IX. Monolithic Microwave Integrated Circuits

A. N. Downey, G. E. Ponchak, and R. R. Romanofsky

A. General Description

Monolithic Microwave Integrated Circuits (MMICs) are used in satellite systems that require smaller, less expensive circuits or when the parasitic reactance inherent in hybrid integrated circuits degrades the circuit performance, typically in the upper microwave and the millimeter-wave spectrum. Examples of systems that use MMICs are receivers and transmitters for communications, phased-array antennas where small size and uniform circuit performance are required, and sensors and radars that operate at high frequencies. The types of circuits required for each of these systems are illustrated by examining the simple receiver and transmitter systems shown in Figures 3-38 and 3-39, respectively. In both schematics, a phase shifter—which may be placed in either the local oscillator (LO), the RF, or the IF portion of the system—has been added to make the system perform as if each circuit were coupled to a single radiating element of a phased-array antenna. For non-phased-array applications, the schematic is unchanged except for the removal of the phase shifter. A photograph of a completely monolithic 30-GHz receiver is shown in Figure 3-40. Although the high level of circuit integration illustrated in Figure 3-40 decreases the packaging and interconnect costs, this integration is not necessary or common. Instead, each function of the system is typically fabricated on an individual die to permit the optimization of the material system and device type for each application. Regardless of the level of circuit interconnection, the reliability of the system is dependent on the continuous operation of each circuit.

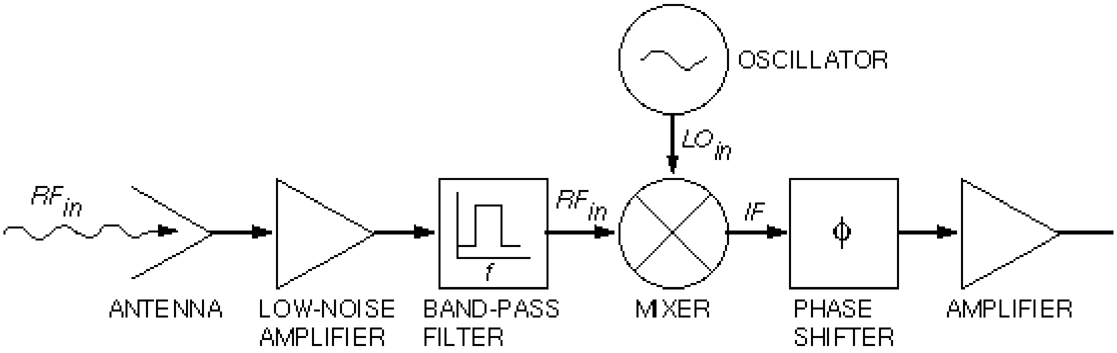

Figure 3-38. Schematic of microwave receiver.

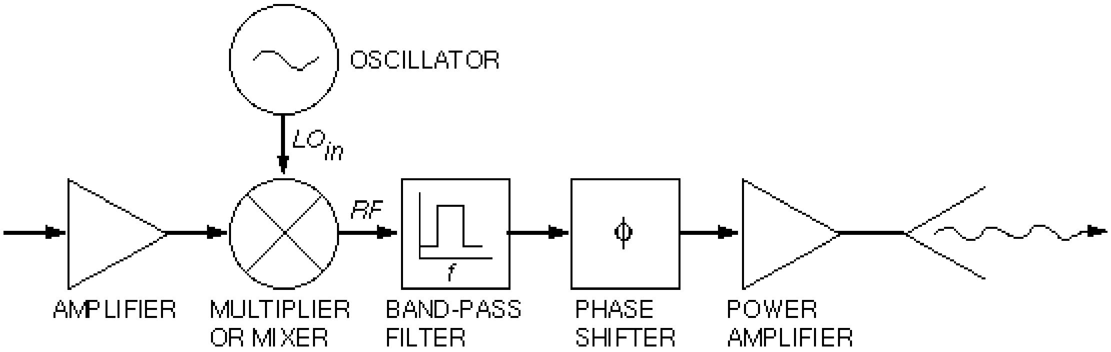

Figure 3-39. Schematic of microwave transmitter.

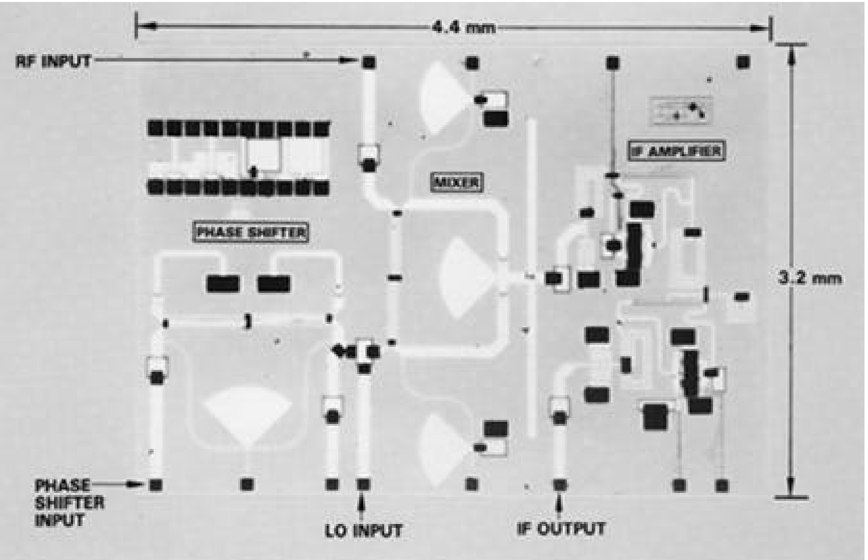

Figure 3-40. 30-GHz MMIC receiver. (Fabricated by Hughes Aircraft Company for NASA Lewis Research Center.)

This is understood by examining the receiver circuit shown in Figure 3-38. The input (RF) signal typically has a very low power level that may be close to the noise floor. The low-noise amplifier (LNA) amplifies the received signal while at the same time introduces very little new noise. If the gain of the LNA is sufficiently large, the noise contributions of the rest of the system will be small since the noise created by later circuits is divided by the gain of the LNA. Thus, the LNA gain and noise figure, the measure of noise added by the LNA, determine the receiver noise characteristics. If the receiver has poor noise characteristics, it will not be able to receive weak signals. The signal may then pass through a narrow-band filter and into the mixer. The LO generates a signal that is also fed into the mixer. The mixer combines the two signals through a nonlinear device, such as a MESFET or diode, and generates a signal at the intermediate frequency (IF) of $f_{RF} - f_{LO}$ or $f_{LO} - f_{RF}$ and harmonics of the IF, RF, and LO frequencies. All but the desired IF components must be filtered out. The conversion efficiency of the mixer is usually dependent on the LO drive power. In addition, a variation in the LO frequency will cause a shift in the IF that may cause the signal to be attenuated in the narrow-band filters that are part of the mixer. If the system is to be associated with a phased-array antenna, the direction and shape of the main beam radiated or received by the antenna is dependent on the relative phase shift and power level of each transmitter (and receiver). The relative phase of each radiating element is set by the phase shifter. Thus, if the phase shift through the circuit varies because of unexpected conditions, the efficiency of the entire antenna will degrade. It is thus seen that a parametric shift by any of the components may cause the entire system to fail.

The phase shifter, local oscillator, and mixer circuits are common to the receiver and transmitter with the exception of a shift in the design frequency. The real difference between the two systems is in relation to the amplifiers. As described above for the receiver, the LNA must be capable of amplifying a weak signal sufficiently for the mixer to work and for the noise contributions of the rest of the system to be minimized, while at the same time introducing as little new noise as possible. In a transmitter, the critical performance specifications are the amount of power transmitted and the efficiency of the circuit. Thus, the power amplifier must be able to provide gain to a very strong signal.

In early MMICs, all of the circuits were made from GaAs MESFETs, impact ionization avalanche transit time (IMPATT) diodes, and varactor diodes, but as GaAs technology matured, HBTs, HEMTs, and PHEMTs have found increasing use in niche applications. Table 3-4 identifies the devices now most commonly used in each of the circuits. Although most MMIC failures originate at one of the active devices, some reliability concerns relating to each specific circuit will be presented below. A more detailed discussion of reliability problems related to specific devices or components is provided in Chapter 4.

Table 3-4. Matrix of solid-state devices and their applications in MMICs.

Device	Varactor Diode	Schottky Diode	PIN Diode	HBT	MESFET	HEMT	PHEMT
Low-noise amplifier, GHz							
f < 12				√	√	√	√
12 < f < 26						√	√
f > 26						√	√
Power amplifiers, GHz							
f < 12				√	√	√	√
12 < f < 26				√	√	√	√
f > 26						√	√
Mixers, GHz							
f < 12		√		√	√	√	√
12 < f < 26		√		√		√	√
f > 26		√		√		√	√
Oscillators, GHz							
f < 12	√			√	√	√	√
12 < f < 26				√	√	√	√
f > 26				√	√	√	√
Multipliers, GHz							
f < 12		√			√	√	√
12 < f < 26		√				√	√
f > 26		√				√	√
Analog phase shifter	√						
Switched-line phase shifter			√	√	√	√	√

B. Amplifiers

Both low-noise and power amplifiers are used to increase the power of the RF signal. In almost all systems, this is accomplished by using the transconductance of MESFETs and HEMTs or the current gain of HBTs. The amount of signal increase is called "gain" and is usually given in dB, where gain in dB = 10 log (gain). For example, if the output power is twice the input power, the amplifier has 3 dB of gain. Typically, the input power and the output power are also specified in dB, permitting the output power to equal the sum of the input power and the gain. This ideal operation of an amplifier is accurate for low power levels. Unfortunately, as power levels increase, the amplifier becomes nonlinear. In the nonlinear region of operation, the output power is less than the sum of the input power and the amplifier gain in the linear region, or it can be stated that the amplifier gain is lower in the nonlinear region. Figure 3-41 shows a typical amplifier characteristic. The point at which the output power drops by 1 dB from the linearly extrapolated value is called the 1-dB compression point [1]. This value separates small-signal or linear amplifiers from large-signal or power amplifiers. Note that this is also the criteria used to differentiate small-signal and large-signal transistors, since a transistor can be viewed as a simple, unmatched amplifier. This differentiation is important in determining the failure mechanisms that need to be addressed and the type of reliability tests that should be performed.

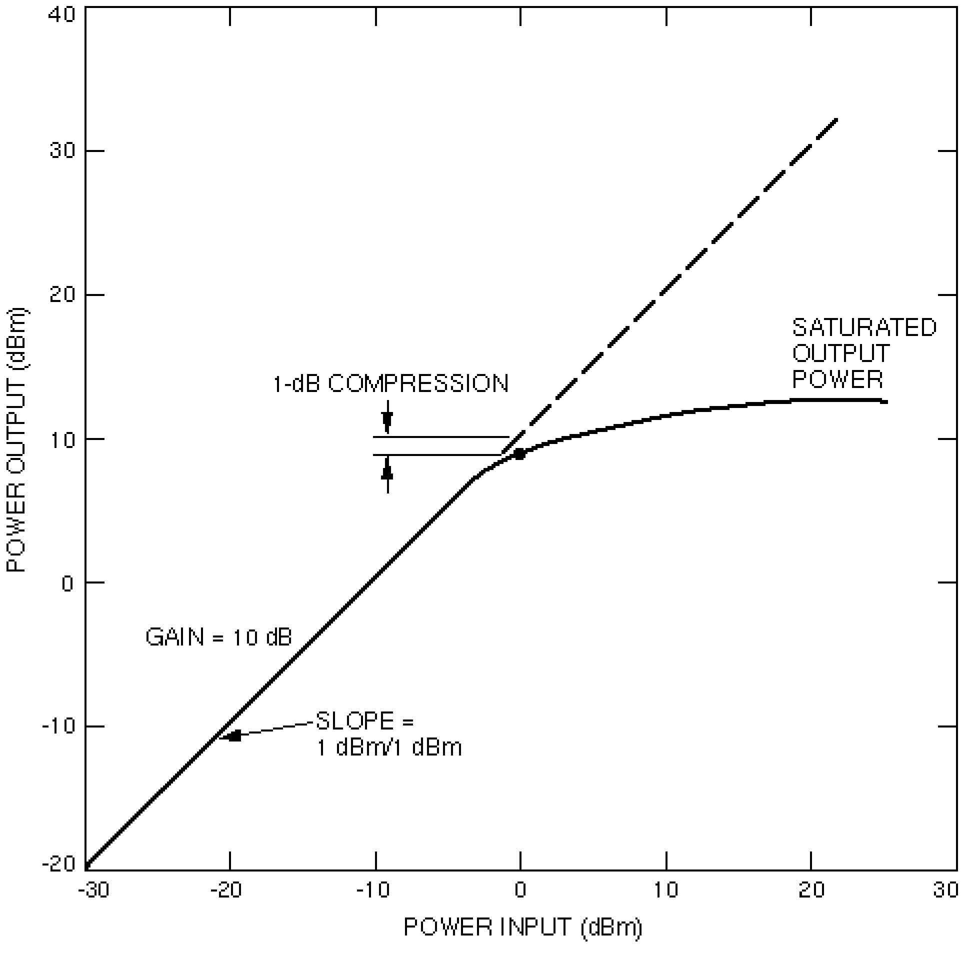

Figure 3-41. Output power as a function of input power for a typical amplifier.

The choice of the bias point is critical in the amplifier operation. If the bias point is chosen so that the output signal from the power amplifier appears as an amplified version of the input signal over the entire period of the voltage wave, the amplifier is called a Class A amplifier. More typical of power amplifiers, the large voltage swing of the input signal will cause the power amplifier to operate at a bias point such that the current is in cutoff or saturation over part of the input signal voltage swing. Thus, over part of the input voltage swing, the output waveform will be zero. If the output signal is zero over half of the period, the amplifier is called Class B. Other classes of amplifiers are based on the amount of time the output signal is at zero voltage. The choice of the bias point or amplifier class determines the linearity and power-added efficiency defined as

$$\eta = \frac{P_{out}^{RF} - P_{in}^{RF}}{P^{DC}}$$

Class A amplifiers are linear, but since they draw dc power over the entire period, they are not efficient. On the other hand, Class B amplifiers are not as linear, but since the amplifier is not drawing dc current over half of the input voltage swing, they have higher efficiency [2].

1. Power Amplifiers

Power amplifiers, by their very nature, must handle high input and output powers. The maximum voltage swing of the input signal is limited by the breakdown voltage of the transistor, and thus transistors with high breakdown voltages are required. The current through each transistor is limited by the resistance in the gate or emitter of FETs and HBTs, respectively, since ohmic losses are converted to heat, which decreases the device’s reliability. To increase the current handling capability of the device, power transistors combine many gates or emitters in parallel. This parallel combination increases the total gate width or emitter area and decreases the resistance, while at the same time increases the difficulty in matching the input impedance of the transistor to the output impedance of the prior stage. In addition, the spacing required between the transistor elements to permit sufficient thermal dissipation creates large devices that are more difficult to maintain with a uniform voltage [3]. To dissipate the heat from the transistors, power amplifiers are fabricated on thin wafers, less than 100 μm thick and typically between 25 and 50 μm, to reduce the thermal path between the transistor’s active region and a good heat sink, such as a metal or diamond carrier. Generally, thermal constraints limit the design and performance of power amplifiers more than frequency constraints. Thus, the efficiency of power amplifiers is one of the most critical specifications, especially in space applications where satellite power is limited, where dissipation of the thermal load requires heat sinks that increase the system weight, and where circuit heating can decrease reliability.

Power amplifiers designed with multiple stages (one stage is one transistor or one parallel combination of transistors) are used to accommodate the thermal constraints, peak-voltage and current constraints, and limited gain available from each transistor. Figure 3-42 shows a 30-GHz power amplifier with three stages that consist of increasingly larger transistors. The number of stages required in the amplifier is dependent on the gain specification and frequency, since transistor output power decreases with increasing frequency. Since the power dividing and combining networks on the MMIC will typically introduce 0.5 to 1 dB of loss and the input impedance of transistors decreases with an increasing number of gate fingers, the degree of power dividing and combining that can be used to increase the power level is limited.

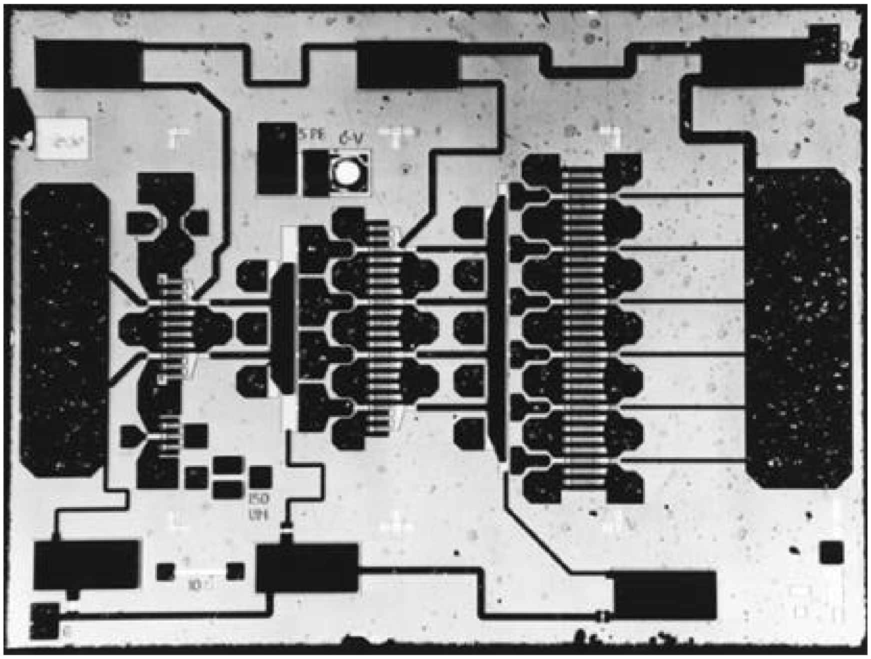

Figure 3-42. 20-GHz high-power amplifier. (Fabricated by Texas Instruments under contract to NASA Lewis Research Center.)

Models for microwave devices, both active and passive, are usually derived from *S* parameters measured on a vector network analyzer. These models are good for low-power circuit designs, but transistors exhibit significant nonlinearity or a power dependence at high-power levels. Therefore, nonlinear models based on load–pull measurements are required for high-power designs. In addition, the nonlinearity of the power transistors creates intermodulation distortion (IMD), which is power at frequencies other than the input frequency: $2f_{RF}$, $3f_{RF}$, etc. Failure to account for these frequency terms in the matching circuit can lead to signal distortion, oscillations, lower efficiency, and package resonances. IMD is specified as the ratio of the power at the IMD frequency to the power at the desired frequency and is usually given in dB. The ability of engineers to design a power amplifier will depend on the availability of good nonlinear models.

Besides thermal-stress-related problems, power amplifiers exhibit some unique failure mechanisms, such as hot-electron trapping, which is covered in Chapter 4. Electromigration and metal diffusion must also be addressed in power amplifier designs due to the large currents and high voltages used during operation.

2. Low-Noise Amplifiers

Since low-noise amplifiers are used on the front end of receivers, they are designed to handle very low power levels. Thus, the thermal problems and high bias currents and voltages that affect power amplifier reliability are generally not a concern for LNA designers. The most important criterion in specifying or measuring an LNA’s performance is the noise figure, and since HEMTs and PHEMTs have the lowest noise figure, they are used in almost all LNAs. To minimize the noise figure, small gate lengths and low parasitic gate and source resistances are required [4]. Thus, state-of-the-art LNAs are usually comprised of 0.1 to 0.25 μm gate-length HEMTs or PHEMTs, and the reliability concerns—such as gate metal sinking and ohmic contact diffusion (see Chapter 4)—arising from small gate lengths and corresponding small channel thicknesses are the most important.

To decrease the noise figure of the system, it is important to reduce the circuit losses, especially before the first stage of the LNA. This includes the package feed losses and transmission line losses from the antenna since they introduce noise into the system before the LNA. Besides reducing the circuit losses, noise can be reduced by operating the amplifier at lower temperatures and lower bias currents and voltages. Lastly, the noise figure of the LNA is dependent on the matching circuits, which are designed with an input matching network that minimizes the noise figure and an output matching network to maximize the gain. The optimum input matching network can be found through noise parameter measurements of the HEMT. From these measurements, an equivalent circuit model of the HEMT that includes noise sources can be generated.

C. Mixers

Mixers convert an input signal at one frequency to an output signal at another frequency to permit filtering, phase shifting, or some other data processing operation at a frequency more easily implemented by the circuits. For example, a system may require the data to be received at W-band, 75 to 110 GHz, but W-band filters have a low *Q* or a high loss, which degrades the receiver noise characteristics. Therefore, it may be advantageous to shift the received signal’s frequency to a lower value where low-loss filters are possible. Ideally, this operation is accomplished without degrading the input signal's amplitude or introducing additional noise.

Frequency conversion is accomplished by devices with nonlinear I–V characteristics. Early mixers were all made with diodes, but MESFETs, HEMTs, and PHEMTs have been used more as the technologies have matured. Consider first the diode mixers that can be represented in the simple diagram shown in Figure 3-43. The nonlinearity of the diode I–V characteristics is given in Equation (3-6) of Section 3-II, which is plotted in Figure 3-44. If two voltage signals, labeled the LO and the RF signals, are placed across the diode terminals, the output current shown in Figure 3-44 can be represented by [5]

$$i(v) = I_0 + A\sum_{j=1}^{\infty} B_j\left[V_{LO}\sin(\omega_{LO}t) + V_{RF}\sin(\omega_{RF}t)\right]^j$$

which, upon performance of some trigonometry, can be shown to yield signals with frequencies of

$$f_0 = mf_{RF} \pm nf_{LO}$$

Usually, the desired output frequency is $f_{RF} - f_{LO}$ and this frequency is called the intermediate frequency or the IF.

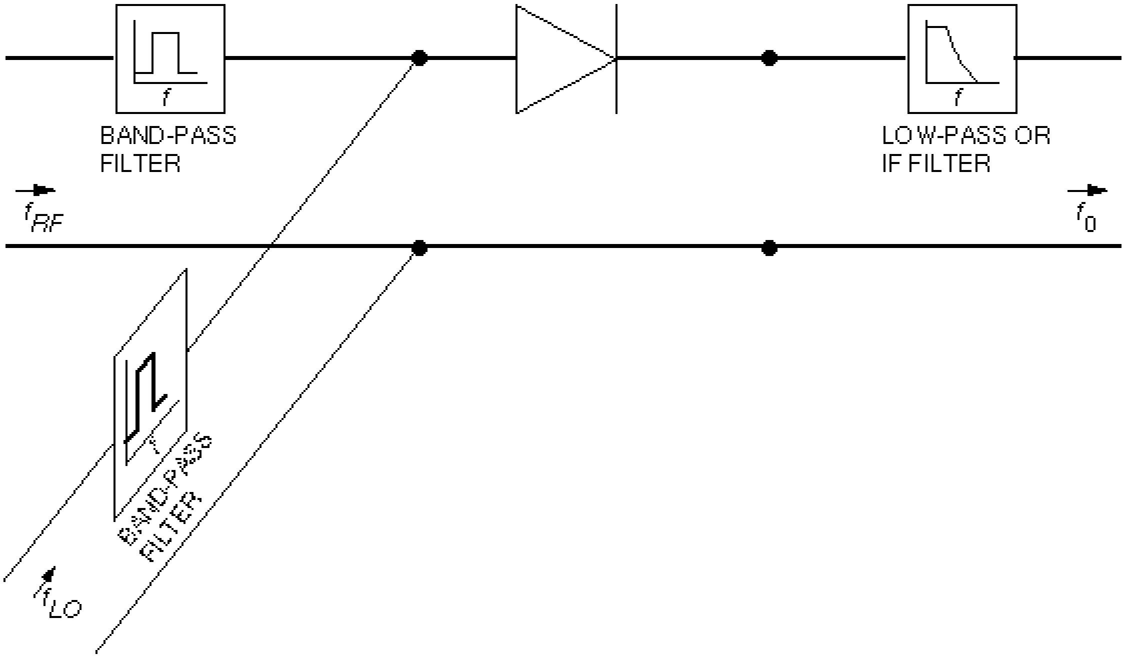

Figure 3-43. Schematic of simple mixer.

A figure of merit for mixers is the ratio of the IF power to the RF power, which is called the conversion loss and is usually specified in dB. There are several contributions to conversion loss. The first is due to poor impedance matching at the RF and IF ports. The second is due to the I–V characteristics of the diode junction, which, if optimized, yields a minimum conversion loss contribution of 3 dB while the remaining half of the power is converted to other frequencies, primarily the image frequency, which is $f_{RF} - 2f_{LO}$ or $2f_{LO} - f_{RF}$. The final contribution to conversion loss is due to the diode parasitics [3] and is given by

$$loss(dB) = 10\log\left(1 + \frac{R_{ohm} + R_{chan}}{R_j} + \left(\omega C_j\right)^2 \left(R_{ohm} + R_{chan}\right) R_j\right)$$

where the parameters are shown in Figure 3-11(c). This loss is minimized when R_j equals $1/\omega C_j$, but since C_j is dependent on the LO power, the conversion loss of a diode mixer is strongly dependent on the LO power and typically decreases to a minimum value with increasing LO power. In addition to optimizing the LO drive power, the cutoff frequency of the diode should be at least 10 times greater than the RF, LO frequency, and IF.

Most FET mixers rely on the nonlinearity of the transconductance by applying the LO and RF signal to the gate of the FET and extracting the IF from the drain. The advantage of FET mixers is that the transistor provides gain that yields mixers with conversion gain instead of conversion loss. The disadvantage of FET mixers is that they also amplify low-frequency noise, 1/*f* noise, which can be converted to a frequency in the desired spectrum.

Note that power at all frequencies—whether from the LO, the RF, noise, or LO instability—applied across the nonlinear device generates power at other frequencies. The elimination of these noise-generated signals and the harmonic frequencies of the RF and LO is critical to the system performance. Many design configurations are possible from the simple single-ended mixers that require only one diode or FET (Figure 3-43) to those requiring up to eight diodes. The more complex circuits use symmetry to cancel frequency components that are not desired and to help eliminate noise created by amplitude variations in the LO. The disadvantage of the mixers with more diodes or FETs is the need for more LO power, which is difficult to obtain at higher frequencies. Reliability problems associated with mixers relate to the generation of harmonics that can cause oscillations in other circuits or the package, distortion of the signal created by harmonics, 1/*f* noise, and device burnout.

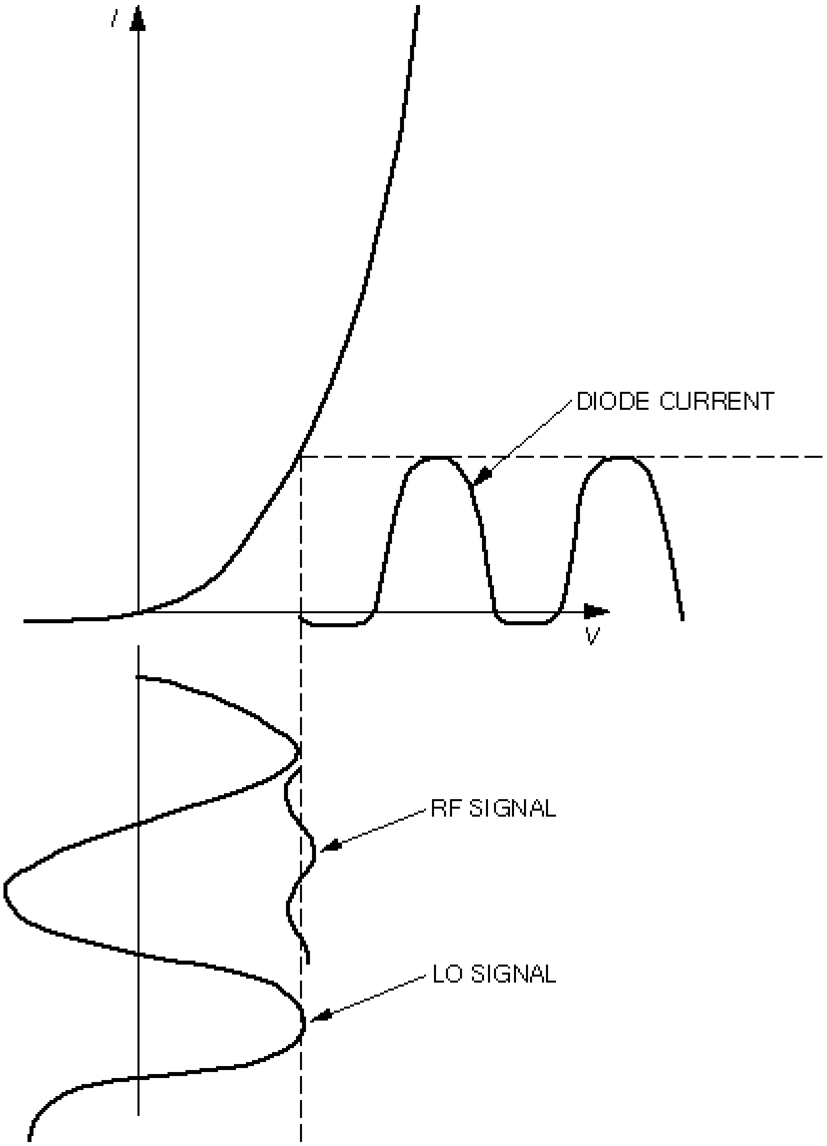

Figure 3-44. Mixer diode I–V characteristics.

D. Oscillators

Oscillators generate microwave energy for communications, radars, and navigation systems. For example, modulators, superheterodyne receivers, and phased-locked loops depend on a good microwave source to function. In principal, any amplifier could be made into an oscillator by providing positive feedback to the input terminals so that the reflection coefficient of the amplifier is greater than one. More often than not, this is accidentally done by amplifier designers. Therefore, an oscillator is basically an LNA with a feedback loop that introduces delay-of-integer multiples of 2π. The choice of the load and terminating impedance to achieve this condition should also guarantee the proper oscillation frequency and maximize the efficiency or RF power delivered to a load. In general, there are two types of oscillators: fixed-frequency oscillators designed to operate at a single frequency and variable-frequency oscillators or voltage-controlled oscillators (VCOs) with tuning circuits that change the oscillation frequency. The schematic of a simple oscillator is shown in Figure 3-45. It consists of a transistor with feedback between the gate and drain, an output matching circuit, and a resonant structure on the input. Oscillator performance specifications or figures of merit that affect the system reliability include phase noise and thermal stability.

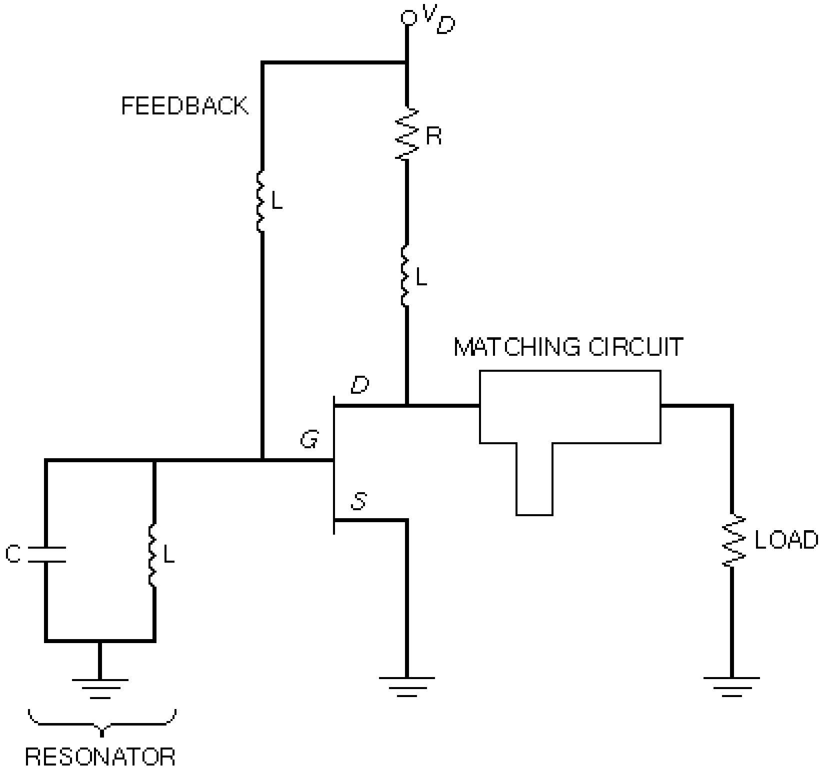

Figure 3-45. Schematic of oscillator.

The phase noise of an oscillator is a measure of the short-term instability of the generated RF signal and is critical in radar applications and digital communication systems where phase noise degrades the system bit error rate (BER). To describe the phase noise, consider a general signal described by

$$V(t) = [V_0 + P(t)] * \sin(\omega_0 t + \phi(t))$$

where $P(t)$ is the amplitude noise term and $f(t)$ is the phase noise term. For $|P(t)| << V_O$ and $|f(t)| << 1$ rad,

$$V(t) \approx V_0 \sin(\omega_0 t) + V_0 \phi(t) \cos(\omega_0 t) + P(t) \sin(\omega_0 t)$$

where the first term represents the desired RF signal and the last two terms represent the amplitude modulated RF signal due to the phase and amplitude noise, respectively. In practice, the phase noise manifests itself as continuous energy sidebands around the carrier in the frequency domain. For the usual case when the phase noise is significantly greater than the amplitude noise, the spectrum around the desired carrier frequency is symmetric.

Noise can be generated by several mechanisms. The first is associated with the kinetic energy of electrons, which is proportional to the temperature of the materials, and thus it is usually called thermal noise. Thermal noise is essentially uniform in magnitude across the entire frequency spectrum, or it is very broad band, which is why it is also referred to as “white” noise, since white light is broad band. The second type of noise is proportional to $1/f$ and is frequently called “flicker” noise because of historical observations of the plate current in vacuum tubes. Flicker noise in active solid-state devices is due to the generation and recombination of carriers at the semiconductor surface [4].

The power spectral density (rad^2/Hz) of phase fluctuations is proportional to the rms phase deviation squared, which results in the spectral slopes of the white and flicker noise becoming twice as steep. It has been shown that the power spectral density decreases at 9 dB/octave where flicker noise dominates, at 6 dB /octave up to the feedback loop half-power bandwidth, and at 0 dB/octave up to the system filter bandwidth, as shown in Figure 3-46 [6]. The figure of merit most often used in oscillators is the ratio of the single sideband noise power per hertz to the carrier signal power at a specific offset frequency.

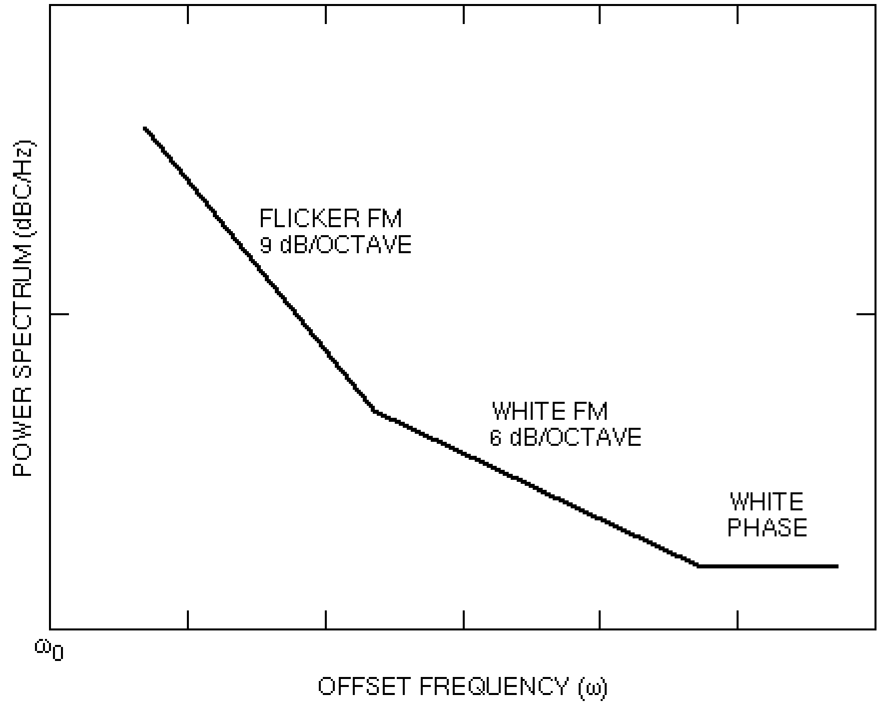

Figure 3-46. Power spectrum for a typical oscillator.

To minimize phase noise, high-*Q* resonators are required to lock in the frequency of the oscillator by providing a reflection coefficient greater than one over a very narrow bandwidth, and transistors with low 1/*f* noise are required. For MMICs, the development of high-*Q* resonators is the more difficult of the two to obtain since thin-film circuit elements on thin GaAs substrates have high conductor loss. HBTs have low 1/*f* noise and are thus frequently used in oscillators. Thermal drift can change the transistor characteristics and cause a shift in the oscillation frequency or cause the circuit to stop oscillating. Temperature compensation can be built in through the use of varactor diodes or other controllable elements with sensors and control circuits.

E. Phase Shifters

Phase shifters are used to impart a repeatable and controllable change of phase to a microwave signal with no effect on the signal's amplitude. Although they are usually associated with phased-array antennas, where they are used to control the beam shape and direction, they are also used in communication systems, radar systems, and microwave instrumentation. Two methods are commonly used to change the phase in MMICs. The first method switches the signal between a short and a long length of transmission line to develop a phase shift of $\beta\ell$ where β is the propagation constant of the transmission line and ℓ is the differential transmission line length. This type of phase shifter is called a switched-line phase shifter and is a true time-delay phase shift. The second method changes the reactance of a transmission line, which changes the propagation constant along the line. The implementation of MMIC phase shifters is broadly characterized as either reflection type or transmission type.

1. Reflection-Type Phase Shifters

Reflection-type phase shifters are one-terminal devices that rely on the reflection of the microwave signal from a termination (e.g., short, open, or other impedance) that has an ideal reflection coefficient with a magnitude of one. An example of a reflection-type phase shifter that employs a switch to add a length of line before the reflective load is shown in Figure 3-47. The resultant transmission-line/termination combination yields a phase shift of $2\beta\ell$ plus a phase that is due to the difference in the termination reactance. Typically, the switches are PIN diodes or MESFETs. Alternatively, an analog phase shifter can be made by removing the switch and replacing the termination load with a varactor diode. Reflection-type phase shifters are primarily used in reflect-array radar applications, or with a coupler to form a transmission-type phase shifter.

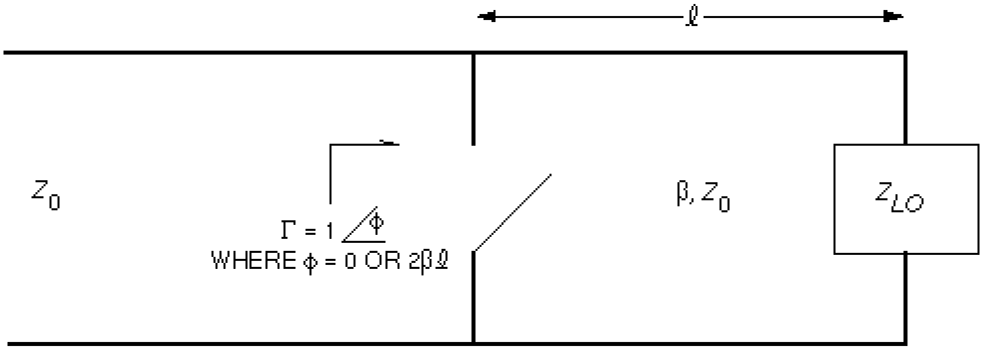

Figure 3-47. Schematic of reflective-type phase shifter.

2. Transmission-Type Phase Shifters

Transmission-type phase shifters are two-terminal devices that change the phase of the input signal as it passes through the circuit. There are three commonly used MMIC implementations of transmission-type phase shifters: hybrid coupled, loaded line, and switched line. The hybrid coupled phase shifters use a reflection-type phase shifter with a coupler to separate the input port from the output port, yielding a two-terminal device. Figure 3-48 is a photograph of an analog hybrid-coupled phase shifter that uses a Lange coupler and varactor diodes.

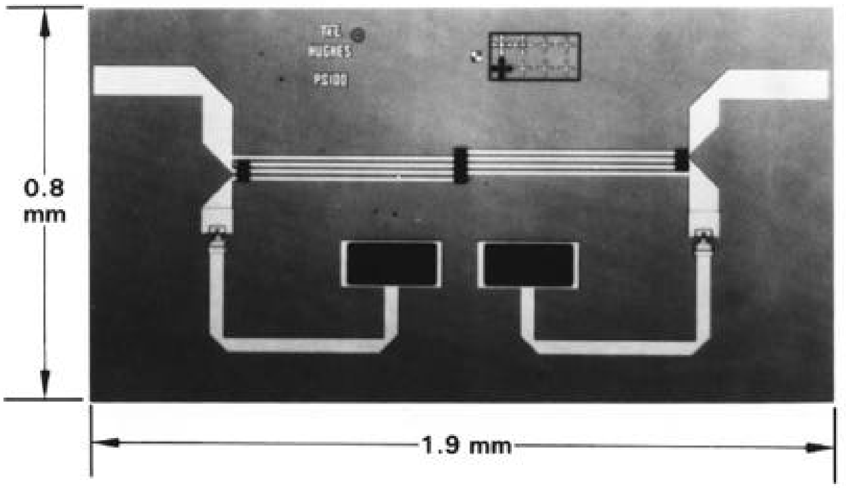

Figure 3-48. Analog phase shifter comprised of a varactor-tuned reflective load and a Lange coupler. (Fabricated by Hughes Aircraft Co. under a contract to NASA Lewis Research Center.)

Figure 3-49 shows a phase shifter comprised of both a switched-line section and a loaded-line section; the schematics of the individual phase-shifter elements are shown in Figure 3-50. The loaded-line phase shifter shown in Figure 3-50(a) is typically comprised of two identical sets of reactive elements separated by about a quarter-wavelength transmission line so that reflections from the reactive elements cancel at the input terminal of the phase shifter. Phase shift is generated by changing the loading on the transmission line and therefore changing its propagation constant, which is approximated by $\beta = \sqrt{LC}$. This type of phase shifter is used for phase shifts less than about 45 deg. The switched-line section is the most straightforward of all. It offers a true time-delay phase shift by switching between two different lengths of transmission line.

Phase shifters are not usually high-power circuits, and therefore the reliability concerns associated with high-power circuits do not need to be considered. Furthermore, the phase shift created by the switch-line type of phase shifters is dependent on transmission line lengths only, and they are therefore very stable over time and temperature. Parametric drift of the active components in the analog type of phase shifters normally translates directly into a phase-state degradation. The active devices used for switching elements (PIN diodes and MESFETs) may also suffer from parametric drift, but this usually manifests itself as a degradation in the insertion loss of the circuit and not the phase shift. Finally, burnout of the switches or varactors must be avoided by proper device and system design.

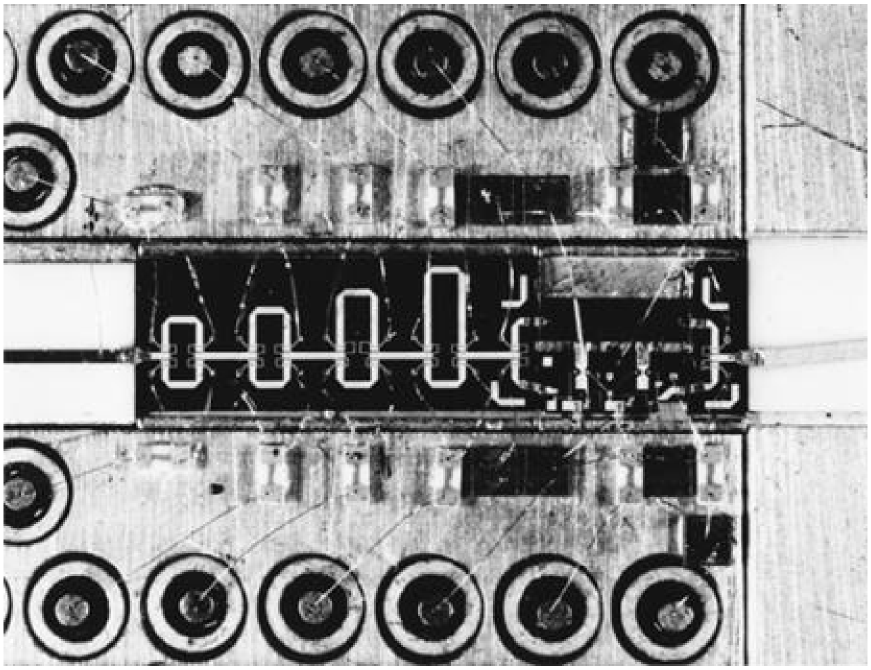

Figure 3-49. Phase shifter comprised of loaded-line and switched-line sections. (Fabricated by Honeywell under a contract to NASA Lewis Research Center.)

References

[1] I. Bahl and P. Bhartia, *Microwave Solid State Circuit Design*, John Wiley & Sons, New York, 1988.

[2] T. T. Ha, *Solid-State Microwave Amplifier Design*, John Wiley & Sons, New York, 1981.

[3] *Monolithic Microwave Integrated Circuits: Technology & Design*, R. Goyal, Editor, Artech House, Norwood, MA, 1989.

[4] S. M. Sze, *Physics of Semiconductor Devices*, John Wiley & Sons, New York, 1981.

[5] C. M. Howell and S. J. Parisi, *Principles, Applications and Selection of Receiving Diodes*, MA-COM Application Note AG314.

[6] D. B. Leeson, “A Simple Model of Feedback Oscillator Noise Spectrum,” *IEEE Proceedings Letters*, Feb. 1966, pp. 329–330.

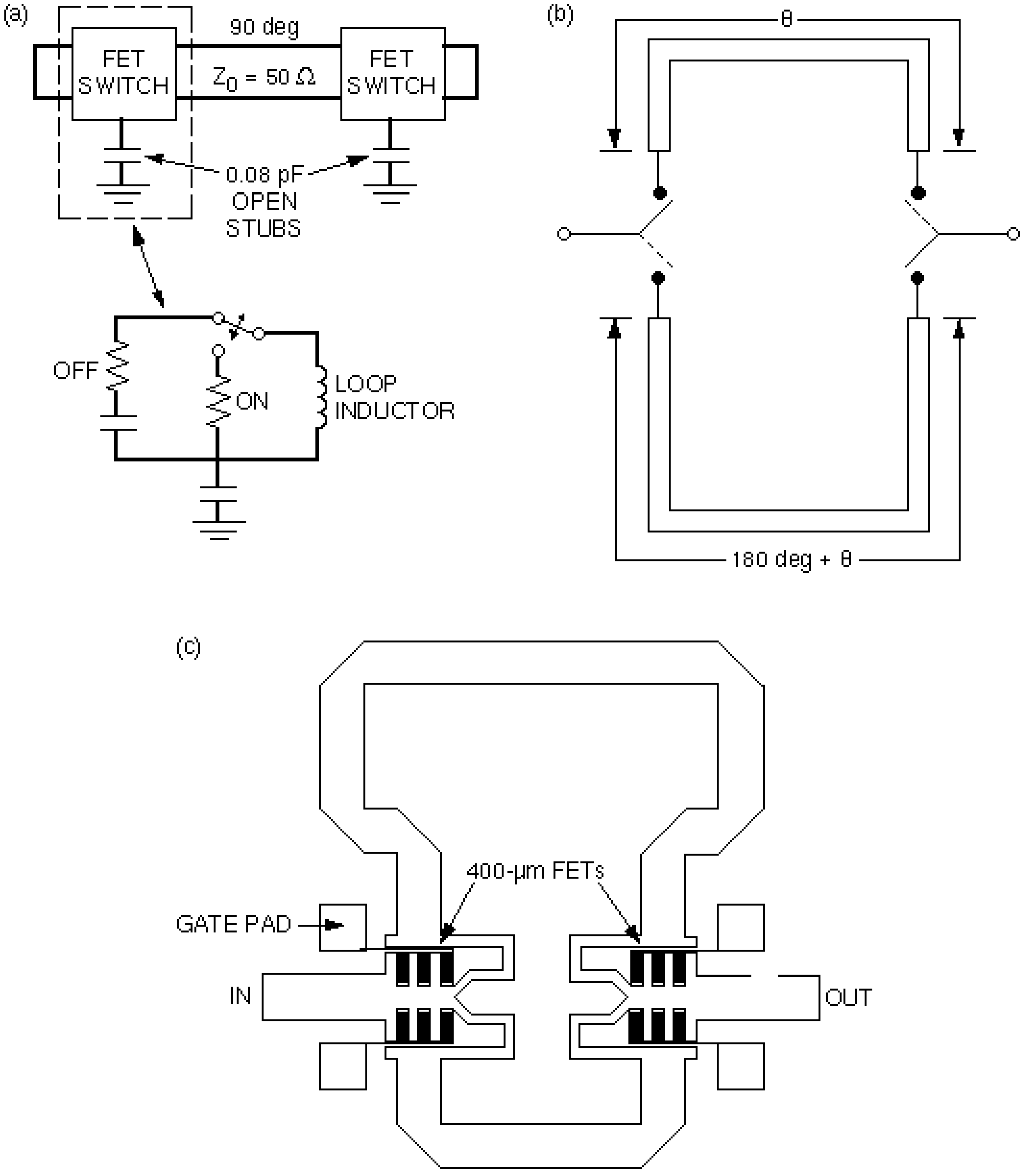

Figure 3-50. 22.5-deg phase-shifter elements: (a) schematic of circuit that is implemented by a loaded-line section; (b) schematic of 180-deg bit with switched lines; (c) layout of 180-deg switched-line bit using four series FET switches.