstruction filter.

The hold response also has a pole centered on the sample clock, which completely removes the clock. Nevertheless, most reconstruction applications refer to clock attenuation as a figure of merit. Now that the function of a reconstruction filter is understood, one can be designed.

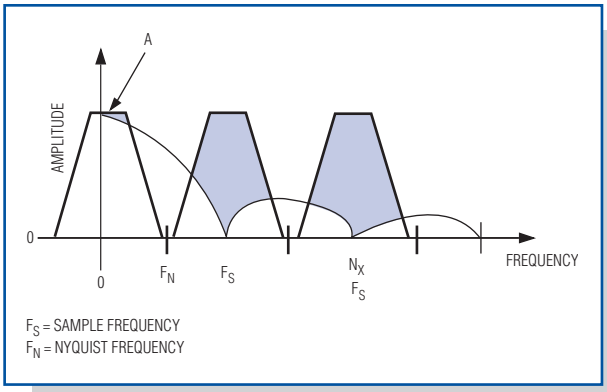


Figure 5. A typical DAC output spectrum is shown in terms of the sampling (F_S) and Nyquist (F_N) frequencies.

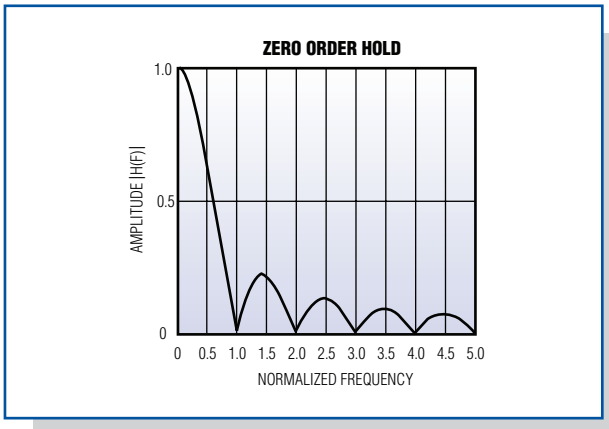


Figure 6. The "hold" function of a DAC produces a (sinx)/x response with nulls at multiples of the sampling frequency.

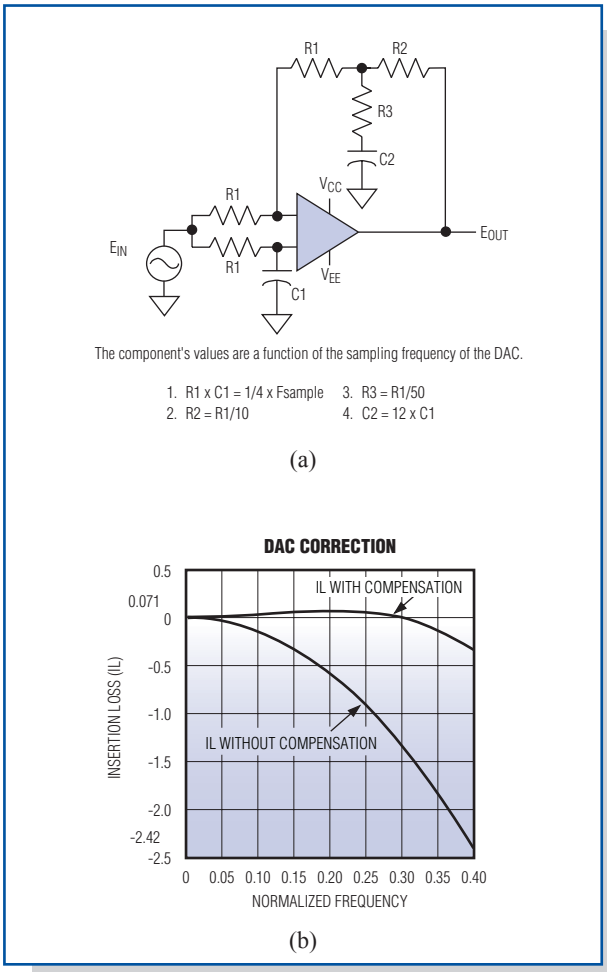


Figure 7. A DAC output (b) is shown with and without the (sinx)/x correction provided by an amplitude equalizer circuit (a).

The most common requirement for NTSC/PAL reconstruction is an attenuation of >20dB at 13.5MHz and >40dB at 27MHz, where ω_c depends on the applicable video standard. A 3-pole Butterworth with Sallen-Key configuration is chosen for two reasons. First, its gain (+2) drives a back-terminated cable. Second, the group-delay variation can be adjusted to optimize performance without a delay equalizer. (**Figures 8a–8d** show NTSC and PAL designs, including gain and group-delay characteristics.) These applications usually include digital amplitude correction for the DAC, which can easily be added, if necessary.

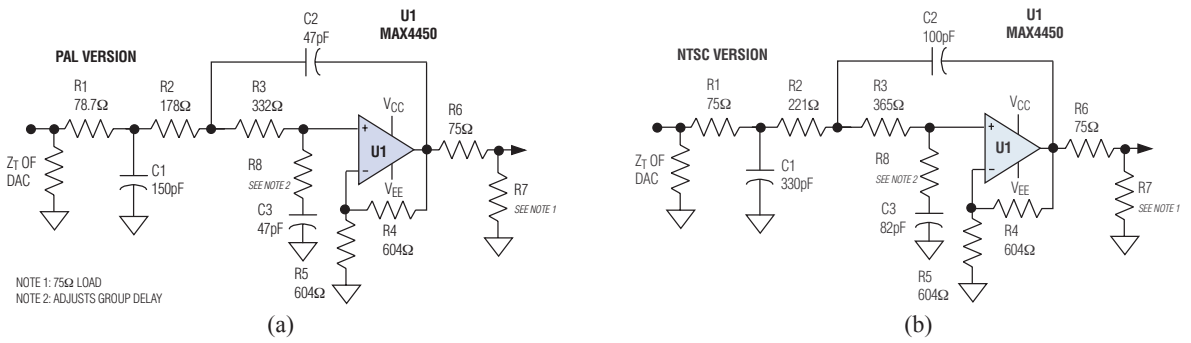
Illustrating a circuit for XGA, a 20MHz, 3-pole Butterworth filter in the Sallen-Key configuration includes the Figure 8 circuit for amplitude correction (**Figures 9a** and **9b**). Complementing the antialiasing filter of Figure 4, this filter has a gain of +2 to drive a back-terminated 75Ω coaxial cable.

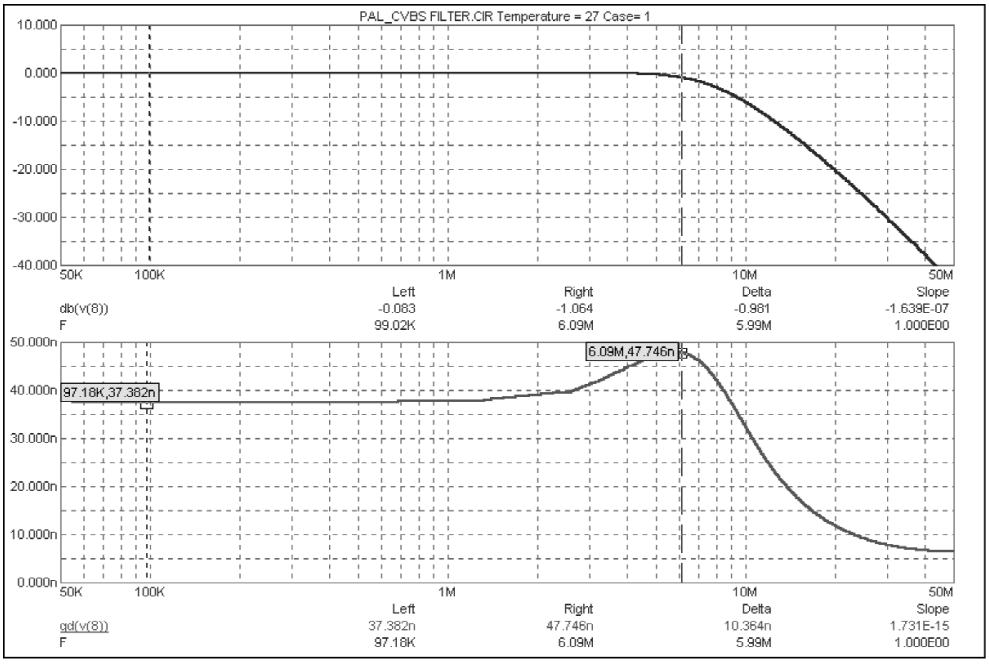
The last application is a reconstruction filter for HDTV. Based on the templates in the SMPTE 274 and 296M, it has a center frequency of $\omega_c = 0.4 \times F_S = 29.7$MHz. Amplitude correction for the DAC is usually included, but group-delay compensation must be added. The resulting 30MHz, 5-pole Sallen-Key filter (**Figure 10**) has >40dB attenuation at 74.25MHz, as well as a group-delay stage with a +2 gain to drive a back-terminated 75Ω coaxial cable.

Practical aspects of active video-filter design

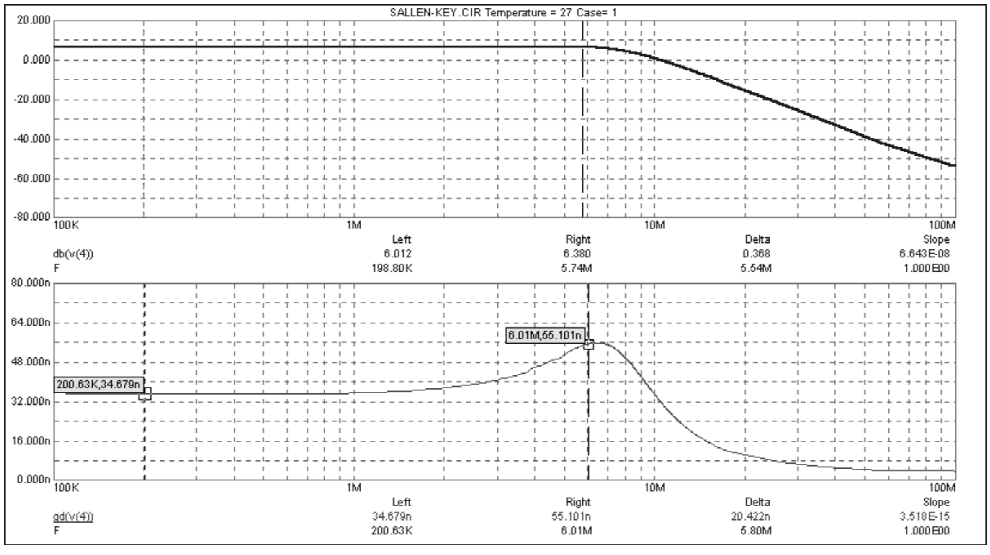
Whether filters are designed by hand, with the aid of software, or with a combination of these approaches, the actual response may not be exactly what is wanted. One cause is the discrepancy between a calculated response and the actual response obtained using standard component values.

That error can be minimized by choosing standard (5%) capacitor values and deriving resistor values from them. The reason is practical—capacitors with 1% or 2% tolerance can be acquired, but only at 5% values, although





(c)



(d)

Figure 8. For reconstruction filters with group-delay adjustment, the PAL version (a) has the amplitude and group-delay responses shown in (c), and the NTSC version (b) has the amplitude and group-delay responses shown in (d).

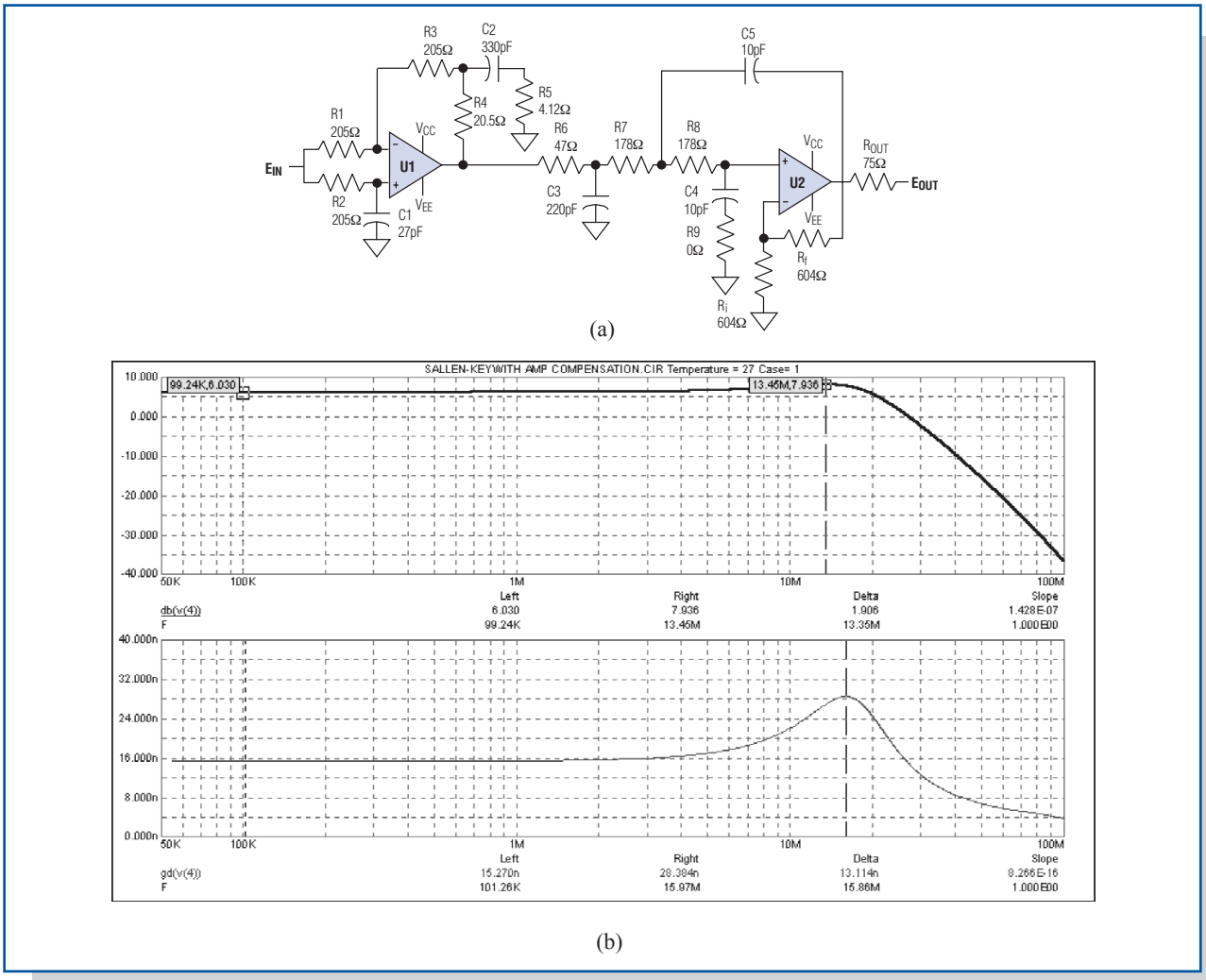


Figure 9. This 3-pole, 20MHz Butterworth filter for XGA reconstruction (a) includes (sinx)/x compensation. Its output response is shown in (b).

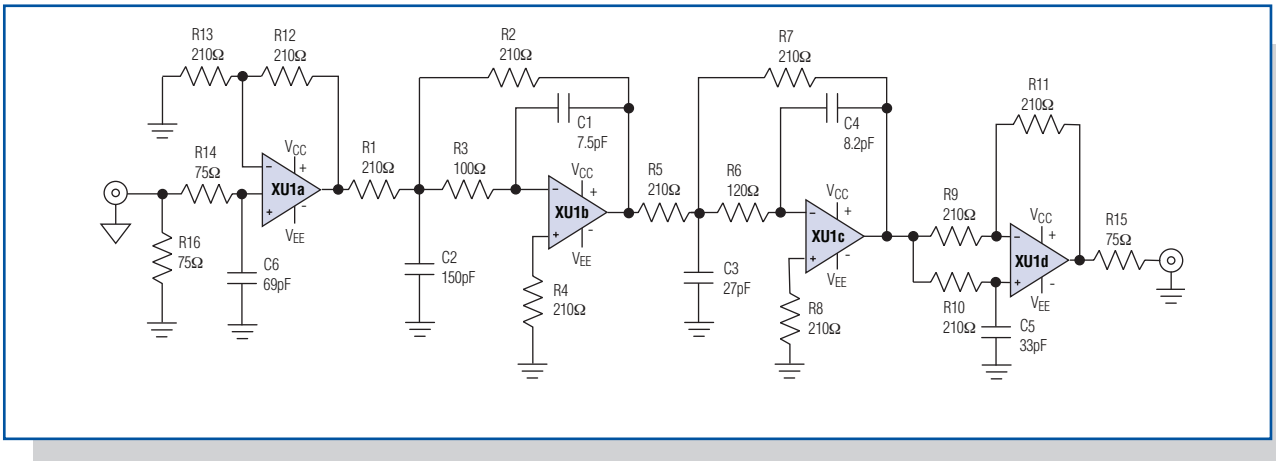


Figure 10. A 5-pole, 30MHz reconstruction filter for HDTV includes amplitude correction for the DAC.

resistors that combine 1% values with 1% tolerance are available. Such components give the best approximation and the most precise amplitude response.

Once built, a filter can be unstable and oscillate. In that case, short the input to ground and see if it continues to oscillate. If it stops, the impedance is too high. Lowering the design impedance should eliminate the oscillation. If it continues, note whether the oscillation is near the filter cutoff frequency or just below. In that case, the oscillation is probably due to components or parasitics. If the oscillation is above the cutoff frequency, it is probably due to the op amp or the circuit layout.

Good layout seems an art, but it is based on a few simple principles. It is important to have a clean supply voltage and a solid ground, meaning filtration with low-ESR capacitors and sometimes with a regulator. The loop formed by the bypass-capacitor connections must be small, or the resulting parasitic inductance will resonate with the capacitance. A good ground plane is essential to good analog design, but as bandwidth increases, it may add parasitic capacitance that can detune the filter. To avoid that problem, remove the ground plane beneath the offending part(s) and traces.

References

[1]A filter response drops by -3dB at the cutoff frequency.

[2]Taylor and Williams, Electronic Filter Design Handbook, McGraw Hill, ISBN 0-07-070441-4.

[3]Ibid.

[4]The 4:2:2 sampling originally indicated the number of times the color subcarrier was oversampled. ITU-601 replaced the subcarrier frequency with 3.375MHz. The 4:2:2 is sampled at 13.5MHz and 6.75MHz, respectively.

[5]For the noninverting case, $R_f/R_i = 0$. For the inverting case, $R_f/R_i = 1$.

[6]E.J. Kennedy, Operational Amplifier Circuits: Theory and Applications.

[7]Defined by H.W. Bode in Network Analysis and Feedback Amplifier Design (D. Van Nostrand, Princeton NJ, 1945).

[8]Taylor and Williams, op. cit.

[9]MAX4450/51 data sheet is available at www.maxim-ic.com.

[10]The Sinc function in mathematics is (sinx)/x.

A similar article appeared in the June 2003 issue of the EDN.

Automatic fan control techniques

Trends in cooling high-speed chips

Cooling fans are an important part of thermal management for high-powered chips (such as CPUs, FPGAs, and GPUs) and systems. Unfortunately, their use can sometimes raise a system's acoustic noise level to the point where it is objectionable to the user. By measuring temperature and adjusting fan speed accordingly, the fan's speed and noise level can be minimized when temperature is low, but increased under worst-case conditions to prevent damage. This article describes two techniques for automatically controlling a cooling fan's speed.

High-speed chips tend to run hot. As they get faster, they get hotter. New generation high-speed digital chips use smaller processes that allow the supply voltage to be reduced, which helps somewhat, but the number of transistors increases faster than the supply voltage decreases. Power levels, therefore, continue to rise.

As chip temperature increases, performance can suffer. Parameters shift, maximum operating frequencies decrease, and timing can fall out of specification. From the user's point of view, the product is no longer operating properly when this occurs. The first reason for cooling high-speed chips, therefore, is to maintain good performance for the longest possible operating time and over the widest possible range of environmental conditions. The maximum allowable temperature for a high-speed chip to meet its parametric specifications depends on the process and how the chip is designed (how "close to the edge" the chip is operating), among other factors. Typical maximum die temperature values range from +90°C to +130°C.

Beyond the point where performance degradation begins, excessive die temperature causes catastrophic damage to chips. The maximum die temperature limit is usually well over +120°C and depends on such factors as process, package, and duration of high-temperature conditions. High-speed chips are, therefore, cooled to avoid reaching a temperature that could both degrade performance and cause irreparable damage.

A single cooling technique is rarely used with high-speed chips. Instead, combinations of techniques are generally necessary to ensure high performance and continued reliability. Heat sinks, heat pipes, fans, and clock throttling are commonly employed to cool high-speed chips. The last two, fans and clock throttling, can help solve the heat problem, but introduce problems of their own.

Fans can dramatically reduce the temperature of a high-speed chip, but they also generate a great deal of acoustic noise. The noise from a full-speed cooling fan is annoying to some consumers and is also becoming a target of government agencies concerned about the long-term effects of noise in the workplace. Fan noise can be reduced significantly by varying the fan's speed based on temperature; the fan can turn slowly (and very quietly) when temperature is low, and can speed up as temperature increases.

Clock throttling—reducing clock speed to reduce power dissipation—works by reducing system performance. When throttling the clock, the system continues to function, but at a reduced speed. Clearly, in high-performance systems, throttling should be done only when it is absolutely necessary—that is, when the temperature reaches the point where functionality is about to be lost.

Controlling fan speed or clock throttling based on temperature requires that the temperature of the high-speed chip is measured first. This can be done by placing a temperature sensor close to the target chip—either directly next to it or, in some cases, under it or on the heat sink. The temperature measured this way corresponds to that of the high-speed chip, but can be significantly lower (up to around 30°C), and the difference between measured temperature and die temperature increases as the power dissipation increases. Therefore, the temperature of the circuit board or heat sink must be correlated to the die temperature of the high-speed chip.

A better alternative is possible with a number of high-speed chips. Many CPUs, graphics chips, FPGAs, and other high-speed ICs include a "thermal diode", which is actually a diode-connected bipolar transistor, on the die. Using a remote-diode temperature sensor connected to this thermal diode, the temperature of the high-speed IC's die can be measured directly with excellent accuracy. This not only eliminates the large temperature gradients involved in measuring temperature outside the target IC's package, but it also eliminates the long thermal time constants, from several seconds to minutes, that cause delays in responding to die temperature changes.

The need for fan control forces the designer to make several key choices. The first choice is the method of adjusting the fan's speed. A common method of

adjusting the speed of a brushless DC fan is to regulate the power-supply voltage of the fan. This approach works well for power-supply voltages as low as about 40% of the nominal value. There is a drawback. If the power-supply voltage is varied using a linear pass device, the efficiency is poor. Better efficiency can be obtained using a switch-mode power supply for the fan, but this increases cost and component count.

Another popular fan-speed control technique is to power the fan with a low-frequency PWM signal, usually in the range of about 30Hz, whose duty cycle is varied to adjust the fan's speed. This is inexpensive because a single, small pass transistor can be used. It is efficient because the pass transistor is used as a switch. A disadvantage of this approach, however, is that it can make the fan somewhat noisier because of the pulsed nature of the power supply. The PWM waveform's fast edges cause the fan's mechanical structure to move (somewhat like a badly designed loudspeaker), which can easily be audible.

Another fan-control design choice is whether the fan's speed is measured as part of the control scheme. In addition to power and ground, many fans are available with a third wire that provides a "tachometer" signal to the fan-control circuitry. The tachometer output produces a specified number of pulses (two pulses, for example) for each revolution of the fan. Some fan-control circuits use this tachometer waveform as a feedback signal that allows the fan's voltage or PWM duty cycle to be adjusted to give a desired RPM. A simpler approach ignores any tachometer signal and simply adjusts the fan's drive to speed up or slow down with no speed feedback. Speed control using this method is less precise, but cost is lower and at least one feedback loop is removed, simplifying the control system.

In some systems, it is important to limit the change rate of the fan speed. This is most critical when the system is in close proximity to users. Simply switching a fan on and off or changing speed immediately as temperature changes is acceptable in some environments. When users are nearby, however, sudden changes in fan noise are apparent and annoying. Limiting the rate of change of the fan's drive signal to an acceptable value (e.g., 1% per second) ensures that the acoustic effects of fan control are minimized. The fan speed still changes, but it does so without attracting attention.

The fan-control profile is another important design variable. Typically, the fan is off below a specific threshold temperature and then begins to spin at a slow rate (for example, 40% of full speed) once the threshold is exceeded. As temperature increases, the fan's drive increases linearly with temperature until it reaches 100% drive. The best slope depends on system requirements. A more rapid slope results in somewhat more consistent chip temperature, but fan speed has more variation as power dissipation changes from one moment to the next. If highest performance is the goal, the starting temperature and the slope should be chosen so that the fan reaches full speed before the die temperature is high enough to initiate clock throttling.

Implementing fan-control circuitry can be done in several ways. A variety of remote temperature sensors with up to five sensing channels is available that can detect the die temperature of the high-speed chip and transmit temperature data to a microcontroller. Fan-speed regulators with multiple channels of fan-tachometer monitoring can provide reliable control of fan RPM or supply voltage based on commands from an external microcontroller. For low cost and simple implementation, ICs are available with temperature sensing and automatic fan control included in a single package. Sensor/controllers also normally include overtemperature detection for clock throttling and system shutdown, thereby protecting the high-speed chips from catastrophic failure due to overheating.

Examples of two such ICs, one with DC drive and one with PWM drive, are shown in **Figures 1** and **2**. The IC in Figure 1 senses remote temperature and controls fan speed based on that temperature. It produces a DC supply voltage for the fan through an internal power transistor. Figure 2 shows an IC that performs a similar function, but drives the fan with a PWM waveform through an external pass transistor. Both include complete thermal fault monitoring with overtemperature outputs, which can be used to shut down the system if the high-speed chip gets too hot.

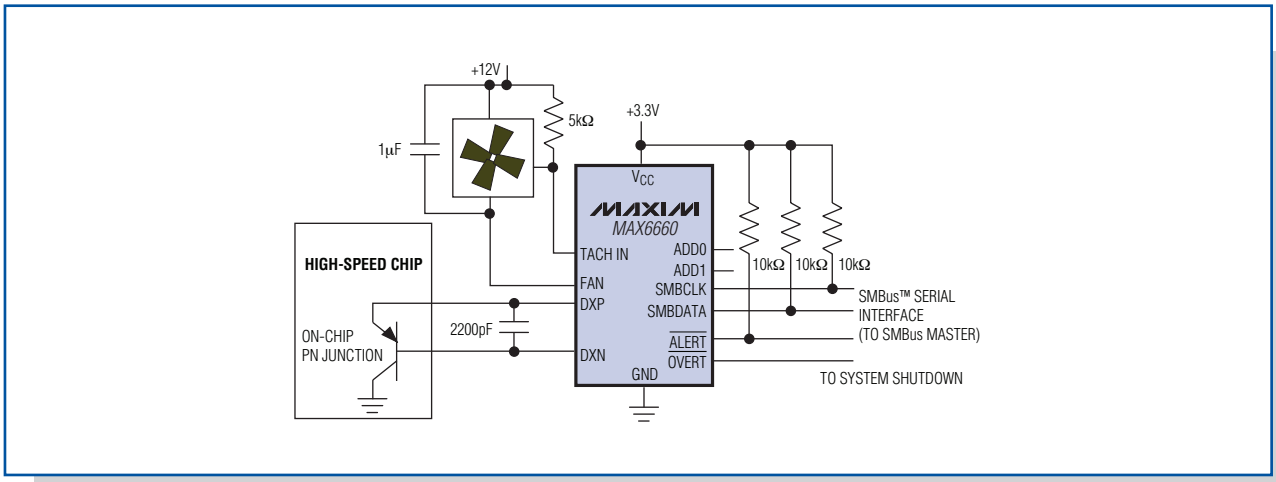


Figure 1. Linear (DC-output) temperature sensor and automatic fan-speed controller. Fan speed is controlled automatically based on the temperature of the high-speed chip. Tachometer feedback from the fan allows the fan controller to regulate fan speed directly. System shutdown output prevents the high-speed chip from reaching destructive temperatures.

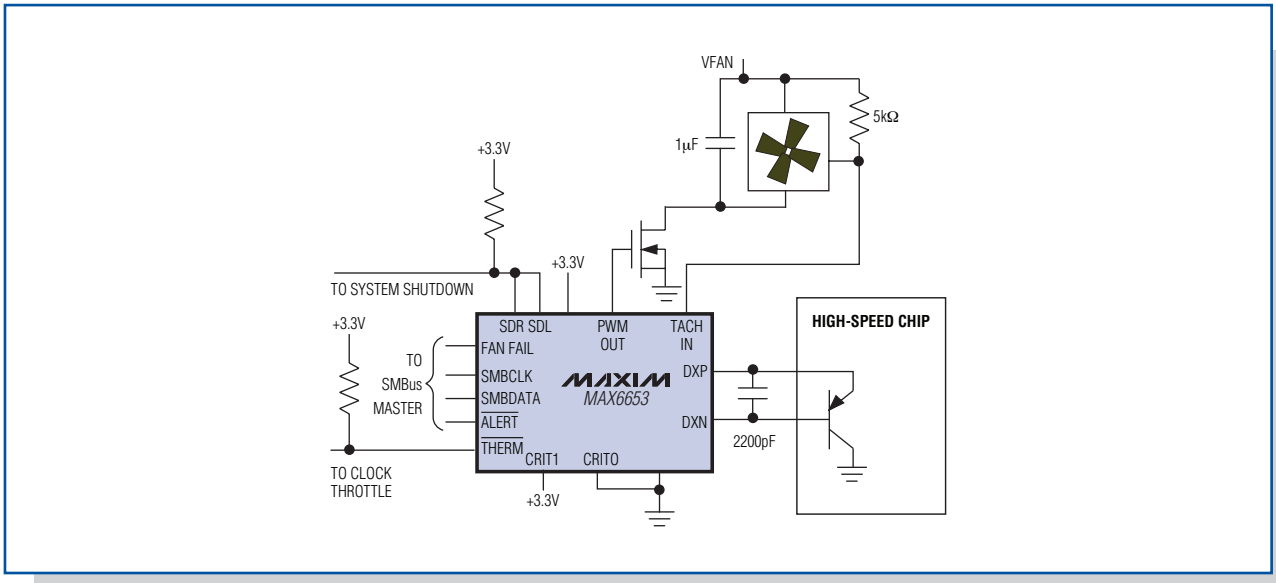


Figure 2. PWM-output temperature sensor and automatic fan-speed controller. Fan speed is controlled automatically based on temperature. Clock throttle and system shutdown outputs prevent a high-speed chip from reaching destructive temperatures. CRIT0 and CRIT1 pins can be strapped to supply or ground to select default shutdown-temperature thresholds, ensuring protection even when system software hangs.

SMBus is a trademark of Intel Corporation.

Selecting power management ICs for cellular handsets

A new generation of sleek, feature-laden cell phones now flourishes on retail shelves. These latest handsets are not only smaller than previous generations, but are also capable of PDA functions, delivering email, instant messaging, and web browsing on a larger, more colorful display. Some models include FM radios, MP3 players, or even a decent digital camera. Despite all this functionality, consumer expectations demand enduring battery life without an increase in size. Forcing more into a smaller box while draining less power has propelled power-management design into a critical role. To meet the challenge, analog IC manufacturers continue to develop smaller, higher performance power-supply solutions.

Power-management ICs

At the heart of most wireless handsets beats a power-management IC (PMIC). The PMIC handles most of the power-supply requirements and other blocks such as interface or audio. Leading analog semiconductor manufacturers provide PMICs as full custom, semicustom, and/or standard products, usually using a 5V submicron BiCMOS process optimized for mixed-signal and power supplies. Any block that is not already integrated elsewhere in the handset is a candidate for integration in the PMIC; some moderation is advised, however, as shown in **Figure 1**. There are reasons not to integrate a particular block: 1) the block may be cheaper or smaller if designed using a different process; 2) the block may change from one design to the next due to advances in technology or changing customer requirements; 3) the block may not be common across platforms; 4) the block may present an aggressive design challenge/risk to the schedule; and 5) the block may be ill-suited for integration due to performance reasons such as noise coupling.

Nonetheless, the reasons to integrate as many blocks as possible into the PMIC are obvious cost and size savings, especially when the integration is common for a significant volume of handsets. Risks may be managed by gradually increasing the integration over successive designs.

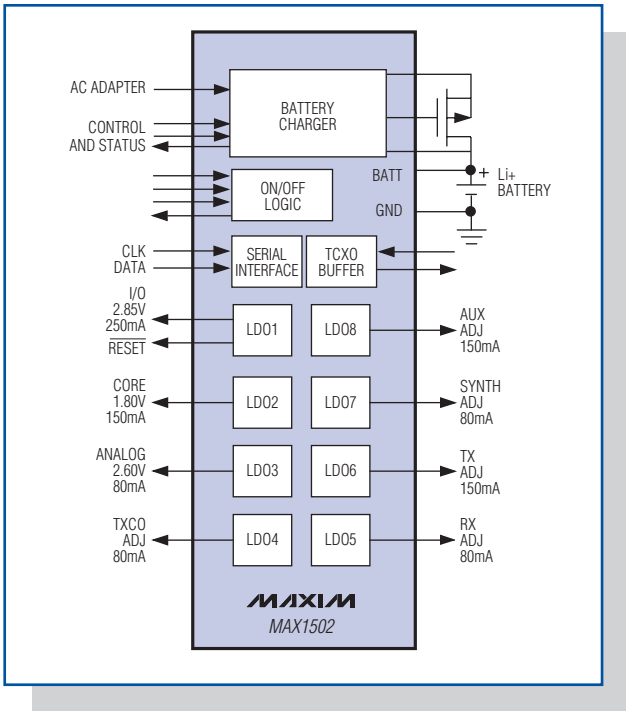


Figure 1. An example of low-risk, constrained integration, the MAX1502 standard-product power-management IC integrates only the most common blocks required to support popular CDMA chipsets.

Low-dropout linear regulators

Cellular handsets typically require from 5 to 12 individual low-dropout (LDO) linear regulators. This number is high, not because there are this many individual voltages, but because the LDOs also act as on/off switches with power-supply rejection ratio (PSRR) to prevent noise coupling. Most LDOs are integrated in the PMIC, but a few discrete units remain due to PCB layout/routing, the noise sensitivity of specific components like the voltage-controlled oscillator, or the need to power a nonstandard block such as an integrated digital camera. In discrete form, the single 150mA LDO in a SOT23 package was a popular choice for many years. However, today's newer packages, submicron processes, and better designs offer higher performance in a smaller size. You can now get a single 300mA LDO in a SOT23, dual 150mA LDOs in a SOT23, or a single 120mA LDO in a tiny SC70 with both standard and extra-low-noise ($10\mu V_{RMS}$ and 85dB PSRR) variants. Furthermore, a modern UCSP™ provides the tiniest size possible, while the new QFN package permits the largest die size and provides the highest thermal conductivity in a plastic 3mm x 3mm footprint. QFN packages thus enable higher current LDOs and more LDOs per package. There are even triple, quadruple, and quintuple LDOs, which reduce the demarcation between a discrete implementation and a PMIC.

UCSP is a trademark of Maxim Integrated Products, Inc.

Stepdown (buck) converters for the processor core

As simple and small as the LDO is, its major weakness is efficiency, especially when powering a low-voltage circuit. As PDA and Internet functionality have increased in cell phones, more powerful processors are applied with ever-decreasing supply rails ranging from 1.8V down to 0.9V. To keep battery current in check, an efficient step-down switching regulator should power the core. Low cost, small size, high-efficiency, low-quiescent (standby) current, and fast transient response are design concerns. Reconciling this difficult grouping of characteristics demands experienced analog design and some ingenuity. Today, only the leading analog semiconductor manufacturers offer appropriate buck converters in small SOT23 packages with 1MHz or faster switching to allow small external inductors and ceramic capacitors.

Stepdown (buck) converters for RF power amplifier

In the mature Japanese cell phone market, buck converters are also used to power the CDMA radio's power amplifier (PA) with varying V_{CC} supply voltage in response to varying distance from the base station. When multiplied by the transmit-probability density function, a buck converter can save an average of 40mA to 65mA of battery current. The amount of current saved depends on the number of output voltage steps, the PA's characteristics, and whether transmitting voice or data in an urban or suburban environment. Based upon Japanese success and pioneering efforts at one European WCDMA innovator, now Korean, American, and other European cell phone makers are testing or designing with switching regulators. Buck converter requirements are very small size, low cost, low output ripple, and high efficiency. Again, SOT23 converters are a good choice. To keep dropout as low as possible, a discrete low-$R_{DS(ON)}$ P-channel MOSFET is usually used to power the PA directly from the battery when transmitting at high power. To further reduce overall size, the newest stepdown converters (e.g., the MAX8500 family) integrate this additional FET.

LEDs, LEDs, and more LEDs

For wireless handsets with color displays, white LEDs now dominate backlighting applications due to circuit simplicity and very high reliability. Efficiency exceeds that of halogen but does not yet equal that of fluorescent bulbs. Modern designs typically utilize three or four white LEDs for the main display, two white LEDs for the subdisplay (in the case of a clamshell design), and six or more white or colored LEDs behind the keyboard. If a camera is integrated, there may be four more white LEDs for flash/strobe and mpeg movie light. That totals up to 16 LEDs in one handset, all of which need to be driven with a constant current.

Years ago, the first generation of color handsets in Japan used inefficient 1.5x charge-pumps and ballast resistors. (That solution effectively discarded the 40mA they struggled to save by adding a buck converter to the PA.) Today, most designs use an inductor-based boost converter for much higher efficiency. The newest 1x/1.5x charge-pumps, however, obtain similar high efficiency without the inductor, but with more wires to the LEDs. Because the market for white LED power is so large, there are an overwhelming number of ICs available. Concerns are high efficiency, small external components, low input ripple (to avoid coupling noise into other circuits), a simple dimming interface, and any other feature that reduces cost or adds reliability, such as output overvoltage protection. Some PMICs include a white LED power supply, but usually do not power multiple displays or the camera strobe, or they may be inefficient or switch too slowly. This would require large inductors and capacitors and generate large input ripple. Adding a discrete IC to work in conjunction with the PMIC or finding a highly integrated discrete solution (such as the MAX1582, **Figure 2**) is often required.

Battery charging

Nearly every wireless handset charges its three NiMH cells or single Li+ cell with a simple linear charger. Most often the charger is integrated in the PMIC, although the current-sense resistor and/or pass transistor may be external for simplicity. There are many options for keeping thermal dissipation in check: 1) charge at C/4 or slower; 2) make the wall adapter output somewhat resistive so most of the voltage drop occurs there; 3) use pulse-charging and a current-limited wall adapter; 4) include feedback to regulate the wall adapter so that the voltage drop across the pass transistor is constant; or 5) add a constant thermal-control loop that throttles back the current to maintain a constant die temperature, which is only practical when the pass transistor is on the PMIC. Discrete charger ICs may offer some flexibility, but the benefits are largely reduced for cell phones because an integrated charger can be easily reprogrammed through the PMIC's serial interface to accommodate varying battery chemistry or capacity.

A similar article appeared in the May 2003 issue of Connecting Industry (UK).

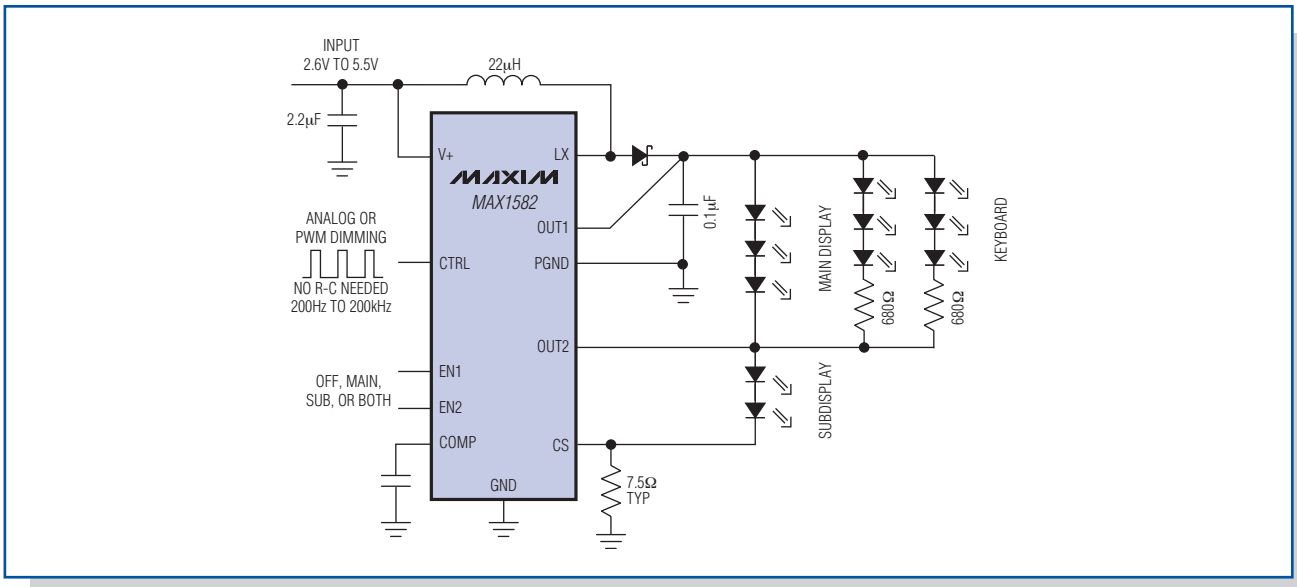


Figure 2. Because LED backlighting is one of the highest consumers of battery power in a handset, the MAX1582 utilizes a high-efficiency boost converter topology to illuminate the main display, subdisplay, and keyboard.

Base station RF power amplifier biasing

Power amplifiers used in base stations require biasing for proper RF performance. This article explains the two classes of biasing that are prevalent in the RF industry, analyzes their characteristics, and shows implementations with existing ICs. The power device of choice for base station amplifiers today is the lateral DMOS (LDMOS) MOSFET, used in this article to illustrate biasing techniques. However, as future generations of devices become available, such as GaN FETs, HFETs, or SiC devices, they too will benefit from the following implementations.

RF classes and biasing

LDMOS amplifiers used in RF circuits exhibit varying degrees of nonlinearity, depending on the DC-bias level upon which the input RF waveshape rides. That is, while maintaining a constant RF gating signal, the output current's (I_{out}) harmonic content varies as the DC bias at the gate of an LDMOS device (**Figure 1**) changes. The harmonic content of the LDMOS amplifier's current is important because, in the RF load, it creates power interference with the local bandwidth (in-band interference) or with adjacent bandwidths (out-of-band interference).

The best linearity occurs when the output current tracks the input voltage—a 360° conduction angle. Operating the MOSFET in this manner (i.e., class-A operation) creates less distortion than when biasing it in any other way. From a power-dissipation perspective, however, class-A operation is least desirable because it consumes the most DC current.

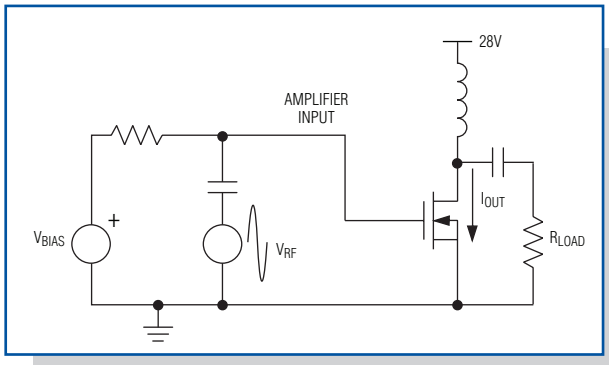


Figure 1. LDMOS device gating is shown with an uncontrolled DC bias.

At high RF power, given a nominal power-supply voltage of 28V, the DC power dissipated in the amplifier is prohibitive. For this reason, RF engineers use class-AB biasing in the last stage of an amplifier chain, while they favor class-A operation in the preceding stages where power dissipation is smaller by orders of magnitude. In class-AB stages, the output current does not track the input voltage entirely, and thus the amplifier's conduction angle is lower than 360°.

Distortion of the RF signal in class AB is more significant than in class A. The spectrum of this distortion is wider and more densely populated than that of class A. However class-AB power dissipation is lower because the average current into the amplifier is lower. In short, the basis for choosing a given class of commercial RF amplifiers is a tradeoff between linearity and efficiency.

Biasing requirements and LDMOS behavior

Biasing requires managing the DC content in the LDMOS current across temperature and supply variation. The ultimate objective is to ensure that the amplifier RF gain, as well as its distortion levels, varies within limits consistent with requirements. In this respect, proper biasing can assist linearization techniques to minimize distortion.

The equation governing LDMOS's gain is $I_{out} = K (V_{gs} - V_{th})^2$, where K is a constant reflecting gain due to electron mobility and V_{th} is the FET's threshold. Both K and V_{th} are temperature dependent. In **Figure 2**, LDMOS characteristics are shown across temperature. In class AB, designers tend to operate the bias to the left of the crossover region where the gain has a positive temperature coefficient. In class A, operation occurs to the right of the crossover region.

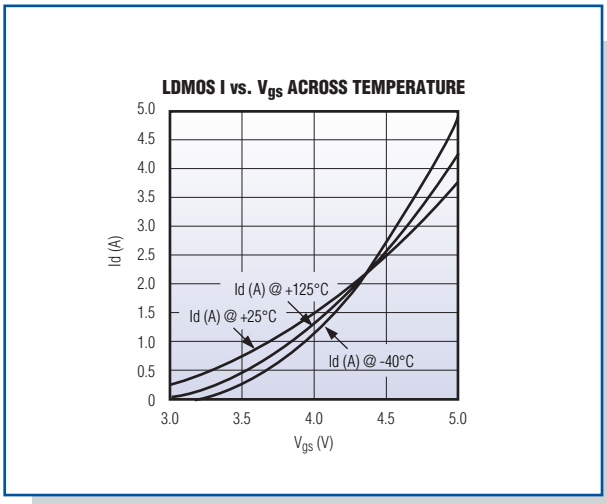


Figure 2. LDMOS characteristics are shown across temperature.

Controlling class-A and class-AB bias with the DS1847

Figure 3 shows a DS1847 dual, temperature-controlled variable resistor controlling the gate of an LDMOS amplifier. The DS1847's internal temperature sensor provides a temperature reading to its look-up tables. These look-up tables adjust the IC's two 256-position variable resistors so the amplifier's gate receives the proper bias voltage. The user programs the look-up tables to generate a constant LDMOS-amplifier output current. Refer to Figure 2 (or to manufacturer-specific data curves) for LDMOS characteristics. By using the two resistors to attenuate the reference voltage, a temperature-insensitive voltage is maintained.

A similar article appeared in the November 2003 issue of Wireless Design & Development.

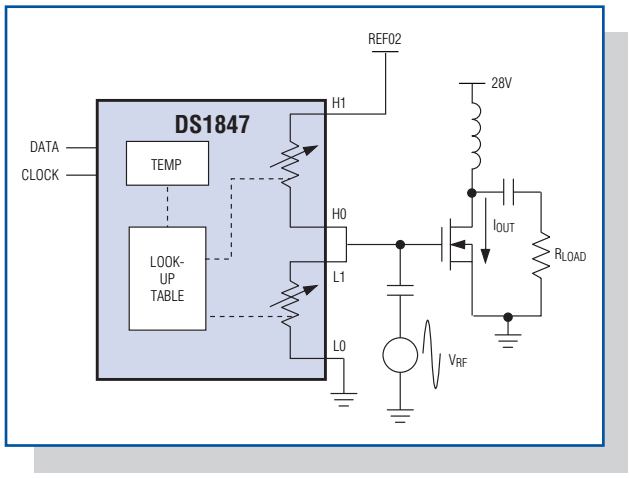


Figure 3. DS1847 dual, temperature-controlled variable resistor controls the gate of an LDMOS amplifier.

FREE SAMPLES AND TECHNICAL INFORMATION

FAX: 408-222-1770

www.maxim-ic.com/samples

Contact us at 1-800-998-8800 (Toll Free)
www.maxim-ic.com/samples

Please send me a sample of:

_____ _____

_____ _____

_____ _____

_____ _____

(Limit is 8 part numbers, 2 samples each.)

Request free information about Maxim and our products.

❒ Please send me **Future Issues of the Maxim *Engineering Journal*.**

Please complete the information below.

Name _____ Title _____

Company _____ Department _____

Address _____

City _____ State/Province _____

Zip Code _____ Country _____

Telephone _____

E-mail Address _____

My application is _____ My end product is _____

EJ-49 US 11/03

Request Design Guides from Our Library

❒	1-Wire® Products
❒	A/D Converters
❒	Audio
❒	Battery Management
❒	Communications
❒	D/A Converters
❒	Digital Potentiometers
❒	Displays
❒	Fiber Cable
❒	Flat-Panel Displays
❒	High-Speed ADCs & DACs
❒	High-Speed Interconnect
❒	High-Speed Microcontrollers
❒	Interface
❒	Low-Power Notebook
❒	µP Supervisors
❒	Multiplexers and Switches
❒	Op Amps and Comparators
❒	Power Supplies
❒	Real-Time Clocks
❒	Signal Conditioners
❒	System Timing and Control
❒	Temperature Sensors
❒	Video
❒	Voltage References
❒	Wireless

DALLAS SEMICONDUCTOR MAXIM

www.maxim-ic.com

Maxim Integrated Products, Inc.
120 San Gabriel Drive
Sunnyvale, CA 94086

Presorted Standard
U.S. Postage
PAID
Greenfield, OH
Permit No. 270