Chapter 10

Oscillators and Synthesizers

Just say in public that oscillators are one of the most important, fundamental building blocks in radio technology and you will immediately be interrupted by someone pointing out that *tuned-RF* (*TRF*) receivers can be built without any form of oscillator at all. This is certainly true, but it shows how some things can be taken for granted. What use is any receiver without signals to receive? All intentionally transmitted signals trace back to some sort of signal generator—an oscillator or frequency synthesizer. In contrast with the TRF receivers just mentioned, a modern, all-mode, feature-laden MF/HF transceiver may contain in excess of a dozen RF oscillators and synthesizers, while a simple QRP CW transmitter may consist of nothing more than a single oscillator. (This chapter was written by David Stockton, GM4ZNX. Frederick J. Telewski, WA7TZY, also contributed to the Frequency Synthesizers section.)

In the 1980s, the main area of progress in the performance of radio equipment was the recognition of receiver intermodulation as a major limit to our ability to communicate, with the consequent development of receiver front ends with improved ability to handle large signals. So successful was this campaign that other areas of transceiver performance now require similar attention. One indication of this is any equipment review receiver dynamic range measurement qualified by a phrase like "limited by oscillator phase noise." A plot of a receiver's effective selectivity can provide another indication of work to be done: An IF filter's high-attenuation region may appear to be wider than the filter's published specifications would suggest—almost as if the filter characteristic has grown sidebands! In fact, in a way, it has: This is the result of local-oscillator (LO) or synthesizer *phase noise* spoiling the receiver's overall performance. Oscillator noise is the prime candidate for the next major assault on radio performance.

The sheer number of different oscillator circuits can be intimidating, but their great diversity is an illusion that evaporates once their underlying pattern is seen. Almost all RF oscillators share one fundamental principle of operation: an amplifier and a filter operate in a loop (**Fig 10.1**). There are plenty of filter types to choose from:

- LC
- Quartz crystal and other piezoelectric materials
- Transmission line (stripline, microstrip, troughline, open-wire, coax and so on)
- Microwave cavities, YIG spheres, dielectric resonators
- Surface-acoustic-wave (SAW) devices

Should any new forms of filter be invented, it's a safe guess that they will also be applicable to oscillators. There is an equally large range of amplifiers to choose from:

- Vacuum tubes of all types
- Bipolar junction transistors
- Field effect transistors (JFET, MOSFET, GaAsFET, in all their varieties)
- Gunn diodes, tunnel diodes and other negative-resistance generators

It seems superfluous to state that anything that can amplify can be used in an oscillator, because of the well-known propensity of all prototype amplifiers to oscillate! The choice of amplifier is widened further by the option of using single- or multiple-stage amplifiers and discrete devices versus integrated circuits. Multiply all of these options with those of filter choice and the resulting set of combinations is very large, but a long way from complete. Then there are choices of how to couple the amplifier into the filter and the filter into the amplifier. And then there are choices to make in the filter section: Should it be tuned by variable capacitor, variable inductor or some form of sliding cavity or line?

Despite the number of combinations that are possible, a manageably small number of types will cover all but very special requirements. Look at an oscillator circuit and "read" it: What form of filter—*resonator*—does it use? What form of amplifier? How have the amplifier's input and output been coupled into the filter? How is the filter tuned? These are simple, easily answered questions that put oscillator types into appropriate categories and make them understandable. The questions themselves may make more

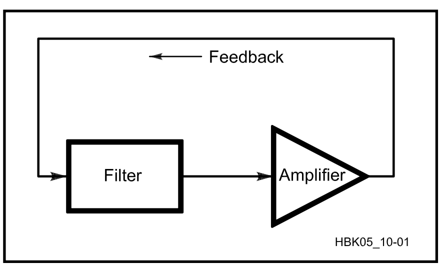


Fig 10.1—Reduced to essentials, an oscillator consists of a filter and an amplifier operating in a feedback loop.

sense if we understand the mechanics of oscillation, in which *resonance* plays a major role.

HOW OSCILLATORS WORK

Maintained Resonance

The pendulum, a good example of a resonator, has been known for millennia and understood for centuries. It is closely analogous to an electronic resonator, as shown in **Fig 10.2**. The weighted end of the pendulum can store energy in two different forms: The *kinetic energy* of its motion and the *potential energy* of it being raised above its rest position. As it reaches its highest point at the extreme of a swing, its velocity is zero for an instant as it reverses direction. This means that it has, at that instant, no kinetic energy, but because it is also raised above its rest position, it has some extra potential energy.

At the center of its swing, the pendulum is at its lowest point with respect to gravity and so has lost the extra potential energy. At the same time, however, it is moving at its highest speed and so has its greatest kinetic energy. Something interesting is happening: The pendulum's stored energy is continuously moving between potential and kinetic forms. Looking at the pendulum at intermediate positions shows that this movement of energy is smooth. Newton provided the keys to understanding this. It took his theory of gravity and laws of motion to explain the behavior of a simple weight swinging on the end of a length of string and calculus to perform a quantitative mathematical analysis. Experiments had shown the period of a pendulum to be very stable and predictable. Apart from side effects of air drag and friction, the length of the period should not be affected by the mass of the weight, nor by the amplitude of the swing.

A pendulum can be used for timing events, but its usefulness is spoiled by the action of drag or friction, which eventually stops it. This problem was overcome by the invention of the *escapement*, a part of a clock mechanism that senses the position of the pendulum and applies a small push in the right direction and at the right time to maintain the amplitude of its swing or oscillation. The result is a mechanical oscillator: The pendulum acts as the filter, the escapement acts as the amplifier and a weight system or wound-up spring powers the escapement.

Electrical oscillators are closely analogous to the pendulum, both in operation and in development. The voltage and current in the tuned circuit—often called *tank circuit* because of its energy-storage ability—both vary sinusoidally with time and are 90° out of phase. There are instants when the current is zero, so the energy stored in the inductor must be zero, but at the same time the voltage across the capacitor is at its peak, with all of the circuit's energy stored in the electric field between the capacitor's plates. There are also instants when the voltage is zero and the current is at a peak, with no energy in the capacitor. Then, all of the circuit's energy is stored in the inductor's magnetic field.

Just like the pendulum, the energy stored in the electrical system is swinging smoothly between two forms; electric field and magnetic field. Also like the pendulum, the tank circuit has losses. Its conductors have resistance, and the capacitor dielectric and inductor core are imperfect. Leakage of electric and magnetic fields also occurs, inducing currents in neighboring objects and just plain radiating energy off into space as radio waves. The amplitudes of the oscillating voltage and current decrease steadily as a result. Early intentional radio transmissions, such as those of Heinrich Hertz's experiments, involved abruptly dumping energy into a tuned circuit and letting it oscillate, or *ring*, as shown in **Fig 10.3**. This was done by applying a spark to the resonator. Hertz's resonator was a gapped ring, a good choice for radiating as much of the energy as possible. Although this looks very different from the LC tank of Fig 10.2, it has inductance *distributed* around its length and capacitance distributed across it and across its gap, as opposed to the *lumped* L and C values in Fig 10.2. The gapped ring therefore works just the same as the LC tank in terms of oscillating voltages and currents. Like the pendulum and the LC tank, its period, and

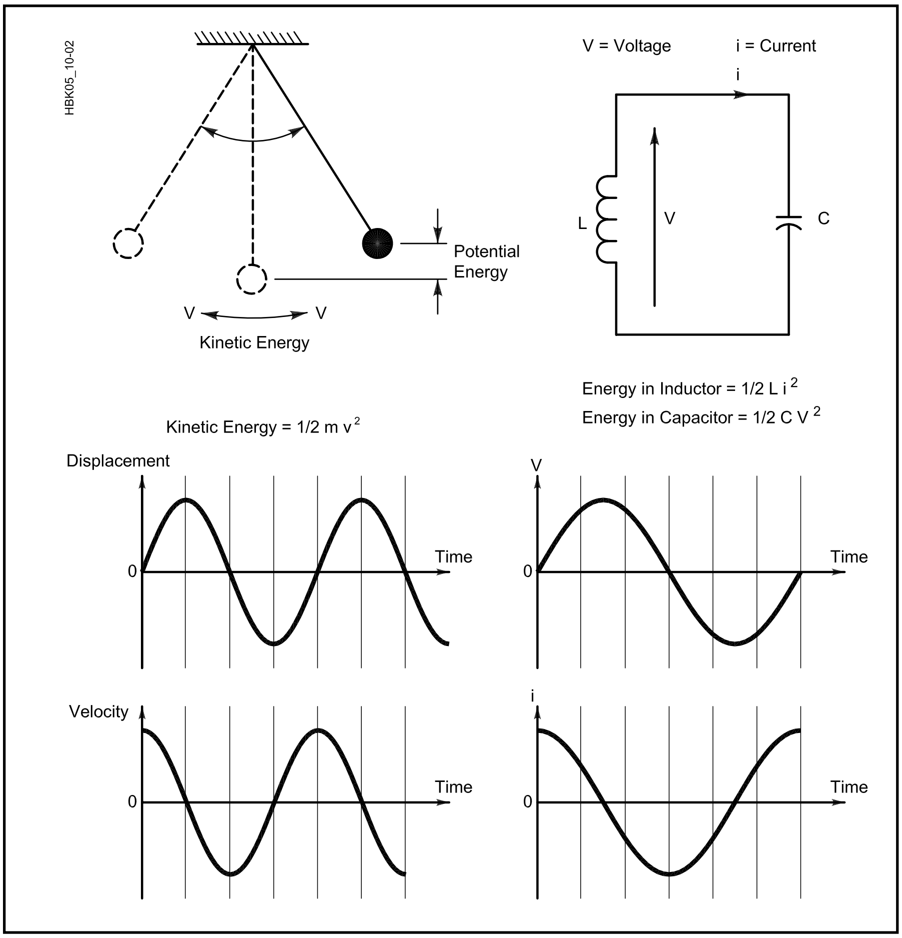


Fig 10.2—A resonator lies at the heart of every oscillatory mechanical and electrical system. A mechanical resonator (here, a pendulum) and an electrical resonator (here, a tuned circuit consisting of L and C in parallel) share the same mechanism: the regular movement of energy between two forms—potential and kinetic in the pendulum, electric and magnetic in the tuned circuit. Both of these resonators share another trait: Any oscillations induced in them eventually die out because of losses—in the pendulum, due to drag and friction; in the tuned circuit, due to the resistance, radiation and inductance. Note that the curves corresponding to the pendulum's displacement vs velocity and the tuned circuit's voltage vs current, differ by one quarter of a cycle, or 90°.

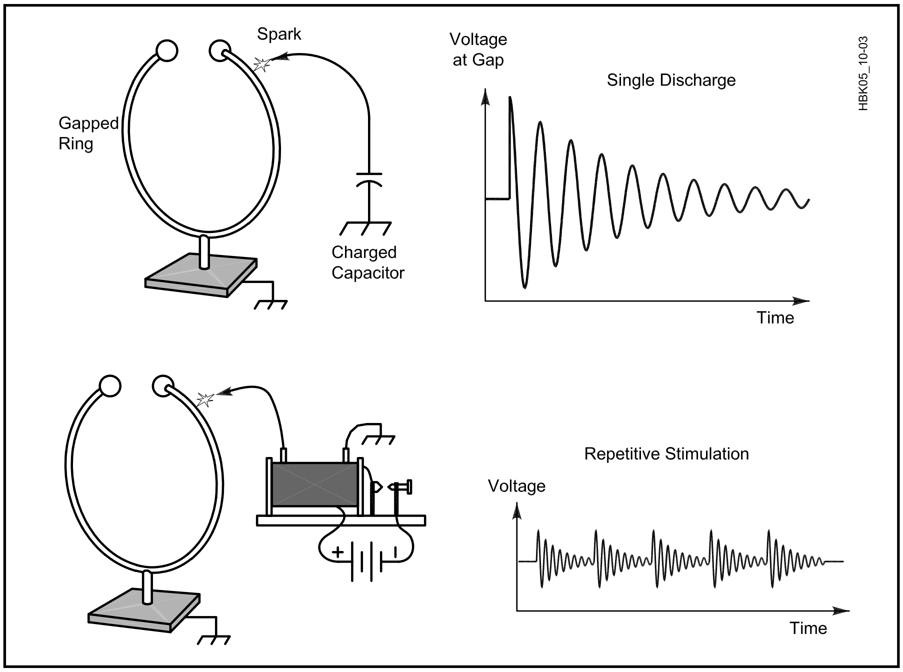


Fig 10.3—Stimulating a resonance, 1880s style. Shock-exciting a gapped ring with high voltage from a charged capacitor causes the ring to oscillate at its resonant frequency. The result is a damped wave, each successive alternation of which is weaker than its predecessor because of resonator losses. Repetitively stimulating the ring produces trains of damped waves, but oscillation is not continuous.

therefore the frequency at which it oscillates, is independent of the magnitude of its excitation.

Making a longer-lasting signal with the Fig 10.3 arrangement merely involves repeating the sparks. The problem is that a truly continuous signal cannot be made this way. The sparks cannot be applied often enough or always timed precisely enough to guarantee that another spark re-excites the circuit at precisely the right instant. This arrangement amounts to a crude spark transmitter, variations of which served as the primary means of transmission for the first generation of radio amateurs. The use of damped waves is now entirely forbidden by international treaty because of their great impurity. Damped waves look a lot like car-ignition waveforms and sound like car-ignition interference when received.

What we need is a *continuous wave* (*CW*) oscillation—a smooth, sinusoidal signal of constant amplitude, without phase jumps, a "pure tone." To get it, we must add to our resonator an equivalent of the clock's escapement—a means of synchronizing the application of energy and a fast enough system to apply just enough energy every cycle to keep each cycle at the same amplitude.

Amplification

A sample of the tank's oscillation can be extracted, amplified and reinserted. The gain can be set to exactly compensate the tank losses and perfectly maintain the oscillation. The amplifier usually need only give low gain, so active devices can be used in oscillators not far below their unity (unity = 1) gain frequency. The amplifier's output must be lightly coupled into the tank—the aim is just to replace lost energy, not forcibly drive the tank. Similarly, the amplifier's input should not heavily load the tank. It is a good idea to think of *coupling* networks rather than *matching* networks in this application, because a matched impedance extracts the maximum available energy from a source, and this would certainly spoil an oscillator.

Fig 10.4A shows the block diagram of an oscillator. Certain conditions must be met for oscillation. The criteria that separate oscillator loops from stable loops are often attributed to Barkhausen by those aiming to produce an oscillator and to Nyquist by those aiming for amplifier stability, although they boil down to the same boundary. Fig 10.4B shows the loop broken and a test signal inserted. (The loop can be broken anywhere; the amplifier input just happens to be the easiest place to do it.) The criterion for oscillation says that at a frequency at which the phase shift around the loop is exactly zero, the net gain around the loop must equal or exceed unity (that is, one). (Later, when we design phase-locked loops, we must revisit this concept in order to check that *those* loops *cannot* oscillate.) Fig 10.4C shows what happens to the amplitude of an oscillator if the loop gain is made a little higher or lower than one.

The loop gain has to be precisely one if we want a stable amplitude. Any inaccuracy will cause the amplitude to grow to clipping or shrink to zero, making the oscillator useless. Better accuracy will only slow, not stop this process. Perfect precision is clearly impossible, yet there are enough working oscillators in existence to prove that we are missing something important. In an amplifier, nonlinearity is a nuisance, leading to signal distortion and intermodulation, yet nonlinearity is what makes stable oscillation possible. All of the vacuum tubes and transistors used in oscillators tend to reduce their gain at higher signal levels. With such components, only a tiny change in gain can shift the loop's operation between amplitude growth and shrinkage. Oscillation stabilizes at that level at which the gain of the active device sets the loop gain at exactly one.

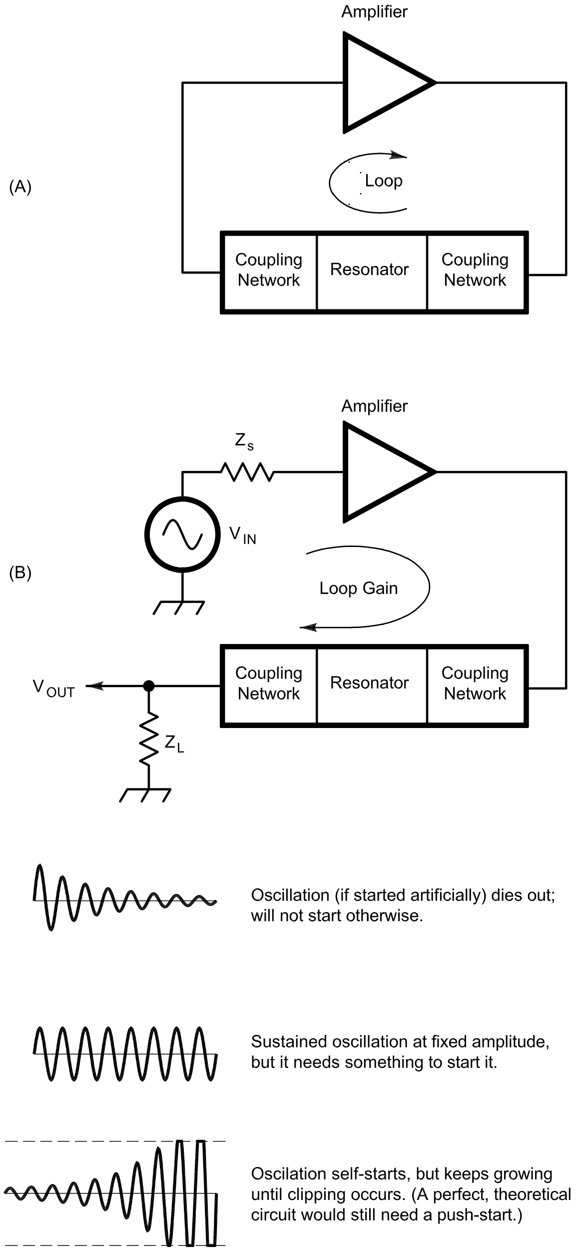
Another gain-stabilization technique involves biasing the device so that once some level is reached, the device starts to turn off over part of each cycle. At higher levels, it cuts off over more of each cycle. This effect reduces the effective gain quite strongly, stabilizing the amplitude. This badly distorts the signal (true in most common oscillator circuits) in the amplifying device, but provided the amplifier is lightly coupled to a high-Q resonant tank, the signal in the tank should not be badly distorted.

Many radio amateurs now have some form of circuit-analysis software, usually running on a PC. Attempts to analyze oscillators by this means often fail by predicting growing or shrinking amplitudes, and often no signal at all, in circuits that are known to work. Computer analysis of oscillator circuits can be done, but it requires a sophisticated program with accurate, nonlinear, RF-valid models of the devices used, to be able to predict operating amplitude. Often even these programs need some special tricks to get their modeled oscillators to start. Such software is likely to be priced higher than most private users can justify, and it still doesn't replace the need for the user to understand the circuit. With that understanding, some time, some parts and a little patience will do the job, unassisted.

Textbooks give plenty of coverage to the frequency-determining mechanisms of oscillators, but the amplitude-determining mechanism is rarely covered. It is often not even mentioned. There is a

HBK05_10-04

Fig 10.4—The bare-bones oscillator block diagram of Fig 10.1 did not include two practical essentials: Networks to couple power in and out of the resonator (A). Breaking the loop, inserting a test signal and measuring the loop's overall gain (B) allows us to determine whether the system can oscillate, sustain oscillation or clip (C).



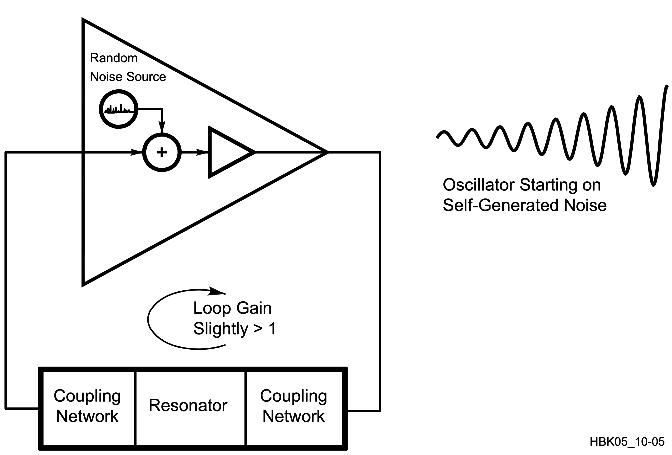


Fig 10.5—An oscillator with noise. Real-world amplifiers, no matter how quiet, generate some internal noise; this allows real-world oscillators to self-start.

good treatment in Clarke and Hess, *Communications Circuits: Analysis and Design*.

Start-Up

Perfect components don't exist, but if we could build an oscillator from them, we would naturally expect perfect performance. We would nonetheless be disappointed. We could assemble from our perfect components an oscillator that exactly met the criterion for oscillation, having slightly excessive gain that falls to the correct amount at the target operating level and so is capable of sustained, stable oscillation. But being capable of something is not the same as doing it, for there is another stable condition. If the amplifier in the loop shown in Fig 10.4 has no input signal, and is perfect, it will give no output! No signal returns to the amplifier's input via the resonator, and the result is a sustained and stable *lack* of oscillation. Something is needed to start the oscillator.

This fits the pendulum-clock analogy: A wound-up clock is stable with its pendulum at rest, yet after a push the system will sustain oscillation. The mechanism that drives the pendulum is similar to a Class C amplifier: It does not act unless it is driven by a signal that exceeds its threshold. An electrical oscillator based on a Class C amplifier can sometimes be kicked into action by the turn-on transient of its power supply. The risk is that this may not always happen, and also that should some external influence stop the oscillator, it will not restart until someone notices the problem and cycles the power. This can be very inconvenient!

A real-life oscillator whose amplifier does not lose gain at low signal levels can self-start due to noise. **Fig 10.5** shows an oscillator block diagram with the amplifier's noise shown, for our convenience, as a second input that adds with the true input. The amplifier amplifies the noise. The resonator filters the output noise, and this signal returns to the amplifier input. The importance of having slightly excessive gain until the oscillator reaches operating amplitude is now obvious. If the loop gain is slightly above one, the recirculated noise must, within the resonator's bandwidth, be larger than its original level at the input. More noise is continually summed in as a noise-like signal continuously passes around the loop, undergoing amplification and filtering as it does. The level increases, causing the gain to reduce. Eventually, it stabilizes at whatever level is necessary to make the net loop gain equal to one.

So far, so good. The oscillator is running at its proper level, but something seems very wrong. It is not making a proper sine wave; it is recirculating and filtering a noise signal. It can also be thought of as a Q multiplier with a controlled (high) gain, filtering a noise input and amplifying it to a set level. Narrow-band filtered noise approaches a true sine wave as the filter is narrowed to zero width. What this means is that we cannot make a true sine-wave signal—all we can do is make narrow-band filtered noise as an approximation to one. A high-quality, low-noise oscillator is merely one that does tighter filtering. Even a kick-started Class-C-amplifier oscillator has noise continuously entering its circulating signal, and so behaves similarly.

A small-signal gain greater than one is absolutely critical for reliable starting, but having too much gain can make the final operating level unstable. Some oscillators are designed around limiting amplifiers to make their operation predictable. AGC systems have also been used, with an RF detector and dc amplifier used to servo-control the amplifier gain. It is notoriously difficult to design reliable crystal oscillators that can be published or mass-produced without having occasional individuals refuse to start without some form of shock.

Mathematicians have been intrigued by "chaotic systems" where tiny changes in initial conditions can yield large changes in outcome. The most obvious example is meteorology, but much of the necessary math was developed in the study of oscillator start-up, because it is a case of chaotic activity in a simple system. The equations that describe oscillator start-up are similar to those used to generate many of the popular, chaotic fractal images.

PHASE NOISE

Viewing an oscillator as a filtered-noise generator is relatively modern. The older approach was to think of an oscillator making a true sine wave with an added, unwanted noise signal. These are just different ways of visualizing the same thing: They are equally valid views, which are used interchangeably, depending which best makes some point clear. Thinking in terms of the signal-plus-noise version, the noise surrounds the carrier, looking like sidebands and so can also be considered to be equivalent to random-noise FM and AM on the ideal sine-wave signal. This gives us a third viewpoint. Strangely, these noise sidebands are called *phase noise*. If we consider the addition of a noise voltage to a sinusoidal voltage, we

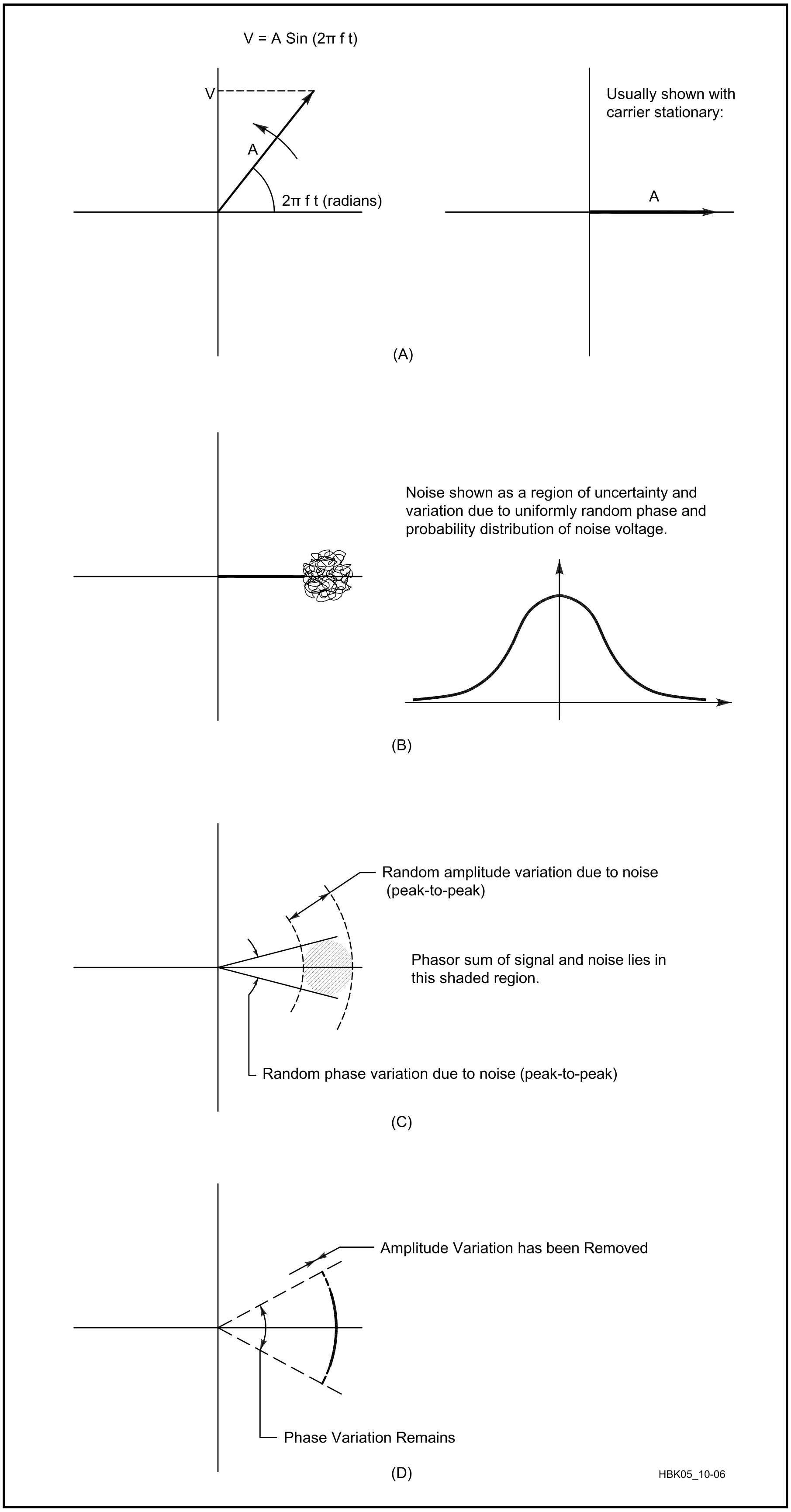


Fig 10.6—At A, a phasor diagram of a clean (ideal) oscillator. Noise creates a region of uncertainty in the vector's length and position (B). AM noise varies the vector's length; PM noise varies the vector's relative angular position (C) Limiting a signal that includes AM and PM noise strips off the AM and leaves the PM (D).

must take into account the phase relationship. A *phasor diagram* is the clearest way of illustrating this. **Fig 10.6A** represents a clean sine wave as a rotating vector whose length is equal to the peak amplitude and whose frequency is equal to the number of revolutions per second of its rotation. Moving things are difficult to depict on paper, so phasor diagrams are usually drawn to show the dominant signal as stationary, with other components drawn relative to this.

Noise contains components at many frequencies, so its phase with respect to the dominant, theoretically pure signal—the "carrier"—is random. Its amplitude is also random. Noise can only be described in statistical terms because its voltage is constantly and randomly changing, yet it does have an average amplitude that can be expressed in rms volts. Fig 10.6B shows noise added to the carrier phasor, with the noise represented as a fuzzy, uncertain region in which the sum phasor wanders randomly. The phase of the noise is uniformly random—no direction is more likely than any other—but the instantaneous magnitude of the noise obeys a probability distribution like that shown, higher values being progressively rarer. Fig 10.6C shows how the extremities of the noise region can be considered as extremes of phase and amplitude variation from the normal values of the carrier.

Phase modulation and frequency modulation are closely related. Phase is the integral of frequency, so phase modulation resembles frequency modulation in which deviation decreases with increasing modulating frequency. Thus, there is no need to talk of "frequency noise" because *phase noise* already covers it.

Fig 10.6C clearly shows AM noise as the random variation of the length of the sum phasor, yet "amplitude noise" is rarely discussed. The oscillator's amplitude control mechanism acts to reduce the AM noise by a small amount, but the main reason is that the output is often fed into some form of limiter that strips off AM components just as the limiting IF amplifier in an FM receiver removes any AM on incoming signals. The limiter can be obvious, like a circuit to convert the signal to logic levels, or it can be implicit in some other function. A diode ring mixer may be driven by a sine-wave LO of moderate power, yet this signal drives the diodes hard on and hard off, approximating square-wave switching. This is a form of limiter, and it removes the effect of any AM on the LO. Fig 10.6D shows the result of passing a signal with noise through a limiting amplifier. For these reasons, AM noise sidebands are rarely a problem in oscillators, and so are normally ignored. There is one subtle problem to beware of, however. If a sine wave drives some form of switching circuit or limiter, and the threshold is offset from the signal's mean voltage, any level changes will affect the exact switching times and cause *phase jitter*. In this way, AM sidebands are translated into PM sidebands and pass through the limiter. This is usually called *AM-to-PM conversion* and is a classic problem of limiters.

Effects of Phase Noise

You would be excused for thinking that phase noise is a recent discovery, but all oscillators have always produced it. Other changes have elevated an unnoticed characteristic up to the status of a serious impairment. Increased crowding and power levels on the ham bands, allied with greater expectations of receiver performance as a result of other improvements, have made phase noise more noticeable, but the biggest factor has been the replacement of VFOs in radios by frequency synthesizers. It is a major task to develop a synthesizer that tunes in steps fine enough for SSB and CW while competing with the phase-noise performance of a reasonable-quality LC

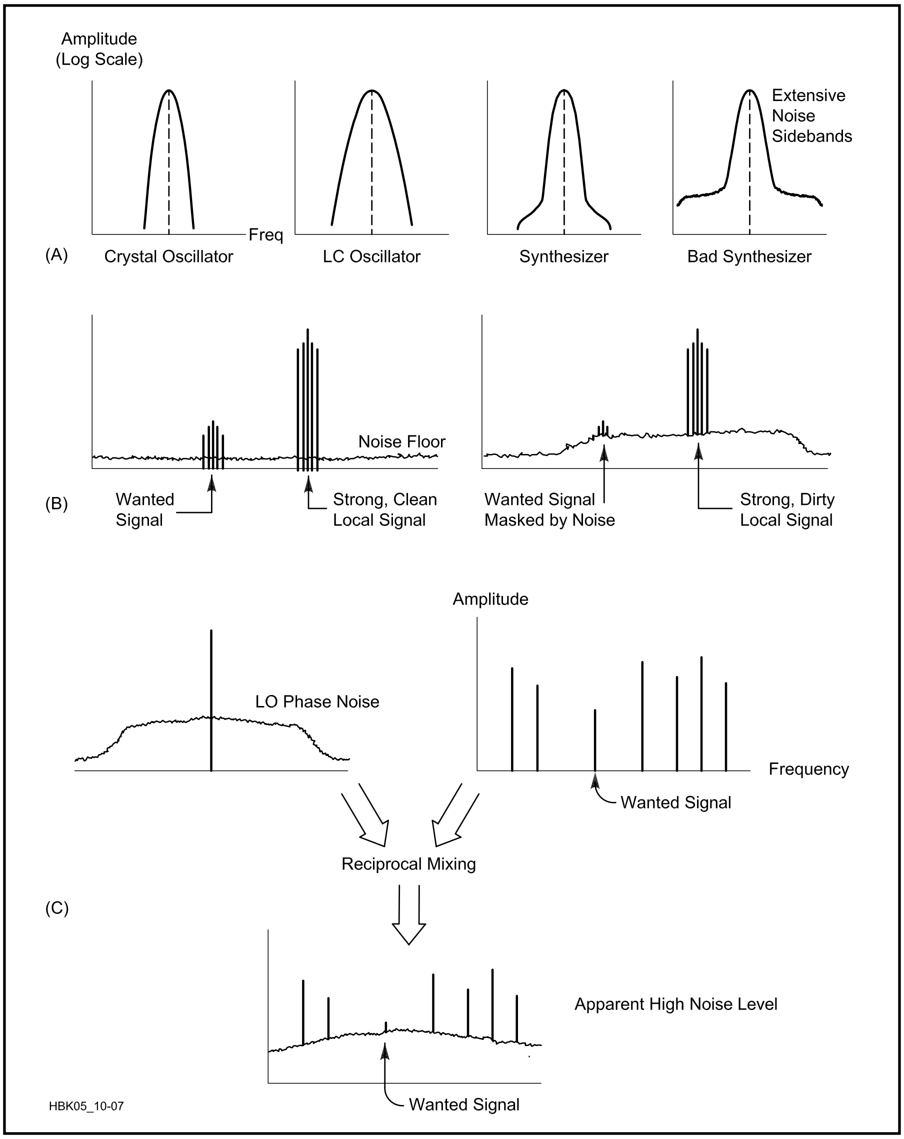


Fig 10.7—The effects of phase noise. At A, the relative phase-noise spectra of several different signal-generation approaches. At B, how transmitted phase noise degrades the weak-signal performance even of nearby receivers with phase-quiet oscillators, raising the effective noise floor. What is perhaps most insidious about phase noise in Amateur Radio communication is that its presence in a receiver LO can allow strong, clean transmitted signals to degrade the receiver's noise floor just as if the transmitted signals were dirty. This effect, reciprocal mixing, is shown at C.

VFO. Many synthesizers have fallen far short of that target. Phase noise is worse in higher-frequency oscillators and the trend towards general-coverage, upconverting structures has required that local oscillators operate at higher and higher frequencies. **Fig 10.7A** shows sketches of the relative phase-noise performance of some oscillators. The very high Q of the quartz crystal in a crystal oscillator gives it the potential for much lower phase noise than LC oscillators. A medium-quality synthesizer has close-in phase noise performance approaching that of a crystal oscillator, while further from the carrier frequency, it degrades to the performance of a modest-quality LC oscillator. There may be a small bump in the noise spectrum at the boundary of these two zones. A bad synthesizer can have extensive noise sidebands extending over many tens (hundreds in extreme cases) of kilohertz at high levels.

Phase noise on a transmitter's local oscillators passes through its stages, is amplified and fed to the antenna along with the intentional signal. The intentional signal is thereby surrounded by a band of noise. This radiated noise exists in the same proportion to the transmitter power as the phase noise was to the oscillator power if it passes through no narrow-band filtering capable of limiting its bandwidth. This radiated noise makes for a noisier band. In bad cases, nearby stations can be unable to receive over many tens of kilohertz. Fig 10.7B illustrates the difference between clean and dirty transmitters.

The effects of receiver-LO phase noise are more complicated, but at least it doesn't affect other stations' reception. The process is called *reciprocal mixing*.

Reciprocal Mixing

This is an effect that occurs in all mixers, yet despite its name, reciprocal mixing is an LO, not a mixer, problem. Imagine that the outputs of two supposedly unmodulated signal generators are mixed together and the mixer output is fed into an FM receiver. The receiver produces sounds, indicating that the resultant signal is modulated nonetheless. Which signal generator is responsible? This is a trick question, of course. A moment spent thinking about how a change in the frequency of either input signal affects the output signal will show that FM or PM on either input reaches the output. The best answer to the trick question is therefore "either or both."

The modulation on the mixer output is the combined modulations of the inputs. This means that modulating a receiver's local oscillator is indistinguishable from using a clean LO and having the same modulation present on *all* incoming signals. (This is also true for AM provided that a fully linear multiplier is used as the mixer, but mixers are commonly driven into switching, which strips any AM off the LO signal. This is the chief reason why the phase component of oscillator noise is more important than any AM component.)

The word *indistinguishable* is important in the preceding paragraph. It does not mean that the incoming signals are themselves modulated, but that the signals in the receiver IF and the noise in the IF, sound exactly as if they were. What really happens is that the noise components of the LO are extra LO signals that are offset from the carrier frequency. Each of them mixes other signals that are appropriately offset from the LO carrier into the receiver's IF. Noise is the sum of an infinite number of infinitesimal components spread over a range of frequencies, so the signals it mixes into the IF are spread into an infinite number of small replicas, all at different frequencies. This amounts to scrambling these other signals into noise. It is tedious to look at the effects of receiver LO phase noise this way. The concept of reciprocal mixing gives us an easier, alternative view that is much more digestible and produces identical results.

A poor oscillator can have significant noise sidebands extending out many tens of kilohertz on either side of its carrier. This is the same, as far as the signals in the receiver IF are concerned, as if the LO were clean and every signal entering the mixer had these noise sidebands. Not only will the wanted signal (and its noise sidebands) be received, but the noise sidebands added by the LO to signals near, but outside, the receiver's IF passband will overlap it. If the band is busy, each of many signals present will add its set of noise sidebands—and the effect is cumulative. This produces the appearance of a high background-noise level on the band. Many hams tend to accept this, blaming "conditions."

Hams now widely understand reception problems due to intermodulation, and almost everyone knows to apply RF attenuation until the signal gets no clearer. Intermodulation is a nonlinear effect, and the levels of the intermod products fall by greater amounts than the reduction in the intermodulating signals. The net result is less signal, but with the intermodulation products dropped still further. This improvement reaches a limit when more attenuation pushes the desired signal too close to the receiver's noise floor.

Reciprocal mixing is a linear process, and the mixer applies the same amount of noise "deviation" to incoming signals as that present on the LO. Therefore the ratio of noise-sideband power to signal power is the same for each signal, and the same as that on the LO. Switching in RF attenuation reduces the power of signals entering the mixer, but the reciprocal mixing process still adds the noise sidebands at the same *relative* power to each. Therefore, no reception improvement results. Other than building a quieter oscillator, the only way of improving things is to use narrow preselection to band-limit the receiver's input response and reduce the number of incoming signals subject to reciprocal mixing. This reduces the number of times the phase noise sidebands get added into the IF signal. Commercial *tracking preselectors*—selective front-end circuits that tune in step with a radio's band changes and tuning, are expensive, but one that is manually tuned would make a modest-sized home-brew project and could also help reduce intermodulation effects. When using a good receiver with a linear front end and a clean LO, amateurs accustomed to receivers with poor phase-noise performance report the impression is of a seemingly emptier band with gaps between signals—and then they begin to find readable signals in some of the gaps.

Fig 10.7C shows how a noisy oscillator affects transmission and reception. The effects on reception are worst in Europe, on 40 m, at night. Visitors from North America, and especially Asia, are usually shocked by the levels of background noise. In ITU Region 1, the Amateur Radio 40-m allocation is 7.0 to 7.1 MHz; above this, ultra-high-power broadcasters operate. The front-end filters in commercial ham gear are usually fixed band-pass designs that cover the wider 40-m allocations in the other regions. This allows huge signals to reach the mixer and mix large levels of LO phase noise into the IF. Operating collocated radios, on the same band, in a multioperator contest, requires linear front ends, preselection and state-of-the-art phase-noise performance. Outside of amateur circles, only warship operation is more demanding, with kilowatt transmitters and receivers sharing antennas on the same mast.

A Phase Noise Demonstration

Healthy curiosity demands some form of demonstration so the scale of a problem can be judged "by ear" before measurements are attempted. We need to be able to measure the noise of an oscillator alone (to aid in the development of quieter ones) and we also need to be able to measure the phase noise of the oscillators in a receiver (a transmitter can be treated as an oscillator). Conveniently, a receiver contains most of the functions needed to demonstrate its own phase noise.

No mixer has perfect port-to-port isolation, and some of its local-oscillator signal leaks through into the IF. If we tune a general-coverage receiver, with its antenna disconnected, to exactly 0 Hz, the local oscillator is exactly at the IF center frequency, and the receiver acts as if it is tuned to a very strong unmodulated carrier. A typical mixer might give only 40 dB of LO isolation and have an LO drive power of at least 10 mW. If we tune away from 0 Hz, the LO carrier tunes away from the IF center and out of the passband. The apparent signal level falls. Although this moves the LO carrier out of the IF passband, some of its noise sidebands will not be, and the receiver will respond to this energy as an incoming noise signal. To the receiver operator, this sounds like a rising noise floor as the receiver is tuned toward 0 Hz. To get good noise floor at very low frequencies, some professional/military receivers, like the Racal RA1772, use very carefully balanced mixers to get as much port-to-port isolation as possible, and they also may switch a crystal notch filter into the first mixer's LO feed.

This demonstration cannot be done if the receiver tunes amateur bands only. As it is, most general-coverage radios inhibit tuning in the LF or VLF region. It could be suggested by a cynic that how low manufacturers allow you to tune is an indication of how far they think their phase-noise sidebands could extend!

The majority of amateur transceivers with general-coverage receivers are programmed not to tune below 30 to 100 kHz, so means other than the "0 Hz" approach are needed to detect LO noise in these radios. Because reciprocal mixing adds the LO's sidebands to clean incoming signals, in the same proportion to the incoming carrier as they exist with respect to the LO carrier, all we need do is to apply a strong, clean signal wherever we want within the receiver's tuning range. This signal's generator must have lower phase noise than the radio being evaluated. A general-purpose signal generator is unlikely to be good enough; a crystal oscillator is needed.

It's appropriate to set the level into the receiver to about that of a strong broadcast carrier, say S9 + 40 dB. Set the receiver's mode to SSB or CW and tune around the test signal, looking for an increasing noise floor (higher hiss level) as you tune closer towards the signal, as shown in **Fig 10.8**. Switching in a narrow CW filter allows you to hear noise closer to the carrier than is possible with an SSB filter. This is also the technique used to measure a receiver's effective selectivity, and some equipment reviewers kindly publish their plots in this format. *QST* reviews, done by the ARRL Lab, often include the results of specific phase-noise measurements.

Measuring Receiver Phase Noise

There are several different ways of measuring phase noise, offering different tradeoffs between convenience, cost and effort. Some methods suit oscillators in isolation, others suit them in-situ (in their radios).

If you're unfamiliar with noise measurements, the units involved may seem strange. One reason for this is that a noise signal's power is spread over a frequency range, like an infinite number of infinitesimal sinusoidal components. This can be thought of as similar to painting a house. The area that a gallon of paint can cover depends on how thinly it's spread. If someone asks how much paint was used on some part of a wall, the answer would have to be in terms of paint volume per square foot. The wall can be considered to be an infinite number of points, each with an infinitesimal amount of paint applied to it. The question of what volume of paint has been applied at some specific point is unanswerable. With noise, we must work in terms of *power density*, of watts per hertz. We therefore express phase-noise level as a ratio of the carrier power to the noise's power density. Because of the large ratios involved, expression in decibels is convenient. It has been a convention to use *dBc* to mean "decibels with respect to the carrier."

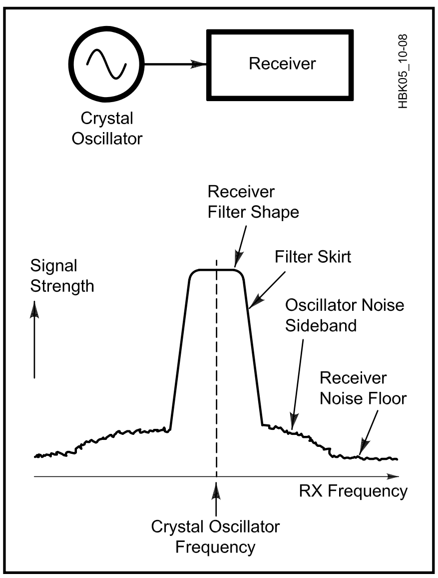


Fig 10.8—Tuning in a strong, clean crystal-oscillator signal can allow you to hear your receiver's relative phase-noise level. Listening with a narrow CW filter switched in allows you to get better results closer to the carrier.

For phase noise, we need to work in terms of a standard bandwidth, and 1 Hz is the obvious candidate. Even if the noise is measured in a different bandwidth, its equivalent power in 1 Hz can be easily calculated. A phase-noise level of –120 dBc in a 1-Hz bandwidth (often written as *–120 dBc/Hz*) translates into each hertz of the noise having a power of 10^{-12} of the carrier power. In a bandwidth of 3 kHz, this would be 3000 times larger.

The most convenient way to measure phase noise is to buy and use a commercial phase noise test system. Such a system usually contains a state-of-the-art, low-noise frequency synthesizer and a low-frequency spectrum analyzer, as well as some special hardware. Often, a second, DSP-based spectrum analyzer is included to speed up and extend measurements very close to the carrier by using the Fast Fourier Transform (FFT). The whole system is then controlled by a computer with proprietary software. With a good system like this costing about $100,000, this is not a practical method for amateurs, although a few fortunate individuals have access to them at work. These systems are also overkill for our needs, because we are not particularly interested in determining phase-noise levels very close to and very far from the carrier.

It's possible to make respectable receiver-oscillator phase-noise measurements with less than $100 of parts and a multimeter. Although it's time-consuming, the technique is much more in keeping with the amateur spirit than using a $100k system! An ordinary multimeter will produce acceptable results, a meter capable of indicating "true rms" ac voltages is preferable because it can give correct readings on sine waves *and* noise. **Fig 10.9** shows the setup. Measurements can only be made around the frequency of the crystal oscillator, so if more than one band is to be tested, crystals must be changed, or else a set of appropriate oscillators is needed. The oscillator should produce about +10 dBm (10 mW) and be *very* well shielded. (To this end, it's advisable to build the oscillator into a die-cast box and power it from internal batteries. A noticeable shielding improvement results even from avoiding the use of an external power switch; a reed-relay element inside the box can be positioned to connect the battery when a small permanent magnet is placed against a marked place outside the box.)

Likewise, great care must be taken with attenuator shielding. A total attenuation of around 140 dB is needed, and with so much attenuation in line, signal leakage can easily exceed the test signal that

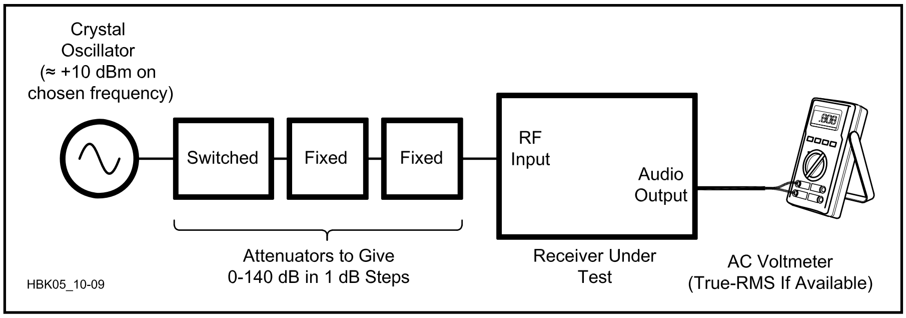


Fig 10.9—Setup for measuring receiver-oscillator phase noise.

reaches the receiver. It's not necessary to be able to switch-select all 140 dB of the attenuation, nor is this desirable, as switches can leak. (The 1995 and older editions of this *Handbook* contained a step attenuator that's satisfactory.) All of the attenuators' enclosure seams must be soldered. A pair of boxes with 30 dB of fixed attenuation each are needed to complete the set. With 140 dB of attenuation, coax cable leakage is also a problem. The only countermeasure against this is to minimize all cable lengths and to interconnect test-system modules with BNC plug-to-plug adapters (UG-491As) where possible.

Ideally, the receiver could simply be tuned across the signal from the oscillator and the response measured using its signal-strength (S) meter. Unfortunately, receiver S meters are notoriously imprecise, so an equivalent method is needed that does not rely on the receiver's AGC system.

The trick is not to measure the response to a fixed level signal, but to measure the changes in applied signal power needed to give a fixed response. Here is a step-by-step procedure based on that described by John Grebenkemper, KI6WX, in March and April 1988 *QST*:

1. Connect the equipment as shown in Fig 10.9, but with the crystal oscillator off. Set the step attenuator to maximum attenuation. Set the receiver for SSB or CW reception with its narrowest available IF filter selected. Switch out any internal preamplifiers or RF attenuators. Select AGC off, maximum AF and RF gain. It may be necessary to reduce the AF gain to ensure the audio amplifier is at least 10 dB below its clipping point. The ac voltmeter or an oscilloscope on the AF output can be used to monitor this.

2. To measure noise, it is important to know the bandwidth being measured. A true-RMS ac voltmeter measures the power in the noise reaching it. To calculate the noise density, we need to divide by the receiver's *noise bandwidth*. The receiver's –6-dB IF bandwidth can be used as an approximation, but purists will want to plot the top 20 dB of the receiver's bandwidth on linear scales and integrate the area under it to find the width of a rectangle of equal area and equal height. This accounts properly for the noise in the skirt regions of the overall selectivity. (The very rectangular shape of common receiver filters tends to minimize the error of just taking the approximation.)

Switch on the test oscillator and set the attenuators to give an AF output above the noise floor and below the clipping level with the receiver peaked on the signal. Tune the receiver off to each side to find the frequencies at which the AF voltage is half that at the peak. The difference between these is the receiver's –6-dB bandwidth. High accuracy is not needed: 25% error in the receiver bandwidth will only cause a 1-dB error in the final result. The receiver's published selectivity specifications will be close enough. The benefit of integration is greater if the receiver has a very rounded, low-ringing or low-order filter.

3. Retune the receiver to the peak. Switch the oscillator off and note the noise-floor voltage. Turn the oscillator back on and adjust the attenuator to give an AF output voltage 1.41 times (3 dB) larger than the noise floor voltage. This means that the noise power and the test signal power at the AF output are equal—a value that's often called the *MDS* (*minimum discernible signal*) of a receiver. Choosing a test-oscillator level at which to do this test involves compromise. Higher levels give more accurate results where the phase noise is high, but limit the lowest level of phase noise that can be measured because better receiver oscillators require a greater input signal to produce enough noise to get the chosen AF-output level. At some point, either we've taken all the attenuation out and our measurement range is limited by the test

Table 10.1
SSB Phase Noise of ICOM IC-745 Receiver Section

Oscillator output power = –3 dBm (0.5 mW)
Receiver bandwidth (Δf) = 1.8 kHz
Audio noise voltage = –0.070 V
Audio reference voltage (V_0) = 0.105 V
Reference attenuation (A_0) = 121 dB

Offset Frequency (kHz)	*Attenuation (A_1) (dB)*	*Audio V_1 (volts)*	*Audio V_2 (volts)*	*Ratio V_2/V_1*	*SSB Phase Noise (dBc/Hz)*
4	35	0.102	0.122	1.20	–119
5	32	0.104	0.120	1.15	–122*
6	30	0.104	0.118	1.13	–124*
8	27	0.100	0.116	1.16	–127*
10	25	0.106	0.122	1.15	–129*
15	21	0.100	0.116	1.16	–133*
20	17	0.102	0.120	1.18	–137
25	14	0.102	0.122	1.20	–140
30	13	0.102	0.122	1.20	–141
40	10	0.104	0.124	1.19	–144
50	8	0.102	0.122	1.20	–146
60	6	0.104	0.124	1.19	–148
80	4	0.102	0.126	1.24	–150
100	3	0.102	0.126	1.24	–151
150	3	0.102	0.124	1.22	–151
200	0	0.104			–154
250	0	0.100			–154
300	0	0.98			–154
400	0	0.96			–154
500	0	0.96			–154
600	0	0.97			–154
800	0	0.96			–154
1000	0	0.96			–154

*Asterisks indicate measurements possibly affected by receiver overload (see text).

oscillator's available power, or we overload the receiver's front end, spoiling the results.

Record the receiver frequency at the peak, (f_0), the attenuator setting (A_0) and the audio output voltage (V_0). These are the carrier measurements against which all the noise measurements will be compared.

4. Now you must choose the offset frequencies—the spacings from the carrier—at which you wish to make measurements. The receiver's skirt selectivity will limit how close to the carrier noise measurements can be made. (Any measurements made too close in are valid measurements of the receiver selectivity, but because the signal measured under these conditions is sinusoidal and not noise like, the corrections for noise density and noise bandwidth are not appropriate.) It is difficult to decide where the filter skirt ends and the noise begins, and what corrections to apply in the region of doubt and uncertainty. A good practical approach is to listen to the audio and tune away from the carrier until you can't distinguish a tone in the noise. The ear is superb at spotting sine tones buried in noise, so this criterion, although subjective, errs on the conservative side.

Tune the receiver to a frequency offset from f_0 by your first chosen offset and adjust the attenuators to get an audio output voltage as close as possible to V_0. Record the total attenuation, A_1 and the audio output voltage, V_1. The SSB phase noise (qualified as *SSB* because we're measuring the phase noise on only one side of the carrier, whereas some other methods cannot segregate between upper and lower noise sidebands and measure their sum, giving *DSB* phase noise) is now easy to calculate:

$$L(f) = A_1 - A_0 10 \log(BW_{noise})$$

where

L(f) = SSB phase noise in dBc/Hz

BWnoise = receiver noise bandwidth, Hz.

5. It's important to check for overload. Decrease the attenuation by 3 dB, and record the new audio output voltage, V_2. If all is well, the output voltage should increase by 22% (1.8 dB); if the receiver is operating nonlinearly, the increase will be less. (An 18% increase is still acceptable for the overall accuracy we want.) Record V_2/V_1 as a check: a ratio of 1.22:1 is ideal, and anything less than 1.18:1 indicates a bad measurement.

If too many measurements are bad, you may be overdriving the receiver's AF amplifier, so try reducing the AF gain and starting again back at Step 3. If this doesn't help, reducing the RF gain and starting again at Step 3 should help if the compression is occurring late in the IF stages.

6. Repeat Steps 4 and 5 at all the other offsets you wish to measure. If measurements are made at increments of about half the receiver's bandwidth, any discrete (non-noise) spurs will be found. A noticeable tone in the audio can indicate the presence of one of these. If it is well clear of the noise, the measurement is valid, but the noise bandwidth correction should be ignored, giving a result in dBc.

Table 10.1 shows the results for an ICOM IC-745 as measured by KI6WX, and

Transmitter Phase-Noise Measurement in the ARRL Lab

Here is a brief description of the technique used in the ARRL Lab to measure transmitter phase noise. The system essentially consists of a direct-conversion receiver with very good phase-noise characteristics. As shown in Fig B, we use an attenuator after the transmitter, a Mini-Circuits ZAY-1 mixer, a Hewlett-Packard 8640B signal generator, a band-pass filter, an audio-frequency low-noise amplifier and a spectrum analyzer (HP 8563E) to make the measurements.

The transmitter signal is mixed with the output of the signal generator, and signals produced in the mixing process that are not required for the measurement process are filtered out. The spectrum analyzer then displays the transmitted phase-noise spectrum. The 100 mW output of the HP 8640B is barely enough to drive the mixer—the setup would work better with 200 mW of drive. To test the phase noise of an HP 8640B, we use a second '8640B as a reference source. It is quite important to be sure that the phase noise of the reference source is lower than that of the signal under test, because we are really measuring the combined phase-noise output of the signal generator and the transmitter. It would be quite embarrassing to publish phase-noise plots of the reference generator instead of the transmitter under test! The HP 8640B has much cleaner spectral output than most transmitters.

A sample phase-noise plot for an amateur transceiver is shown in Fig A. It was produced with the test setup shown in Fig B. Measurements from multiple passes are taken and averaged. A 3-Hz video bandwidth also helps average and smooth the plots. These plots do not necessarily reflect the phase-noise characteristics of all units of a particular model.

The log reference level (the top horizontal line on the scale in the plot) represents –60 dBc/Hz. It is common in industry to use a 0-dBc log reference, but such a reference level would not allow measurement of phase-noise levels below –80 dBc/Hz. The actual measurement bandwidth used on the spectrum analyzer is 100 Hz, but the reference is scaled for a 1-Hz bandwidth. This allows phase-noise levels to be read directly from the display in dBc/Hz. Because each vertical division represents 10 dB, the plot shows the noise level between –60 dBc/Hz (the top horizontal line) and –140 dBc/Hz (the bottom horizontal line). The horizontal scale is 2 kHz per division. The offsets shown in the plots are 2 through 20 kHz.

What Do the Phase-Noise Plots Mean?

Although they are useful for comparing different radios, plots can also be used to calculate the amount of interference you may receive from a nearby transmitter

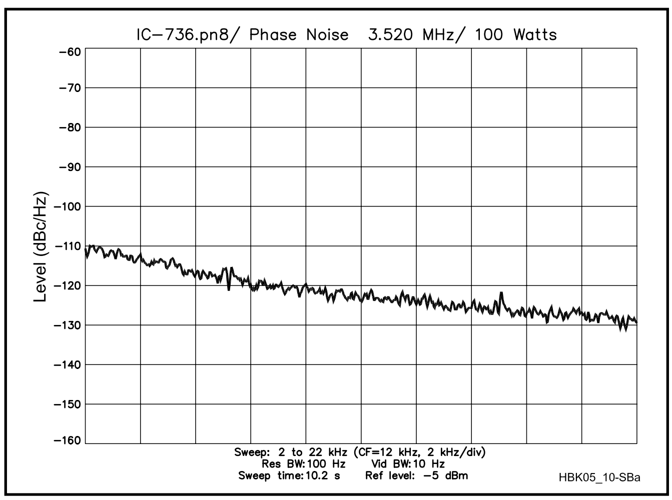


Fig A—Sample phase noise plot for an amateur transceiver.

Fig 10.10 shows this data in graphic form. His oscillator power was only –3 dBm, which limited measurements to offsets less than 200 kHz. More power might have allowed noise measurements to lower levels, although receiver overload places a limit on this. This is not important, because the real area of interest has been thoroughly covered. When attempting phase-noise measurements at large offsets, remember that any front-end selectivity, before the first mixer, will limit the maximum offset at which LO phase-noise measurement is possible.

MEASURING OSCILLATOR AND TRANSMITTER PHASE NOISE

Measuring the composite phase noise of a receiver's LO requires a clean test oscillator. Measuring the phase noise of an incoming signal, whether from a single oscillator or an entire transmitter, requires the use of a clean receiver, with lower phase noise than the source under test. The sidebar, "Transmitter Phase-Noise Measurement in the ARRL Lab," details the method used to measure composite noise (phase noise and amplitude noise, the practical effects of which are indistinguishable on the air) for *QST* Product Reviews. Although targeted at measuring high power signals from entire transmitters, this approach can be used to measure lower-level signals simply by changing the amount of input attenuation used.

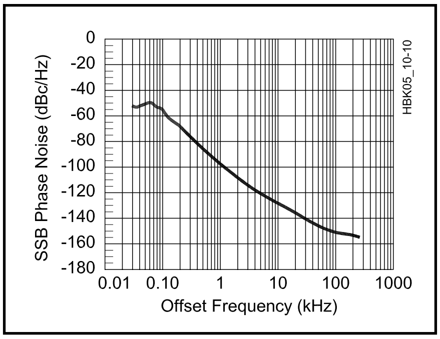


Fig 10.10—The SSB phase noise of an ICOM IC-745 transceiver (serial number 01528) as measured by KI6WX.

At first, this method—using a low-frequency spectrum analyzer and a low-phase-noise signal generator—looks unnecessarily elaborate. A growing number of radio amateurs have acquired good-quality spectrum analyzers for their shacks since older model Tektronix and Hewlett-Packard instruments have started to appear on the surplus market at affordable prices. The obvious question is, "Why not just use one of these to view the signal and read phase-noise levels directly off the screen?" Reciprocal mixing is the problem. Very few spectrum analyzers have clean enough local oscillators not to completely swamp the noise being measured. Phase-noise measurements involve the measurement of low-level components very close to a large carrier, and that carrier will mix the noise sidebands of *the analyzer's LO* into its IF. Some way of notching out the carrier is needed, so that the analyzer need only handle the noise sidebands. A crystal filter could be designed to do the job, but this would be expensive, and one would be needed for every different oscillator

with known phase-noise characteristics. An approximation is given by

$A_{QRM} = NL + 10 \times \log(BW)$

where

A_{QRM} = Interfering signal level, dBc
NL = noise level on the receive frequency, dBc
BW = receiver IF bandwidth, in Hz

For instance, if the noise level is –90 dBc/Hz and you are using a 2.5-kHz SSB filter, the approximate interfering signal will be –56 dBc. In other words, if the transmitted signal is 20 dB over S9, and each S unit is 6 dB, the interfering signal will be as strong as an S3 signal.

The measurements made in the ARRL Lab apply only to transmitted signals. It is reasonable to assume that the phase-noise characteristics of most transceivers are similar on transmit and receiver because the same oscillators are generally used in local-oscillator (LO) chain.

In some cases, the receiver may have better phase-noise characteristics than the transmitter. Why the possible difference? The most obvious reason is that circuits often perform less than optimally in strong RF fields, as anyone who has experienced RFI problems can tell you. A less obvious reason results from the way that many high-dynamic-range receivers work. To get good dynamic range, a sharp crystal filter is often placed immediately after the first mixer in the receive line. This filter removes all but a small slice of spectrum for further signal processing. If the desired filtered signal is a product of mixing an incoming signal with a noisy oscillator, signals far away from the desired one can end up in this slice. Once this slice of spectrum is obtained, however, unwanted signals cannot be reintroduced, no matter how noisy the oscillators used in further signal processing. As a result, some oscillators in receivers don't affect phase noise.

The difference between this situation and that in transmitters is that crystal filters are seldom used for reduction of phase noise in transmitting because of the high cost involved. Equipment designers have enough trouble getting smooth, click-free break-in operation in transceivers without having to worry about switching crystal filters in and out of circuits at 40-wpm keying speeds!—*Zack Lau, W1VT, ARRL Laboratory Engineer*

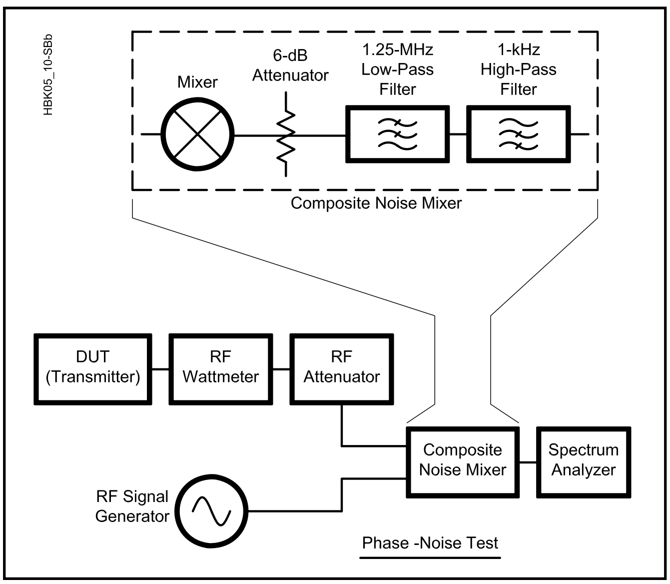


Fig B—ARRL transmitter phase-noise measurement setup.

frequency to be tested. The alternative is to build a direct-conversion receiver using a clean LO like the Hewlett-Packard HP8640B signal generator and spectrum-analyze its "audio" output with an audio analyzer. This scheme mixes the carrier to dc; the LF analyzer is then ac-coupled, and this removes the carrier. The analyzer can be made very sensitive without overload or reciprocal mixing being a problem. The remaining problem is then keeping the LO—the HP8640B in this example—at exactly the carrier frequency. 8640s are based on a shortened-UHF-cavity oscillator and can drift a little. The oscillator under test will also drift. The task is therefore to make the 8640B track the oscillator under test. For once we get something for free: The HP8640B's FM input is dc coupled, and we can use this as an electronic fine-tuning input. As a further bonus, the 8640B's FM deviation control acts as a sensitivity control for this input. We also get a phase detector for free, as the mixer output's dc component depends on the phase relationship between the 8640B and the signal under test (remember to use the dc coupled port of a diode ring mixer as the output). Taken together, the system includes everything needed to create a crude phase-locked loop that will automatically track the input signal over a small frequency range. **Fig 10.11** shows the arrangement.

The oscilloscope is not essential for operation, but it is needed to adjust the system. With the loop unlocked (8640B FM input disconnected), tune the 8640 off the signal frequency to give a beat at the mixer output. Adjust the mixer drive levels to get an undistorted sine wave on the scope. This ensures that the mixer is not being overdriven. While the loop is off-tuned, adjust the beat to a frequency within the range of the LF spectrum analyzer and use it to measure its level, "A_c" in dBm. This represents the carrier level and is used as the reference for the noise measurements. Connect the FM input of the signal generator, and switch on the generator's dc FM facility. Try a deviation range of 10 kHz to start with. When you tune the signal generator toward the input frequency, the scope will show the falling beat frequency until the loop jumps into lock. Then it will display a noisy dc level. Fine tune to get a mean level of 0 V. (This is a very-low-bandwidth, very-low-gain loop. Stability is not a problem; careful loop design is not needed. We actually want as slow a loop as possible; otherwise, the loop would track and cancel the slow components of the incoming signal's phase noise, preventing their measurement.)

When you first take phase-noise plots, it's a good idea to duplicate them at the generator's next lower FM-deviation range and check for any differences in the noise level in the areas of interest. Reduce the FM deviation range until you find no further improvement. Insufficient FM deviation range makes the loop's lock range narrow, reducing the amount of drift it can compensate. (It's sometimes necessary to keep gently trimming the generator's fine tune control.)

Set up the LF analyzer to show the noise. A sensitive range and 100-Hz resolution bandwidth are appropriate. Measure the noise level, "A_n" in dBm. We must now calculate the noise density that this represents. Spectrum-analyzer filters are normally *Gaussian*-shaped and bandwidth-specified at their –3-dB points. To avoid using integration to find their true-noise power bandwidth, we can reasonably assume a value of 1.2 × BW. A spectrum analyzer logarithmically compresses its IF signal ahead of the detectors and averaging filter. This affects the statistical distribution of noise voltage and causes the analyzer to read low by 2.5 dB. To produce the same scale as the ARRL Lab photographs, the analyzer reference level must be set to –60 dBc/Hz, which can be calculated as:

$$A_{ref} = A_c - 10\log(1.2 \times BW) + 62.5\text{ dBm} \quad (2)$$

where

A_{ref} = analyzer reference level, dBm
A_c = carrier amplitude, dBm

This produces a scale of –60 dBc/Hz at the top of the screen, falling to –140 dBc/Hz at the bottom. The frequency scale is 0 to 20 kHz with a resolution bandwidth (BW in the above equation) of 100 Hz. This method combines the power of *both* sidebands and so measures DSB phase noise. To calculate the equivalent SSB phase noise, subtract 3 dB for noncoherent noise (the general "hash" of phase noise) and subtract 6 dB for coherent, discrete components (that is, single-frequency spurs). This can be done by setting the reference level 3 to 6 dB higher.

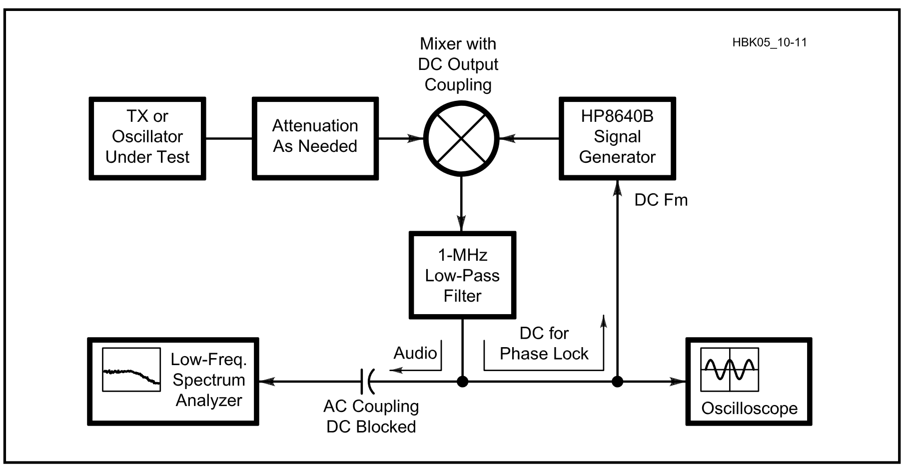


Fig 10.11—Arrangement for measuring phase noise by directly converting the signal under test to audio. The spectrum analyzer views the signal's noise sidebands as audio; the signal's carrier, converted to dc, provides a feedback signal to phase-lock the Hewlett-Packard HP8640B signal generator to the signal under test.

Low-Cost Phase Noise Testing

All that expensive equipment may seem far beyond the means of the average Amateur Radio experimenter. With careful shopping and a little more effort, alternative equipment can be put together for pocket money. (All of the things needed—parts for a VXO, a surplus spectrum analyzer and so on—have been seen on sale cheap enough to total less than $100.) The HP8640B is good and versatile, but for use at one oscillator frequency, you can build a VXO for a few dollars. It will only cover one oscillator frequency, but a VXO can provide even better phase-noise performance than the 8640B.

As referenced at the end of this chapter, *Pontius* has also demonstrated that signal-source phase-noise measurements can be accurately obtained without the aid of expensive equipment.

OSCILLATOR CIRCUITS AND CONSTRUCTION

LC Oscillator Circuits

The LC oscillators used in radio equipment are usually arranged as *variable frequency oscillators* (*VFOs*). Tuning is achieved by either varying part of the capacitance of the resonator or, less commonly, by using a movable magnetic core to vary the inductance. Since the early days of radio, there has been a huge quest for the ideal, low-drift VFO. Amateurs and professionals have spent immense effort on this pursuit. A brief search of the literature reveals a large number of designs, many accompanied by claims of high stability. The quest for stability has been solved by the development of low-cost frequency synthesizers, which can give crystal-controlled stability. Synthesizers have other problems though, and the VFO still has much to offer in terms of signal cleanliness, cost and power consumption, making it attractive for home construction. No one VFO circuit has any overwhelming advantage over any other—component quality, mechanical design and the care taken in assembly are much more important.

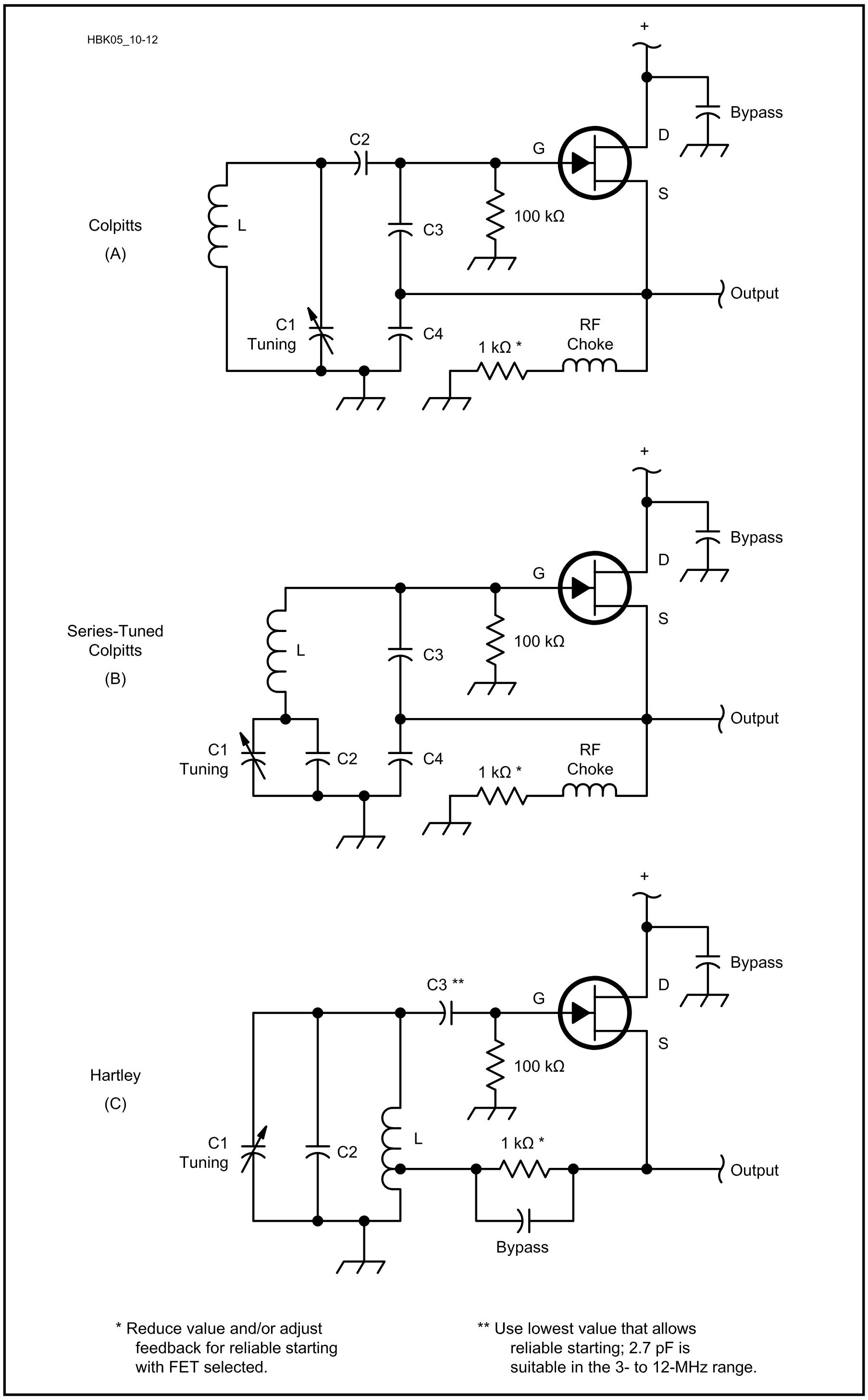


Fig 10.12—The Colpitts (A), series-tuned Colpitts (B) and Hartley (C) oscillator circuits. Rules of thumb: C3 and C4 at A and B should be equal and valued such that their X_c = 45 Ω at the operating frequency; for C2 at A, X_c = 100 Ω. For best stability, use C0G or NP0 units for all capacitors associated with the FETs' gates and sources. Depending on the FET chosen, the 1-kΩ source-bias-resistor value shown may require adjustment for reliable starting.

Fig 10.12 shows three popular oscillator circuits stripped of any unessential frills so they can be more easily compared. The original Colpitts circuit (Fig 10.12A) is now often referred to as the *parallel-tuned Colpitts* because its series-tuned derivative (Fig 10.12B) has become common. All three of these circuits use an amplifier with a voltage gain less than unity, but large current gain. The N-channel JFET source follower shown appears to be the most popular choice nowadays. In the parallel-tuned Colpitts, C3 and C4 are large values, perhaps 10 times larger than typical values for C1 and C2. This means that only a small fraction of the total tank voltage is applied to the FET, and the FET can be considered to be only lightly coupled into the tank. The FET is driven by the sum of the voltages across C3 and C4, while it drives the voltage across C4 alone. This means that the tank operates as a resonant, voltage-step-up transformer compensating for the less-than-unity-voltage-gain amplifier. The resonant circuit consists of L, C1, C2, C3 and C4. The resonant frequency can be calculated by using the standard formulas for capacitors in series and parallel to find the resultant capacitance effectively connected across the inductor, L, and then use the standard formula for LC resonance:

$$f = \frac{1}{2\pi\sqrt{LC}} \tag{3}$$

where
f = frequency in hertz
L = inductance in henries
C = capacitance in farads.

For a wide tuning range, C2 must be kept small to reduce the effect of C3 and C4 swamping the variable capacitor C1.

The series-tuned Colpitts circuit works in much the same way. The difference is that the variable capacitor, C1, is positioned so that it is well-protected from being swamped by the large values of C3 and C4. In fact, *small* values of C3, C4 would act to limit the tuning range. Fixed capacitance, C2, is often added across C1 to allow the tuning range to be reduced to that required, without interfering with C3 and C4, which set the amplifier coupling. The series-tuned Colpitts has a reputation for better stability than the parallel-tuned original. Note how C3 and C4 swamp the capacitances of the amplifier in both versions.

The Hartley is similar to the parallel-tuned Colpitts, but the amplifier source is tapped up on the tank inductance instead of the tank capacitance. A typical tap placement is 10 to 20% of the total turns up from the “cold” end of the inductor. (It’s usual to refer to the lowest-signal-voltage end of an inductor as *cold* and the other, with the highest signal voltage as *hot.*) C2 limits the tuning range as required; C3 is reduced to the minimum value that allows reliable starting. This is necessary because the Hartley’s lack of the Colpitts’s capacitive divider would otherwise couple the FET’s capacitances to the tank more strongly than in the Colpitts, potentially affecting the circuit’s frequency stability.

In all three circuits, there is a 1 kΩ resistor in series with the source bias choke. This resistor does a number of desirable things. It spoils the Q of the inevitable low-frequency resonance of the choke with the tank tap circuit. It reduces tuning drift due to choke impedance and winding capacitance variations. It also protects against spurious oscillation due to internal choke resonances. Less obviously, it acts to stabilize the loop gain of the built-in AGC action of this oscillator. Stable operating conditions act to reduce frequency drift. Some variations of these circuits may be found with added resistors providing a dc bias to stabilize the system quiescent current. More elaborate still are variations characterized by a constant-current source providing bias. This can be driven from a separate AGC detector system to give very tight level control. The gate-to-ground clamping diode (1N914 or similar) long used by radio amateurs as a means of avoiding gate-source conduction has been shown by Ulrich Rohde, KA2WEU, to degrade oscillator phase-noise performance, and its use is virtually unknown in professional circles.

Fig 10.13 shows some more VFOs to illustrate the use of different devices. The triode Hartley shown includes *permeability tuning*, which has no sliding contact like that to a capacitor’s rotor and can be made reasonably linear by artfully spacing the coil turns. The slow-motion drive can be done with a screw thread. The disadvantage is that special care is needed to avoid backlash and eccentric rotation of the core. If a nonrotating core is used, the slides have to be carefully designed to prevent motion in unwanted directions. The Collins Radio Company made extensive use of tube-based permeability tuners, and a semiconductor version can still be found in a number of Ten-Tec radios.

Vacuum tubes cannot run as cool as competitive semiconductor circuits, so care is needed to keep the tank circuit away from the tube heat. In many amateur and commercial vacuum-tube oscillators, oscillation drives the tube into grid current at the positive peaks, causing rectification and producing a negative grid bias. The oscillator thus runs in Class C, in which the conduction angle reduces as the signal amplitude increases until the amplitude stabilizes. As in the FET circuits of Fig 10.12, better stability and phase-

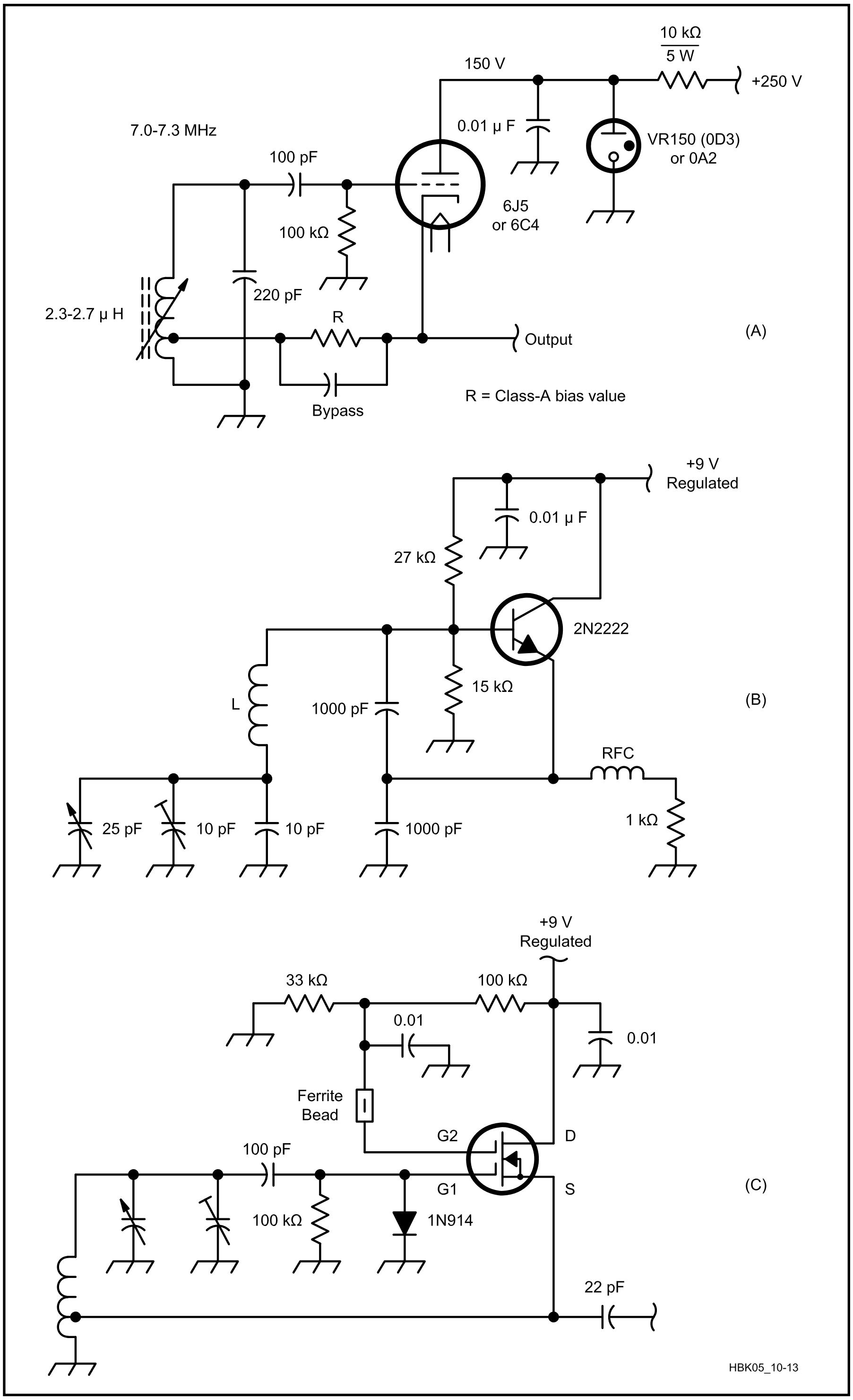


Fig 10.13—Three more oscillator examples: at A, a triode-tube Hartley; at B, a bipolar junction transistor in a series-tuned Colpitts; at C, a dual-gate MOSFET Hartley.

noise performance can be achieved in a vacuum-tube oscillator by moving it out of true Class C—that is, by including a bypassed cathode-bias resistor (the resistance appropriate for Class A operation is a good starting value). A small number of people still build vacuum-tube radios partly to be different, partly for fun, but the semiconductor long ago achieved dominance in VFOs.

The voltage regulator (*VR*) tube shown in Fig 10.13A has a potential drawback: It is a gas-discharge device with a high striking voltage at which it starts to conduct and a lower extinguishing voltage, at which it stops conducting. Between these extremes lies a region in which decreasing voltage translates to increasing current, which implies negative resistance. When the regulator strikes, it discharges any capacitance connected across it to the extinguishing voltage. The capacitance then charges through the source resistor until the tube strikes again, and the process repeats. This *relaxation oscillation* demonstrates how negative resistance can cause oscillation. The oscillator translates the resultant sawtooth modulation of its power supply into frequency and amplitude variation. Because of VR tubes' ability to support relaxation oscillation, a traditional rule of thumb is to keep the capacitance directly connected across a VR tube to 0.1 µF or less. A value much lower than this can provide sufficient bypassing in Fig 10.13A because the dropping resistor acts as a decoupler, increasing the bypass's effectiveness.

There is a related effect called *squegging*, which can be loosely defined as oscillation on more than one frequency at a time, but which may also manifest itself as RF oscillation interrupted at an AF rate, as in a superregenerative detector. One form of squegging occurs when an oscillator is fed from a power supply with a high source impedance. The power supply charges up the decoupling capacitor until oscillation starts. The oscillator draws current and pulls down the capacitor voltage, starving itself of power until oscillation stops. The oscillator stops drawing current, and the decoupling capacitor then recharges until oscillation restarts. The process, the low-frequency cycling of which is a form of relaxation oscillation, repeats indefinitely. The oscillator output can clearly be seen to be pulse modulated if an oscilloscope is used to view it at a suitable time-base setting. This fault is a well-known consequence of poor design in battery-powered radios. As dry cells become exhausted, their internal resistance rises quickly and circuits they power can begin to misbehave. In audio stages, such misbehavior may manifest itself in the *putt-putt* sound of the slow relaxation oscillation called *motorboating*.

Compared to the frequently used JFET, bipolar transistors, Fig 10.13B, are relatively uncommon in oscillators because their low input and output impedances are more difficult to couple into a high-Q tank without excessively loading it. Bipolar devices do tend to give better sample-to-sample amplitude uniformity for a given oscillator circuit, however, as JFETs of a given type tend to vary more in their characteristics.

The dual-gate MOSFET, Fig 10.13C, is very rarely seen in VFO circuits. It imposes the cost of the components needed to bias and suppress VHF parasitic oscillation at the second gate, and its inability to generate its own AGC through gate-source conduction forces the addition of a blocking capacitor, resistor and diode at the first gate for amplitude control.

VFO Components and Construction

Tuning Capacitors and Reduction Drives

As most commercially made radios now use frequency synthesizers, it has become increasingly difficult to find certain key components needed to construct a good VFO. The slow-motion drives, dials, gearboxes, associated with names like Millen, National, Eddystone and Jackson are no longer available. Similarly, the most suitable silver-plated variable capacitors with ball races at both ends, although still made, are not generally marketed and are expensive. Three approaches remain: Scavenge suitable parts from old equipment; use tuning diodes instead of variable capacitors—an approach that, if uncorrected through phase locking, generally degrades stability and phase-noise performance; or use two tuning capacitors, one with a capacitance range $^1/_5$ to $^1/_{10}$ that of the other, in a bandset/bandspread approach.

Assembling a variable capacitor to a chassis and its reduction drive to a front panel can result in *backlash*—an annoying tuning effect in which rotating the capacitor shaft deforms the chassis and/or panel rather than tuning the capacitor. One way of minimizing this is to use the reduction drive to support the capacitor, and use the capacitor to support the oscillator circuit board.

Fixed Capacitors

Traditionally, silver-mica fixed capacitors have been used extensively in oscillators, but their temperature coefficient is not as low as can be achieved by other types, and some silver micas have been known to behave erratically. Polystyrene film has become a proven alternative. One warning is worth noting: Polystyrene capacitors exhibit a permanent change in value should they ever be exposed to temperatures much over 70°C; they do not return to their old value on cooling. Particularly suitable for oscillator construction are the low-temperature-coefficient ceramic capacitors, often described as *NP0* or *C0G* types. These names are actually temperature-coefficient codes. Some ceramic capacitors are available with deliberate, controlled temperature coefficients so that they can be used to compensate for other causes of frequency drift with temperature. For example, the code N750 denotes a part with a temperature coefficient of –750 parts per million per degree Celsius. These parts are now somewhat difficult to obtain, so other methods are needed.

In a Colpitts circuit, the two large-value capacitors that form the voltage divider for the active device still need careful selection. It would be tempting to use any available capacitor of the right value, because the effect of these components on the tank frequency is reduced by the proportions of the capacitance values in the circuit. This reduction is not as great as the difference between the temperature stability of an NP0 ceramic part and some of the low-cost, decoupling-quality X7R-dielectric ceramic capacitors. It's worth using low-temperature coefficient parts even in the seemingly less-critical parts of a VFO circuit—even as decouplers. Chasing the cause of temperature drift is more challenging than fun. Buy critical components like high-stability capacitors from trustworthy sources.

Inductors

Ceramic coil forms can give excellent results, as can self-supporting air-wound coils (Miniductor). If you use a magnetic core, make it powdered iron, never ferrite, and support it stably. Stable VFOs have been made using toroidal cores, but again, ferrite must be avoided. Micrometals mix number 6 has a low temperature coefficient and works well in conjunction with NP0 ceramic capacitors. Coil forms in other materials have to be assessed on an individual basis.

A material's temperature stability will not be apparent until you try it in an oscillator, but you can apply a quick test to identify those nonmetallic materials that are lossy enough to spoil a coil's Q. Put a sample of the coil-form material into a microwave oven along with a glass of water and cook it about 10 s on low power.

Do not include any metal fittings or ferromagnetic cores. Good materials will be completely unaffected; poor ones will heat and may even melt, smoke, or burst into flame. (This operation is a fire hazard if you try more than a tiny sample of an unknown material. Observe your experiment continuously and do not leave it unattended.)

Wes Hayward, W7ZOI, suggests annealing toroidal VFO coils after winding. Roy Lewallen, W7EL reports achieving success with this method by boiling his coils in water and letting them cool in air.

Voltage Regulators

VFO circuits are often run from locally regulated power supplies, usually from resistor/Zener diode combinations. Zener diodes have some idiosyncrasies that could spoil the oscillator. They are noisy, so decoupling is needed down to audio frequencies to filter this out. Zener diodes are often run at much less than their specified optimum bias current. Although this saves power, it results in a lower output voltage than intended, and the diode's impedance is much greater, increasing its sensitivity to variations in input voltage, output current and temperature. Some common Zener types may be designed to run at as much as 20 mA; check the data sheet for your diode family to find the optimum current.

True Zener diodes are low-voltage devices; above a couple of volts, so-called Zener diodes are actually avalanche types. The temperature coefficient of these diodes depends on their breakdown voltage and crosses through zero for diodes rated at about 5 V. If you intend to use nothing fancier than a common-variety Zener, designing the oscillator to run from 5 V and using a 5.1-V Zener, will give you a free advantage in voltage-versus-temperature stability. There are some diodes available with especially low temperature coefficients, usually referred to as *reference* or *temperature-compensated diodes*. These usually consist of a series pair of diodes designed to cancel each other's temperature drift. Running at 7.5 mA, the 1N829A gives 6.2 V ±5% and a temperature coefficient of just ±5 parts per million (ppm) maximum per degree Celsius. A change in bias current of 10% will shift the voltage less than 7.5 mV, but this increases rapidly for greater current variation. The 1N821A is a lower-grade version, at ±100 ppm/°C. The LM399 is a complex IC that behaves like a superb Zener at 6.95 V, ±0.3 ppm/°C. There are also precision, low-power, three-terminal regulators designed to be used as voltage references, some of which can provide enough current to run a VFO. There are comprehensive tables of all these devices between pages 334 and 337 of Horowitz and Hill, *The Art of Electronics*, 2nd ed.

Oscillator Devices

The 2N3819 FET, a classic from the 1960s, has proven to work well in VFOs, but, like the MPF102 also long-popular with ham builders, it's manufactured to wide tolerances. Considering an oscillator's importance in receiver stability, you should not hesitate to spend a bit more on a better device. The 2N5484, 2N5485 and 2N5486 are worth considering; together, their transconductance ranges span that of the MPF102, but each is a better-controlled subset of that range. The 2N5245 is a more recent device with better-than-average noise performance that runs at low currents like the 2N3819. The 2N4416/A, also available as the plastic-cased PN4416, is a low-noise device, designed for VHF/UHF amplifier use, which has been featured in a number of good oscillators up to the VHF region. Its low internal capacitance contributes to low frequency drift. The J310 (plastic; the metal-cased U310 is similar) is another popular JFET in oscillators.

The 2N5179 (plastic, PN5179 or MPS5179) is a bipolar transistor capable of good performance in oscillators up to the top of the VHF region. Care is needed because its absolute-maximum collector-emitter voltage is only 12 V, and its collector current must not exceed 50 mA. Although these characteristics may seem to convey fragility, the 2N5179 is sufficient for circuits powered by stabilized 6-V power supplies.

VHF-UHF devices are not really necessary in LC VFOs because such circuits are rarely used above 10 MHz. Absolute frequency stability is progressively harder to achieve with increasing frequency, so free-running oscillators are rarely used to generate VHF-UHF signals for radio communication. Instead, VHF-UHF radios usually use voltage-tuned, phase-locked oscillators in some form of synthesizer. Bipolar devices like the BFR90 and MRF901, with f_Ts in the 5-GHz region and mounted in stripline packages, are needed at UHF.

Integrated circuits have not been mentioned until now because few specific RF-oscillator ICs exist. Some consumer ICs—the NE602, for example—include the active device(s) necessary for RF oscillators, but often this is no more than a single transistor fabricated on the chip. This works just as a single discrete device would, although using such on-chip devices may result in poor isolation from the rest of the circuits on the chip. There is one specialized LC-oscillator IC: the Motorola MC1648. This device has been made since the early 1970s and is a surviving member of a long-obsolete family of emitter-coupled-logic (ECL) devices. Despite the MC1648's antiquity, it has no real competition and is still widely used in current military and commercial equipment. Market demand should force its continued production for some more years to come. Its circuitry is complex for an oscillator, with a multitransistor oscillator cell controlled by a detector and amplifier in an on-chip ALC system. The MC1648's first problem is that the ECL families use only about a 1-V swing between logic levels. Because the oscillator is made using the same ECL-optimized semiconductor manufacturing processes and circuit design techniques, this same limitation applies to the signal in the oscillator tuned circuit. It is possible to improve this situation by using a tapped or transformer-coupled tank circuit to give improved Q, but this risks the occurrence of the device's second problem.

Periodically, semiconductor manufacturers modernize their plants and scrap old assembly lines used to make old products. Any surviving devices then must undergo some redesign to allow their production by the new processes. One common result of this is that devices are shrunk, when possible, to fit more onto a wafer. All this increases the f_T of the transistors in the device, and such evolution has rendered today's MC1648s capable of operation at much higher frequencies than the specified 200-MHz limit. This allows higher-frequency use, but great care is needed in the layout of circuits using it to prevent spurious oscillation. A number of old designs using this part have needed reengineering because the newer parts generate spurious oscillations that the old ones didn't, using PC-board traces as parasitic tuned circuits.

The moral is that a UHF-capable device requires UHF-cognizant design and layout even if the device will be used at far lower frequencies. **Fig 10.14** shows the MC1648 in a simple circuit and with a tapped resonator. These more complex circuits have a greater risk of presenting a stray resonance within the device's operating range, risking oscillation at an unwanted frequency. This device is *not* a prime choice for an HF VFO because the physical size of the variable capacitor and the inevitable lead lengths, combined with the need to tap-couple to get sufficient Q for good noise performance, makes spurious oscillation difficult to avoid. The MC1648 is really intended for tuning-diode control in phase-locked loops operating at VHF. This difficulty is inherent in wideband devices, especially oscillator circuits connected to their tank by a single "hot" terminal, where there is simply no isolation between the amplifier's input and output paths. Any resonance in the associated circuitry can control the frequency of oscillation.

The popular NE602 mixer IC has a built-in oscillator and can be found in many published circuits. This device has separate input and output pins to the tank and has proved to be quite tame. It's still a relatively new part and may not have been "improved" yet (so far, it has progressed from the NE602 to the NE602A, the A version affording somewhat higher dynamic range than the original NE602). It might be a good idea for anyone laying out a board using one to take a little extra care to keep PCB traces short in the oscillator section to build in some safety margin so that the board can be used reliably in the future. Experienced (read: "bitten") professional designers know that their designs are going to be built for possibly more than 10 years and have learned to make allowances for the progressive improvement of semiconductor manufacture.

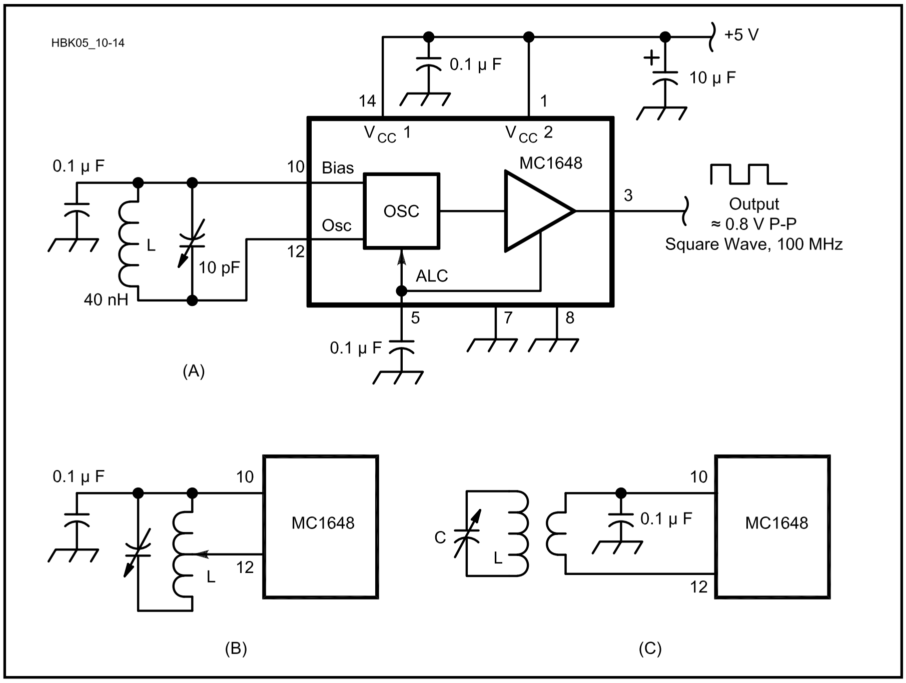


Fig 10.14—One of the few ICs ever designed solely for oscillator service, the ECL Motorola MC1648 (A) requires careful design to avoid VHF parasitics when operating at HF. Keeping its tank Q high is another challenge; B and C show means of coupling the IC's low-impedance oscillator terminals to the tank by tapping up on the tank coil (B) or with a link (C).

Three High-Performance HF VFOs

The G3PDM Vackar VFO

The Vackar VFO shown in **Fig 10.15** was developed over 20 years ago by Peter Martin, G3PDM, for the Mark II version of his high-dynamic-range receiver. This can be found in the Radio Society of Great Britain's *Radio Communication Handbook*, with some further comments on the oscillator in RSGB's *Amateur Radio Techniques*. This is a prime example of an oscillator that has been successfully optimized for maximum frequency stability. Not only does it work extremely well, but it still represents the highest stability that can be achieved. Its developer commented on a number of points, which also apply to the construction of different VFO circuits.

• Use a genuine Vackar circuit, with

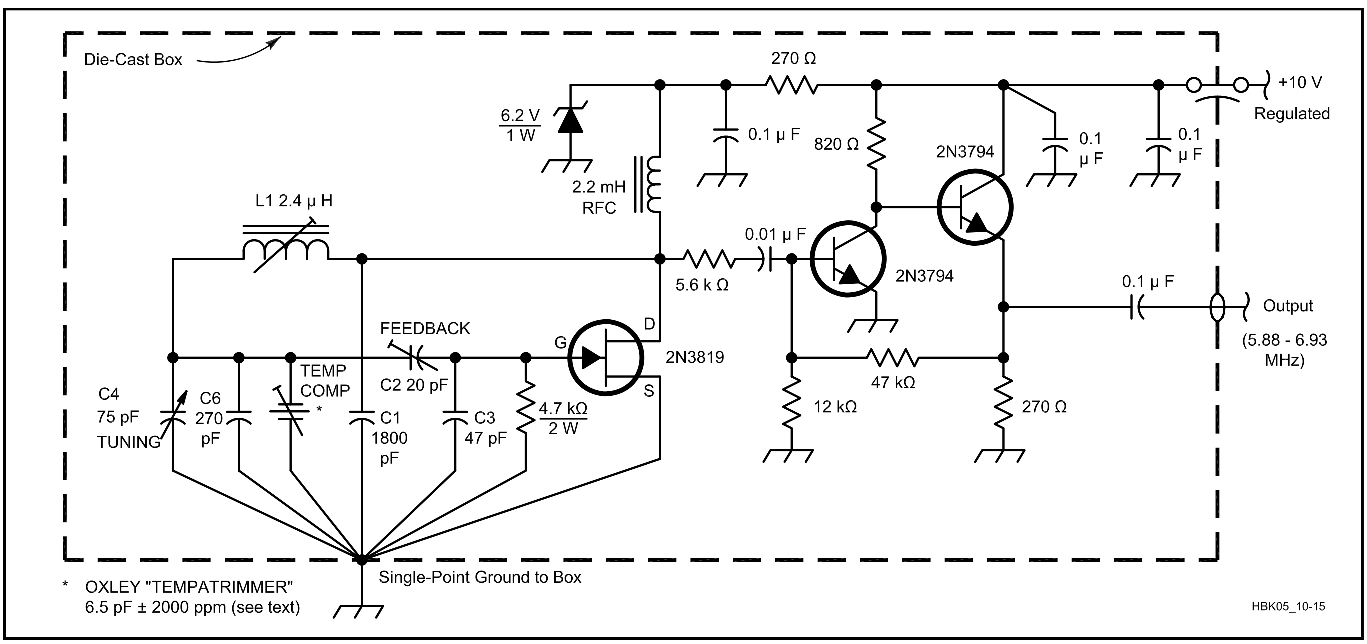


Fig 10.15—G3PDM's Vackar VFO has proved popular and successful for two decades. The MPF102 can be used as a substitute for the 2N3819. Generally, VFOs can be adapted to work at other frequencies (within the limits of the active device). To do so, compute an adjustment factor: f_{old} / f_{new}. Multiply the value of each frequency determining or feedback L or C by the factor. As frequency increases, it may help to increase feedback even more than indicated by the factor.

C1/(C4+C6) and C3/C2 = 6.

- Use a strong box, die-cast or milled from solid metal.
- Use a high-quality variable capacitor (double ball bearings, silver plated).
- Adjust feedback control C2, an air-dielectric trimmer, so the circuit just oscillates.
- Thoroughly clean all variable capacitors in an ultrasonic bath.
- Use an Oxley "Tempatrimmer," a fixed capacitor whose temperature coefficient is variable over a wide range, for adjustable temperature compensation. (The "Thermatrimmer" is a lower cost, smaller range alternative.)
- C1, C3 and C6 are silver-mica types glued to a solid support to reduce sensitivity to mechanical shock.
- The gate resistor is a 4.7-kΩ, 2-W carbon-composition type to minimize self-heating.
- The buffer amplifier is essential.
- Circuits using an added gate-ground diode seem to suffer increased drift.
- Its power supply should be well-regulated. Make liberal use of decoupling capacitors to prevent unintentional feedback via supply rails.
- Single-point grounding for the tank and FET is important. This usually means using one mounting screw of the tuning capacitor.
- The inductor used a ceramic form with a powdered-iron core mounted on a spring-steel screw.
- Short leads and thick solid wire (#16 to #18 gauge) are essential for mechanical stability.

Operating in the 6-MHz region, this oscillator drifted 500 Hz during the first minute as it warmed up and then drifted at 2 Hz in 30 minutes. It must be stressed that such performance does not indicate a wonderful circuit so much as care in construction, skillful choice of components, artful mechanical layout and diligence in adjustment. The 2-W gate resistor may seem strange, but 20 years ago many components were not as good as they are today. A modern, low-inductance component of low temperature coefficient and more modest power rating should be fine.

The buffer amplifier is an important part of a good oscillator system as it serves to prevent the oscillator seeing any impedance changes in the circuit being driven. This could otherwise affect the frequency.

The Oxley Tempatrimmer used by G3PDM was a rare and expensive component, used commercially only in very high-quality equipment, and it may no longer be made. An alternative temperature compensator can be constructed from currently available components, as will be described shortly. G3PDM also referred to his "standard mallet test," where a thump with a wooden hammer produced an average shift of 6 Hz, as an illustration of the benefits of solid mechanical design.

The K7HFD Low-Noise Oscillator

The other high performance oscillator example, shown in **Fig 10.16**, was designed for low-noise performance by Linley Gumm, K7HFD, and appears on page 126 of ARRL's *Solid State Design for the Radio Amateur*. Despite its publication in the homebrewer's bible, this circuit seems to have been overlooked by many builders. It uses no unusual components and looks simple, yet it is a subtle and sophisticated circuit. It represents the antithesis of G3PDM's VFO: In the pursuit of low noise sidebands, a number of design choices have been made that will degrade the stability of frequency over temperature.

The effects of oscillator noise have already been covered, and Fig 10.7 shows the effect of limiting on the signal from a noisy oscillator. Because AM noise sidebands can get translated into PM noise sidebands by imperfect limiting, there is an advantage to stripping off the AM as early as possible, in the oscillator itself. An ALC system in the oscillator will counteract and cancel only the AM components within its bandwidth, but an oscillator based on a limiter will do this over a broad bandwidth. K7HFD's oscillator uses a differential pair of bipolar transistors as a limiting amplifier. The dc bias voltage at the bases and the resistor in the common emitter path to ground establishes a controlled dc bias current. The ac voltage between the bases switches this current between the two collectors. This applies a rectangular pulse of current into link winding L2, which drives the tank. The output impedance of the collector is high in both the current on and current off states. Allied with the small number of turns of the link winding, this presents a very high impedance to the tank circuit, which minimizes the damping of the tank Q. The input impedance of the limiter is also quite high and is applied across only a one-turn tap of L1, which similarly minimizes the effect on the tank Q. The input transistor base is driven into conduction only on one peak of the tank waveform. The output transformer has the inverse of the current pulse applied to it, so the output is not a low distortion sine wave, although the output harmonics will not be as extensive as simple theory would suggest because the circuit's high output impedance allows stray capacitances to attenuate high-frequency components. The low-frequency transistors used also act to reduce the harmonic power.

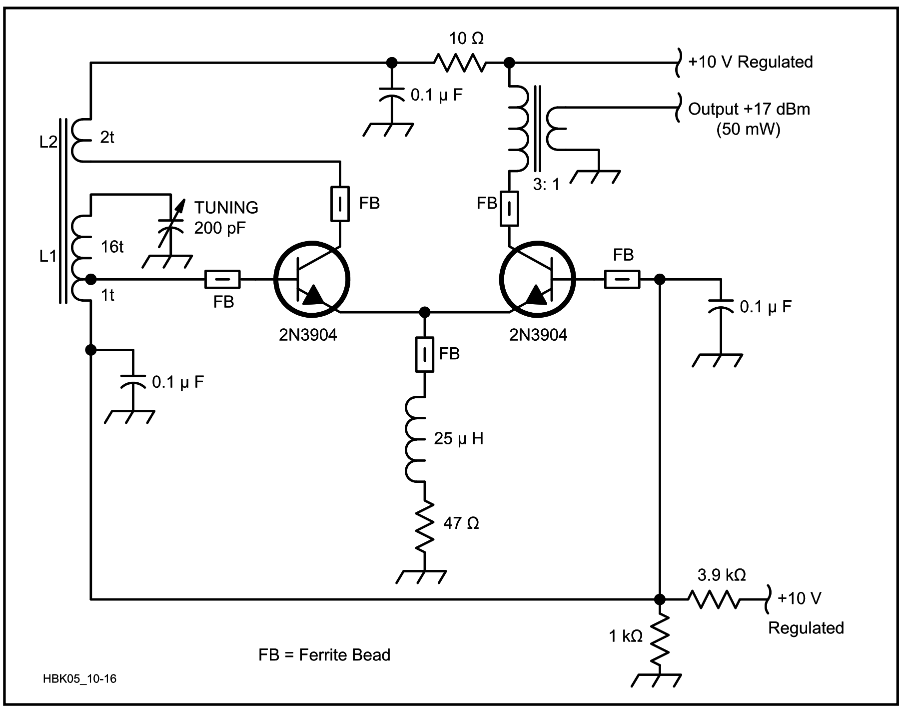


Fig 10.16—This low-noise oscillator design by K7HFD operates at an unusually high power level to achieve a high C/N (carrier-to-noise) ratio. Need other frequencies? See Fig 10.15 (caption) for a frequency-scaling technique.

With an output of +17 dBm, this is a power oscillator, running with a large dc input power, so appreciable heating results that can cause temperature-induced drift. The circuit's high-power operation is a deliberate ploy to create a high signal-to-noise ratio by having as high a signal power as possible. This also reduces the problem of the oscillator's broadband noise output. The limitation on the signal level in the tank is the transistors' base-emitter-junction breakdown voltage. The circuit runs with a few volts peak-to-peak across the one-turn tap, so the full tank is running at over 50 V P-P. The single easiest way to damage a bipolar transistor is to reverse bias the base-emitter junction until it avalanches. Most devices are only rated to withstand 5 V applied this way, the current needed to do damage is small, and very little power is needed. If the avalanche current is limited to less than that needed to perform immediate destruction of the transistor, it is likely that there will be some degradation of the device, a reduction in its bandwidth and gain along with an increase in its noise. These changes are irreversible and cumulative. Small, fast signal diodes have breakdown voltages of over 30 V and less capacitance than the transistor bases, so one possible experiment would be to try the effect of adding a diode in series with the base of each transistor and running the circuit at even higher levels.

The amplitude must be controlled by the drive current limit. The voltage on L2 must never allow the collector of the transistor driving it to go into saturation, if this happens, the transistor presents a very low impedance to L2 and badly loads the tank, wrecking the Q and the noise performance. The circuit can be checked to verify the margin from saturation by probing the hot end of L2 and the emitter with an oscilloscope. Another, less obvious, test is to vary the power-supply voltage and monitor the output power. While the circuit is under current control, there is very little change in output power, but if the supply is low enough to allow saturation, the output power will change significantly.

The use of the 2N3904 is interesting, as it is not normally associated with RF oscillators. It is a cheap, plain, general-purpose type more often used at dc or audio frequencies. There is evidence that suggests some transistors that have good noise performance *at RF* have worse noise performance at low frequencies, and that the low-frequency noise they create can modulate an oscillator, creating noise sidebands. Experiments with low-noise audio transistors may be worthwhile, but many such devices have very low f_T and high junction capacitances. In the description of this circuit in *Solid State Design for the Radio Amateur*, the results of a phase-noise test made using a spectrum analyzer with a crystal filter as a preselector are given. Ten kilohertz out from the carrier, in a 3-kHz measurement bandwidth, the noise was over 120 dB below the carrier level. This translates into better than –120 – 10 log (3000), which equals –154.8 dBc/Hz, SSB. At this offset, –140 dBc is usually considered to be excellent. This is state-of-the art performance by today's standards—in a 1977 publication.

A JFET Hartley VFO

Fig 10.17 shows an 11.1-MHz version of a VFO and buffer closely patterned after that used in 7-MHz transceiver designs published by Roger Hayward, KA7EXM, and Wes Hayward, W7ZOI (The Ugly Weekender) and Roy Lewallen, W7EL (the Optimized QRP Transceiver). In it, a 2N5486 JFET Hartley oscillator drives the two-2N3904 buffer attributed to Lewallen. This version diverges from the originals in that its JFET uses source bias (the bypassed 910-Ω resistor) instead of a gate-clamping diode and is powered from a low-current 7-V regulator IC instead of a Zener diode and dropping resistor. The 5-dB pad sets the buffer's output to a level appropriate for "Level 7" (+7-dBm-LO) diode ring mixers.

The circuit shown was originally built with a gate-clamping diode, no source bias and a 3-dB output pad. Rebiasing the oscillator as shown increased its output by 2 dB without degrading its frequency stability (200 to 300-Hz drift at power up, stability within ±20 Hz thereafter at a constant room temperature).

Temperature Compensation

The general principle for creating a high-stability VFO is to use components with minimal temperature coefficients in circuits that are as insensitive as possible to changes in components' secondary characteristics. Even after careful minimization of the causes of temperature sensitivity, further improvement can still be desirable. The traditional method was to split one of the capacitors in the tank so that it could be apportioned between low-temperature-coefficient parts and parts with deliberate temperature dependency. Only a limited number of different, controlled temperature coefficients are available, so the proportioning between low coefficient and controlled coefficient parts was varied to "dilute" the temperature sensitivity of a part more sensitive than needed. This was a tedious process, involving much trial and error, an undertaking made more complicated by the difficulty of arranging means of heating and cooling the unit being compensated. (Hayward described such a means in December 1993 *QST*.) As commercial and military equipment have been based on frequency synthesizers for some time, supplies of capacitors with controlled temperature sensitivity are drying up. An alternative approach is needed.

A temperature-compensated crystal oscillator (TCXO) is an improved-stability version of a crystal oscillator that is used widely in industry. Instead of using controlled-temperature coefficient capacitors, most TCXOs use a network of thermistors and normal resistors to control the bias of a tuning diode. See **Fig 10.18**. Manufacturers measure the temperature vs frequency characteristic of sample oscillators, and use a computer program to calculate the optimum normal resistor values for production. This can reliably achieve at least a tenfold improvement in stability. We are not interested in mass manufacture, but the idea of a thermistor tuning a varactor is worth stealing. The parts involved are likely to be available for a long time.

Browsing through component suppliers' catalogs shows ready availability of 4.5- to 5-kΩ-bead thermistors intended for temperature-compensation purposes, at less than a dollar each. **Fig 10.19** shows a circuit based on this form of temperature compensation. **Fig 10.20** illustrates a practical VCO using a tuning diode. Commonly available thermistors have negative temperature coefficients, so as temperature rises, the voltage at the counterclockwise (ccw) end of R8 increases, while that at the clockwise (cw) end drops. Somewhere near the center, there is no change. Increasing the voltage on the tuning diode decreases its capacitance, so settings toward R8's ccw end simulate a negative-temperature-coefficient capacitor; toward its clockwise end, a positive-temperature-coefficient part. Choose R1 to pass 8.5 mA from whatever supply voltage is available to the 6.2-V reference diode, D1. The 1N821A/1N829A-family diode used has a very low temperature coefficient and needs 7.5 mA bias for best performance; the bridge takes 1 mA. R7 and R8 should be good-quality multiturn trimmers. D2 and C1 are best determined by trial and error. Practical components aren't known well enough to rely on analytical models. Choose the least capacitance that provides enough compensation range. This reduces the noise added to the oscillator. (It is possible, though tedious, to solve for the differential varactor voltage with respect to R2 and R5, via differential calculus and circuit theory. The equations in Hayward's 1993 article can then be

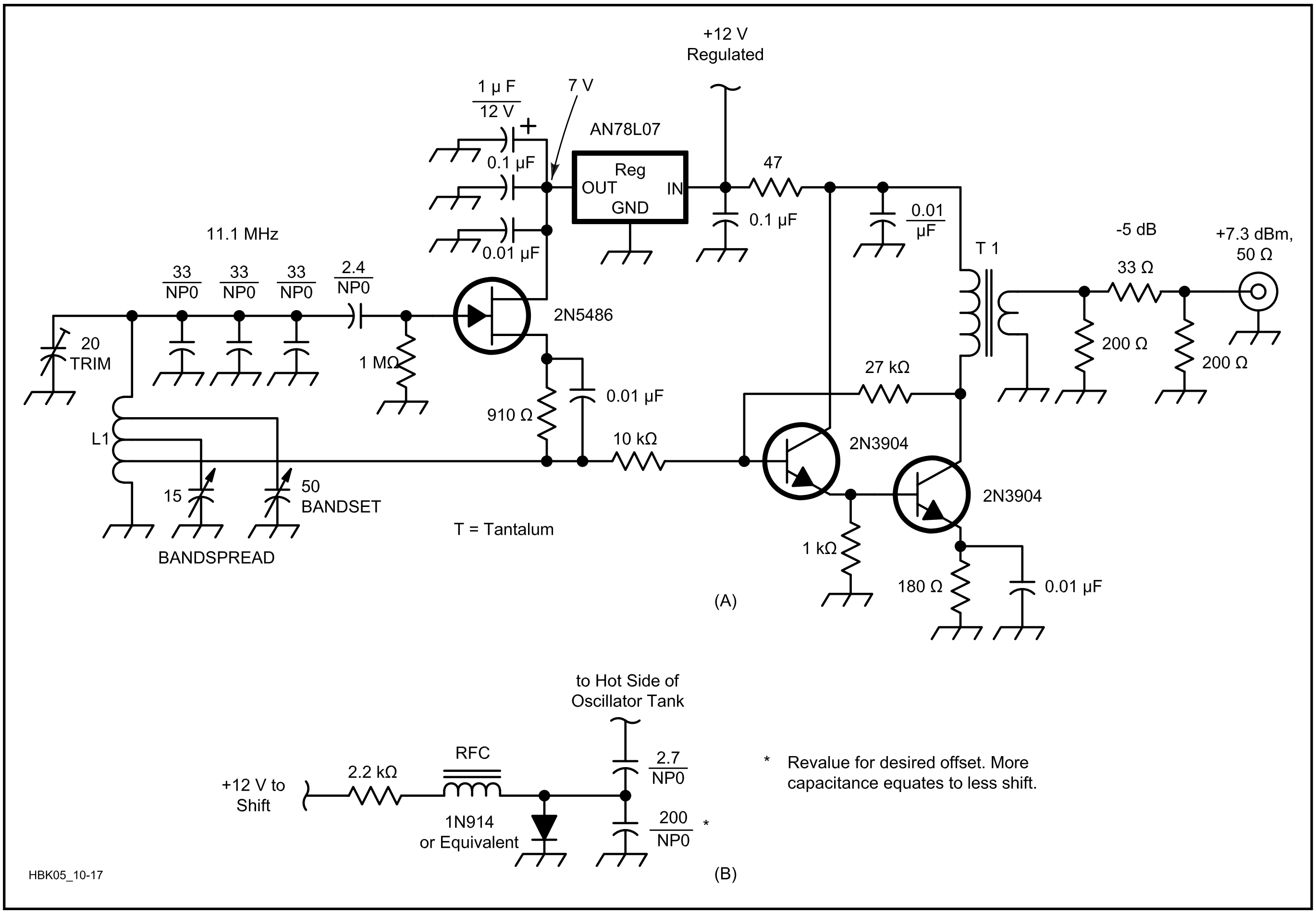


Fig 10.17—Incorporating ideas from KA2WEU, KA7EXM, W7ZOI and W7EL, the oscillator at A achieves excellent stability and output at 11.1 MHz without the use of a gate-clamping diode, as well as end-running the shrinking availability of reduction drives through the use of bandset and bandspread capacitors. L1 consists of 10 turns of B & W #3041 Miniductor (#22 tinned wire, $^5/_8$ inch in diameter, 24 turns per inch). The source tap is $2^1/_2$ turns above ground; the tuning-capacitor taps are positioned as necessary for bandset and bandspread ranges required. T1's primary consists of 15 turns of #28 enameled wire on an FT-37-72 ferrite core; its secondary, 3 turns over the primary. B shows a system for adding fixed TR offset that can be applied to any LC oscillator. The RF choke consists of 20 turns of #26 enameled wire on an FT-37-43 core. Need other frequencies? See Fig 10.15 (caption) for a frequency-scaling technique.

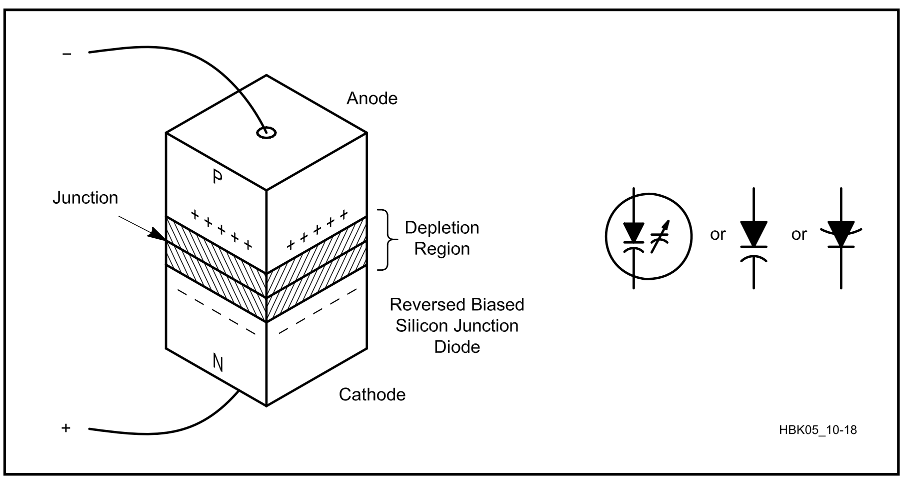


Fig 10.18—A modern voltage-controlled oscillator (VCO) uses the voltage-variable characteristic of a diode PN junction for tuning.

modified to accommodate the additional capacitors formed by D2 and C1.) Use a single ground point near D2 to reduce the influence of ground currents from other circuits. Use good-quality metal-film components for the circuit's fixed resistors.

The circuit requires two adjustments, one at each of two different temperatures, and achieving them requires a stable frequency counter that can be kept far enough from the radio so that the radio, not the counter, is subjected to the temperature extremes. (Using a receiver to listen to the oscillator under test can speed the adjustments.) After connecting the counter to the oscillator to be corrected, run the radio containing the oscillator and compensator in a room-temperature, draft-free environment until the oscillator's frequency

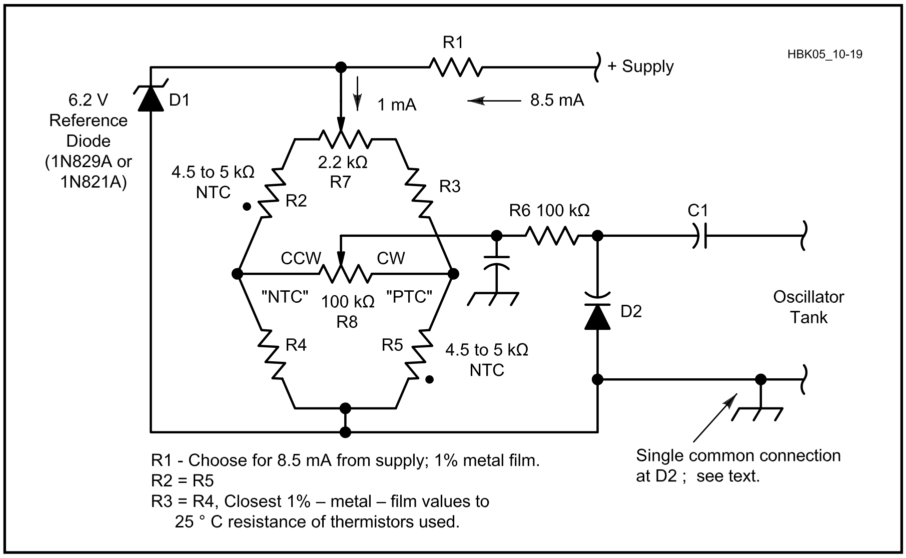


Fig 10.19—Oscillator temperature compensation has become more difficult because of the scarcity of negative-temperature-coefficient capacitors. This circuit, by GM4ZNX, uses a bridge containing two identical thermistors to steer a tuning diode for drift correction. The 6.2-V Zener diode used (a 1N821A or 1N829A) is a temperature-compensated part; just any 6.2-V Zener will not do.

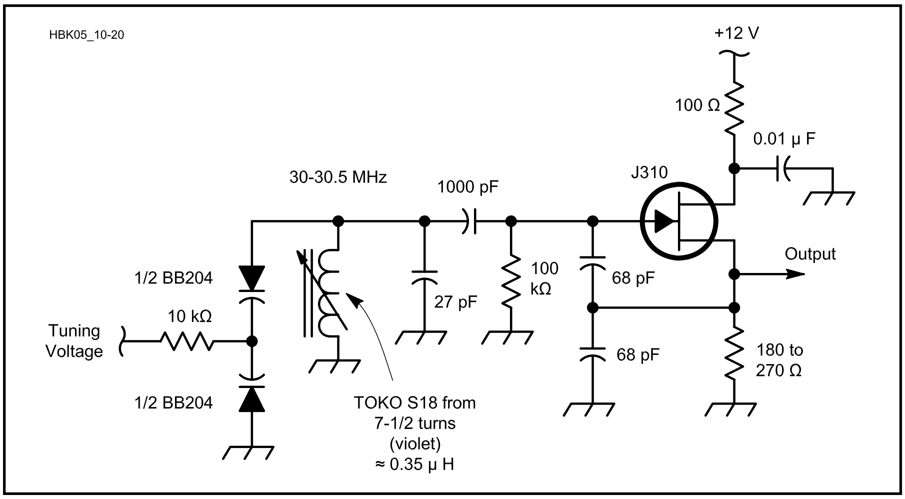


Fig 10.20—A practical VCO. The tuning diodes are halves of a BB204 dual, common-cathode tuning diode (capacitance per section at 3 V, 39 pF) or equivalent. The ECG617, NTE617 and MV104 are suitable dual-diode substitutes, or use pairs of 1N5451s (39 pF at 4 V) or MV2109s (33 pF at 4 V).

reaches its stable operating temperature rise over ambient. Lock its tuning, if possible. Adjust R7 to balance the bridge. This causes a drop of 0 V across R8, a condition you can reach by winding R8 back and forth across its range while slowly adjusting R7. When the bridge is balanced and 0 V appears across R8, adjusting R8 causes no frequency shift. When you've found this R7 setting, leave it there, set R8 is to the exact center of its range and record the oscillator frequency.

Run the radio in a hot environment and allow its frequency to stabilize. Adjust R8 to restore the frequency to the recorded value. The sensitivity of the oscillator to temperature should now be significantly reduced between the temperatures at which you performed the adjustments. You will also have somewhat improved the oscillator's stability outside this range.

For best results with any temperature-compensation scheme, it's important to group all the oscillator and compensator components in the same box, avoiding differences in airflow over components. A good oscillator should not dissipate much power, so it's feasible, even advisable, to mount all of the oscillator components in an unventilated box. In the real world, temperatures change, and if the components being compensated and the components doing the compensating have different thermal time constants, a change in temperature can cause a temporary change in frequency until the slower components have caught up. One cure for this is to build the oscillator in a thick-walled metal box that's slow to heat or cool, and so dominates and reduces the possible rate of change of temperature of the circuits inside. This is sometimes called a *cold oven*.

Shielding and Isolation

Oscillators contain inductors running at moderate power levels and so can radiate strong enough signals to cause interference with other parts of a radio, or with other radios. Oscillators are also sensitive to radiated signals. Effective shielding is therefore important. A VFO used to drive a power amplifier and antenna (to form a simple CW transmitter) can prove surprisingly difficult to shield well enough. Any leakage of the power amplifier's high-level signal back into the oscillator can affect its frequency, resulting in a poor transmitted note. If the radio gear is in the station antenna's near field, sufficient shielding may be even more difficult. The following rules of thumb continue to serve ham builders well:

- Use a complete metal box, with as few holes as possible drilled in it, with good contact around surface(s) where its lid(s) fit(s) on.
- Use feedthrough capacitors on power and control lines that pass in and out of the VFO enclosure, and on the transmitter or transceiver enclosure as well.
- Use *buffer amplifier* circuitry that amplifies the signal by the desired amount *and* provide sufficient attenuation of signal energy flowing in the reverse direction. This is known as *reverse isolation* and is a frequently overlooked loophole in shielding. Figs 10.15 and 10.17 include buffer circuitry of proven performance. As another (and higher-cost) option, consider using a high-speed buffer-amplifier IC (such as the LM6321N by National Semiconductor, a part that combines the high input impedance of an op amp with the ability to drive 50-Ω loads directly up into the VHF range).
- Use a mixing-based frequency-generation scheme instead of one that operates straight through or by means of multiplication. Such a system's oscillator stages can operate on frequencies with no direct frequency relationship to its output frequency.

- Use the time-tested technique of running your VFO at a *subharmonic* of the output signal desired—say, 3.5 MHz in a 7-MHz transmitter—and *multiply* its output frequency in a suitably nonlinear stage for further amplification at the desired frequency.

Quartz Crystals in Oscillators

Because crystals afford Q values and frequency stabilities that are orders of magnitude better than those achievable with LC circuits, fixed-frequency oscillators usually use quartz-crystal resonators. Master references for frequency counters and synthesizers are always based on crystal oscillators.

So glowing is the executive summary of the crystal's reputation for stability that newcomers to radio experimentation naturally believe that the presence of a crystal in an oscillator will force oscillation at the frequency stamped on the can. This impression is usually revised after the first few experiences to the contrary! There is no sure-fire crystal oscillator circuit (although some are better than others); reading and experience soon provide a learner with plenty of anecdotes to the effect that:

- Some circuits have a reputation of being temperamental, even to the point of not always starting.
- Crystals sometimes mysteriously oscillate on unexpected frequencies.

Even crystal manufacturers have these problems, so don't be discouraged from building crystal oscillators. The occasional uncooperative oscillator is a nuisance, not a disaster, and it just needs a little individual attention. Knowing how a crystal behaves is the key to a cure.

Quartz and the Piezoelectric Effect

Quartz is a crystalline material with a regular atomic structure that can be distorted by the simple application of force. Remove the force, and the distorted structure springs back to its original form with very little energy loss. This property allows *acoustic waves*—sound—to propagate rapidly through quartz with very little attenuation, because the velocity of an acoustic wave depends on the elasticity and density (mass/volume) of the medium through which the wave travels.

If you heat a material, it expands. Heating may cause other characteristics of a material to change—such as elasticity, which affects the speed of sound in the material. In quartz, however, expansion and change in the speed of sound are very small and tend to cancel, which means that the transit time for sound to pass through a piece of quartz is very stable.

The third property of this wonder material is that it is *piezoelectric*. Apply an electric field to a piece of quartz and the crystal lattice distorts just as if a force had been applied. The electric field applies a force to electrical charges locked in the lattice structure. These charges are captive and cannot move around in the lattice as they can in a semiconductor, for quartz is an insulator. A capacitor's dielectric stores energy by creating physical distortion on an atomic or molecular scale. In a piezoelectric crystal's lattice, the distortion affects the entire structure. In some piezoelectric materials, this effect is sufficiently pronounced that special shapes can be made that bend *visibly* when a field is applied.

Consider a rod made of quartz. Any sound wave propagating along it eventually hits an end, where there is a large and abrupt change in acoustic impedance. Just as when an RF wave hits the end of an unterminated transmission line, a strong reflection occurs. The rod's other end similarly reflects the wave. At some frequency, the phase shift of a round trip will be such that waves from successive round trips exactly coincide in phase and reinforce each other, dramatically increasing the wave's amplitude. This is *resonance*.

The passage of waves in opposite directions forms a standing wave with antinodes at the rod ends. Here we encounter a seeming ambiguity: not just one, but a *family* of different frequencies, causes standing waves—a family fitting the pattern of $^1/_2$, $^2/_3$, $^5/_2$, $^7/_2$ and so on, wavelengths into the length of the rod. And this *is* the case: A quartz rod *can* resonate at any and all of these frequencies.

The lowest of these frequencies, where the crystal is $^1/_2$ wavelength long, is called the *fundamental* mode. The others are named the third, fifth, seventh and so on, *overtones*. There is a small phase-shift error during reflection at the ends, which causes the frequencies of the overtone modes to differ slightly from odd integer multiples of the fundamental. Thus, a crystal's third overtone is very close to, but not exactly, three times, its fundamental frequency. Many people are confused by overtones and harmonics. Harmonics are additional signals at *exact* integer multiples of the fundamental frequency. Overtones are not signals at all; they are additional resonances that can be exploited if a circuit is configured to excite them.

The crystals we use most often resonate in the 1- to 30-MHz region and are of the *AT cut, thickness shear* type, although these last two characteristics are rarely mentioned. A 15-MHz-fundamental crystal of this type is about 0.15 mm thick. Because of the widespread use of reprocessed war-surplus, pressure-mounted *FT-243* crystals, you may think of crystals as small rectangles on the order of a half inch in size. The crystals we commonly use today are discs, etched and/or doped to their final dimensions, with metal electrodes deposited directly on the quartz. A crystal's diameter does not directly affect its frequency; diameters of 8 to15 mm are typical.

AT cut is one of a number of possible standard designations for the orientation at which a crystal disc is sawed from the original quartz crystal. The crystal lattice atomic structure is asymmetric, and the orientation of this with respect to the faces of the disc influences the crystal's performance. *Thickness shear* is one of a number of possible orientations of the crystal's mechanical vibration with respect to the disc. In this case, the crystal vibrates perpendicularly to its thickness. This is not easy to visualize, and diagrams don't help much, but **Fig 10.21** is an attempt at illustrating this. Place a moist bathroom sponge between the palms of your hands, move one hand up and down, and you'll see thickness shear in action.

There is a limit to how thin a disc can be made, given requirements of accuracy and price. Traditionally, fundamental-mode crystals have been made up to 20 MHz, although 30 MHz is now common at a

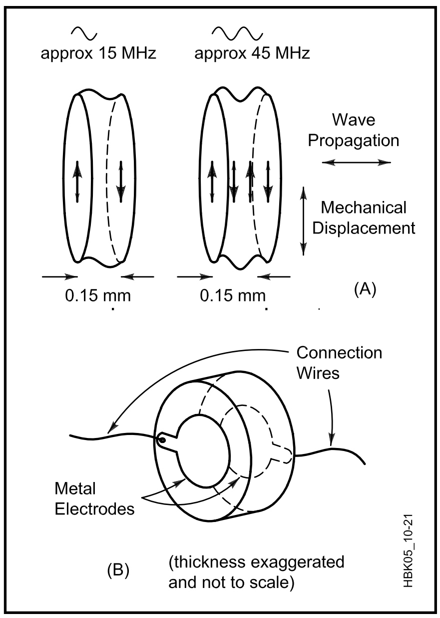


Fig 10.21—Thickness-shear vibration at a crystal's fundamental and third overtone (A); B shows how the modern crystals commonly used by radio amateurs consist of etched quartz discs with electrodes deposited directly on the crystal surface.

moderately raised price. Using techniques pioneered in the semiconductor industry, crystals have been made with a central region etched down to a thin membrane, surrounded by a thick ring for robustness. This approach can push fundamental resonances to over 100 MHz, but these are more lab curiosities than parts for everyday use. The easy solution for higher frequencies is to use a nice, manufacturably thick crystal on an overtone mode. All crystals have all modes, so if you order a 28.060-MHz, third-overtone unit for a little QRP transmitter, you'll get a crystal with a fundamental resonance somewhere near 9.353333 MHz, but its manufacturer will have adjusted the thickness to plant the third overtone exactly on the ordered frequency. An accomplished manufacturer can do tricks with the flatness of the disc faces to make the wanted overtone mode a little more active and the other modes a little less active. (As some builders discover, however, this does not *guarantee* that the wanted mode is the most active!)

Quartz's piezoelectric property provides a simple way of driving the crystal electrically. Early crystals were placed between a pair of electrodes in a case. This gave amateurs the opportunity to buy surplus crystals, open them and grind them a little to reduce their thickness, thus moving them to higher frequencies. The frequency could be reduced very slightly by loading the face with extra mass, such as by blackening it with a soft pencil. Modern crystals have metal electrodes deposited directly onto their surfaces (Fig 10.21B), and such tricks no longer work.

The piezoelectric effect works both ways. Deformation of the crystal produces voltage across its electrodes, so the mechanical energy in the resonating crystal can also be extracted electrically by the same electrodes. Seen electrically, at the electrodes, the mechanical resonances look like electrical resonances. Their Q is very high. A Q of 10,000 would characterize a *poor* crystal nowadays; 100,000 is often reached by high-quality parts. For comparison, a Q of over 200 for an LC tank is considered good.

Accuracy

A crystal's frequency accuracy is as outstanding as its Q. Several factors determine a crystal's frequency accuracy. First, the manufacturer makes parts with certain tolerances: ±200 ppm for a low-quality crystal for use as in a microprocessor clock oscillator, ±10 ppm for a good-quality part for professional radio use. Anything much better than this starts to get expensive! A crystal's resonant frequency is influenced by the impedance presented to its terminals, and manufacturers assume that once a crystal is brought within several parts per million of the nominal frequency, its user will perform fine adjustments electrically.

Second, a crystal ages after manufacture. Aging could give increasing or decreasing frequency; whichever, a given crystal usually keeps aging in the same direction. Aging is rapid at first and then slows down. Aging is influenced by the care in polishing the surface of the crystal (time and money) and by its holder style. The cheapest holder is a soldered-together, two-part metal can with glass bead insulation for the connection pins. Soldering debris lands on the crystal and affects its frequency. Alternatively, a two-part metal can be made with flanges that are pressed together until they fuse, a process called *cold-welding*. This is much cleaner and improves aging rates roughly fivefold compared to soldered cans. An all-glass case can be made in two parts and fused together by heating in a vacuum. The vacuum raises the Q, and the cleanliness results in aging that's roughly ten times slower than that achievable with a soldered can. The best crystal holders borrow from vacuum-tube assembly processes and have a *getter*, a highly reactive chemical substance that traps remaining gas molecules, but such crystals are used only for special purposes.

Third, temperature influences a crystal. A reasonable, professional quality part might be specified to shift not more than ±10 ppm over 0 to 70°C. An AT-cut crystal has an *S*-shaped frequency-versus-temperature characteristic, which can be varied by slightly changing the crystal cut's orientation. **Fig 10.22** shows the general shape and the effect of changing the cut angle by only a few seconds of arc. Notice how all the curves converge at 25°C. This is because this temperature is normally chosen as the reference for specifying a crystal. The temperature stability specification sets how accurate the manufacturer must make the cut. Better stability may be needed for a crystal used as a receiver frequency standard, frequency counter clock and so on. A crystal's temperature characteristic shows a little hysteresis. In other words, there's a bit of offset to the curve depending on whether temperature is increasing or decreasing. This is usually of no consequence except in the highest-precision circuits.

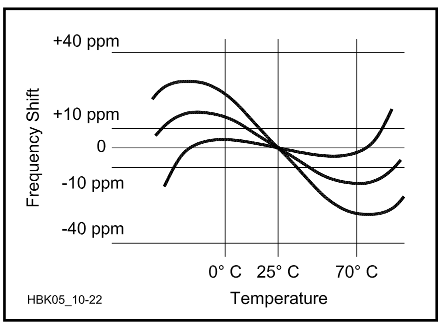


Fig 10.22—Slight changes in a crystal cut's orientation shift its frequency-versus-temperature curve.

It is the temperature of the quartz that is important, and as the usual holders for crystals all give effective thermal insulation, only a couple of milliwatts dissipation by the crystal itself can be tolerated before self-heating becomes troublesome. Because such heating occurs in the quartz itself and does not come from the surrounding environment, it defeats the effects of temperature compensators and ovens.

The techniques shown earlier for VFO for temperature compensation can also be applied to crystal oscillators. An after-compensation drift of 1 ppm is routine and 0.5 ppm is good. The result is a *temperature-compensated crystal oscillator* (*TCXO*). Recently, oscillators have appeared with built-in digital thermometers, microprocessors and ROM look-up tables customized on a unit-by-unit basis to control a tuning diode via a digital-to-analog converter (DAC) for temperature compensation. These *digitally temperature-compensated oscillators* (*DTCXO*s) can reach 0.1 ppm over the temperature range. With automated production and adjustment, they promise to become the cheapest way to achieve this level of stability.

Oscillators have long been placed in temperature-controlled *ovens*, which are typically held at 80°C. Stability of several parts per billion can be achieved over temperature, but this is a limited benefit as aging can easily dominate the accuracy. These are usually called *oven-controlled crystal oscillators* (*OCXOs*).

Fourth, the crystal is influenced by the impedance presented to it by the circuit in which it is used. This means that care is needed to make the rest of an oscillator circuit stable, in terms of impedance and phase shift.

Gravity can slightly affect crystal resonance. Turning an oscillator upside down usually produces a small frequency shift, usually much less than 1 ppm; turning the oscillator back over reverses this. This effect is quantified for the highest-quality reference oscillators.

The Equivalent Circuit of a Crystal

Because a crystal is a passive, two-terminal device, its electrical appearance is that of an impedance that varies with

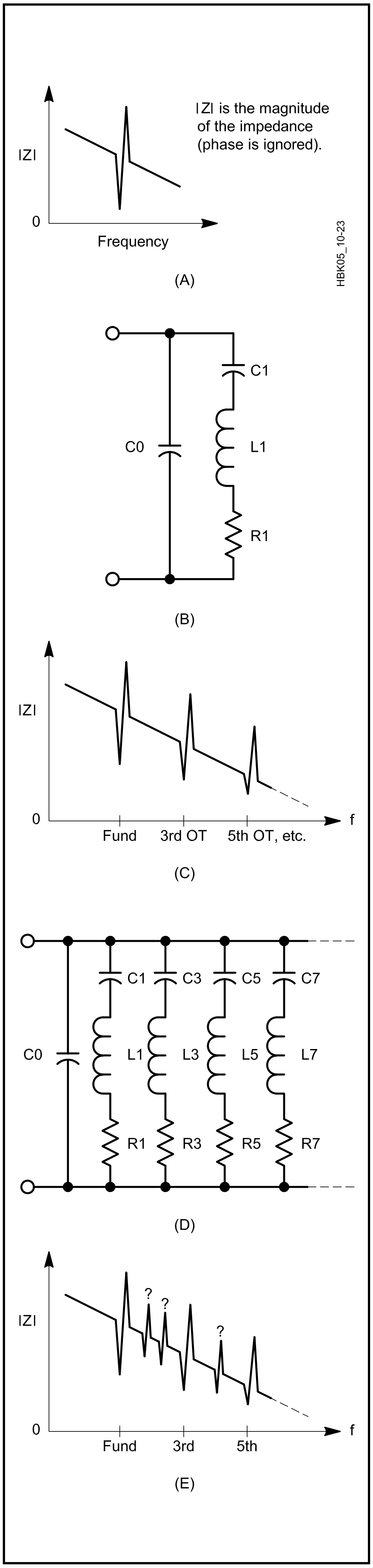


Fig 10.23—Exploring a crystal's impedance (A) and equivalent circuit (B) through simplified diagrams. C and D extend the investigation to include overtones; E, to spurious responses not easily predictable by theory or controllable through manufacture. A crystal may oscillate on any of its resonances under the right conditions.

frequency. **Fig 10.23A** shows a very simplified sketch of the magnitude (phase is ignored) of the impedance of a quartz crystal. The general trend of dropping impedance with increasing frequency implies capacitance across the crystal. The sharp fall to a low value resembles a series-tuned tank, and the sharp peak resembles a parallel-tuned tank. These are referred to as series and parallel resonances. Fig 10.23B shows a simple circuit that will produce this impedance characteristic. The impedance looks purely resistive at the exact centers of both resonances, and the region between them has impedance increasing with frequency, which looks inductive.

C1 (sometimes called *motional capacitance*, C_m, to distinguish it from the lumped capacitance it approximates) and L1 (*motional inductance,* L_m) create the series resonance, and as C0 and R1 are both fairly small, the impedance at the bottom of the dip is very close to R1. At parallel resonance, L1 is resonating with C1 and C0 in series, hence the higher frequency. The impedance of the parallel tank is immense, the terminals are connected to a capacitive tap, which causes them to see only a small fraction of this, which is still a very large impedance. The overtones should not be neglected, so Figs 10.23C and 10.23D include them. Each overtone has series and parallel resonances and so appears as a series tank in the equivalent circuit. C0 again provides the shifted parallel resonance.

This is still simplified, because real-life crystals have a number of spurious, unwanted modes that add yet more resonances, as shown in Fig 10.23E. These are not well controlled and may vary a lot even between crystals made to the same specification. Crystal manufacturers work hard to suppress these spurs and have evolved a number of recipes for shaping crystals to minimize them. Just where they switch from one design to another varies from manufacturer to manufacturer.

Always remember that the equivalent circuit is just a representation of crystal behavior and does not represent circuit components actually present. Its only use is as an aid in designing and analyzing circuits using crystals. **Table 10.2** lists typical equivalent-circuit values for a variety of crystals. It is impossible to build a circuit with 0.026 to 0.0006-pF capacitors; such values would simply be swamped by strays. Similarly, the inductor must have a Q orders of magnitude better than is practically achievable, and impossibly low stray C in its winding.

The values given in Table 10.2 are nothing more than rough guides. A crystal's frequency is tightly specified, but this still allows inductance to be traded for capacitance. A good manufacturer could hold these characteristics within a ±25% band or could vary them over a 5:1 range by special design. Similarly marked parts from different sources vary widely in motional inductance and capacitance.

Quartz is not the only material that behaves in this way, but it is the best. Resonators can be made out of lithium tantalate and a group of similar materials that have lower Q, allowing them to be *pulled* over a larger frequency range in VCXOs. Much more common, however, are ceramic resonators based on the technology of the well-known ceramic IF filters. These have much lower Q than quartz and much poorer frequency precision. They serve mainly as clock resonators for cheap microprocessor systems in which every last cent must be saved. A ceramic resonator could be used as the basis of a wide range, cheap VXO, but its frequency stability would not be as good as a good LC VFO.

Crystal Oscillator Circuits

Crystal oscillator circuits are usually categorized as series- or parallel-mode types, depending on whether the crystal's low- or high-impedance resonance comes

Table 10.2
Typical Equivalent Circuit Values for a Variety of Crystals

Crystal Type	*Series L*	*Series C (pF)*	*Series R (Ω)*	*Shunt C (pF)*
1-MHz fundamental	3.5 H	0.007	340	3.0
10-MHz fundamental	9.8 mH	0.026	7	6.3
30-MHz third overtone	14.9 mH	0.0018	27	6.2
100-MHz fifth overtone	4.28 mH	0.0006	45	7.0

into play at the operating frequency. The series mode is now the most common; parallel-mode operation was more often used with vacuum tubes. **Fig 10.24** shows a basic series-mode oscillator. Some people would say that it is an overtone circuit, used to run a crystal on one of its overtones, but this is not necessarily true. The tank (L-C1-C2) tunes the collector of the common-base amplifier. C1 is larger than C2, so the tank is tapped in a way that transforms to a lower impedance, decreasing signal voltage, but increasing current. The current is fed back into the emitter via the crystal. The common-base stage provides a current gain of less than unity, so the transformer in the form of the tapped tank is essential to give loop gain. There are *two* tuned circuits, the obvious collector tank and the series-mode one "in" the crystal. The tank kills the amplifier's gain away from its tuned frequency, and the crystal will only pass current at the series resonant frequencies of its many modes. The tank resonance is much broader than any of the crystal's modes, so it can be thought of as the crystal setting the frequency, but the tank selecting which of the crystal's modes is active. The tank could be tuned to the crystal's fundamental, or one of its overtones.

Fundamental oscillators can be built without a tank quite successfully, but there is always the occasional one that starts up on an overtone or spurious mode. Some simple oscillators have been known to change modes while running (an effect triggered by changes in temperature or loading) or to not always start in the same mode! A series-mode oscillator should present a low impedance to the crystal at the operating frequency. In Fig 10.24, the tapped collector tank presents a transformed fraction of the 1-kΩ collector load resistor to one end of the crystal, and the emitter presents a low impedance to the other. To build a practical oscillator from this circuit, choose an inductor with a reactance of about 300 Ω at the wanted frequency and calculate C1 in series with C2 to resonate with it. Choose C1 to be 3 to 4 times larger than C2. The amplifier's quiescent ("idling") current sets the gain and hence the operating level. This is not easily calculable, but can be found by experiment. Too little quiescent current and the oscillator will not start reliably; too much and the transistor can drive itself into saturation. If an oscilloscope is available, it can be used to check the collector waveform; otherwise, some form of RF voltmeter can be used to allow the collector voltage to be set to 2 to 3 V RMS. 3.3 kΩ would be a suitable starting point for the emitter bias resistor. The transistor type is not critical; 2N2222A or 2N3904 would be fine up to 30 MHz; a 2N5179 would allow operation as an overtone oscillator to over 100 MHz (because of the low collector voltage rating of the 2N5179, a supply voltage lower than 12 V is required). The ferrite bead on the base gives some protection against parasitic oscillation at UHF.

If the crystal is shorted, this circuit should still oscillate. This gives an easy way of adjusting the tank; it is even better to temporarily replace the crystal with a small-value (tens of ohms) resistor to simulate its *equivalent series resistance* (ESR), and adjust L until the circuit oscillates close to the wanted frequency. Then restore the crystal and set the quiescent current. If a lot of these oscillators were built, it would sometimes be necessary to adjust the current individually due to the different equivalent series resistance of individual crystals. One variant of this circuit has the emitter connected directly to the C1/C2 junction, while the crystal is a decoupler for the transistor base (the existing capacitor and ferrite bead not being used). This works, but with a greater risk of parasitic oscillation.

We commonly want to trim a crystal

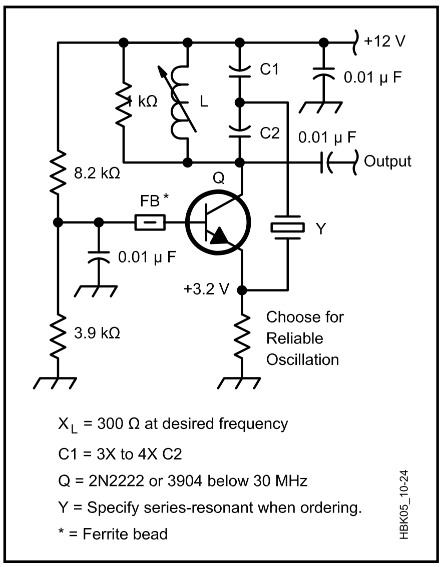


Fig 10.24—A basic series-mode crystal oscillator. A 2N5179 can be used in this circuit if a lower supply voltage is used; see text.

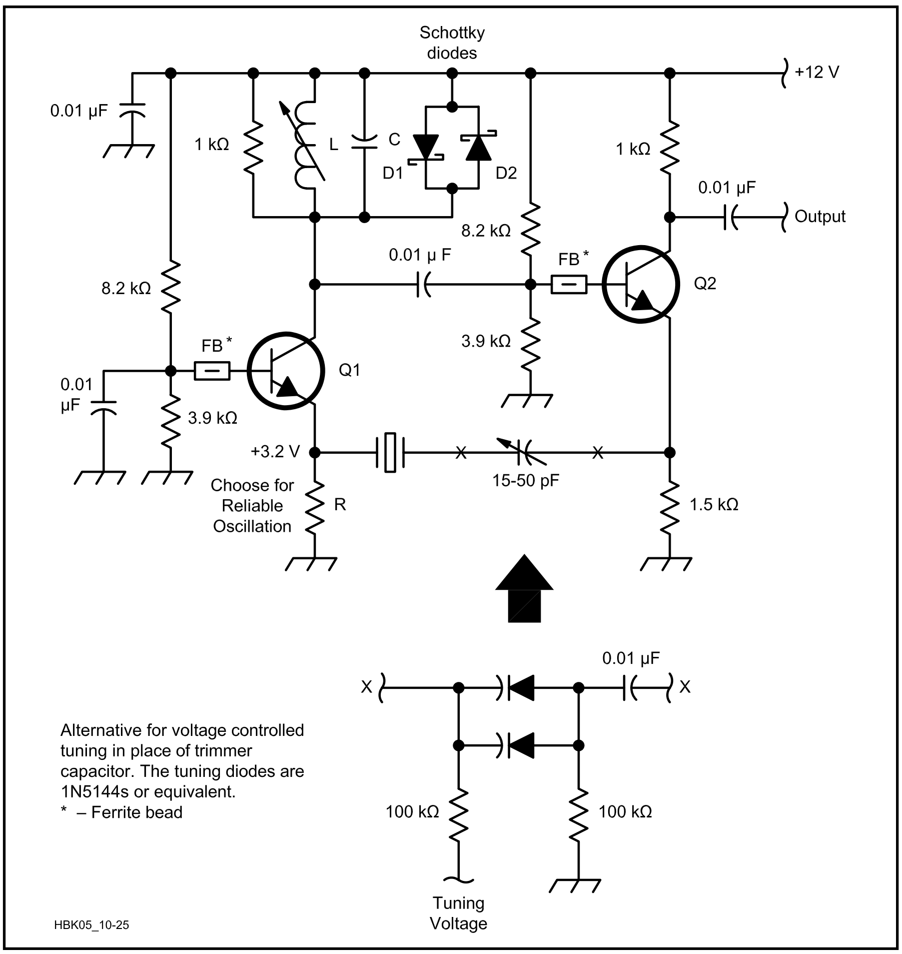


Fig 10.25—A Butler crystal oscillator.

oscillator's frequency. While off-tuning the tank a little will pull the frequency slightly, too much detuning spoils the mode control and can stop oscillation (or worse, make the circuit unreliable). The answer to this is to add a trimmer capacitor, which will act as part of the equivalent series tuned circuit, in series with the crystal. This will shift the frequency in one way only, so the crystal frequency must be respecified to allow the frequency to be varying around the required value. It is common to specify a crystal's frequency with a standard load (30 pF is commonly specified), so that the manufacturer grinds the crystal such that the series resonance of the specified mode is accurate when measured with a capacitor of this value in series. A 15- to 50-pF trimmer can be used in series with the crystal to give fine frequency adjustment. Too little capacitance can stop oscillation or prevent reliable starting. The Q of crystals is so high that marginal oscillators can take several seconds to start!

This circuit can be improved by reducing the crystal's driving impedance with an emitter follower as in **Fig 10.25**. This is the *Butler* oscillator. Again the tank controls the mode to either force the wanted overtone or protect the fundamental mode. The tank need not be tapped because Q2 provides current gain, although the circuit is sometimes seen with C split, driving Q2 from a tap. The position between the emitters offers a good, low-impedance environment to keep the crystal's in-circuit Q

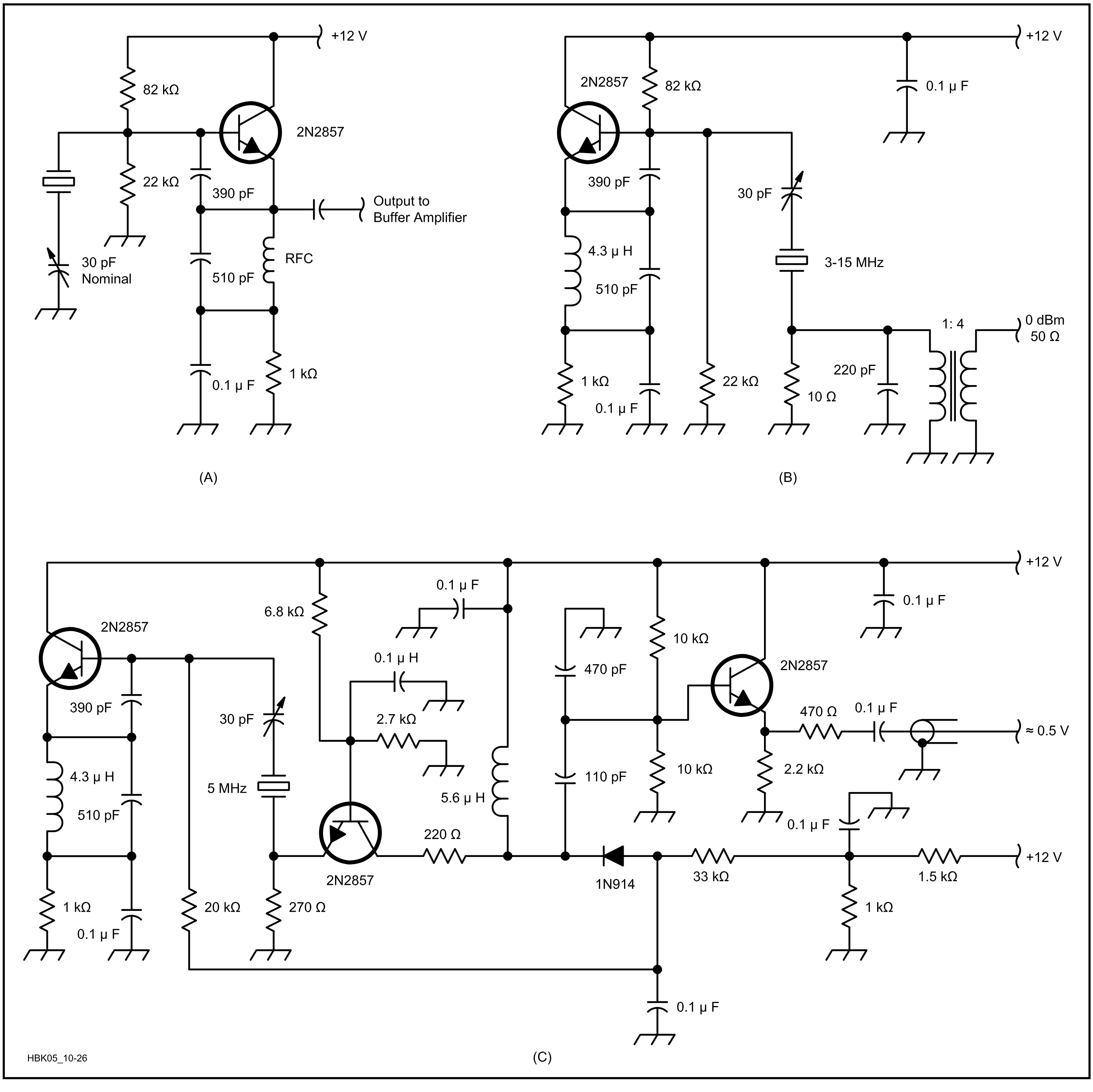


Fig 10.26—The crystal in the series-tuned Colpitts oscillator at A operates in its series-resonant mode. B shows KA2WEU's low-noise version, which uses the crystal as a filter and features high harmonic suppression. The circuit at C builds on the B version by adding a common-base output amplifier and ALC loop.

high. R, in the emitter of Q1, is again selected to give reliable oscillation. The circuit has been shown with a capacitive load for the crystal, to suit a unit specified for a 30-pF load. An alternative circuit to give electrical fine tuning is also shown. The diodes across the tank act as limiters to stabilize the operating amplitude and limit the power dissipated in the crystal by clipping the drive voltage to Q2. The tank should be adjusted to peak at the operating frequency, not used to trim the frequency. The capacitance in series with the crystal is the proper frequency trimmer.

The Butler circuit works well, and has been used in critical applications to 140 MHz (seventh-overtone crystal, 2N5179 transistor). Although the component count is high, the extra parts are cheap ones. Increasing the capacitance in series with the crystal reduces the oscillation frequency but has a progressively diminishing effect. Decreasing the capacitance pulls the frequency higher, to a point at which oscillation stops; before this point is reached, start-up will become unreliable. The possible amount of adjustment, called *pulling range*, depends on the crystal; it can range from less than ten to several hundred parts per million. Overtone crystals have much less pulling range than fundamental crystals on the same frequency; the reduction in pulling is roughly proportional to the square of the overtone number.

Low-Noise Crystal Oscillators

Fig 10.26A shows a crystal operating in its series mode in a series-tuned Colpitts circuit. Because it does not include an LC tank to prevent operation on unwanted modes, this circuit is intended for fundamental mode operation only and relies on that mode being the most active. If the crystal is ordered for 30-pF loading, the frequency trimming capacitor can be adjusted to compensate for the loading of the capacitive divider of the Colpitts circuit. An unloaded crystal without a trimmer would operate slightly off the exact series resonant frequency in order to create an inductive impedance to resonate with the divider capacitors. Ulrich Rohde, KA2WEU, in Fig 4-47 of his book *Digital PLL Frequency Synthesizers—Theory and Design*, published an elegant alternative method of extracting an output signal from this type of circuit, shown in Fig 10.26B. This taps off a signal from the current in the crystal itself. This can be thought of as using the crystal as a band-pass filter for the oscillator output. The RF choke in the emitter keeps the emitter bias resistor from loading the tank and degrading the Q. In this case (3-MHz operation), it has been chosen to resonate close to 3 MHz with the parallel capacitor (510 pF) as a means of forcing operation on the wanted mode. The 10-Ω resistor and the transformed load impedance will reduce the in-circuit Q of the crystal, so a further development substituted a common base amplifier for the resistor and transformer. This is shown in Fig 10.26C. The common-base amplifier is run at a large quiescent current to give a very low input impedance. Its collector is tuned to give an output with low harmonic content and an emitter follower is used to buffer this from the load. This oscillator sports a simple ALC system, in which the amplified and rectified signal is used to reduce the bias voltage on the oscillator transistor's base. This circuit is described as achieving a phase noise level of –168 dBc/Hz a few kilohertz out from the carrier. This may seem far beyond what may ever be needed, but frequency multiplication to high frequencies, whether by classic multipliers or by frequency synthesizers, multiplies the deviation of any FM/PM sidebands as well as the carrier frequency. This means that phase noise worsens by 20 dB for each tenfold multiplication of frequency. A clean crystal oscillator and a multiplier chain is still the best way of generating clean microwave signals for use with narrow-band modulation schemes.

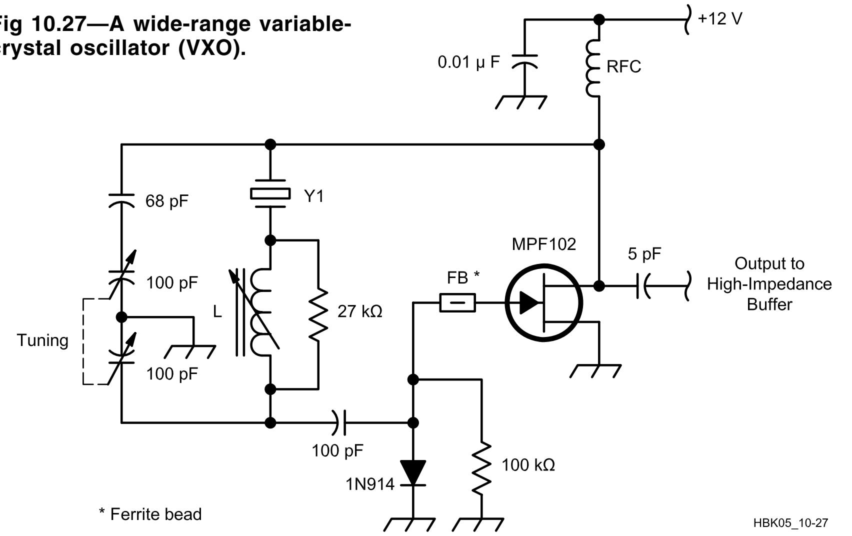
It has already been mentioned that overtone crystals are much harder to pull than fundamental ones. This is another way of saying that overtone crystals are less influenced by their surrounding circuit, which is helpful in a frequency-standard oscillator like this one. Even though 5 MHz is in the main range of fundamental-mode crystals and this circuit will work well with them, an overtone crystal has been used. To further help stability, the power dissipated in the crystal is kept to about 50 µW. The common-base stage is effectively driven from a higher impedance than its own input impedance, under which conditions it gives a very low noise figure.

VXOs

Some crystal oscillators have frequency trimmers. If the trimmer is replaced by a variable capacitor as a front-panel control, we have a *variable crystal oscillator* (*VXO*): a crystal-based VFO with a narrow tuning range, but good stability and noise performance. VXOs are often used in small, simple QRP transmitters to tune a few kilohertz around common calling frequencies. Artful constructors, using optimized circuits and components, have achieved 1000-ppm tuning ranges. Poor-quality "soft" crystals are more pullable than high-Q ones. Overtone crystals are not suited to VXOs. For frequencies beyond the usual limit for fundamental mode crystals, use a fundamental unit and frequency multipliers.

ICOM and Mizuho made some 2-m SSB transceivers based on multiplied VXO local oscillators. This system is simple and can yield better performance than many expensive synthesized radios. SSB filters are available at 9 or 10.7 MHz, to yield sufficient image rejection with a single conversion. Choice of VXO frequency depends on whether the LO is to be above or below signal frequency and how much multiplication can be tolerated. Below 8 MHz multiplier filtering is difficult. Above 15 MHz, the tuning range per crystal narrows. A 50-200 kHz range per crystal should work with a modern front-end design feeding a good 9-MHz IF, for a

Fig 10.27—A wide-range variable-crystal oscillator (VXO).



contest quality 2-m SSB Receiver.

The circuit in **Fig 10.27** is a JFET VXO from Wes Hayward, W7ZOI, and Doug DeMaw, W1FB, optimized for wide-range pulling. Published in *Solid State Design for the Radio Amateur*, many have been built and its ability to pull crystals as far as possible has been proven. Ulrich Rohde, KA2WEU, has shown that the diode arrangement as used here to make signal-dependent negative bias for the gate confers a phase-noise disadvantage, but oscillators like this that pull crystals as far as possible need any available means to stabilize their amplitude and aid start-up. In this case, the noise penalty is worth paying. This circuit can achieve a 2000-ppm tuning range with amenable crystals. If you have some overtone crystals in your junk box whose fundamental frequency is close to the wanted value, they are worth trying.

This sort of circuit doesn't necessarily stop pulling at the extremes of the possible tuning range, sometimes the range is set by the onset of undesirable behavior such as jumping mode or simply stopping oscillating. L was a 16-µH slug-tuned inductor for 10-MHz operation. It is important to minimize the stray and interwinding capacitance of L since this dilutes the range of impedance presented to the crystal.

One trick that can be used to aid the pulling range of oscillators is to tune out the C0 of the equivalent circuit with an added inductor. **Fig 10.28** shows how. L is chosen to resonate with C0 for the individual crystal, turning it into a high-impedance parallel-tuned circuit. The Q of this circuit is orders of magnitude less than the Q of the true series resonance of the crystal, so its tuning is much broader. The value of C0 is usually just a few picofarads, so L has to be a fairly large value considering the frequency it is resonated at. This means that L has to have low stray capacitance or else it will self-resonate at a lower frequency. The tolerance on C0 and the variations of the stray C of the inductor means that individual adjustment is needed. This technique can also work wonders in crystal ladder filters.

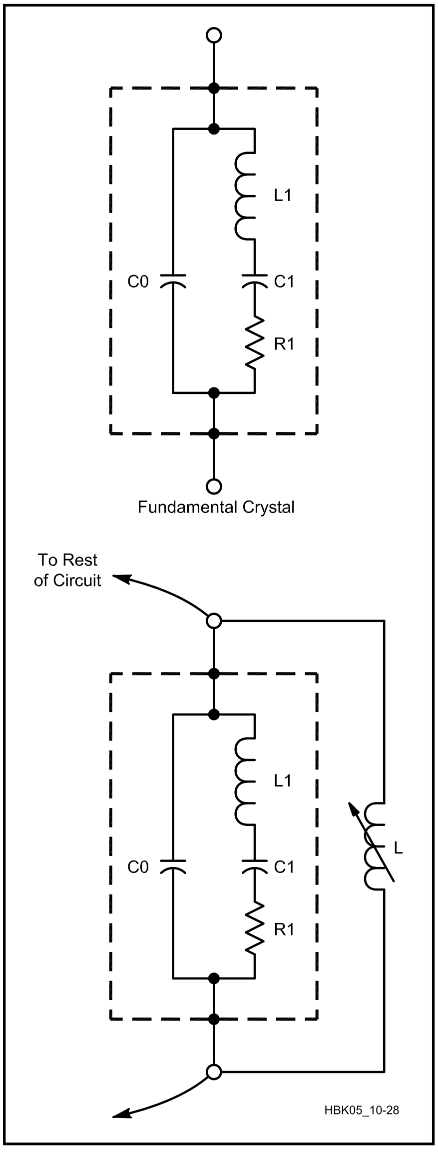


Fig 10.28—Using an inductor to "tune out" C0 can increase a crystal oscillator's pulling range.

Logic-Gate Crystal Oscillators

A 180° phase-shift network and an inverting amplifier can be used to make an oscillator. A single stage RC low-pass network cannot introduce more than 90° of phase shift and only approaches that as the signal becomes infinitely attenuated. Two stages can only approach 180° and then have immense attenuation. It takes three stages to give 180° and not destroy the loop gain. **Fig 10.29A** shows the basic form of the *phase-shift* oscillator; Fig 10.29B is an example the phase-shift oscillator commonly used in commercial Amateur Radio transceivers as an audio sidetone oscillator. The frequency-determining network of an RC oscillator has a Q of less than one, which is a massive disadvantage compared to an LC or crystal resonator. The Pierce crystal oscillator is a converted phase-shift oscillator, with the crystal taking the place of one series resistor, Fig 10.29C. At the exact series resonance, the crystal looks resistive, so by suitable choice of the capacitor values, the circuit can be made to oscillate. The crystal has a far steeper phase/frequency relationship than the rest of the network, so the crystal is the dominant controller of the frequency. The Pierce circuit is rarely seen in this full form. Instead, a cut-down version has become the most common circuit for crystal-clock oscillators for digital systems. Fig 10.29D shows this minimalist Pierce, using a logic inverter as the amplifier. R_{bias} provides dc negative feedback to bias the gate into its linear region. At first sight it appears that this arrangement should not oscillate, but the crystal is a resonator (not a simple resistor) and oscillation occurs offset from the series resonance, where the crystal appears inductive, which makes up the missing phase shift. This is one circuit that *cannot* oscillate exactly on the crystal's series resonance. This circuit is included in many microprocessors and other digital ICs that need a clock. It is also the usual circuit inside the miniature clock oscillator cans. It is not a very reliable circuit, as operation is dependent on the crystal's equivalent series resistance and the output impedance of the logic gate, occasionally the logic device or the crystal needs to be changed to start oscillation, sometimes playing with the capacitor values is necessary. It is doubtful whether these circuits are ever designed—values (the two capacitors) seem to be arrived at by experiment. Once going, the circuit is reliable, but its drift and noise are moderate, not good—acceptable for a clock oscillator. The commercial packaged oscillators are the same, but the manufacturers have handled the production foibles on a batch-by-batch basis.

RC Oscillators

Plenty of RC oscillators are capable of operating to several megahertz. Some of these are really constant-current-into-capacitor circuits, which are easier to make in silicon. Like the phase-shift oscillator above, the timing circuit Q is less than one, giving very poor noise performance that's unsuitable even to the least demanding radio application. One example is the oscillator section of the CD4046 phase-locked-loop IC. This oscillator has poor stability over temperature, large batch-to-batch variation and a wide variation in its voltage-to-frequency relationship. It is not recommended that this sort of oscillator is used at RF in radio systems. (The '4046 phase detector section is very useful, however, as we'll see later.) These oscillators are best suited to audio applications.

VHF AND UHF OSCILLATORS

A traditional way to make signals at higher frequencies is to make a signal at a lower frequency (where oscillators are easier) and multiply it up to the wanted range. Multipliers are still one of the easiest ways of making a clean UHF/microwave signal. The design of a multiplier depends on whether the multiplication factor is an odd or even number. For odd

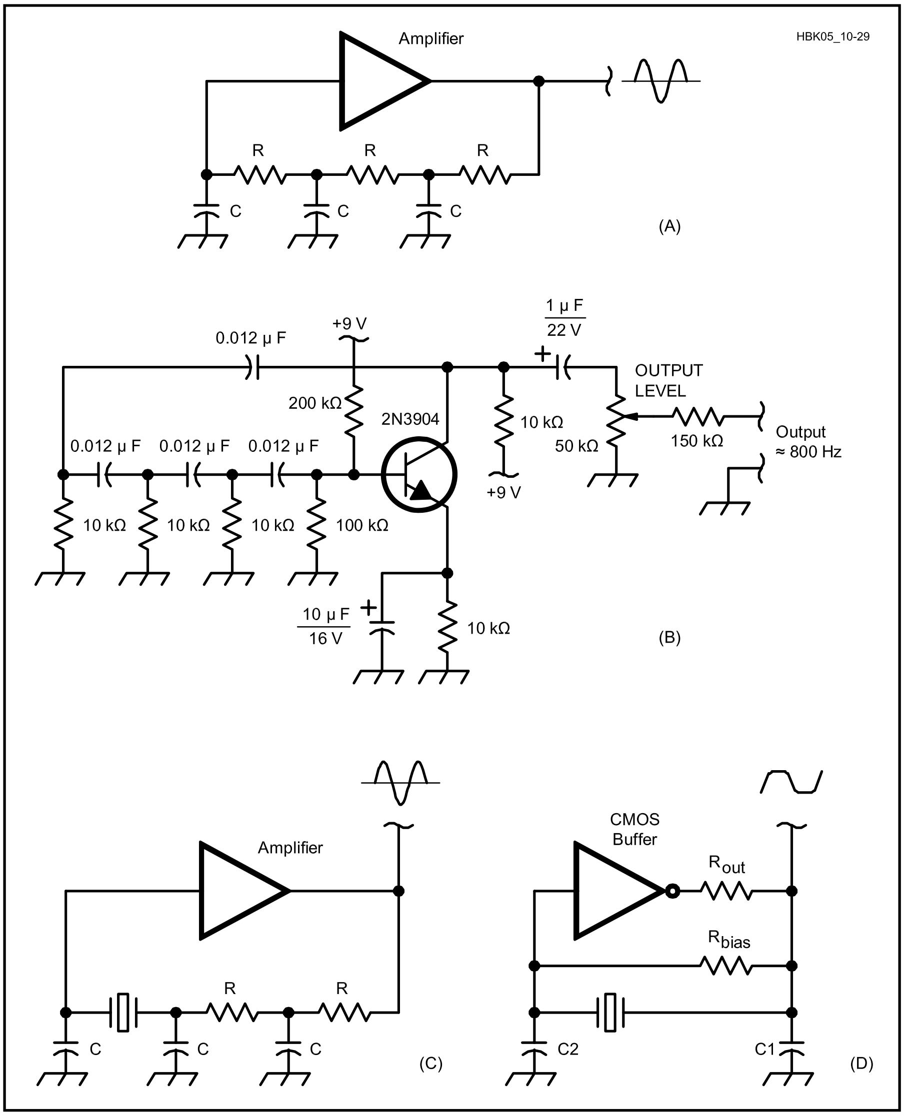


Fig 10.29—In a phase-shift oscillator (A) based on logic gates, a chain of RC networks—three or more—provide the feedback and phase shift necessary for oscillation, at the cost of low Q and considerable loop loss. Many commercial Amateur Radio transceivers have used a phase-lead oscillator similar to that shown at B as a sidetone generator. Replacing one of the resistors in A with a crystal produces a Pierce oscillator (C), a cut-down version of which (D) has become the most common clock oscillator configuration in digital systems.

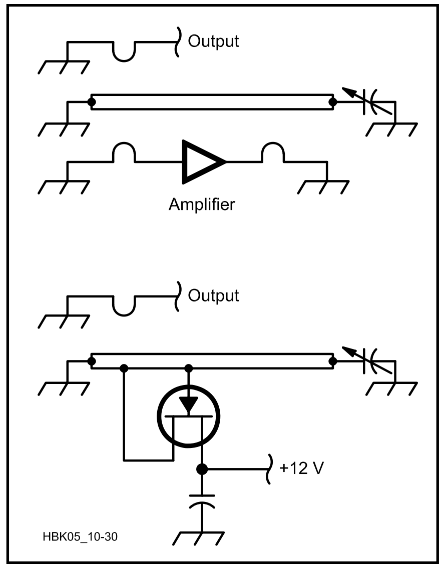


Fig 10.30—Oscillators that use transmission-line segments as resonators. Such oscillators are more common than many of us may think, as Fig 10.31 reveals.

multiplication, a Class-C biased amplifier can be used to create a series of harmonics; a filter selects the one wanted. For even multiplication factors, a full-wave-rectifier arrangement of distorting devices can be used to create a series of harmonics with strong even-order components, with a filter selecting the wanted component. At higher frequencies, diode-based passive circuits are commonly used. Oscillators using some of the LC circuits already described can be used in the VHF range. At UHF different approaches become necessary.

Fig 10.30 shows a pair of oscillators based on a resonant length of line. The first one is a return to basics: a resonator, an amplifier and a pair of coupling loops. The amplifier can be a single bipolar or FET device or one of the monolithic microwave integrated circuit (MMIC) amplifiers. The second circuit is really a Hartley, and one was made as a test oscillator for the 70-cm band from a 10-cm length of wire suspended 10 mm over an unetched PC board as a ground plane, bent down and soldered at one end, with a trimmer at the other end. The FET was a BF981 dual-gate device used as a source follower.

No free-running oscillator will be stable enough on these bands except for use with wideband FM or video modulation and AFC at the receiver. Oscillators in this range are almost invariably tuned with tuning diodes controlled by phase-locked-loop synthesizers, which are themselves controlled by a crystal oscillator.

There is one extremely common UHF oscillator that is rarely applied intentionally. The answer to this riddle is a configuration that is sometimes deliberately built as a useful wide-tuning oscillator covering say, 500 MHz to 1 GHz—and is also the modus operandi of a very common form of *spurious* VHF/UHF oscillation in circuitry intended to process lower-frequency signals! This oscillator has no generally accepted name. It relies on the creation of a small negative resistance in series with a series resonant LC tank. **Fig 10.31A** shows the circuit in its simplest form. This circuit is well-suited to construction with printed-circuit inductors. Common FR4 glass-epoxy board is lossy at these frequencies; better performance can be achieved by using the much more expensive glass-Teflon board. If you can get surplus offcuts of this type of material, it has many uses at UHF and microwave, but it is difficult to use, as the adhesion between the copper and the substrate can be poor. A high-UHF transistor with a 5-GHz f_T like the BFR90 is suitable; the base inductor can be 30 mm of 1-mm trace folded into a hairpin shape (inductance, less than 10 nH). Analyzing this circuit using a comprehensive model of the UHF transistor reveals that the emitter presents an impedance that is small, resistive and negative to the outside world. If this is large enough to more than cancel the effective series resistance of the emitter tank, oscillation will occur. Fig 10.31B shows a very basic emitter-follower circuit with some capacitance to ground on both the input and output. If the capacitor shunting the input is a distance away from the transistor, the trace can look like an inductor—and small capacitors at audio and low RF can look like very good decouplers at hundreds of MHz. The length to the capacitor shunting the output will behave as a series resonator at a frequency where it is a 1/4 wavelength long. This circuit is the same as in Fig 10.31A; it, too, can oscillate. The semiconductor manufacturers have steadily improved their small-signal transistors to give better gain and bandwidth so that any transistor circuit where all three electrodes find themselves decoupled to ground at UHF may oscillate at several hundred megahertz. The upshot of this is that there is no longer any branch of electronics where RF design and layout techniques can be safely ignored. A circuit must not just be designed to do what it *should* do, it must also be designed so that it cannot do what it *should not* do.

There are three ways of taming such a circuit; adding a small resistor, perhaps 50 to 100 Ω in the collector lead, close to the transistor, or adding a similar resistor in the base lead, or by fitting a ferrite bead over the base lead under the transistor. The resistors can disturb dc conditions, depending on the circuit and its operating currents. Ferrite beads have the advantage that they can be easily added to existing equipment and have no effect at dc and low frequencies. Beware of some electrically conductive ferrite materials that can short transistor leads. If an HF oscillator uses beads to prevent any risk of spurs (Fig 10.16), the beads should be anchored with a spot of adhesive to prevent movement that can cause small frequency shifts. Ferrite beads of Fair-Rite no. 43 material are especially suitable for this purpose; they are specified in terms of impedance, not inductance. Ferrites at frequencies above their normal usable range become very lossy and can make a lead look not inductive, but like a few tens of ohms, resistive.

Microwave Oscillators

Low-noise microwave signals are still best made by multiplying a very-low-noise HF crystal oscillator, but there are a number of oscillators that work directly at microwave frequencies. Such oscillators can be based on resonant lengths of stripline or microstrip, and are simply scaled-down versions of UHF oscillators, using microwave transistors and printed striplines on a low-loss substrate, like alumina. Techniques for printing metal traces

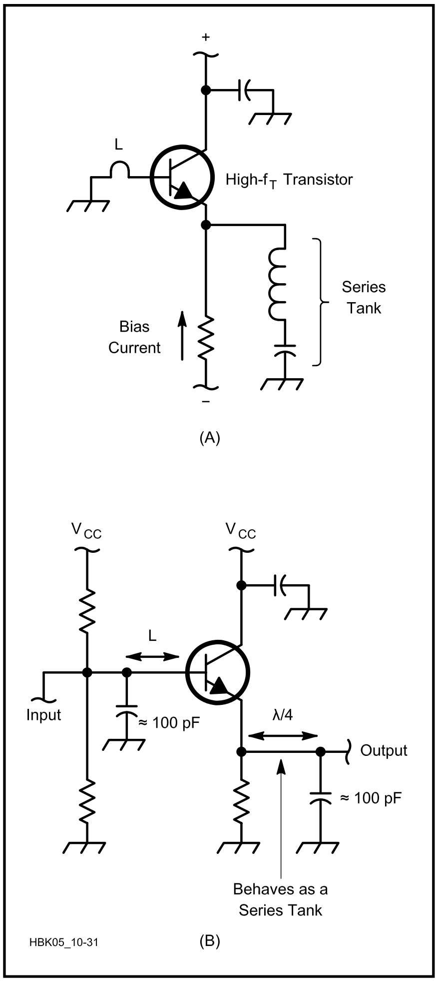


Fig 10.31—High device gain at UHF and resonances in circuit board traces can result in spurious oscillations even in non-RF equipment.

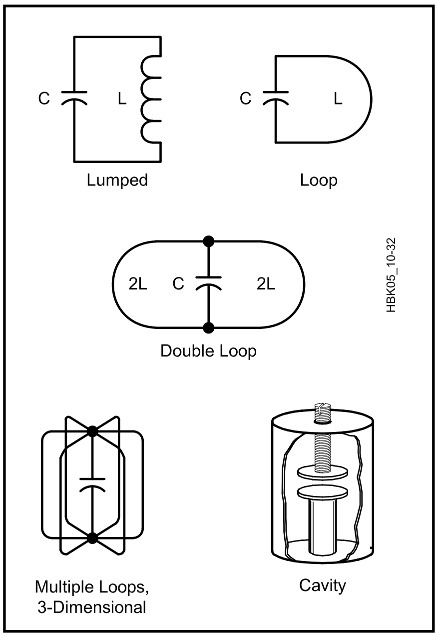


Fig 10.32—Evolution of the cavity resonator.

on substrates to form filters, couplers, matching networks and so on, have been intensively developed over the past two decades. Much of the professional microwave community has moved away from waveguide and now uses low-loss coaxial cable with a solid-copper shield—*semi-rigid cable*—to connect circuits made flat on ceramic or Teflon based substrates. Semiconductors are often bonded on as unpackaged chips, with their bond-wire connections made directly to the traces on the substrate. At lower microwave frequencies, they may be used in standard surface-mount packages. From an Amateur Radio viewpoint, many of the processes involved are not feasible without access to specialized furnaces and materials. Using ordinary PC-board techniques with surface-mount components allows the construction of circuits up to 4 GHz or so. Above this, structures get smaller and accuracy becomes critical; also PC-board materials quickly become very lossy.

Older than stripline techniques and far more amenable to home construction, cavity-based oscillators can give the highest possible performance. The dielectric constant of the substrate causes stripline structures to be much smaller than they would be in free air, and the lowest-loss substrates tend to have very high dielectric constants. Air is a very low-loss dielectric with a dielectric constant of 1, so it gives high Q and does not force excessive miniaturization. **Fig 10.32** shows a series of structures used by G. R. Jessup, G6JP, to illustrate the evolution of a cavity from a tank made of lumped components. All cavities have a number of different modes of resonance, the orientation of the currents and fields are shown in **Fig 10.33**. The cavity can take different shapes, but that shown has proven to suppress unwanted modes well. The gap need not be central and is often right at the top. A screw can be fitted through the top, protruding into the gap, to adjust the frequency.

To make an oscillator out of a cavity, an amplifier is needed. Gunn and tunnel diodes have regions in their characteristics where their current *falls* with increasing bias voltage. This is negative resistance. If such a device is mounted in a loop in a cavity and bias applied, the negative resistance can more than cancel the effective loss resistance of the cavity, causing oscillation. These diodes are capable of operating at extremely high frequencies and were discovered long before transistors were developed that had any gain at microwave frequencies.

A *Gunn-diode* cavity oscillator is the basis of many of the Doppler-radar modules used to detect traffic or intruders. **Fig 10.34** shows a common configuration. The coupling loop and coax output connector could be replaced with a simple aperture to couple into waveguide or a mixer cavity. **Fig 10.35** shows a transistor cavity oscillator version using a modern microwave transistor. FET or bipolar devices can be used. The two coupling loops are completed by the capacitance of the feedthroughs.

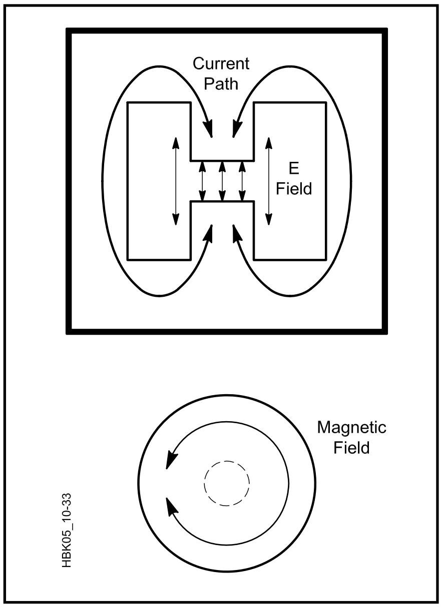


Fig 10.33—Currents and fields in a cavity.

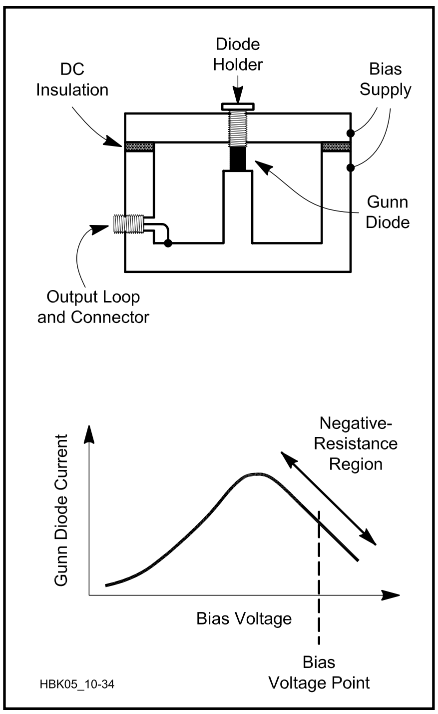


Fig 10.34—A Gunn diode oscillator uses negative resistance and a cavity resonator to produce radio energy.

The *dielectric-resonator oscillator* (*DRO*) may soon become the most common microwave oscillator of all, as it is used in the downconverter of satellite TV receivers. The dielectric resonator itself is a ceramic cylinder, like a miniature hockey puck, several millimeters in diameter. The ceramic has a very high dielectric constant, so the surface (where ceramic meets air in an abrupt mismatch) reflects electromagnetic waves and makes the ceramic body act as a resonant cavity. It is mounted on a substrate and coupled to the active device of the oscillator by a stripline that runs past it. At 10 GHz, a FET made of gallium arsenide (GaAsFET), rather than silicon, is normally used. The dielectric resonator elements are made at frequencies appropriate to mass applications like satellite TV. The set-up charge to manufacture small quantities at special frequencies is likely to be prohibitive for the foreseeable future. The challenge with these devices is to devise new ways of using oscillators on industry standard frequencies. Their chief attraction is their low cost in large quantities and compatibility with microwave stripline (microstrip) techniques. Frequency stability and Q are competitive with good cavities, but are inferior to that achievable with a crystal oscillator and chain of frequency multipliers. Satellite TV downconverters need free running oscillators with less than 1 MHz of drift at 10 GHz.

There are a number of thermionic (vacuum-tube) microwave sources, klystrons, magnetrons and *backwards wave oscillators* (*BWO*s). Available devices are either very old or designed for very high power.

The *yttrium-iron garnet* (*YIG*) oscillator was developed for a wide tuning range as a solid-state replacement for the BWO, and many of them can be tuned over more than an octave. They are complete, packaged units that appear to be a heavy block of metal with low-frequency connections for power supplies and tuning, and an SMA connector for the RF output. The manufacturer's label usually states the tuning range and often the power supply voltages. This is very helpful because, with new units priced in the kilodollar range, it is important to be able to identify surplus units. The majority of YIGs are in the 2- to 18-GHz region, although units down to 500 MHz and up to 40 GHz are occasionally found. In this part of the spectrum there is no octave-tunable device that can equal their cleanliness and stability. A 3-GHz unit drifting less than 1 kHz per second gives an idea of typical stability. This seems very poor—until we realize that this is 0.33 ppm per second. Nevertheless, any YIG application involving narrow-band modulation will usually require some form of frequency stabilizer.

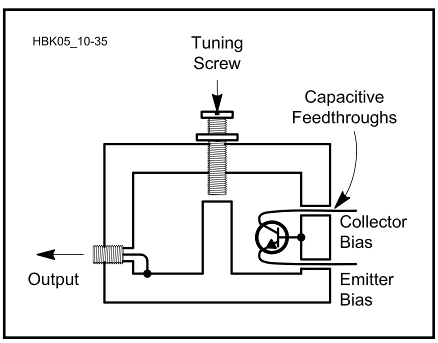


Fig 10.35—A transistor can also directly excite a cavity resonator.

Good quality, but elderly, RF spectrum analyzers have found their way into the workshops of a number of dedicated constructors. A 0- to 1500-MHz analyzer usually uses a 2- to 3.5-GHz YIG as its first local oscillator, and its tuning circuits are designed around it. A reasonable understanding of YIG oscillators will help in troubleshooting and repair.

YIG spheres are resonant at a frequency controlled not only by their physical dimensions, but also by any applied magnetic field. A YIG sphere is carefully oriented within a coupling loop connected to a negative-resistance device and the whole assembly is placed between the poles of an electromagnet. Negative-resistance diodes have been used, but transistor circuits are now common. The support for the YIG sphere often contains a thermostatically controlled heater to reduce temperature sensitivity.

The first problem with a magnetically tuned oscillator is that magnetic fields, especially at low frequencies, are extremely difficult to shield; the tuning will be influenced by any local fields. Varying fields will cause frequency modulation. The magnetic core must be carefully designed to be all-enclosing in an attempt at self-shielding and then one or more nested mu-metal cans are fitted around everything. It is still important to site the unit away from obvious sources of magnetic noise, like power transformers. Cooling fans are less obvious sources of fluctuating magnetic fields, since some are 20 dB worse than a well designed 200-W 50/60-Hz transformer.

The second problem is that the oscillator's internal tuning coils need significant current from the power supply to create strong fields. This can be eased by adding a permanent magnet as a fixed "bias" field, but the bias will shift as the magnet ages. The only solution is to have a coil with many turns, hence high inductance, which will require a high supply voltage to permit rapid tuning, thus increasing the power consumption. The usual compromise is to have dual coils: One with many turns allows slow tuning over a wide range; a second with much fewer turns allows fast tuning or FM over a limited range. The main coil can have a sensitivity in the 20 MHz/mA range; and the "FM" coil, perhaps 500 kHz/mA.

The frequency/current relationship can have excellent linearity. **Fig 10.36** shows the construction of a YIG oscillator.

For some insight into present and future trends in oscillator applications, see the sidebar beginning on page 52.

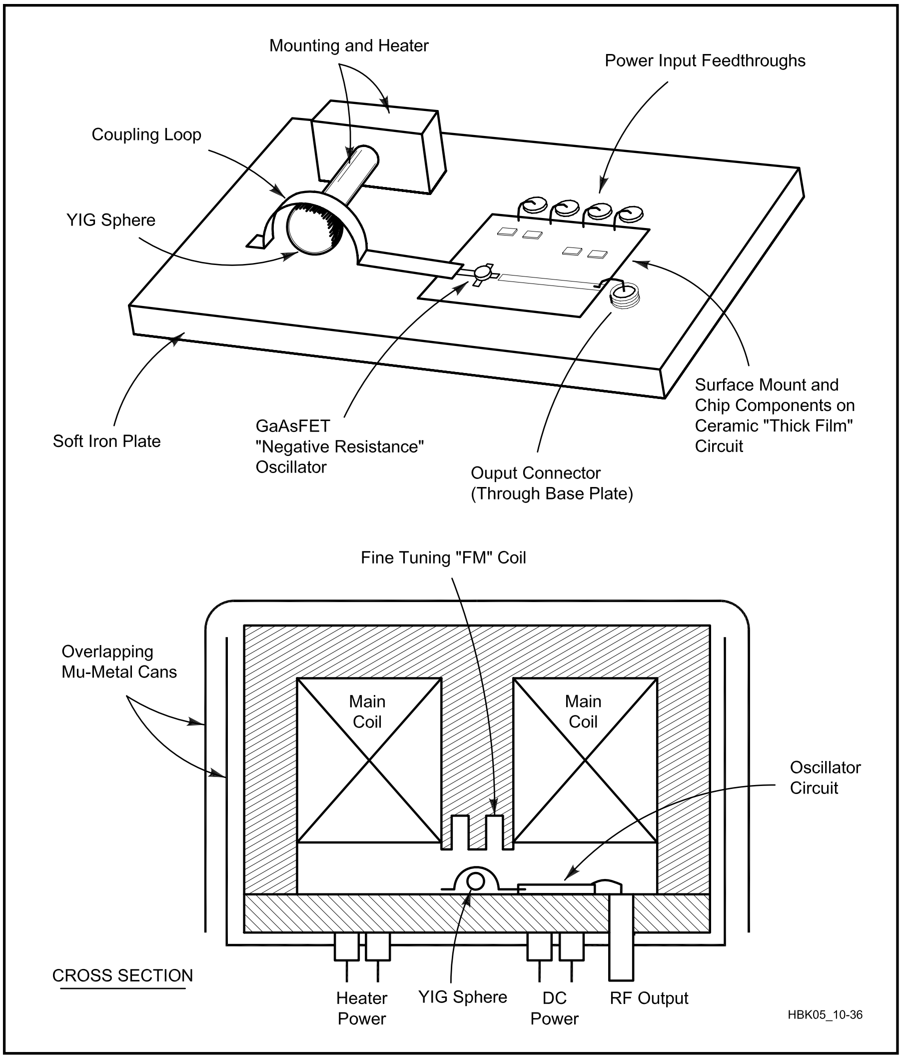


Fig 10.36—A yttrium-iron-garnet (YIG) sphere serves as the resonator in the sweep oscillators used in many spectrum analyzers.

Frequency Synthesizers

Like many of our modern technologies, the origins of Frequency Synthesis can be traced back to WW II. The driving force was the desire for stable, rapidly switchable and accurate frequency control technology to meet the demands of narrowband, frequency-agile HF communications systems without resorting to large banks of switched crystals. Early synthesizers were cumbersome and expensive, and therefore their use was limited to the most sophisticated communications systems. With the help of the same technologies that have taken computers from "rooms" to the palms of our hands, the role of frequency synthesis has gone far beyond its original purpose and has become one of the most enabling technologies in modern communications equipment.

Just about every communications device manufactured today, be it a handheld transceiver, cell phone, pager, AM/FM entertainment radio, scanner, television, HF communications equipment, or test equipment contains a synthesizer. Synthesis is the technology that allows an easy interface with both computers and microprocessors. It provides amateurs with many desirable features, such as the feel of an analog knob with 10-Hz frequency increments, accuracy and stability determined by a single precision crystal oscillator, frequency memories, and continuously variable splits. Now reduced in size to only a few small integrated circuits, frequency synthesizers have also replaced the cumbersome chains of frequency multipliers and filters in VHF, UHF and microwave equipment, giving rise to many of the highly portable communications devices we use today. Frequency synthesis has also had a major impact in lowering the cost of modern equipment, as well as reducing manufacturing complexity.

Frequency synthesizers have been categorized in two general types, *direct synthesizers* (not to be confused with direct digital synthesis) and *indirect synthesizers*. The architecture of the direct types consists mainly of multipliers, dividers, mixers, filters and copious amounts of shielding. They are also cumbersome and expensive, and have all but disappeared from the market. The indirect varieties make use of programmable dividers and phase-lock loops, thereby considerably reducing their complexity. There are a number of variations on the indirect theme. They include programmable division, variable or dual modulus division, and fractional N. Other techniques that have been developed since the direct/indirect classification are direct digital synthesis (DDS) and rate multiplier synthesis. Synthesizers used in equipment today are usually a hybrid of the indirect technique and some of the latter developments. There are entire textbooks devoted to treating these various methods in depth, and as such it is not practical to discuss each one in detail in this handbook. We will focus on the indirect approach, based on phase-locked loops. This approach came to dominate frequency synthesis long ago, and is still dominant today.

Phase-Locked Loops

To understand the indirect synthesizer, we need to understand the phase-lock loop. The principle of the *phase-locked loop* (*PLL*) synthesizer is very simple. An oscillator can be built to cover the required frequency range, so what is needed is a system to keep its tuning correct. This is done by continuously comparing the phase of the oscillator to a stable reference, such as a crystal oscillator and using the result to steer the tuning. If the oscillator frequency is too high, the phase of its output will start to lead that of the reference by an increasing amount. This is detected and is used to decrease the frequency of the oscillator. Too low a frequency causes an increasing phase lag, which is detected and used to increase the frequency of the oscillator. In this way, the oscillator is locked into a fixed-phase relationship with the reference, which means that their frequencies are held exactly equal.

The oscillator is now under control, but is locked to a fixed frequency. **Fig 10.37** shows the next step. The phase detector does not simply compare the oscillator frequency with the reference oscillator; both signals have now been passed through frequency dividers. Advances in digital integrated circuits have made frequency dividers up to microwave frequencies (over 10 GHz) commonplace. The divider on the reference path has a fixed division factor, but that in the VCO path is programmable, the factor being entered digitally as a (usually) binary word. The phase detector is now operating at a lower frequency, which is a submultiple of both the reference and output frequencies.

The phase detector, via the loop amplifier, steers the oscillator to keep both its inputs equal in frequency. The reference frequency, divided by M, is equal to the output frequency divided by N. The output frequency equals the reference frequency × N/M. N is programmable (and an integer), so this synthesizer is capable of tuning in steps of F_{ref}/M.

As an example, to make a 2-m radio covering 144 to 148 MHz with a 10.7-MHz IF, we need a local oscillator covering 154.7-158.7 MHz. If the channel spacing is 20 kHz, then F_{ref}/M is 20 kHz, so N is 7735 to 7935. There is a free choice of F_{ref} or M, but division by a round binary number is easiest. Crystals are readily available for 10.24 MHz, and one of these (divided by 512) will give 20 kHz at low cost. ICs are readily available containing most of the circuitry necessary for the reference oscillator (less the crystal), with the programmable divider and the phase detector.

The phase detector in this example compares the relative timing of 5-V pulses (CMOS logic level) at 20 kHz. Inevitably, the phase detector output will contain strong components at 20 kHz and its harmonics in addition to the wanted "steering" signal. The loop filter must block these unwanted signals; otherwise the loop amplifier will amplify them and apply them to the VCO, generating unwanted FM. No filtering can be perfect and the VCO is very sensitive, so most synthesizers have measurable sidebands spaced at the phase detector operating frequency (and harmonics) away from the carrier. These are called *reference-frequency sidebands*. (Exactly which is the reference frequency is ambiguous: Do we mean the frequency of the reference oscillator, or the frequency applied to the reference input of the phase detector? You must look carefully at context whenever *reference frequency* is mentioned.)

The loop filter is not usually built as a single block of circuitry, it is often made up of three areas: Some components directly after the phase detector, a shaped frequency response in the loop amplifier and some more components between the loop amplifier and the VCO.

The PLL is like a feedback amplifier, although the "signal" around its loop is represented by frequency in some places, by phase in others and by voltage in others. Like any feedback amplifier, there is the risk of instability and oscillations traveling around the loop, which can be seen as massive FM on the output and a strong ac signal on the tuning line to the VCO. The loop's filtering and gain has to be designed to prevent this.

The following example illustrates the loop action: Imagine that we shift the reference oscillator by 1 Hz. The reference applied to the phase detector would shift 1/M Hz, and the loop would respond by shifting the output frequency to produce a matching shift at the other phase detector port, so the output is shifted by N/M Hz. Imagine now that we apply a very small amount of FM to the reference oscillator. The amount of deviation will be amplified by N/M—but this is only true for low modulating frequencies. If the modulating frequency is increased, eventually the loop filter starts to reduce the gain around the loop, the loop ceases to track the modulation and the deviation at the output falls. This is referred to as the *closed-loop frequency response* or *closed-loop bandwidth*. Poorly designed loops can have poor closed-loop responses. For example, the gain can have large peaks above N/M at some reference modulation frequencies, indicating marginal stability and excessively amplifying any noise at those off-

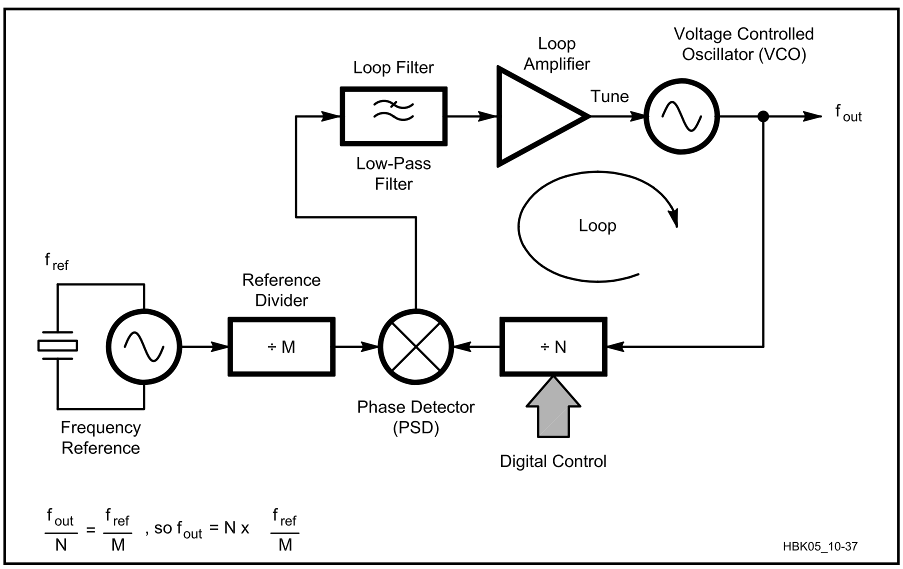


Fig 10.37—A basic phase-locked-loop (PLL) synthesizer acts to keep the divided-down signal from its voltage-controlled oscillator (VCO) phase-locked to the divided-down signal from its reference oscillator. Fine tuning steps are therefore possible without the complication of direct synthesis.

sets from the carrier. The design of the loop filtering and the gain around the loop sets the closed-loop performance.

In a single-loop synthesizer, there is a trade-off between how fine the step size can be versus the performance of the loop. A loop with a very fine step size runs the phase detector at a very low frequency, so the loop bandwidth has to be kept very, very low to keep the reference-frequency sidebands low. Also, the reference oscillator usually has much better phase-noise performance than the VCO, and (within the loop bandwidth) the loop acts to oppose the low-frequency components of the VCO phase-noise sidebands. This very useful cleanup activity is lost when loops have to be narrowed in bandwidth to allow narrow steps. Low-bandwidth loops are slow to respond—they exhibit overly long *settling times*—to the changes in N necessary to change channels. Absolutely everything seems to become impossible in any attempt to get fine resolution from a phase-locked loop.

Single-loop synthesizers are okay for the 2-m FM example for a number of reasons: Channel spacings in this band are not too small, FM is not as critical of phase noise as other modes, 2-m FM is rarely used for weak-signal DXing and the channelization involved does away with the desirability of simulating fast, smooth, continuous tuning. None of these excuses apply to MF/HF radios, however, and ways to circumvent the problems are needed.

A clue was given earlier, by carefully referring to *single*-loop synthesizers. It's possible to use *several* PLLs, or for that matter, some of the other mentioned techniques such as DDS, fractional N, variable- or dual-modulus etc, as components in a larger structure, dividing the frequency of one loop and adding it to the frequency of another that has not been divided. This represents a form of hybrid between the old direct synthesizer and the PLL. A form of PLL containing a mixer in place of the programmable divider can be used to perform the frequency addition.

Let us continue our discussion of phase-lock loop synthesizers by examining the role of each of the component pieces of the system. They are, the VCO, the dividers, prescalers, the phase detector and the loop-compensation amplifier.

Voltage-Controlled Oscillators

Voltage-controlled oscillators are commonly referred to as *VCOs*. (Voltage-tuned oscillator, VTO, more accurately describes the circuitry most commonly used in VCOs, but tradition is tradition! An exception to this is the YIG oscillator, described earlier as sometimes used in UHF and microwave PLLs, which is *current*-tuned.) In all the oscillators described so far, except for permeability tuning, the frequency is controlled or trimmed by a variable capacitor. These are modified for voltage tuning by using a tuning diode. The oscillator section of this chapter shows the schematic symbol and construction detail of a voltage-controlled or *varactor* (tuning) diode as is typically used in a VCO (see Fig 10.18).

When this oscillator is used in a PLL, the PLL adjusts the tuning voltage to completely cancel any drift in the oscillator (provided it does not drift further than it can be tuned) regardless of its cause, so no special care is needed to compensate a PLL VCO for temperature effects. Adjusting the inductor core will not change the frequency; the PLL will adjust the tuning voltage to hold the oscillator on frequency. This adjustment is important, though. It's set so that the tuning voltage neither gets too close to the maximum available, nor too low for acceptable Q, as the PLL is tuned across its full range.

In a high-performance, low noise synthesizer, the VCO may be replaced by a bank of switched VCOs, each covering a section of the total range needed, each VCO having better Q and lower noise because the lossy diodes constitute only a smaller part of the total tank capacity than they would in a single, full-range oscillator. Another method uses tuning diodes for only part of the tuning range and switches in other capacitors (or inductor taps) to extend the range. The performance of wide-range VCOs can be improved by using a large number of varactors in parallel in a high-C, low-L tank.

Programmable Dividers

Designing your own programmable divider is now a thing of the past, as complete systems have been made as single ICs for over 10 years. The only remaining reason to do so is when an ultra-low-noise synthesizer is being designed.

It may seem strange to think of a digital counter as a source of random noise, but digital circuits have propagation delays caused by analog effects, such as the charging of stray capacitances. Digital circuits are composed of the same sorts of components as analog circuits; there are no such things as digital and analog electrons! Signals are made of voltages and currents. The prime difference between analog and digital electronics lies in the different ways meaning is assigned to the magnitude of a signal. The differences in circuit design are consequences of this, not causes.

Thermal noise in the components of a digital circuit, added to the signal voltage will slightly change the times at which thresholds are crossed. As a result, the output of a digital circuit will have picked up some timing *jitter*. Jitter in one or both of the signals applied to a phase detector can be viewed as low-level random phase modulation. Within the loop bandwidth, the phase detector steers the VCO so as to oppose and cancel it, so phase jitter or noise is applied to the VCO in order to cancel out the jitter added by the divider. This makes the VCO noisier. The noise performance of the high-speed CMOS logic normally used in programmable dividers is good. For ultra-low-noise dividers, ECL devices are sometimes used.

As shown in **Fig 10.38A**, a programmable divider consists of a loadable counter and some control circuitry. The counter is designed to count downward. The programmed division factor is loaded into the counter, and the incoming frequency clocks the downward counting. When the count reaches zero, the counter is quickly reloaded with the division factor before the next clock edge. The maximum frequency of operation is limited by the minimum time needed to reliably perform the loading operation. One improvement is to have a circuit to recognize a state a few clock cycles before the zero count is reached, so the reload can be performed at a more leisurely pace during those cycles. This imposes a minimum division ratio, but that is rarely a problem. An output pulse is given during the loading cycle. In this way a frequency is divided by the number loaded.

The maximum input frequency that can be handled is set by the speed of available logic. The synchronous reset cycle forces the entire divider to be made of equally fast logic. Just adding a fast *prescaler*—a fixed-ratio divider ahead of the programmable one—will increase the maximum frequency that can be used, but it also scales up the loop step size. Equal division has to be added to the reference input of the phase detector to restore the step size, reducing the phase detector's operating frequency. The loop bandwidth has to be reduced to restore the filtering needed to suppress reference frequency sidebands. This makes the loop much slower to change frequency and degrades the noise performance.

Variable or Dual-Modulus Prescalers

A plain programmable divider for VHF use would need to be built from ECL devices. It would be expensive, hot and power-hungry. The idea of a fast front end

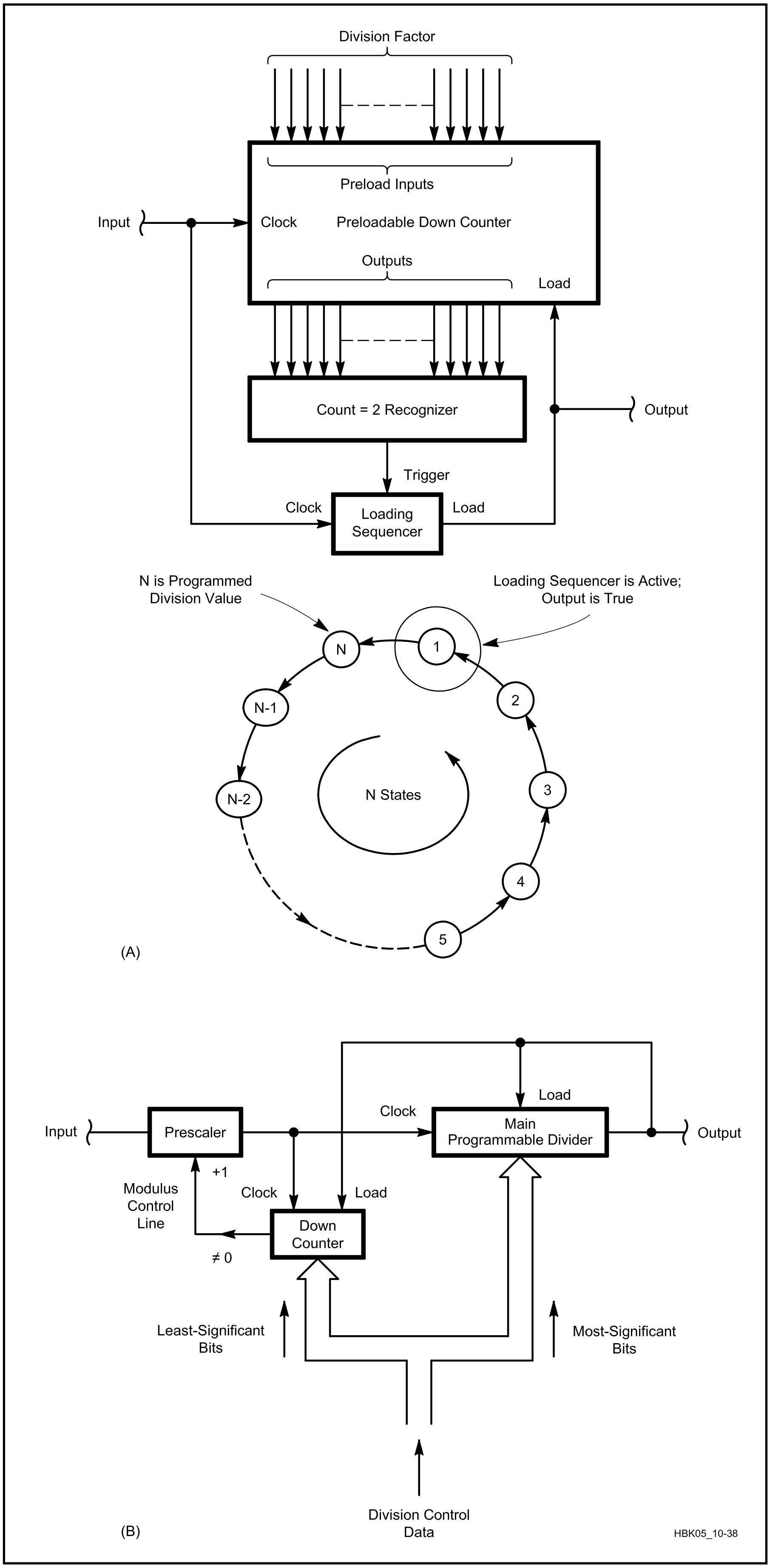


Fig 10.38—(A) Shows the mechanism of a programmable frequency divider. (B) shows the function of a dual-modulus prescaler. The counter is reloaded with N when the count reaches 0 or 1, depending on the sequencer action.

ahead of a CMOS programmable divider would be perfect if these problems could be circumvented. Consider a fast divide-by-ten prescaler ahead of a programmable divider where division by 947 is required. If the main divider is set to divide by 94, the overall division ratio is 940. The prescaler goes through its cycle 94 times and the main divider goes through its cycle once, for every output pulse. If the prescaler is changed to divide by 11 for 7 of its cycles for every cycle of the entire divider system, the overall division ratio is now $[(7 \times 11) + (87 \times 10)] = 947$. At the cost of a more elaborate prescaler and the addition of a slower programmable counter to control it, this prescaler does not multiply the step size and avoids all the problems of fixed prescaling.

Fig 10.38B shows the general block diagram of a dual-modulus prescaled divider. The down counter controls the modulus of the prescaler. The numerical example, just given, used decimal arithmetic, although binary is now usual. Each cycle of the system begins with the last output pulse having loaded the frequency control word, into both the main divider and the prescaler controller. If the part of the word loaded into the prescaler controller is not zero, the prescaler is set to divide by 1 greater than its normal ratio. Each cycle of the prescaler clocks the down counter. Eventually, it reaches zero, and two things happen: The counter is designed to freeze at zero (and it will remain frozen until it is next reloaded) and the prescaler is switched back to its normal ratio, at which it will remain until the next reload. One way of visualizing this is to think of the prescaler as just being a divider of its normal ratio, but with the ability to "steal" a number of input pulses controlled by the data loaded into its companion down counter. Note that a dual-modulus prescaler system has a *minimum* division ratio, needed to ensure there are enough cycles of the prescaler to allow enough input pulses to be stolen.

Dual-modulus prescaler ICs are widely used and widely available. Devices for use to a few hundred megahertz are cheap, and devices in the 2.5-GHz region are commonly available. Common prescaler IC division ratio pairs are: 8-9, 10-11, 16-17, 32-33, 64-65 and so on. Many ICs containing programmable dividers are available in versions with and without built-in prescaler controllers.

Phase Detectors

A phase detector (PSD) produces an output voltage that depends on the phase relationship between its two input signals. If two signals, *in phase on exactly the same frequency*, are mixed together in a conventional diode-ring mixer with a dc-coupled output port, one of the products is direct current (0 Hertz). If the phase relationship between the signals changes, the mixer's dc output voltage changes. With both signals in phase, the output is at its most positive; with the signals 180° out of phase, the output is at its most negative. When the phase difference is 90° (the signals are said to be *in quadrature*), the output is 0 V.

Applying sinusoidal signals to a phase detector causes the detector's output voltage to vary sinusoidally with phase angle, as in **Fig 10.39A**. This nonlinearity is not a problem, as the loop is usually arranged to run with phase differences close to 90°. What might seem to be a more serious complication is that the detector's phase-voltage characteristic repeats every 180°, not 360°. Two possible input phase differences can therefore produce a given output voltage. This turns out not to be a problem, because the two identical voltage points lie on opposing slopes on the detector's output-voltage curve. One direction of slope gives positive feedback (making the loop unstable and driving the VCO away from what otherwise would be the lock angle) over to the other slope, to the true and stable lock condition.

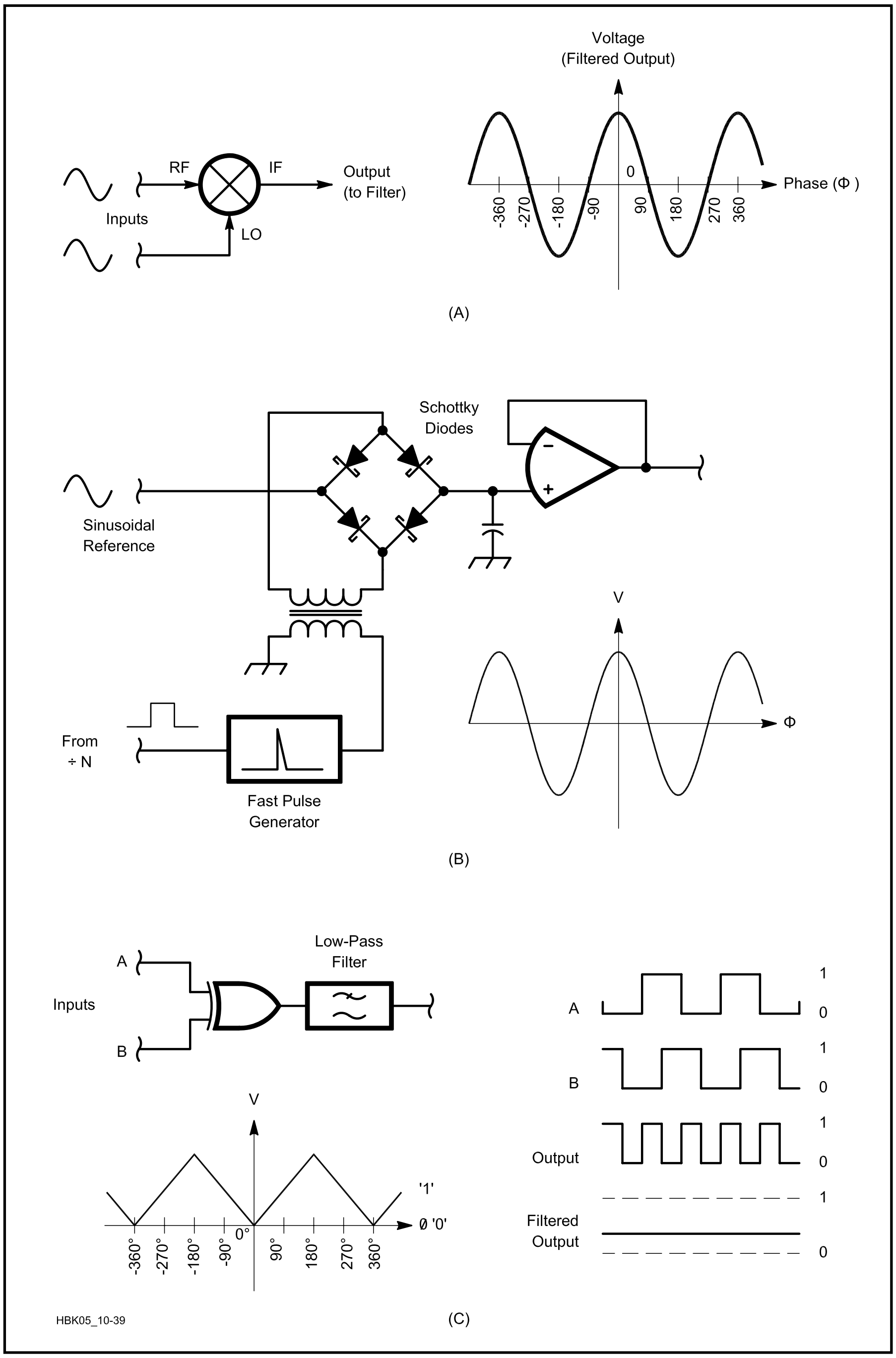


Fig 10.39—Simple phase detectors: a mixer (A), a sampler (B) and an exclusive-OR gate (C).

MD108, SBL-1 and other diode mixers can be used as phase detectors, as can active mixers like the MC1496. Mixer manufacturers make some parts (Mini-Circuits RPD-1 and so on) that are specially optimized for phase-detector service. All these devices can make excellent, low-noise phase detectors but are not commonly used in ham equipment. A high-speed sample-and-hold circuit, based on a Schottky-diode bridge, can form a very low-noise phase detector and is sometimes used in specialized instruments. This is just a variant on the basic mixer; it produces a similar result, as shown in Fig 10.39B.

The most commonly used simple phase detector is just a single exclusive-OR (XOR) logic gate. This circuit gives a logic 1 output only when *one* of its inputs is at logic 1; if both inputs are the same, the output is a logic 0. If inputs A and B in Fig 10.39C are almost in phase, the output will be low most of the time, and its average filtered value will be close to the logic 0 level. If A and B are almost in opposite phase, the output will be high most of the time, and its average voltage will be close to logic 1. This circuit is very similar to the mixer. In fact, the internal circuit of ECL XOR gates is the same transistor tree as found in the MC1496 and similar mixers, with some added level shifting. Like the other simple phase detectors, it produces a cyclic output, but because of the square-wave input signals, produces a triangular output signal. To achieve this circuit's full output-voltage range, it's important that the reference and VCO signals applied operate at a 50% duty cycle.

Phase-Frequency Detectors

All the simple phase detectors described so far are really specialized mixers. If the loop is out of lock and the VCO is far off frequency, such phase detectors give a high-frequency output midway between zero and maximum. This provides no information to steer the VCO towards lock, so the loop remains unlocked. Various solutions to this problem, such as crude relaxation oscillators that start up to sweep the VCO tuning until lock is acquired, or laboriously adjusted "pretune" systems in which a DAC, driven from the divider control data, is used to coarse tune the VCO to within locking range, have been used in the past. Many of these solutions have been superseded, although pretune systems are still used in synthesizers that must change frequency very rapidly.

The *phase-frequency detector* is the usual solution to lock-acquisition problems. It behaves like a simple phase detector over an extended phase range, but its characteristic is not repetitive. Because its output voltage stays high or low, depending on which input is higher in frequency, this PSD can steer a loop towards lock from anywhere in its tuning range. **Fig 10.40** shows the internal logic of the phase-frequency detector in the CD4046 PLL chip. (The CD4046 also contains an XOR PSD). When the phase of one input leads that of the other, one of the output MOSFETs is pulsed on with a duty cycle proportional to the phase difference. Which MOSFET receives drive depends on which input is leading. If both inputs are in phase, the output will include small pulses due to noise, but their effects on the average output voltage will cancel. If one signal is at a higher frequency than the other, its phase will lead by an increasing amount, and the detector's output will be held close to either V_{DD} or ground, depending on which input signal is higher in frequency. To get a usable voltage output, the MOSFET outputs can be terminated in a high-value resistor to $V_{DD}/2$, but the CD4046's output stage was really designed to drive current pulses into a capacitive load, with pulses of one polarity charging the capacitor and pulses of the other discharging it. The capacitor integrates the pulses. Simple phase detectors are normally used to lock their inputs in quadrature, but phase-frequency detectors are used to lock their input signals in phase.

Traditional textbooks and logic-design courses give extensive coverage to avoiding *race hazards* caused by near-simultaneous signals racing through parallel paths in a structure to control a single output. Avoiding such situations is important in many circuits because the outcome is strongly dependent on slight differences in gate speed. The phase-frequency detector is the one circuit whose entire function depends on its built-in ability to *make* both its inputs race. Consequently, it's not easy to home-brew a phase-frequency detector from ordinary logic parts. Fortunately, they are available in IC form, usually combined with other functions. The MC4044 is an old stand-alone TTL phase-frequency detector; the MC12040 is an ECL derivative, and the much faster and compatible MCH 12140 can be used to over 600 MHz. The Hittite HMC403S8G can be used to 1.3 GHz. CMOS versions can be found in the CD4046 PLL chip and in almost all current divider-PLL synthesizer chips.

The race-hazard tendency does cause one problem in phase-frequency detectors: A device's delays and noise, rather than its input signals, control its phase-voltage characteristic in the zero-phase-difference region. This degrades the loop's noise performance and makes its phase-to-voltage coefficient uncertain and variable in a small region, a "dead zone"—unfortunately, in the detector's normal operating range! It's therefore normal to bias operation slightly away from *exact* phase equality to avoid this problem. Fortunately, the newer, faster phase detectors like the MCH 12140 and HMC403S8G tend to minimize this "dead zone" problem.

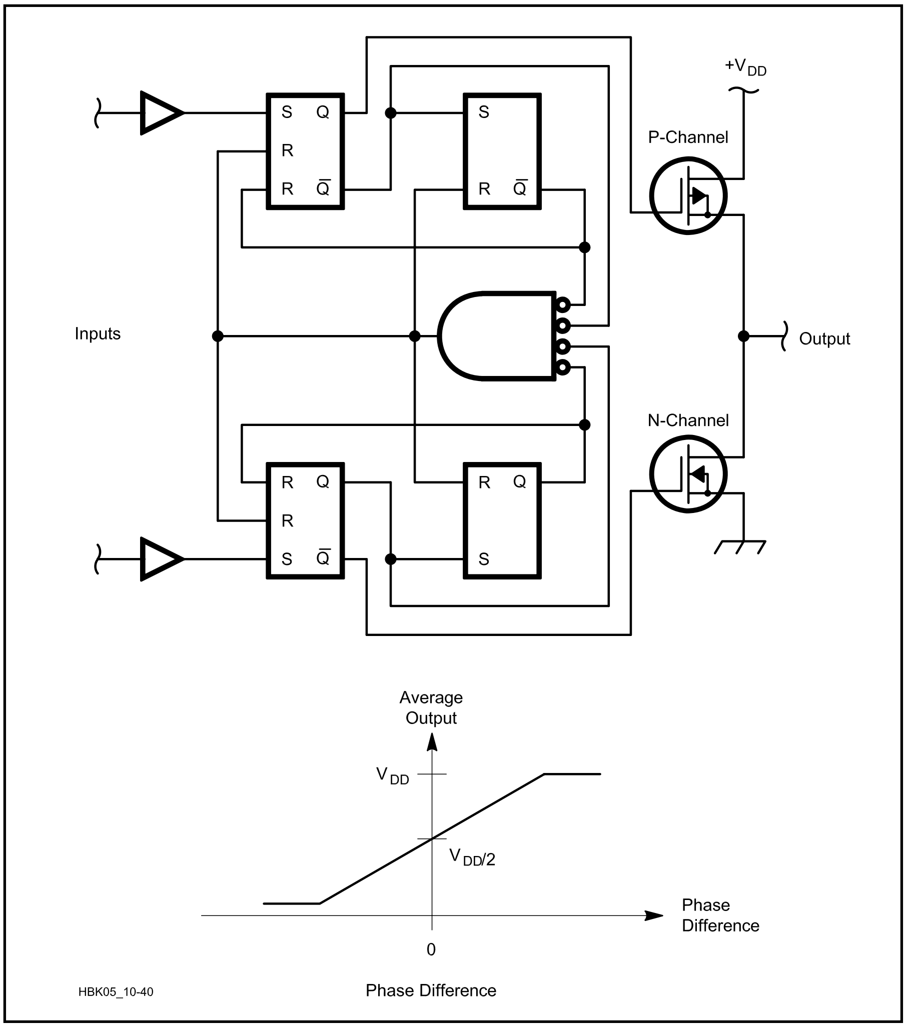


Fig 10.40—Input signals very far off frequency can confuse a simple phase-detector; a phase-frequency detector solves this problem.

The Loop Compensation Amplifier

Phase-locked loops have acquired a troublesome reputation for a variety of reasons. In the past, a number of commercially made radios have included poorly designed synthesizers that produced excessive noise sidebands, or wouldn't lock reliably. Some un-producible designs have been published for home construction that could not be made to work. Many experimenters have toiled over an unstable loop, desperately trying anything to get it to lock stably. So the PLL has earned a shady reputation. Because of all of this aversion therapy, very few amateurs are now prepared to attempt to build a PLL.

Luckily, things are not as bad as they used to be. The proliferation of synthesis and PLLs in contemporary equipment have lead to a number of excellent integrated circuits, as well as fine application notes to support them.

The **first** task is to take action to ensure that all previously discussed components of the loop are functioning properly and stably before attempting to close the loop. This means that the oscillators, dividers and amplifiers must all be unconditionally stable. Any instability or squegging in these components must be dealt with prior to attempting to close the loop.

The **second** task is to produce a closed-loop characteristic that best suits the application. Herein is where the greatest difficulty frequently lies. The selection of the closed-loop bandwidth and phase margin or peaking is usually one of great compromise and thought. The process frequently begins with an educated approximation that is then evaluated and modified. There are many factors to consider, including the degree of reference frequency suppression, switching speed, noise, microphonics and modulation, if any. Once the optimum loop characteristic is established, the next problem is to maintain it with respect to variations of the division resulting from the synthesis, and also gain variations associated with the nonlinear tuning curve of the VCO. While not always required in amateur applications, loops in sophisticated synthesizers frequently employ programmable multiplying DACs to maintain constant loop bandwidths and phase margin. In an example later in this chapter, we will actually describe the process for establishing the loop bandwidth for a simple synthesizer.

The **third** task is to design a loop compensation amplifier (filter) that will ensure stable operation when the loop is closed. The design of this compensation network or filter requires knowledge of the VCO gain and linearity, the phase-detector gain, and the effective division ratio and its variation in the loop. The foregoing requirements imply the necessity for measurement and calculation to have any reasonable hope of producing a successful outcome.

Note carefully that we are designing a *loop*. Trial and error, intuitive component choice, or "reusing" a loop amplifier design from a different synthesizer may lead to a low probability of success. All oscillators are loops, and any loop will oscillate when certain conditions are met. Look again at the RC phase-shift oscillator from which the "Pierce" crystal oscillator is derived (Fig 10.29A). Our PLL is a loop, containing an amplifier and a number of RC sections. It has all the parts needed to make an oscillator. If a loop is unstable and oscillates, there will be an oscillatory voltage superimposed on the VCO tuning voltage. This will produce massive unwanted FM on the output of the PLL, which is absolutely undesirable.

Before you turn to a less demanding chapter, consider this: It's one of nature's jokes that easy procedures often hide behind a terror-inducing facade. The math you need to design a good PLL may look weird when you see it for the first time and seem to involve some obviously impossible concepts, but it ends in a very simple procedure that allows you to calculate the response and stability of a loop. Professional designers must handle these things daily, and because they like differential equations no more than anyone else does, they use a graphical method based on the work of the mathematician, Laplace, to generate PLL designs. (We'll leave proofs in the textbooks, where they belong! Incidentally, if you can get the math required for PLL design under control, as a free gift you will also be ready to do modern filter synthesis and design beam antennas, since the math they require is essentially the same.)

The easy solutions to PLL problems can only be performed during the design phase. We must deliberately design loops with sufficient stability safety margins so that all foreseeable variations in component tolerances, from part-to-part, over temperature and across the tuning range, cannot take the loop anywhere near the threshold of oscillation. More than this, it is important to have adequate stability margin because loops with lesser margins will exhibit amplified phase-noise sidebands, and that is another PLL problem that can be designed out.

At the beginning of this chapter (following the text cite of Fig 10.4A), the criterion for stability/oscillation of a loop was mentioned: Barkhausen's for oscillation or Nyquist's for stability. This time Harry Nyquist is our hero.

A PLL is a negative feedback system, with the feedback opposing the input, this opposition or inversion around the loop amounts to a 180° phase shift. This is the frequency-*independent* phase shift round the loop, but there are also frequency-*dependent* shifts that will add in. These will inevitably give an increasing, lagging phase with increasing frequency. Eventually an extra 180° phase shift will have occurred, giving a total of 360° around the loop at some frequency. We are in trouble if the gain around the loop has not dropped below unity (0 dB) by this frequency.

Note that we are not concerned only with the loop operating frequency. Consider a 30-MHz low-pass filter passing a 21-MHz signal. Just because there is no signal at 30 MHz, our filter is no less a 30-MHz circuit. Here we use the concept of frequency to describe "what would happen *if* a signal was applied at that frequency."

The next sections show us how to use the graphical concepts of poles and zeros to calculate the gain and phase response around a loop. The graphical concepts are quite useful in that they provide insight into the effect of various component choices on the loop performance.

Poles and Zeros

Let's define some terms first. We all think of frequency as stretching from zero to infinity, but let's imagine that additional numbers could be used. This is pure imagination—we can't make signals at *complex* values (with *real* and *imaginary* components) of hertz—but if we try using complex numbers for frequency, some "impossible" things happen. For example, take a simple RC low-pass network shown in **Fig 10.41**. The frequency response of such a network is well known, but its phase response is *not* so well known. The equation shown for the *transfer function*, the "gain" of this simple circuit, is pretty standard, although the *j* is included to make it complete and accurately describe the output in phase as well as magnitude (*j* is an impossible number, the square root of –1, usually called an *imaginary* number and used to represent a 90° phase shift).

Although the equation shown in Fig 10.41 was constructed to show the behavior of the circuit at real-world frequencies, there is nothing to stop us from using it to explore how the circuit would behave at impossible frequencies, just out of curiosity. At one such frequency, the circuit goes crazy. At $f = -j\,2\pi RC$, the denominator of the equation is zero, and the gain is infinite! Infinite gain is a pretty amazing thing to achieve with a passive circuit—but because this can only happen at an impossible frequency, it cannot happen in the real world. This frequency is equal to the network's –3 dB frequency multiplied by *j; it* is called a *pole*. Other circuits, which produce real amplitude responses that start to *rise* with frequency, act similarly, but their crazy effect is called a *zero*—an imaginary frequency at which gain goes to zero. The frequency of a circuit's zero is again *j* times its 3-dB frequency (+3 dB in this case).

These things are clearly impossible, but we can map all the poles and zeros of a complex circuit and look at the patterns to determine if the circuit possesses unwanted gain and/or phase bumps across a frequency range of interest. Poles are associated with 3-dB roll-off frequencies and 45° phase lags, zeros are associated with 3-dB "roll-up" frequencies and 45° phase leads. We can plot the poles and zeros on a two-dimensional chart of real and imaginary frequency. (The names of this chart and its

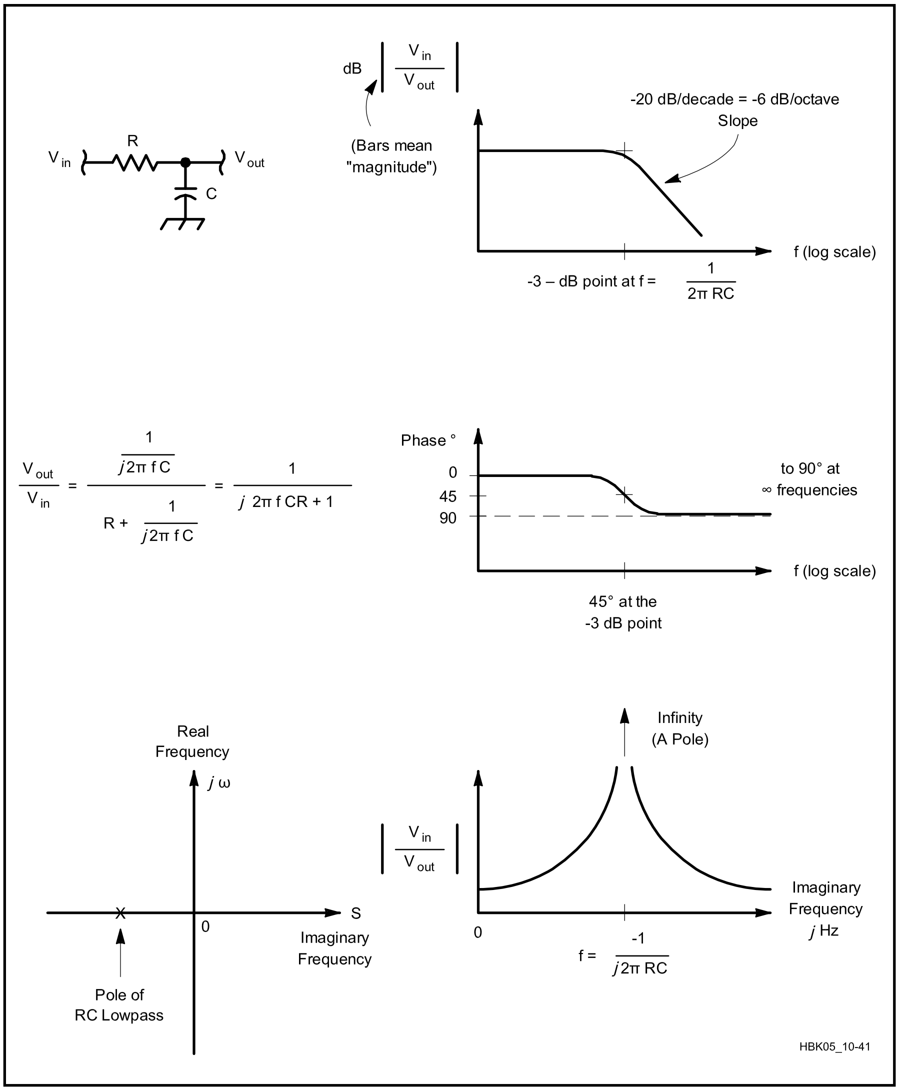


Fig 10.41—A simple RC filter is a "pole."

axes are results of its origins in *Laplace transforms*, which we don't need to touch in this discussion.) The traditional names for the axes are used, but we can add labels with clearer meanings.

What is a Pole?—A pole is associated with a bend in a frequency response plot where attenuation with increasing frequency increases by 6 dB per octave (20 dB per decade; an octave is a 2:1 frequency ratio, a decade is a 10:1 frequency ratio).

There are four ways to identify the existence and frequency of a pole:

1. A downward bend in a gain vs frequency plot, the pole is at the –3 dB point for a single pole. If the bend is more than 6 dB/octave, there must be multiple poles at this frequency.

2. A 90° change in a phase vs frequency plot, where lag increases with frequency. The pole is at the point of 45° added lag. Multiple poles will add their lags, as above.

3. On a circuit diagram, a single pole looks like a simple RC low-pass filter. The pole is at the –3 dB frequency (1/(2πRC) Hz). Any other circuit that gives the same response will produce a pole at the same frequency.

4. In an equation for the transfer function of a circuit, a pole is a theoretical value of frequency, which would make the equation predict infinite gain. This is clearly impossible, but as the value of frequency will either be absolute zero, or will have an imaginary component, it is impossible to make a signal at a pole frequency.

What is a Zero?—A zero is the complement of a pole. Each zero is associated with an upward bend of 6 dB per octave in a gain vs frequency plot. The zero is at the +3 dB point. Each zero is associated with a transition on a phase-versus-frequency plot that *reduces* the lag by 90°. The zero is at the 45° point of this S-shaped transition.

In math, it is a frequency at which the transfer-function equation of a circuit predicts zero gain. This is not impossible in real life (unlike the pole), so zeros can be found at real-number frequencies as well as complex-number frequencies.

On a circuit diagram, a pure zero would need gain that increases with frequency forevermore above the zero frequency. This implies active circuitry that would inevitably run out of gain at some frequency, which implies one or more poles up there. In real-world circuits, zeros are usually not found making gain go up, but rather in conjunction with a pole, giving a gain slope between two frequencies and flat gain beyond them. Real-world zeros are only found chaperoned by a greater or equal number of poles.

Consider a classic RC high-pass filter. The gain increases at 6 dB per octave from 0 Hz (so there must be a zero at 0 Hz) and then levels off at 1/(2πRC) Hz. This leveling off is really a pole, it adds a 6 dB per octave roll off to cancel the roll-up of the zero.

Poles and Zeros in the Loop Amplifier—The loop-amplifier circuit used in the example loop has a blocking capacitor in its feedback path. This means there is no dc feedback. At higher frequencies the reactance of this capacitor falls, increasing the feedback and so reducing the gain. This is an *integrator*. The gain is immense at 0 Hz (dc) and falls at 6 dB per octave. This points to a pole at 0 Hz. This is true whatever the value of the series resistance feeding the signal into the amplifier inverting input and whatever the value of the feedback capacitor. The values of these components scale the gain rather than shape it. The shape of the integrator gain is fixed at –6 dB per octave, but these components allow us to move it to achieve some wanted gain at some wanted frequency.

At high frequencies, the feedback capacitor in an ordinary integrator will have very low reactance, giving the circuit very low gain. In our loop amplifier, there is a resistor in series with the capacitor. This limits how low the gain can go, in other words the integrator's downward (with increasing frequency) slope is leveled off. This resistor and capacitor make a zero, and its +6 dB per octave slope cancels the –6 dB per octave slope created by the pole at 0 Hz.

Recipe for A PLL Pole-Zero Diagram

We can choose a well-tried loop-amplifier circuit and calculate values for

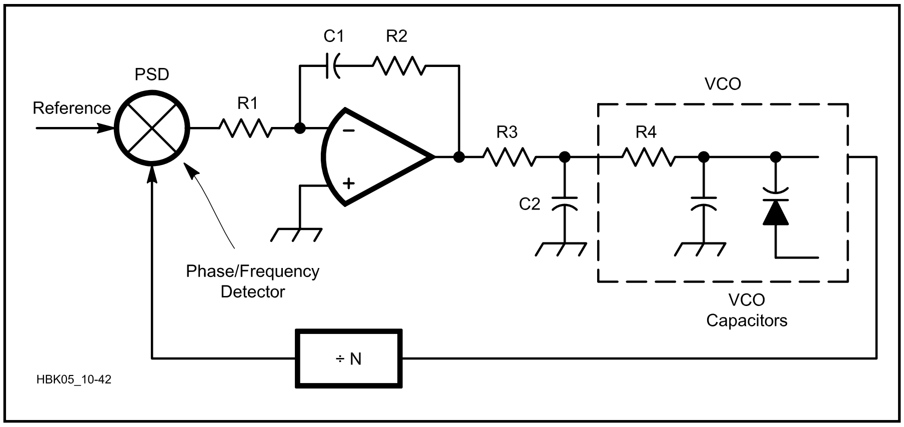


Fig 10.42—A common loop-amplifier/filter arrangement.

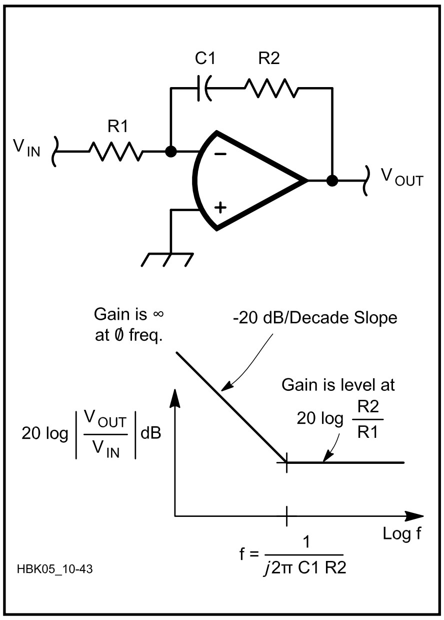


Fig 10.43—Loop-amplifier detail.

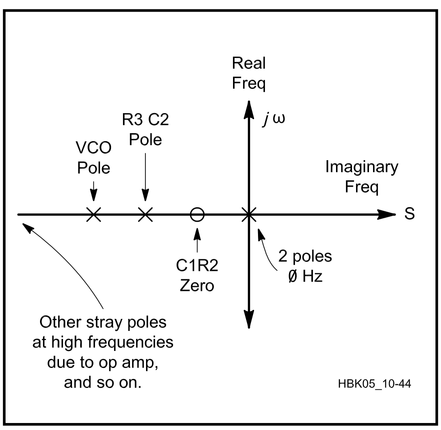


Fig 10.44—The loop's resulting pole-zero "constellation."

its components to make it suit our loop. **Fig 10.42** shows a common synthesizer loop arrangement. The op amp operates as an integrator, with R2 added to level off its falling gain. An integrator converts a dc voltage at its input to a ramping voltage at its output. Greater input voltages yield faster ramps. Reversing the input polarity reverses the direction of the ramp.

The system's phase-frequency detector (connected to work in the right sense) steers the integrator to ramp in the direction that tunes the VCO toward lock. As the VCO approaches lock, the phase detector reduces the drive to the integrator, and the ramping output slows and settles on the right voltage to give the exact, locked output frequency. Once lock is achieved, the phase detector outputs short pulses that "nudge" the integrator to keep the divided VCO frequency exactly locked in phase with the reference. Now we'll take a look around the entire loop and find the poles and zeros of the circuits.

The integrator produces a –20 dB per decade roll off from 0 Hz, so it has a pole at 0 Hz. It also includes R2 to cancel this slope, which is the same as adding a rising slope that exactly offsets the falling one. This implies a zero at $f = 1/(2\pi C1R2)$ Hz, as **Fig 10.43** shows. R3 and C2 make another pole at $f = 1/(2\pi C2R3)$ Hz. A VCO usually includes a series resistor that conveys the control voltage to the tuning diode, which also is loaded by various capacitors. This creates another pole.

The VCO generates frequency, while the phase detector responds to phase. Phase is the integral of frequency, so together they act as another integrator and add another pole at 0 Hz.

What About the Frequency Divider?—The frequency divider is like a simple attenuator in its effect on loop response. Imagine a signal generator that is switched between 10 and 11 MHz on alternate Tuesdays. Imagine that we divide its output by 10. The output of the divider will change between 1 and 1.1 MHz, still on alternate Tuesdays. The divider divides the deviation of the frequency (or phase) modulated signal, but it cannot affect the modulating frequency.

A possible test signal passing around the example loop is in the form of frequency modulation as it passes through the divider, so it is simply attenuated. A divide by N circuit will give $20 \log_{10}$ (N) dB of attenuation (reduction of loop gain). This completes the loop. We can plot all this information on one *s-plane*, **Fig 10.44**, a plot sometimes referred to as a PLL system's *pole-zero constellation*.

Open-Loop Gain and Phase

Just putting together the characteristics of the blocks around the loop, without allowing for the loop itself, gives us the system's *open-loop* characteristic, which is all we really need.

What Is Stability?—Here "open-loop response" means the gain around the loop that would be experienced by a signal at some frequency *if* such a signal were inserted. We do not insert such signals in the actual circuit, but the concept of frequency-dependent gain is still valid. We need to ensure that there is insufficient gain, at *all* frequencies, to ensure that the loop cannot create a signal and begin oscillating. More than this, we want a good safety margin to allow for component variations and because loops that are *close* to instability perform poorly.

The loop gain and phase are doing interesting things on our plots at frequencies in the AF part of the spectrum. This does not mean that there is visible operation at those frequencies. Imagine a bad, unstable loop. It will have a little too much loop gain at some frequency, and unavoidable noise will build up until a strong signal is created. The amplitude will increase until nonlinearities limit it. A 'scope view of the tuning voltage input to the VCO will show a big signal, often in the hertz to tens of kilohertz region, often large enough to drive the loop amplifier to the limits of its output swing, close to its supply voltages. As this big signal is applied to the tuning voltage input to the VCO, it will modulate the VCO frequency across a wide range. The output will look like that of a sweep generator on a spectrum analyzer. What is wrong? How can we fix this loop? Well, the problem may be excessive loop gain, or improper loop time constants. Rather than work directly with the loop time constants, it is far easier to work with the pole and zero frequencies of the loop response. It's

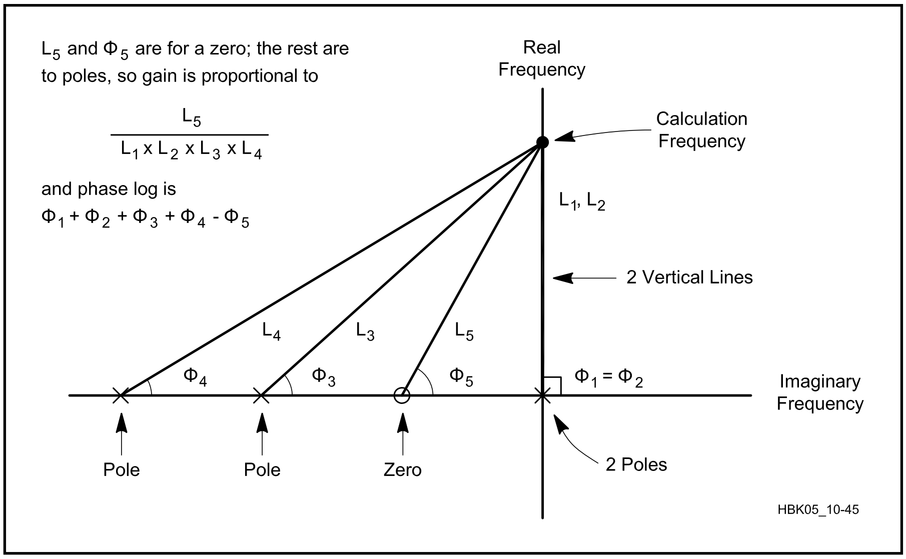


Fig 10.45—Calculating loop gain and phase characteristics from a pole-zero diagram.

just a different view of the same things.

Now imagine a good, stable loop, with an adequate stability safety margin. There will be some activity around the loop: the PSD (phase detector) will demodulate the phase noise of the VCO and feed the demodulated noise through the loop amplifier in such a way as to cancel the phase noise. This is how a good PLL should give *less* phase noise than the VCO alone would suggest. In a good loop, this noise will be such a small voltage that a 'scope will not show it. In fact, connecting test equipment in an attempt to measure it can usually *add* more noise than is normally there.

Finally, imagine a poor loop that is only just stable. With an inadequate safety margin, the action will be like that of a Q multiplier or a regenerative receiver close to the point of oscillation: There will be an amplified noise peak at some frequency. This spoils the effect of the phase detector trying to combat the VCO phase noise and gives the opposite effect. The output spectrum will show prominent bumps of exaggerated phase noise, as the excess noise frequency modulates the VCO.

Now, let's use the pole-zero diagram as a graphic tool to find the system open-loop gain and phase. As we do so, we need to keep in mind that the frequency we have been discussing in designing a loop response is the frequency of a theoretical test signal passing around the loop. In a real PLL, the loop signal exists *in two forms*: As sinusoidal voltage between the output of the phase detector and the input of the VCO, and as a sinusoidal modulation of the VCO frequency in the remainder of the loop.

With our loop's pole-zero diagram in hand, we can pick a frequency at which we want to know the system's open-loop gain and phase. We plot this value on the graph's vertical (Real Frequency) axis and draw lines between it and each of the poles and zeros, as shown in **Fig 10.45**. Next, we measure the lengths of the lines and the "angles of elevation" of the lines. The loop gain is proportional to the product of the lengths of all the lines to zeros *divided by* the product of all the lengths of the lines to the poles. The phase shift around the loop (lagging phase equates to a positive phase shift) is equal to the sum of the pole angles minus the sum of the zero angles. We can repeat this calculation for a number of different frequencies and draw graphs of the loop gain and phase versus frequency.

All the lines to poles and zeros are hypotenuses of right triangles, so we can use Pythagoras's rule and the tangents of angles to eliminate the need for scale drawings. Much tedious calculation is involved because we need to repeat the whole business for each point on our open-loop response plots. This much tedious calculation is an ideal application for a computer.

The procedure we've followed so far gives only *proportional* changes in loop gain, so we need to calculate the loop gain's *absolute* value at some (chosen to be easy) frequency and then relate everything to this. Let's choose 1 Hz as our reference. (Note that it's usual to express angles in radians, not degrees, in these calculations and that this normally renders frequency in peculiar units of *radians per second*. We can keep frequencies in hertz if we remember to include factors of 2π in the right places. A frequency of 1 Hz = 2π radians/second, because 2π radians = 360°.)

We must then calculate the loop's proportional gain at 1 Hz from the pole-zero diagram, so that the constant of proportion can be found. For starters, we need a reasonable estimate of the VCO's *voltage-to-frequency gain*—how much it changes frequency per unit change of tuning voltage. As we are primarily interested in stability, we can just take this number as the slope, in Hertz per volt, at the steepest part of voltage-versus-frequency tuning characteristic, which is usually at the low-frequency end of the VCO tuning range. (You can characterize a VCO's voltage-versus-frequency gain by varying the bias on its tuning diode with an adjustable power supply and measuring its tuning characteristic with a voltmeter and frequency counter.)

The loop divide-by-N stage divides our theoretical modulation—the tuning corrections provided through the phase detector,

$$\text{Open loop gain (dB)} = 20\log\left[\frac{\sqrt{P_1^2+1}\sqrt{P_2^2+1}\sqrt{P_3^2+1}\sqrt{P_4^2+1}\sqrt{Z^2+f^2}}{\sqrt{P_1^2+f^2}\sqrt{P_2^2+f^2}\sqrt{P_3^2+f^2}\sqrt{P_4^2+f^2}\sqrt{Z^2+1}} \times 10^{\frac{\text{unity freq gain}}{20}}\right]$$

$$\text{Phase (lead)} = \tan^{-1}\left(\frac{f}{Z}\right) - \tan^{-1}\left(\frac{f}{P_2}\right) - \tan^{-1}\left(\frac{f}{P_3}\right) - \tan^{-1}\left(\frac{f}{P_4}\right)$$

f is the frequency of the point to be characterized. P_1, P_2, P_3, P_4, Z are the frequencies of the poles and zero (all in the same units!)

Fig 10.46—Pole-zero frequency-response equations capable of handling up to four poles and one zero.

loop amplifier and the VCO tuning diode—by its programmed ratio. The worst case for stability occurs at the divider's lowest N value, where the divider's "attenuation" is least. The divider's voltage-versus-frequency gain is therefore 1/N, which, in decibels, equates to –20 log (N).

The change from frequency to phase has a voltage gain of one at the frequency of 1 radian/second, 1/(2π), which equates to –16 dB, at 1 Hz. The phase detector will have a specified "gain" in volts per radian. To finish off, we then calculate the gain of the loop amplifier, including its feedback network, at 1 Hz.

There is no need even to draw the pole-zero diagram. **Fig 10.46** gives the equations needed to compute the gain and phase of a system with up to four poles and one zero. They can be extended to more singularities and put into a simple computer program. Alternatively, you can type them into your favorite spreadsheet and get printed plots. The computational power needed is trivial, and a listing to run these equations on a Hewlett-Packard HP11C (or similar pocket calculator, using RPN) is available from ARRL HQ for an SASE. (Contact the Technical Department Secretary and ask for the 1995 *Handbook* PLL design program.)

Fig 10.47 shows the sort of gain- and phase-versus-frequency plots obtained from a "recipe" loop design. We want to know where the loop's phase shift equals, or becomes more negative than, –180°. Added to the –180° shift inherent in the loop's feedback (the polarity of which must be negative to ensure that the phase detector drives the VCO toward lock instead of away from it), this will be the point at which the loop itself oscillates—*if* any loop gain remains at this point. The game is to position the poles and zero, and set the unit frequency gain, so that the loop gain falls to below 0 dB before the –180° line is crossed.

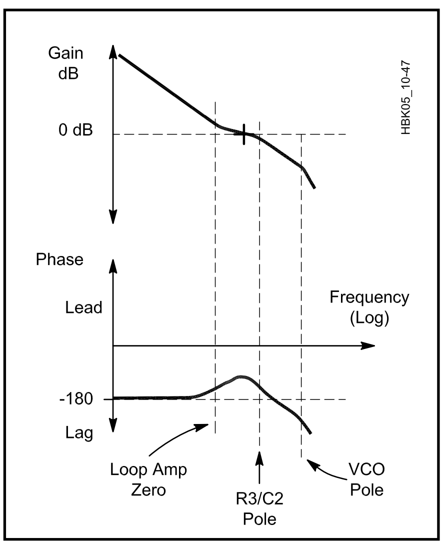


Fig 10.47—An open-loop gain/phase plot.

At the low-frequency end of the Fig 10.47 plots, the two poles have had their effect, so the gain falls at 40 dB/decade, and the phase remains just infinitesimally on the safe side of –180°. The next influence is the zero, which throttles the gain's fall back to 20 dB/decade and peels the phase line away from the –180° line. The zero would eventually bring the phase back to 90° of lag, but the next pole bends it back before that and returns the gain slope back to a fall of 20 dB/decade. The last pole, that attributable to the VCO, pushes the phase over the –180° line.

It's essential that there not be a pole between the 0-Hz pair and the zero, or else the phase will cross the line early. In this example, the R3-C2 pole and the zero do all the work in setting this critical portion of the loop's response. Their frequency spacing should create a phase bump of 30 to 45°, and their particular frequency positions are constrained as a compromise between sufficient loop bandwidth and sufficient suppression of reference-frequency sidebands. We know the frequency of the loop amplifier zero (that contributed by R1 and C1), so all we need to do now is design the loop amplifier to exhibit a frequency response such that the open-loop gain passes through 0 dB at a frequency close to that of the phase-plot bump.

This loop-design description may have been hard to follow, but a loop amplifier can be designed in under 30 minutes with a pocket calculator and a little practice. The high output impedance of commonly used CMOS phase detectors favors loop amplifiers based on FET-input op amps and allows the use of high-impedance RC networks (large capacitors can therefore be avoided in their design). LF356 and TL071 op amps have proven successful in loop-amplifier service and have noise characteristics suited to this environment.

To sum up, the recipe gives a proven pole zero pattern (listed in order of increasing frequency):

1. Two poles at 0 Hz (one is the integrator, the other is implicit in all PLs)
2. A zero controlled by the RC series in the loop amplifier feedback path
3. A simple RC low-pass pole
4. A second RC low-pass pole formed by the resistor driving the capacitive load of the tuning voltage inputs of the VCO

To design a loop: Design the VCO, divider, PSD and find their coefficients to get the loop gain at unit frequency (1 Hz), then, keeping the poles and zeros in the above order, move them about until you get the same 45° phase bump at a frequency close to your desired bandwidth. Work out how much loop-amplifier unit-frequency gain is needed to shift the gain-frequency-frequency plot so that 0-dB loop gain occurs at the center of the phase bump. Then calculate R and C values to position the poles and zero and the feedback-RC values to get the required loop amplifier gain at 1 Hz.

To Cheat—You can scale the example loop to other frequencies: Just take the reactance values at the zero and pole frequencies and use them to scale the components, you get the nice phase plot bump at a scaled frequency. Even so, you still must do the loop-gain design.

Don't forget to do this for your lowest division factor, as this is usually the least stable condition because there is less attenuation. Also, VCOs are usually most sensitive at the low end of their tuning range.

Noise in Phase-Locked Loops

Differences in Q usually make the phase-noise sidebands of a loop's reference oscillator much smaller than those of the VCO. Within its loop bandwidth, a PLL acts to correct the phase-noise components of its VCO and impose those of the reference. Dividing the reference oscillator to produce the reference signal applied to the phase detector also divides the deviation of the reference oscillator's phase-noise sidebands, translating to a 20-dB reduction in phase noise per decade of division, a factor of 20 log (M) dB, where M is the reference divisor. Offsetting this, within its loop bandwidth the PLL acts as a frequency *multiplier*, and this multiplies the deviation, again by 20 dB per decade, a factor of 20 log (N) dB, where N is the loop divider's divisor. Overall, the reference sidebands are increased by 20 log (N/M) dB. Noise in the dividers is, in effect, present at the phase detector input, and so is increased by 20 log (N) dB. Similarly, op-amp noise can be calculated into an equivalent phase value at the input to the phase detector, and this can be increased by 20 log (N) to arrive at the effect it has on the output.

Phase noise can be introduced into a PLL by other means. Any amplifier stages between the VCO and the circuits that follow it (such as the loop divider) will contribute some noise, as will microphonic effects in loop- and reference-filter components (such as those due to the piezoelectric properties of ceramic capacitors

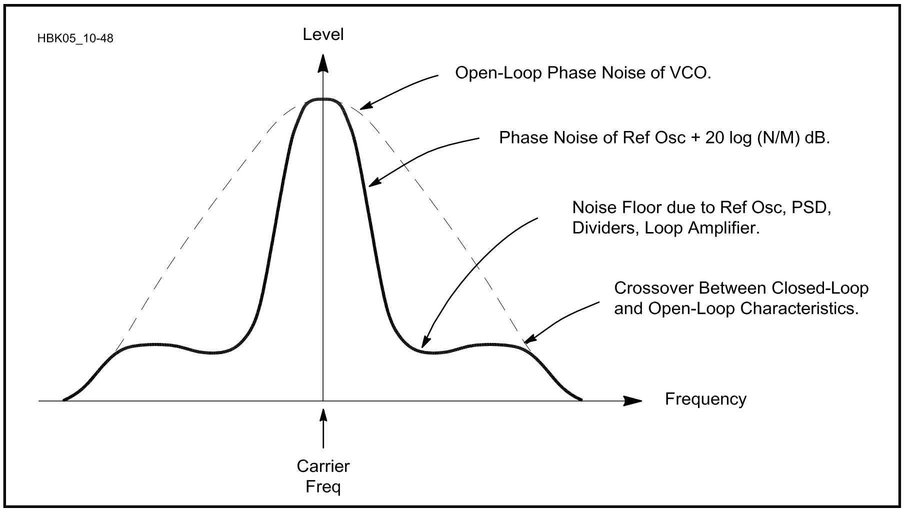


Fig 10.48—A PLL's open- and closed-loop phase-noise characteristics. The noise bumps at the crossover between open- and closed-loop characteristics are typical of PLLs; the severity of the bumps reflects the quality of the system's design.

and the crystal filters sometimes used for reference-oscillator filtering). Noise on the power supply to the system's active components can modulate the loop. The fundamental and harmonics of the system's ac line supply can be coupled into the VCO directly or by means of ground loops.

Fig 10.48 shows the general shape of the PLL's phase noise output. The dashed curve shows the VCO's noise performance when unlocked; the solid curve shows how much locking the VCO to a cleaner reference improves its noise performance. The two noise bumps are a classic characteristic of a phase-locked loop. If the loop is poorly designed and has a low stability margin, the bumps may be exaggerated—a sign of noise amplification due to an overly peaky loop response.

Exaggerated noise bumps can also occur if the loop bandwidth is less than optimum. Increasing the loop bandwidth in such a case would widen the band over which the PLL acts, allowing it to do a better job of purifying the VCO—but this might cause other problems in a loop that's deliberately bandwidth-limited for better suppression of reference-frequency sidebands. Immense loop bandwidth is not desirable, either: Farther away from the carrier, the VCO may be so quiet on its own that widening the loop would make it *worse*.

A BASIC DESIGN EXAMPLE

In this section, we will explore the application of the design principles previously covered, as well as some of the tradeoffs required in a practical design. We will also cover measurement techniques designed to give the builder confidence that the loop design goals have been met. Finally, we will cover some common troubleshooting problems.

As our design example, a synthesized local oscillator chain for a 10-GHz Transverter will be considered. **Fig 10.49** is a simplified block diagram of a 10-GHz converter. This example was chosen as it represents a departure from the traditional multi-stage multiply and filter approach. It permits realization of the oscillator system with two very simple loops and minimal RF hardware. It is also representative of what is achievable with current hardware, and can fit in a space of 2 to 3 square inches. This example is intended as a vehicle to explore the loop design aspects and is not offered as a "construction project." The additional detail required would be beyond the scope of this chapter.

In this example, two synthesized frequencies, 10 GHz and 340 MHz, are required. Since 10.368 GHz is one of the popular traffic frequencies, we will mix this with the 10-GHz LO to produce an IF of 368 MHz. The 368-MHz IF signal will be subsequently mixed with a 340-MHz LO to produce a 28-MHz final IF, which can be fed into the 10-meter input of any amateur transceiver. We will focus our attention on the design of the 10-GHz synthesizer only. Once this is done, the same principles may be applied to the 340-MHz section. Our goal is to attempt to design a low-noise LO system (this implies minimum division) with a loop reference oscillator that is an integer multiple of 10 MHz. Using this technique will allow the entire system to be locked at a later time to a 10-MHz standard for precise frequency control.

Ideally, we would like to start with a low-noise 100-MHz crystal reference and a 10-GHz oscillator divided by 100. It is here that we are already confronted with

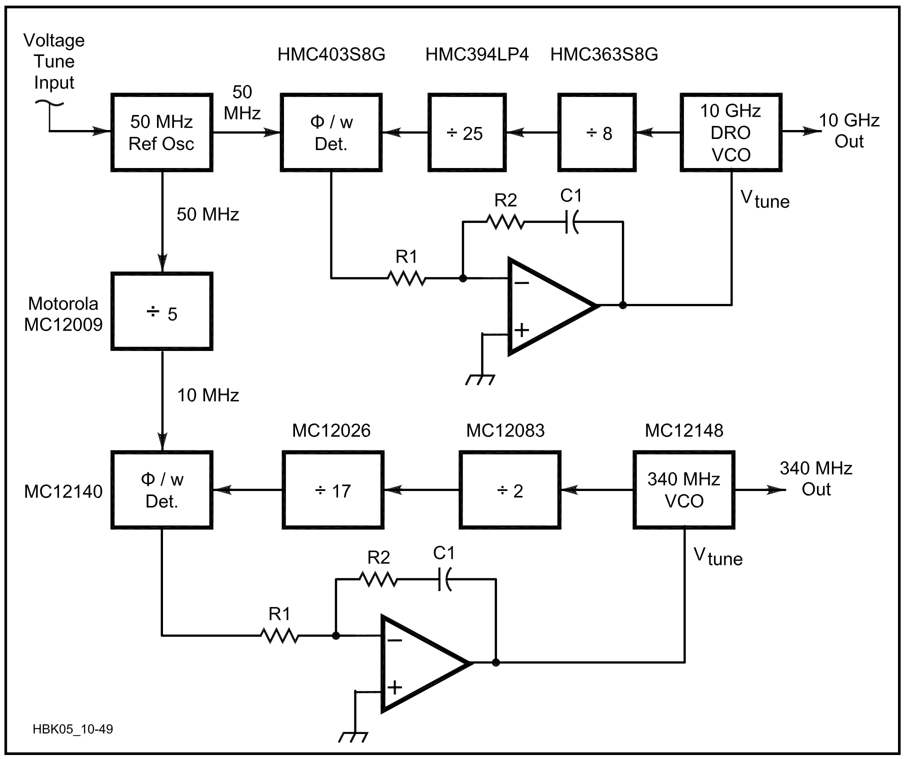


Fig 10.49—A simplified block diagram of a local oscillator, for a 10-GHz converter.

our first practical design tradeoff. We are considering using a line of microwave integrated circuits made by Hittite Microwave. These include a selection of prescalers operating to 12 GHz, a 5-bit counter that will run to 2.2 GHz and a phase/frequency detector that will run up to 1.3 GHz. The 5-bit counter presents the first problem. Our ideal scheme would be to use a divide-by-4 prescaler and enter the 5-bit counter at 2.5 GHz, subsequently dividing by 25 to 100 MHz. The problem is that the 5-bit counter is only rated to 2.2 GHz and not 2.5 GHz. This forces us to use a divide-by-8 prescaler and enter the 5-bit divider at 1.25 GHz. This is well in the range of the divider, but we can no longer use an integer division (ie, 1.25 GHz/12.5 = 100 MHz) to get to 100 MHz. The next easiest option is to reduce the reference frequency to 50 MHz and to let the 5-bit divider run as a divide-by-25, giving us a total division ratio of 200.

Having already faced our first design tradeoff, a number of additional aspects must be considered to minimize the conflicts in the design. **First**, this is a "static" synthesizer—That is, it will not be required to change frequency during operation. As a result, switching time considerations are irrelevant, as are the implications that the switching time would have had on the loop bandwidth. The effects of variable division ratios are also eliminated, as well as any problems associated with nonlinear tuning of the VCOs. **Second**, the reference frequencies (50 MHz and 10 MHz) are very large with respect to any practically desirable loop bandwidths. This makes the requirement of the loop filter to eliminate reference sidebands quite easy to achieve. In fact, reference sideband suppression will more likely be a function of board layout and shielding effectiveness rather than suppression in the loop filter. This also allows placement of the reference suppression poles far out in the pole zero constellation and well outside the loop bandwidth (about 10 × BW3 of the loop). This placement of the reference suppression poles will give us the option of using a "type 2, 2nd order" loop approximation, which will greatly simplify the calculation process. **Third**, based on all of the foregoing tradeoffs, the loop bandwidth can be chosen almost exclusively on the basis of noise performance.

Before we can proceed with the loop calculations we need some additional information: specifically VCO gain, VCO noise performance, divider noise performance, phase detector gain, phase detector noise performance and finally reference noise performance. The phase detector and divider information is available from specification sheets. Now we need to select the 50-MHz reference and the 10-GHz VCO. An excellent choice for a low noise reference is the one described by John Stephensen on page 13 in Nov/Dec 1999 *QEX*. The noise performance of this VCXO is in the order of –160 dBc/Hz at 10-kHz offset at the fundamental frequency. For the 10-GHz VCO, there are many possibilities, including cavity oscillators, YIG oscillators and dielectric resonator oscillators. Always looking for parts that are easily available, useable and economical, salvaging a dielectric resonator from a Ku band LNB (see Jan-Feb 2002 *QEX*, p 3) appears promising. These high-Q oscillators can be fitted with a varactor and tuned over a limited range with good results. The tuning sensitivity of our dielectric resonator VCO is about 10 MHz per volt and the phase noise at 10-kHz offset is –87dBc/Hz. We are now in a position to plot the phase noise of the reference, the dividers, the division effect (46 dB) and the VCO. The phase noise plots are shown in **Fig 10.50**. The noise of the reference is uniformly about 25 dB below the divider and phase detector noise floor. This means that the noise of the phase detector and dividers will be the dominant contribution within the loop bandwidth. The cross over point is at approximately 9 kHz. At frequencies lower than 9 kHz, the loop will reduce the noise of the VCO (eg, at 1 kHz the reduction will be about 20 dB). At frequencies above 9 kHz, the natural noise roll off of the VCO will dominate. At this point, we can also see the effect of having chosen a reference frequency of 50 MHz. Had we been able to use a 100-MHz reference instead of 50 MHz and reduce the division factor from 200 to 100, we could have put the loop bandwidth at about 30 kHz and picked up an additional 6-dB improvement in close-in noise performance. Nevertheless, even with the 6-dB penalty, this will be quite a respectable 10-GHz LO. It should be apparent that a fair amount of thought and effort was required to be able to determine an applicable bandwidth for this loop.

Now that we have some estimate of the noise performance and the required loop bandwidth, it is time to design the loop. The second consideration cited above enabling the application of the "type-2, 2nd order" approximation is now going to pay off. The math for this loop is really quite simple. There are two main variables, the "natural frequency" of the loop, ω_n, (omega-n in radians/second) and the "damping factor" D (delta is dimensionless). The closed-loop bandwidth of the loop is a function of ω_n and D. **Fig 10.51**

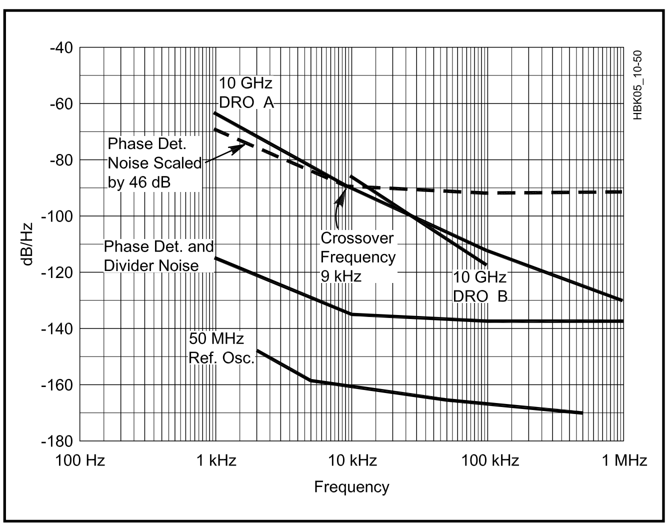


Fig 10.50—Phase noise plots of the 10-GHz transverter design example.

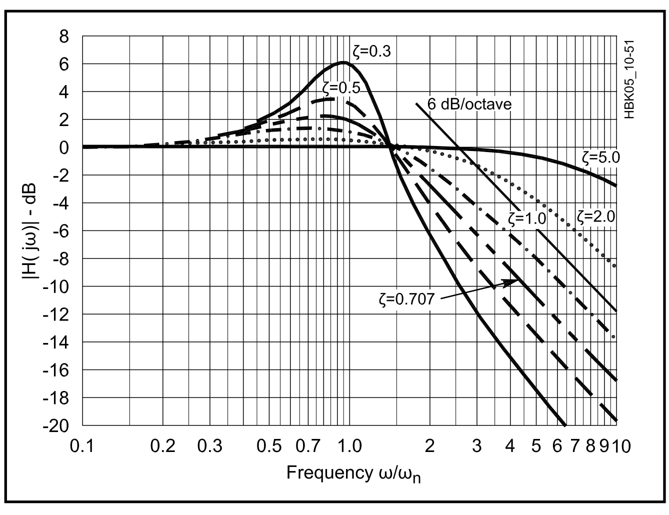


Fig 10.51—Closed-loop frequency response of a second-order loop as a function of the loop's natural frequency and damping factor (ω_n).

shows loop response as a function of ω_n and D. Values of D less than 0.5 are not desirable, as they tend towards instability and poor phase margin. A value for D of 0.707 is referred to as "critically damped" and is favored in applications where settling time is critical (see Gardner). For our application, we will choose a D of 1.0. This will give acceptable phase margin and further simplify the math. For a D of 1, $\omega_n = (6.28 \times BW3) / 2.48$. Substituting the 9-kHz loop bandwidth for BW3 yields 22790 radians per second for ω_n. Our loop filter will take the basic form shown in Fig 10.43. We can compute values for R1, R2 and C1 as follows.

$$R1 = (Kpd \times Kvco)/(N \times \omega_n^2 \times C1)$$

where Kpd is the phase detector gain, 1/6.28 volts/radian in this case, Kvco is the VCO gain in radians/Hz, 6.28×10^7 radians/volt in this case, N is the division factor, 200 in this case. C1 is the feedback capacitor value in Farads. To proceed, we need to start with an estimate for C1. Practically speaking, since odd values of capacitors are more difficult to obtain, let us try a value for C1 of 0.01 µF.

$$R2 = 2 \times D/(\omega_n \times C1)$$

R1 computes as 9626 Ω and R2 is 8775 Ω. Since we will be using a phase/frequency detector with differential outputs, we will have to modify the circuit of Fig 10.43 to become a differential amplifier. We will also add a passive input filter to keep the inputs of the op amp from being stressed by the very fast and short pulses emanating from the phase/frequency detector. We will also add a passive "hash filter" at the output of the op amp to limit the amount of out of band noise delivered to the VCO tuning port. Both of these filters will be designed for a cutoff frequency of 90 kHz (ie,10 times the 9-kHz closed-loop bandwidth, as cited above in the second consideration. Both of these filters will also aid in the rejection of reference frequency sidebands. For the input filter, we simply divide the input resistor in half and add a capacitor to ground. The value of this capacitor is determined as follows:

$$C2 = 4/(62.8 \times BW3 \times R1)$$

$$C2 = 735\ \text{pF}$$

Using 680 or 750 pF will be adequate. The output filter is also quite simple, with one minor stipulation. Op Amps can exhibit stability problems when asked to drive a capacitive load through too low a value of resistor. One very safe way around this is not to use a value for R3 lower than the recommended minimum impedance for full output of the amplifier. For many amplifiers, this value is around 600 Ω. For the output filter, we compute C3 as follows, for R3 = 600 Ω:

$$C3 = 1/(R3 \times 62.8 \times BW3)$$

$$C3 = 2950\ \text{pF}$$

Using 3000 pF will be adequate. This completes the computations for the loop compensation amplifier. The complete amplifier is shown in **Fig 10.52**.

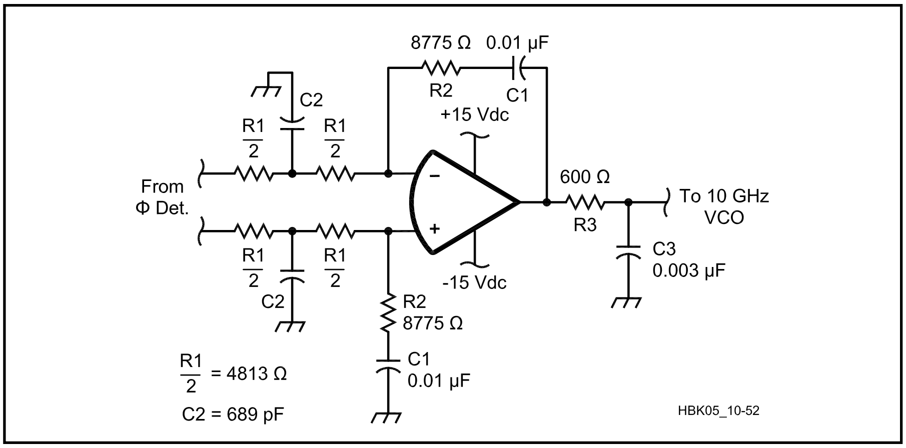


Fig 10.52—A complete loop-compensation amplifier for the design example 10-GHz transverter.

MEASUREMENTS

One of the first things we need to measure when designing a PLL is the VCO gain. The tools we will need include a voltmeter, some kind of frequency measuring device like a receiver or frequency counter and a source of clean variable DC voltage. The setup in **Fig 10.53**, containing one or more 9-V batteries and a 10-turn, 10-kΩ pot will due nicely. One simply varies the voltage some amount and then records the associated frequency change of the VCO. The gain of the VCO is then delta F over delta V. The phase detector gain constant can usually be found on the specification sheets of the components selected.

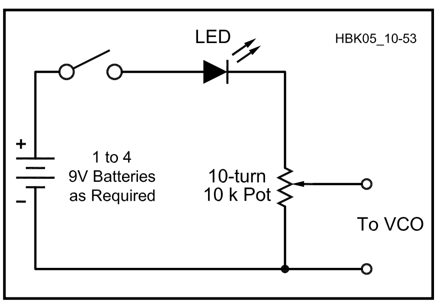


Fig 10.53—Clean, variable d-c voltage source used to measure VCO gain in a PLL design.

Measuring closed loop bandwidth is slightly more complicated. The way it is commonly done in the laboratory is to replace the reference oscillator with a DCFM-able (ie, a signal generator whose FM port is DC coupled) signal generator and then feed the tracking generator output of a low-frequency spectrum analyzer into the DCFM-able generator while observing the spectrum of the tuning voltage on the spectrum analyzer (see **Fig 10.54**). While this approach is quite straightforward, few amateurs have access to or can afford the test equipment to do this. Today however, thanks to the PC sound card-based spectrum analyzer and tracking generator programs (software by Interflex for example), amateurs can measure the closed-loop response of loops that are less than 20 kHz in bandwidth for significantly less than a king's ransom in test equipment! Here's how.

One approach is to build the reference oscillator with some "built-in test equipment" (BITE) already in the design. This BITE takes the form of some means of DCFM-ing the reference oscillator. This is one of the things that make John Stephenson's oscillator (previously mentioned) attractive. The oscillator includes a varactor input for the purpose of phase locking it to a high-stability low-frequency standard oscillator. This same varactor tuning input takes the place of the aforementioned DCFM-able signal generator. Now all we need to complete our test setup is some batteries. First, we need to make sure the input of our sound card is AC coupled (when connected to the VCO tuning port, there will be a DC component present and the sound card may not like this) and that the AC coupling is flat down to at least 10%

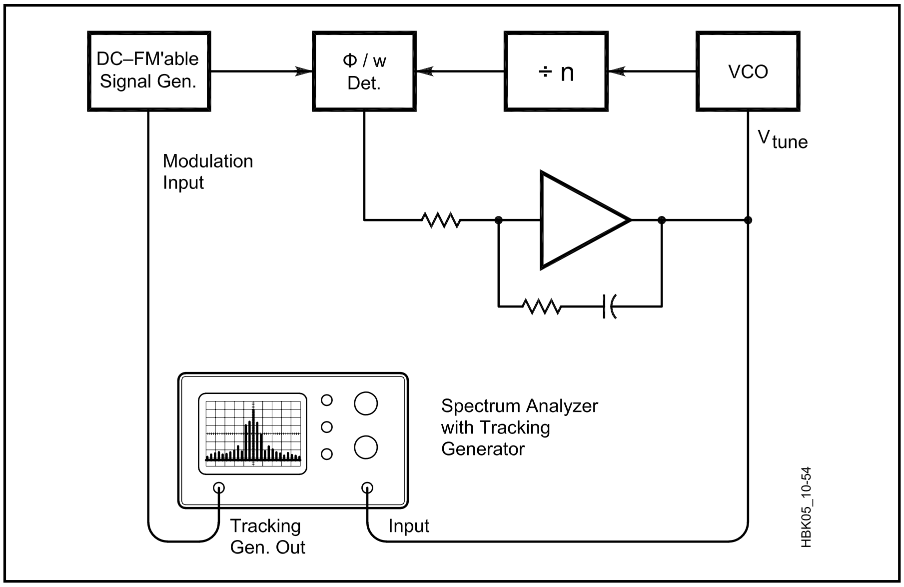


Fig 10.54—Common laboratory setup for measuring the closed-loop bandwidth of loops.

of the natural frequency of the loop. If the sound card is not AC coupled, a capacitor will be needed at the input to block the DC on the tuning voltage.

The next step is to determine the tuning sensitivity of the reference oscillator. This is done in the same manner as was done for the VCO described above. Once this has been established, we will set the audio oscillator for an output voltage that will produce between 100 Hz and 1 kHz of deviation. We will also need to establish that the output of the sound card is DC coupled so that we can place a small battery in series with the audio generator and the tuning port of the reference oscillator. We can provide a DC return on the output by putting a resistive 10-dB attenuator of the appropriate impedance on the output of the sound card and passing the varactor bias through the attenuator to ground. This is required to bias the varactors and also assure that the audio voltage will not drive the varactors into conduction. We can now connect the AC coupled input of the sound card to the VCO tuning voltage and turn on the loop.

Do not be surprised to see signal components at multiples of the power line frequency, as well as the signal from the audio oscillator. The presence of line frequency components is an indication that the loop is doing its job and removing these components from the spectrum of the VCO. We can now "sweep" the audio oscillator and, by plotting the amplitude of the audio oscillators response, determine the closed loop bandwidth of the PLL. In the case of the loop described above, (ω_n = 22790, D = 1.0) we should expect to see slightly over 1 dB of peaking at $0.7\omega_n$ and the 3-dB down point should fall at $2.5\omega_n$. In any event, if the peaking exceeds 3 dB, the loop phase margin is growing dangerously small and steps should be taken to improve it.

COMMON PROBLEMS

Here are some frequently encountered problems in PLL designs:

- The outputs of the phase detector are inverted. This results in the loop going to one or the other rail. The loop cannot possibly lock in this condition. Solution: Swap the phase detector outputs.
- The loop cannot comply with the tuning voltage requirements of the VCO. If the loop runs out of tuning voltage before the required voltage for a lock is reached, the locked condition is not possible. Solution: Re-center the VCO at a lower tuning voltage or increase the rail voltages on the op amp.
- The loop is very noisy and the tuning voltage is very low. The tuning voltage on the varactor diodes should not drop below the RF voltage swing in the oscillator tank circuit. Solution: Adjust the VCO so that the loop locks with a higher tuning voltage.

PLL SYNTHESIZER ICS

Now that we have learned how to deal with the loop design aspects of an indirect synthesizer, it is time to look at just a few of the many PLL synthesizers available today. Simple PLL ICs have been available since the early 1970s, but most of these contain a crude, low-Q VCO and a phase detector. They were intended for general use as tone detectors and demodulators, not as major elements in communication-quality frequency synthesizers.

One device well worth noting in this group, however, is the CD4046, which contains a VCO and a pair of phase detectors. The CD4046's VCO is useless for our purposes, but its phase-frequency detector is quite good. Better yet, the CD4046 is a low-cost part and one of the cheapest ways of getting a good phase-frequency detector. Its VCO-disable pin is a definite design plus. Later CD4046 derivatives, the 74HC4046 (CMOS input levels) and 74HCT4046 (TTL compatible input levels) are usable to much higher frequencies and seem to be more robust, but note that these versions are for +5 V supplies only, whereas the original CD4046 can be used up to 15 V.

Since then, many more complex ICs specifically intended for frequency synthesizers have been introduced by companies like Motorola, National (National LMX series) and others as a result of the growth in popularity of wireless devices. They normally contain a programmable divider, a phase-frequency detector and a reference divider that usually allows a small choice of division ratios. Usually, the buffer amplifier on these parts' reference input is arranged so that it can be used as part of a simple crystal oscillator. This is adequate for modest frequency accuracy, but an independent TCXO or OCXO is better.

Among Motorola's more popular parts is the MC145151. The MC145151 brings all of its division-control bits out to individual pins and needs no sequencing for control. It is the best choice if only a few output frequencies are needed, because they can be programmed via a diode matrix. The MC145151's divide-by-*N* range is 3 to 16383, controlled by a 14-bit word in binary format. The reference can be divided by 8, 128, 256, 512, 1024, 2048, 2410 and 8192. It's possible to operate the MC145151 to 30 MHz, and its phase-frequency detector is similar to that in the CD4046, lacking the MOSFET "Tri-State" output, but with the added benefit of a lock-detector output. The MC145152 is a variant that includes control circuitry for an external dual-modulus prescaler. The choice of external prescaler sets the maximum operating frequency, as well as maximum and minimum division ratios. For some unknown reason, the reference division choices differ from those of the MC145151: 8, 64, 128, 256, 512, 1024, 1160 and 2048. There is also a Tri-State

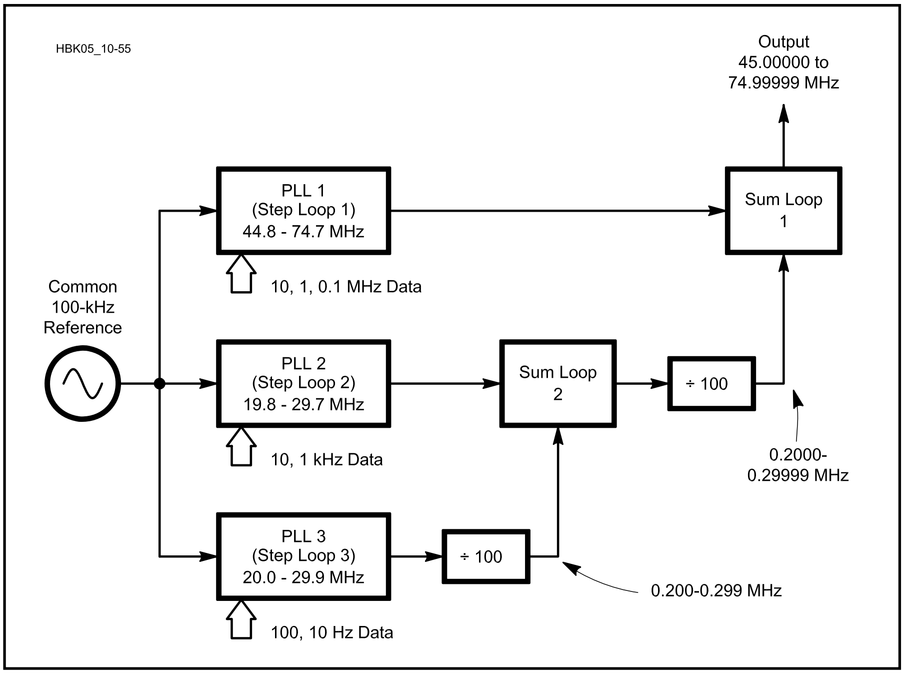


Fig 10.55—A five-loop synthesizer.

phase-frequency detector output.

The MC145146 is simply an MC145152 with its control data formatted as a 4-bit bus writing to eight addresses. This reduces the number of pins required and allows more data to be input, so the reference divider is fully programmable from 3 to 4095.

Multiloop Synthesizers

The trade-off between small step size (resolution) and all other PLL performance parameters has already been mentioned. The *multiloop synthesizer*, a direct synthesizer constructed from two or more PLL synthesizers, is one way to break away from this trade-off. **Fig 10.55**, the block diagram of a five-loop synthesizer, reflects the complexity found in some professional receivers. All of its synthesis loops run with a 100-kHz step size, which allows about a 10-kHz loop bandwidth. Such a system's noise performance, settling time after a frequency change and reference-sideband suppression can all be very good.

Cost concerns generally render such elaboration beyond reach for consumer-grade equipment, so various cut-down versions, all of which involve trade-offs in performance, have been used. For example, a three-loop machine can be made by replacing the lower three loops with a single loop that operates with a 1-kHz step size, accepting the slower frequency stepping determined by the narrow loop bandwidth that this involves.

DIRECT DIGITAL SYNTHESIS

Direct digital synthesis (DDS) is covered in this *Handbook's* **DSP and Software Radio Design** chapter. DDS technology is on the verge of covering all the local oscillator requirements of an HF transceiver. Very economical chipsets that cover partial requirements are available and can be exploited with the help of PLLs. For an example, see **Fig 10.56** and **Table 10.3**. A single loop with a programmable divider could be used in place of the crystal-oscillator bank. The result would be very close to the approach used in the latest generation of commercial Amateur Radio gear. That manufacturers advertise them as using direct digital synthesis has given some people the impression that their synthesizers consist entirely of DDS. In fact, the DDS in modern ham transceivers replaces the lower-significance "interpolation" loops in what otherwise would be regular multiloop synthesizers.

This is not to say that the DDS in our current ham gear is not a great improvement over the size, complexity and cost of what it replaces. Its random noise sidebands are usually excellent, and it can execute fast, clean frequency changes. The latest devices do 32-bit phase arithmetic and so offer over a billion frequency steps, which translates into a frequency resolution of a few *millihertz* with the usual clock rates for our applications—so the old battle for better resolution and low cost has already been won. Direct digital synthesis is not without a few problems, but fortunately, hybrid structures using PLL and DDS can allow one technique to compensate for the weaknesses of the other.

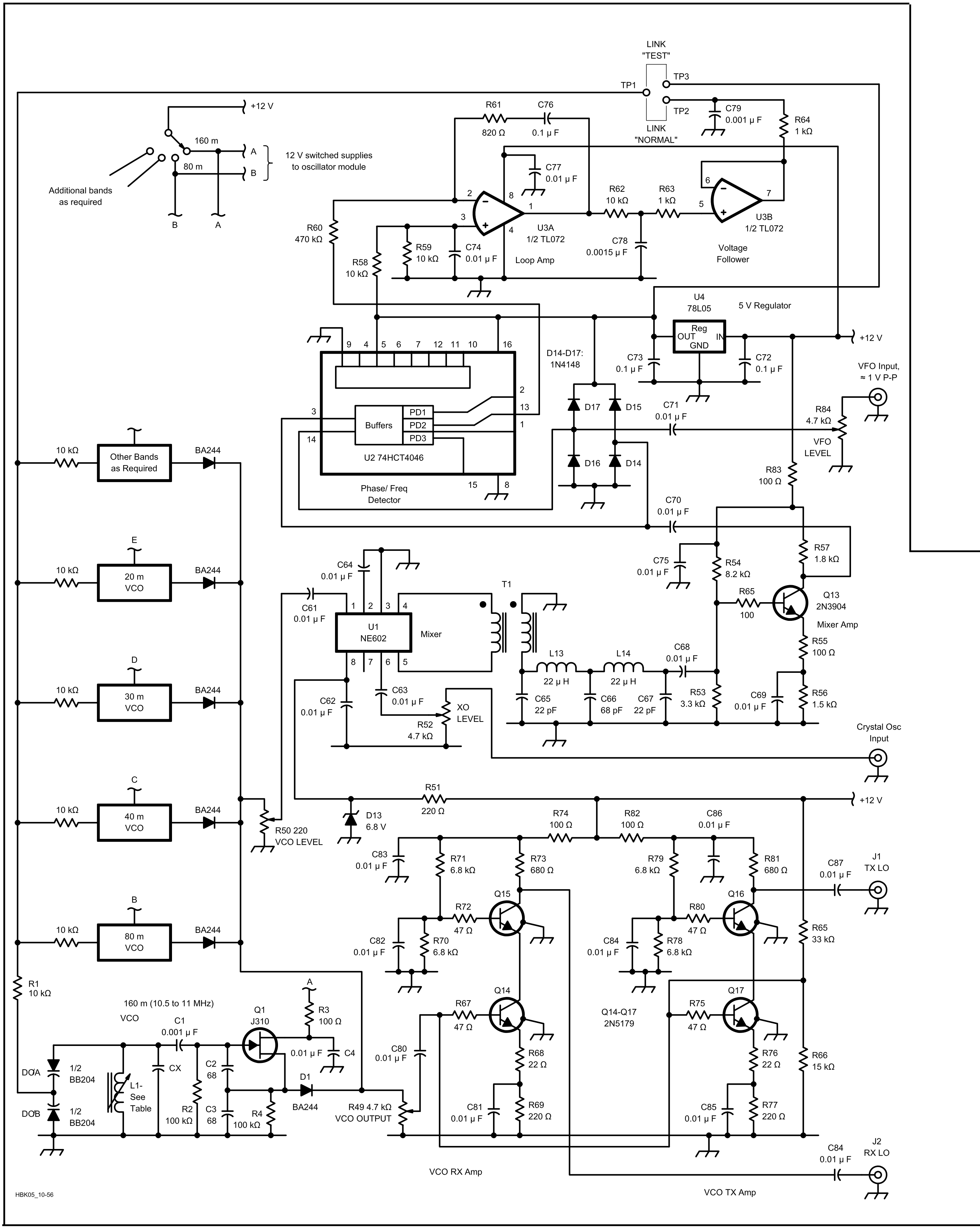
The prime weakness of the direct digital synthesizer is its quantization noise. A DDS cannot construct a perfect sine wave because each sample it outputs must be the nearest available voltage level from the set its DAC can make. So a DDS's output waveform is really a series of steps that only approximate a true sine wave.

We can view a DDS's output as an ideal sine wave plus an irregular "error" waveform. The spectrum of the error waveform is the set of unwanted frequency components found on the output of the DDS. Quantization noise is not the only source of unwanted DAC outputs. DACs can also give large output spikes, called glitches, as they transit from one level to another. In some DACs, glitches cause larger unwanted components than quantization. These components are scattered over the full frequency range passed by the DDS's low-pass filter, and their frequencies shift as the DDS tunes to different frequencies. At some frequencies, a number of components may coincide and form a single, larger component.

A summing loop PLL acts as a tracking filter, with a bandwidth measured in kilohertz. Acting on a DDS's output, a summing loop passes only quantization components close to the carrier. A system designer seeking to minimize the DDS's noise contribution must choose between using a loop bandwidth narrow enough to filter the DDS (thereby reducing the loop's ability to purify its VCO) or a more expensive low "glitch energy" DAC with more bits of resolution (allowing greater loop bandwidth and better VCO-noise control).

Complex ICs containing most of a DDS have been on the market for over 15 years, steadily getting cheaper, adding functions and occasionally taking onboard the functions of external parts (first the sine ROM, now the DAC as well). The Analog Devices AD7008 is an entire DDS (just add a low-pass filter…) on one CMOS chip. It includes a 10-bit DAC (fast enough for the whole system to clock at 50 MHz) and some digital-modulation hardware.

For the home-brewing amateur, DDS's first drawback is that these devices must receive their frequency data in binary form, loaded via a serial port, and this really forces the use of a microprocessor system in the radio. DDS's second drawback for experimenters is that surface-mount packages are becoming the norm for DDS ICs. Offsetting this, the resulting



LINK
"TEST"
TP1
TP3
TP2
LINK
"NORMAL"
C79
0.001 μ F
R64
1 kΩ
+12 V
160 m
A
80 m
B
12 V switched supplies
to oscillator module
Additional bands
as required
B
A
R61
820 Ω
C76
0.1 μ F
C77
0.01 μ F
R62
10 kΩ
R63
1 kΩ
U3A
1/2 TL072
Loop Amp
C78
0.0015 μ F
U3B
1/2 TL072
Voltage
Follower
R60
470 kΩ
R59
10 kΩ
C74
0.01 μ F
R58
10 kΩ
U4
78L05
5 V Regulator
Reg
OUT
IN
GND
C73
0.1 μ F
C72
0.1 μ F
+12 V
D14-D17:
1N4148
D17
D15
D16
D14
C71
0.01 μ F
VFO Input,
≈ 1 V P-P
R84
4.7 kΩ
VFO
LEVEL
Buffers
PD1
PD2
PD3
U2 74HCT4046
Phase/ Freq
Detector
R83
100 Ω
C70
0.01 μ F
10 kΩ
Other Bands
as Required
BA244
E
20 m
VCO
D
30 m
VCO
C
40 m
VCO
B
80 m
VCO
C64
0.01 μ F
C61
0.01 μ F
U1
NE602
Mixer
T1
L13
22 μ H
L14
22 μ H
C68
0.01 μ F
C65
22 pF
C66
68 pF
C67
22 pF
R53
3.3 kΩ
C69
0.01 μ F
R56
1.5 kΩ
C62
0.01 μ F
C63
0.01 μ F
XO
LEVEL
R52
4.7 kΩ
C75
0.01 μ F
R54
8.2 kΩ
R57
1.8 kΩ
R65
100
Q13
2N3904
Mixer Amp
R55
100 Ω
Crystal Osc
Input
R51
220 Ω
+12 V
R50 220
VCO LEVEL
D13
6.8 V
R74
100 Ω
R82
100 Ω
C86
0.01 μ F
C83
0.01 μ F
R71
6.8 kΩ
R73
680 Ω
R79
6.8 kΩ
R81
680 Ω
C87
0.01 μ F
J1
TX LO
Q15
Q16
R72
47 Ω
R80
47 Ω
C82
0.01 μ F
R70
6.8 kΩ
C84
0.01 μ F
R78
6.8 kΩ
R65
33 kΩ
R1
10 kΩ
160 m (10.5 to 11 MHz)
VCO
C1
0.001 μ F
Q1
J310
A
R3
100 Ω
0.01 μ F
C4
Q14
R67
47 Ω
Q14-Q17
2N5179
Q17
R75
47 Ω
C80
0.01 μ F
R68
22 Ω
R76
22 Ω
R66
15 kΩ
D1
DCA
1/2
BB204
DCB
1/2
BB204
L1-
See
Table
CX
C2
68
C3
68
R2
100 kΩ
R4
100 kΩ
BA244
R49 4.7 kΩ
VCO OUTPUT
C81
0.01 μ F
R69
220 Ω
C85
0.01 μ F
R77
220 Ω
C84
0.01 μ F
J2
RX LO
VCO RX Amp
VCO TX Amp
HBK05_10-56

Table 10.3
Band-Specific Component Data for the Summing Loop in Fig 10.56

Band (MHz)	*Output Range (MHz)*	*Crystal Frequency (MHz)*	C_x *(pF)*	*VCO Coil*
1.5-2.0	10.5-11.0	16.0	0	20 turns on Toko 10K-series inductor form (≈4.29 µH)
3.5-4.0	12.5-13.0	18.0	0	16 turns on Toko 10K-series inductor form (≈3.03 µH)
7.0-7.5	16.0-16.5	21.5	0	12 turns on Toko 10K-series inductor form (≈1.85 µH)
10.0-10.5	19.0-19.5	24.5	27	8 turns on Toko 10K-series inductor form (≈0.87 µH)
14.0-14.5	23.0-23.5	28.5	56	8 1/2 turns white Toko S-18-series (≈0.435 µH)
18.0-18.5	27.0-27.5	32.5	39	7 1/2 turns violet Toko S-18-series (≈0.375 µH)
21.0-21.5	30.0-30.5	35.5	27	7 1/2 turns violet Toko S-18-series (≈0.350 µH)
24.5-25.0	33.5-34.0	39.0	22	6 1/2 turns blue Toko S-18-series (≈0.300 µH)
28.0-28.5	37.0-37.5	42.5	22	5 1/2 turns green Toko S-18-series (≈0.245 µH)
28.5-29.0	37.5-38.0	43.0	22	5 1/2 turns green Toko S-18-series (≈0.239 µH)
29.0-29.5	38.0-38.5	43.5	22	5 1/2 turns green Toko S-18-series (≈0.232 µH)
29.5-30.0	38.5-39.0	44.0	22	5 1/2 turns green Toko S-18 series (≈0.227 µH)

The Toko 10K-series forms have four-section bobbins. The VCO-coil windings for 160 through 30 m are therefore split into four equal sections (for example, 5 + 5 + 5 + 5 turns for the 160-m coil).

Fig 10.56—The G3ROO/GM4ZNX summing-loop PLL phase-locks its VCO to the frequency difference between a crystal oscillator and 5.0- to 5.5-MHz VFO (these circuits are not shown; see text). Table 10.3 lists the conversion-crystal frequency, VCO tuned-circuit padding capacitance (C_x) and VCO inductor data required for each band. Any 5.0- to 5.5-MHz VFO capable at least 2.5 mW (4 dBm) output with a 50-Ω load—1 V P-P—can drive the circuit's VFO input.

C_x, C2, C3—NP0 or C0G ceramic, 10% or tighter tolerance.
C65, C66, C67—NP0, C0G or general-purpose ceramic, 10% tolerance or tighter.
D0A, D0B—BB204 dual, common-cathode tuning diode (capacitance per section at 3 V, 39 pF) or equivalent. The ECG617, NTE617 and MV104 are suitable dual-diode substitutes, or use pairs of 1N5451s (39 pF at 4 V) or MV2109s (33 pF at 4V).
D1—BA244 switching diode. The 1N4152 is a suitable substitute.
L1—Variable inductor; see Table 10.3 for value.
L13, L14—22-µH choke, 20% tolerance or better (Miller 70F225AI, 78F270J, 8230-52, 9250-223; Mouser 43LQ225; Toko 144LY-220J [Digi-Key TK4232] suitable).
T1—6 bifilar turns of #28 enameled wire on an FT-37-72 ferrite toroid (≈30 µH per winding).
U2—74HC4046 or 74HCT4046 PLL IC (CD4046 unsuitable; see text).

simplification of an entire synthesizer to a few ICs should cause more people to experiment with them.

Fractional-N Synthesis

A single loop would be capable of any step size if its programmable divider weren't tied to integer numbers. Some designers long ago tried to use such *fractional-N* values by switching the division ratio between two integers, with a duty cycle that set the fractional part. This whole process was synchronized with the divider's operation. Averaged over many cycles, the frequency really did come out as wanted, allowing interpolation between the steps mandated by integer-N division. This approach was largely abandoned because it added huge sidebands (at the fractional frequency and its harmonics) to the loop's output. One such synthesizer design—which has been applied in a large amount of equipment and remains in use—uses a hybrid digital/analog system to compute a sawtooth voltage waveform of just the right amplitude, frequency and phase that can be added to the VCO tuning voltage to cancel the fractional-frequency FM sidebands. This system is complex, however, and like all cancellation processes, it can never provide complete cancellation. It is appreciably sensitive to changes due to tolerances, aging and alignment. Even when applied in a highly developed form using many tight-tolerance components, such a system cannot reduce its fractional frequency sidebands to a level much below –70 to –80 dBc.

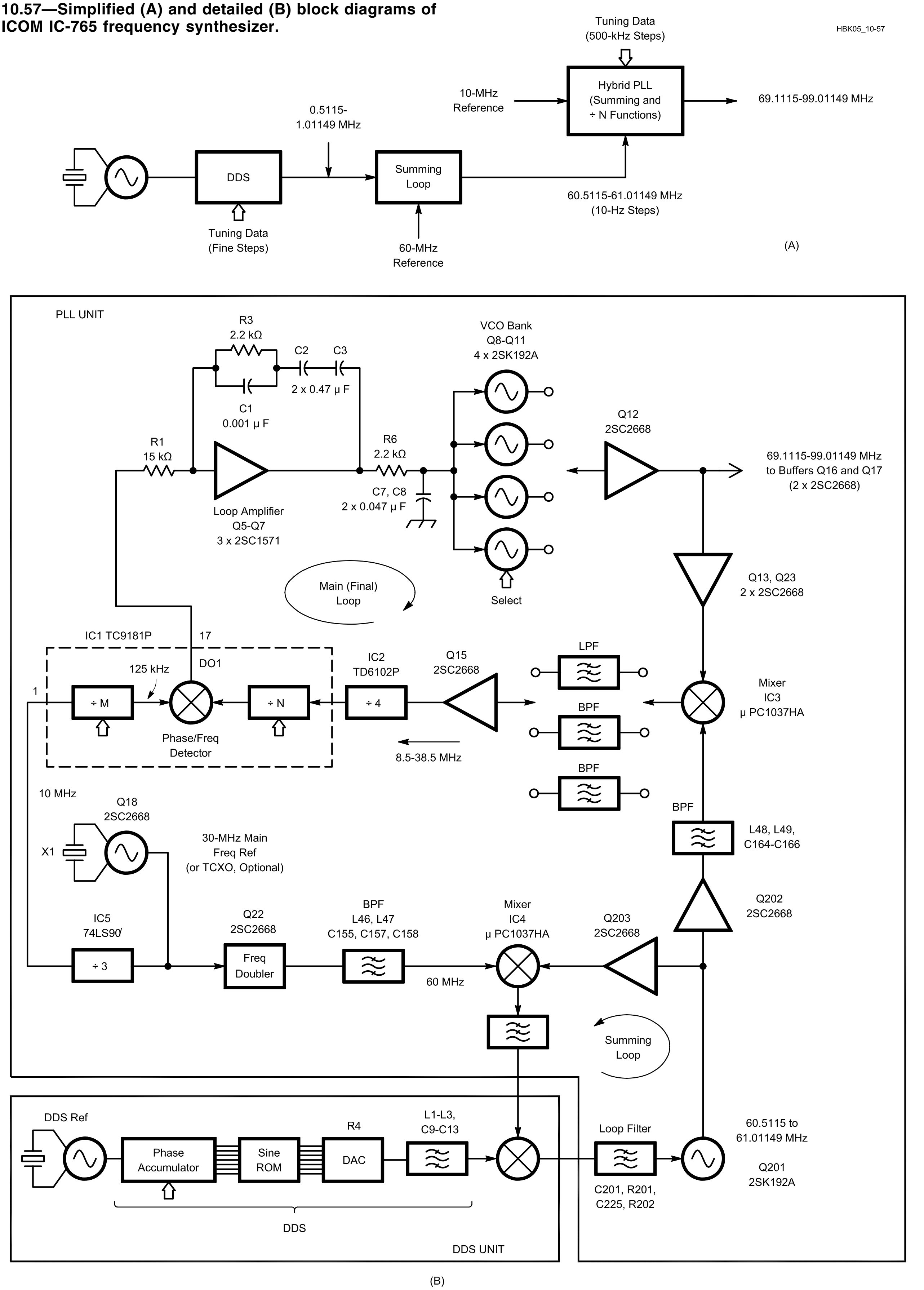
A new approach, which has been described in a few articles in professional and trade journals, does not try to cancel its sidebands at all. Its basic principle is delightfully simple. A digital system switches the programmable divider of a normal single loop around a set of division values. This set of values has two properties: First, its average value is controllable in very small steps, allowing fine interpolation of the integer steps of the loop; second, the FM it applies to the loop is huge, but the resultant sideband energy is strongly concentrated in very-high-order sidebands. The loop cannot track such fast FM and so filters off this modulation!

The key to this elegant approach is that it deliberately shapes the spectrum of its loop's unwanted components such that they're small at frequencies at which the loop will pass them and large at frequencies filterable by the loop. The result is a reasonably clean output spectrum.

EXPLORING THE SYNTHESIZER IN A COMMERCIAL MF/HF TRANSCEIVER

Few people would contemplate building an entire synthesized transceiver, but far more will need to understand enough of a commercially built one to be able to fix it or modify it. Choosing a radio to use as an example was not too difficult. The ICOM IC-765 received high marks for its clean

Fig 10.57—Simplified (A) and detailed (B) block diagrams of the ICOM IC-765 frequency synthesizer.



synthesizer in its *QST* Product Review, so it certainly has a synthesizer worth examining. We are grateful to ICOM America for their permission to reprint the IC-765's schematic in the discussion to follow.

Fig 10.57A shows a simplified block diagram of the IC-765's synthesizer. It contains one DDS and two PLLs. Notice that the DDS has its own frequency-reference oscillator. DDSs usually use binary arithmetic in their phase accumulators, so their step size is equal to their clock (reference) frequency divided by a large, round, binary number. The latest 32-bit machines give a step size of $1/(2^{32})$ of their clock frequency, that is, a ratio of 1/(4,294,967,296). This means that if we want, say, a 10-Hz step size, we must have a peculiar reference frequency so that the increment of the DDS is a submultiple of 10 Hz. This is what ICOM designers chose. One alternative would be to use a convenient reference frequency (say 10 MHz) and accept a strange synthesizer step size. A 32-bit machine clocked at 10 MHz will give a step size of 0.002328 Hz. It would be simple to have the radio's microprocessor select the nearest of these very fine steps to the frequency set by the user, giving the user the appearance of a 10-Hz step size. An error of ±1.2 millihertz is trivial compared to the accuracy of all usual reference oscillators.

Installing the IC-765's optional high-stability 30-MHz TCXO does nothing to improve the accuracy of the radio's DDS section, since it uses its own reference. The effect of the stability and accuracy of the DDS reference on the overall tuned frequency, however, is smaller than that of the main reference.

The DDS runs at a comfortably low frequency (0.5115 to 1.01149 MHz) to ease the demands placed on DAC settling time. The final loop needs a signal near 60 MHz as a summing input. Just mixing the DDS with 60 MHz would require a complex filter to reject the image. The IC-765 avoids this by using a summing loop. The summing loop VCO must only tune from 60.5115 to 61.01149 MHz, plus some margin for aging and temperature, but care is needed in this circuit because, as in the home-brew summing loop described earlier, the loop may latch up if this VCO goes below 60 MHz.

The final loop is a normal PLL with a programmable divider, but instead the VCO is not fed directly into the divider—it is mixed down by the 60.5115- to 61.01149-MHz signal first. The final loop

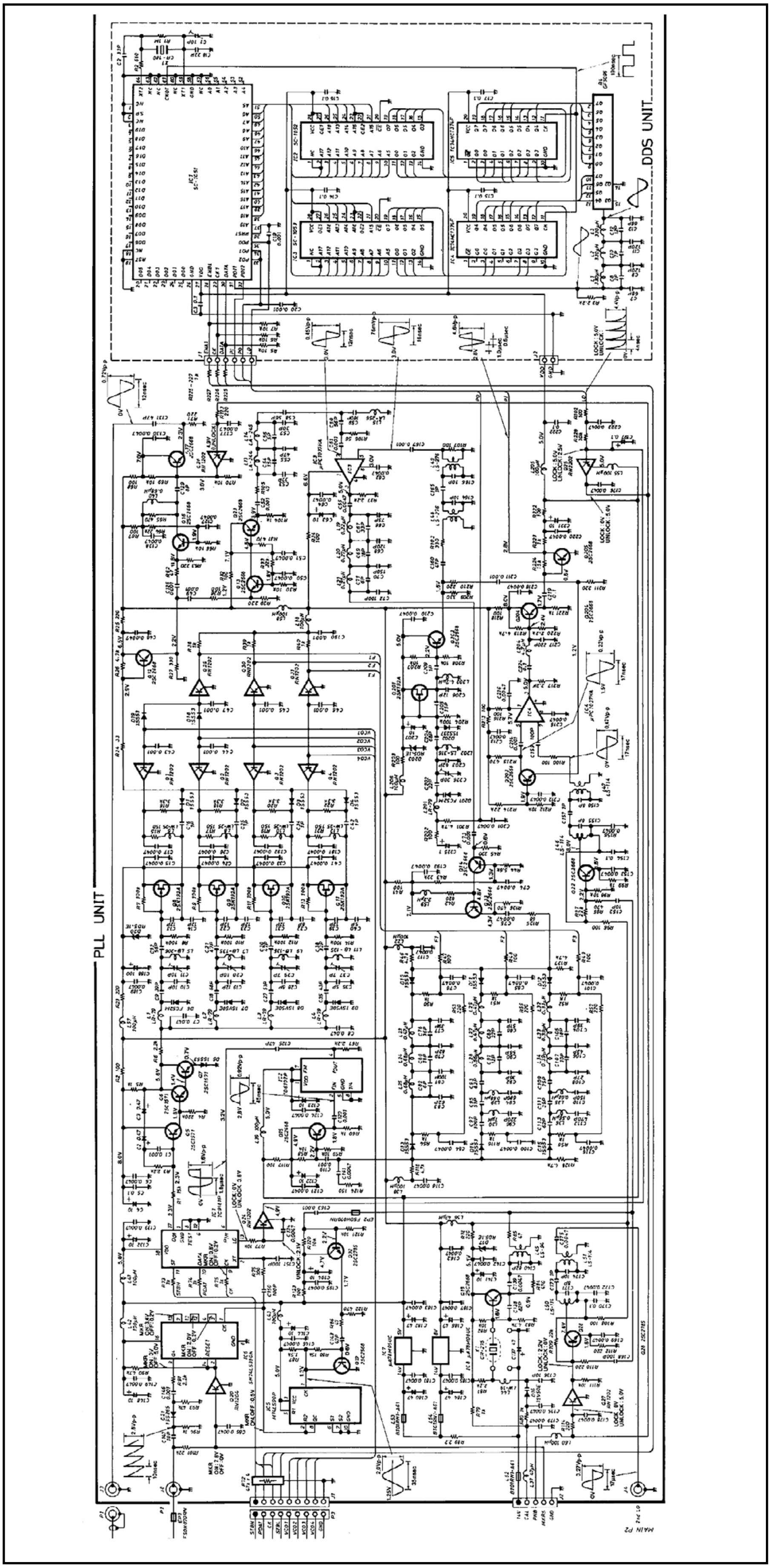


Fig 10.58—The IC-765 frequency synthesizer down to the component level. The key ICs in the radio's DDS (IC1, IC2 and IC3, right) appear to be custom components made especially for this application.

uses a bank of four switched VCOs, each one covering a one fourth of the system's full output range of 69.1115 to 99.01149 MHz. This mix-down feeds an 8.5- to 38.5-MHz signal into the programmable divi-der. To remove unwanted mixer outputs over this range, three different, switched filters are needed before the signal is amplified to drive the divider.

Fig 10.57B shows the synthesizer in greater detail. The programmable divider/PSD chip used (IC1, a TC9181P) will not operate up to 38 MHz, so a prescaler (IC2, a TD6102P) has been added before its input. This is a fixed divide-by-four prescaler, which forces the phase detector to be run at one-fourth of the step size, at 125 kHz. (A dual-modulus prescaler would have been very desirable here because it would allow the PSD to run at 500 kHz, easing the trade-off between reference-frequency suppression and loop bandwidth. Unfortunately, the lowest division factor necessary is too low for the simple application of any of the common ICs with prescaler controllers.)

Fig 10.58 shows the full circuit diagram of the IC-765's synthesizer. The DDS is on a small sub board, DDS unit (far right). IC1, an SC-1051, appears to be a custom IC, incorporating the summing loop's phase accumulator and phase detector. The DDS's sine ROM is split into two parts: IC2, an SC-1052; and IC3, an SC-1053. IC4 and IC5, both 74HCT374s, are high-speed-CMOS flip-flops used as latches to minimize glitches by closely synchronizing all 12 bits of data coming from the ROMs. The DAC is simply a binary-weighted resistor ladder network, R4, which relies on the CMOS latch outputs all switching exactly between the +5 V supply and ground.

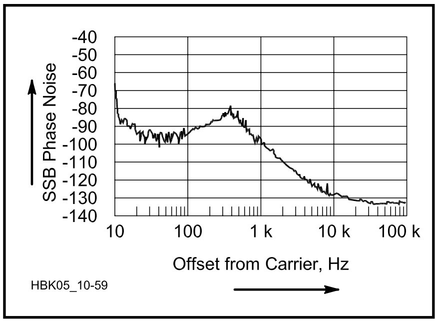


Fig 10.59—The phase noise of an ICOM IC-765 (serial no. 03077) as measured by a Hewlett-Packard phase-noise measurement system.

Look at the output of the main loop phase detector (pin 17 of IC1 on the PLL unit), find Q5, Q6 and Q7 and look at the RC network around them. It's the recipe loop design again! R3 creates the zero, but notice C1 across it. This is a neat and economical way of creating one of the other poles. R6 and C7 create the final pole.

Fig 10.59 shows the phase noise of an IC-765 measured on a professional phase-noise-measurement system. The areas shown in Fig 10.48 can be seen clearly. The slope above 500 Hz is the phase noise of the final VCO and exactly tracks the *Good* curve in KI6WX's *QST* article on the effects of phase noise. Below the noise peak, the phase noise falls to levels 20 dB better than KI6WX's *Excellent* curve. This low-noise area complements the narrow CW filter skirts and the narrow notch filter.

Many technically inclined amateurs may be wondering what could be done to improve the IC-765's synthesizer, or to design an improved one. The peak in the IC-765's phase noise at 500 Hz is quite prominent. It does not appear to be due to a marginally unstable loop design, but rather to a mismatch between the choice of loop bandwidth and the noise performance of the oscillators involved. The DDS's resistor-based DAC is unlikely to perform as well as a purpose-designed

Present and Future Trends in Oscillator Application

In this chapter a wide variety of oscillator types and frequency-synthesis schemes are discussed. Which techniques are the most important today and which ones will likely become important in the future?

As we follow radio technology from the invention of the vacuum tube, we can see a continuous evolution in the types of oscillators deemed most useful at any given time. Broadly speaking, communications systems started out with LC oscillators, and when the need for greater frequency stability arose designers moved to crystal oscillators. As the need to vary frequency became apparent, low-drift VFOs were developed to replace the crystals. As higher frequencies began to be exploited, stability again became the dominant problem. Multi-conversion systems that used low-frequency stable VFOs with crystal oscillators were developed to establish the desired high frequency stability. Today with stability, variability, and programmability all being requirements, frequency-synthesis techniques have been adopted as the norm.

When we look inside today's synthesized transceivers, we usually see only two types of oscillators. The first is a temperature-compensated crystal oscillator (TCXO), whose main purpose is to set the frequency calibration of the transceiver. The second type is the low phase-noise voltage-controlled oscillator (VCO) used in synthesizer phase lock loops.

Today, formerly popular mechanically tuned VFOs have fallen by the wayside, with the exception perhaps being their use in QRP projects and nostalgia radios. Even this could change in the near future in QRP projects, as very low power consumption synthesizers are already employed in cell phones and handhelds.

As we look to the future, we can see several emerging trends. Audio DSP has been around for a number of years, and some of the more contemporary radios are now employing IF DSP. As A-D/D-A technology approaches 16 bits at sample rates exceeding 60 MHz, it becomes feasible to produce an almost entirely DSP 1.8 to 30-MHz transceiver. The only remaining analog RF sections will be input filtering and gain compensation, and the output power amplifier and filtering. All of the traditional local oscillator synthesis hardware, IF filters, IF amplifiers, mixers and demodulation will be done in a DSP engine that resides behind an A-D/D-A converter pair. For the frequency ranges covered by the A-D/D-A pair, the synthesis process will be substantially "hidden." It will likely look like the software for direct digital synthesizer (DDS) minus the DA converter. What kinds of oscillators will be required for this architecture and what parameters will be most important?

Again, two types of oscillators will be necessary. At first glance, it would be reasonable to have a TCXO for frequency accuracy, just as in today's synthesized radios. While this would be perfectly acceptable, another option now exists. Current cell phones are being equipped with low-cost Global Positioning Satellite (GPS) chip sets to meet the federally mandated 911 emergency location requirements. An appropriate GPS chip set could give amateur transceivers frequency accuracy on a par with the cesium frequency standards employed by the GPS system. An additional benefit

12-bit DAC, and it also lets power-supply noise modulate the output. The system's designers appear to have deliberately reduced the loop bandwidth to better filter DDS spurs. If we were free to increase the cost and complexity of the unit, we could buy a high-performance DAC designed for low DDS spurs and trade this improvement off against noise by increasing the loop bandwidth—redesigning the loop poles, zero and gain. To help with the loop bandwidth, and get a 12-dB improvement in any noise from the main loop's phase detector, divider and loop amplifier, we could remove the fixed prescaler (IC2, the TD6102P) and operate the PSD at 500 kHz. This would require the design of a faster programmable divider that can handle 38 MHz.

Finally, the IC-765's VCOs could be improved. Their tuned circuits could be changed to low-L, high-C, multiple-tuning-diode types. If we added a higher-voltage power supply to allow higher diode tuning voltages than the 5.6-V level at which the design now operates (only little current is needed), higher-Q tuning diodes can be used—and we can avoid using them at lower tuning voltages, where tuning-diode Q degrades. This could reduce the phase noise above the noise peak by several decibels. Changing the loop bandwidth would reduce the noise bump's height and move it farther from the carrier.

It's important to keep all of these what-ifs in perspective: The IC-765's synthesizer was one of the best available in amateur MF/HF transceivers when this chapter was written, and its complexity is about half that of old multiloop, non-DDS examples. Like any product of mass production, its design involved trade-offs between cost, performance and component availability. As better components become less expensive, we can expect excellent synthesizer designs like the IC-765's to come down in price and even better synthesizer designs to become affordable in Amateur Radio gear.

BIBLIOGRAPHY AND REFERENCES

Clarke and Hess, *Communications Circuits; Analysis and Design* (Addison-Wesley, 1971; ISBN 0-201-01040-2). Wide coverage of transistor circuit design, including techniques suited to the design of integrated circuits. Its age shows, but it is especially valuable for its good mathematical treatment of oscillator circuits, covering both frequency- and amplitude-determining mechanisms. Look for a copy at a university library or initiate an interlibrary loan.

J. Grebenkemper, "Phase Noise and Its Effects on Amateur Communications," *Part 1*, *QST*, Mar 1988, pp 14-20; *Part 2*, Apr 1988, pp 22-25. Also see Feedback, *QST*, May 1988, p 44. This material also appears in volumes 1 (1991) and 2 (1993) of *The ARRL Radio Buyer's Sourcebook*. Covers the effects of phase noise and details the measurement techniques used in the ARRL Lab. Those techniques are also described in a sidebar earlier in this chapter.

W. Hayward and D. DeMaw, *Solid State Design for the Radio Amateur* (Newington, CT: ARRL, 1986). A good sourcebook of RF design ideas, with good explanation of the reasoning behind design decisions.

"Spectrum Analysis...Random Noise Measurements," Hewlett-Packard Application Note 150-4. Source of correction factors for noise measurements using spectrum analyzers. HP notes should be available through local HP dealers.

U. Rohde, *Digital PLL Frequency Synthesizers* (Englewood Cliffs, NJ: Prentice-

would be that the transceiver would now know its location anywhere in the world. This would be of value to DXpeditioners and contesters!

The second oscillator would most likely take the form of a very low phase noise voltage-controlled crystal oscillator (VCXO). Some variant of the Butler oscillator would be a likely candidate. This oscillator would be disciplined by the output of the GPS chipset or the onboard frequency standard TCXO. The requirement for low phase noise will remain very important, since any clock jitter in the DSP engine or the A-D/D-A clocks will degrade the phase-noise performance of the transceiver.

The possibility of an all-DSP HF radio raises the question of whether or not frequency synthesizers, as we now know them, will go the way of the VFO? For the most demanding applications, such as in signal generators and test equipment, the answer is likely no. The reason is that the most current synthesis techniques offer better performance than is possible with the basic DDS approach, although at additional cost. With respect to HF amateur radio equipment—where cost performance tradeoffs are important—the answer is likely yes.

Now that we have discussed the possibility of a DSP based HF radio, what about VHF, UHF and microwave amateur equipment? It is reasonable to assume that the direct DSP approach could be extended up in frequency as the A-D/D-A technologies improve in speed, however, there are some limiting factors regarding phase noise performance. The present method of using a tunable IF, such as the HF transceiver and a transverter, is likely to remain the practical option for some time to come. What we are likely to see is a different approach to the generation of the LO frequency in transverters. Many of today's transverters employ a crystal oscillator and multiplier chain to produce the LO injection frequency. This architecture has some disadvantages in that the crystal oscillator often has an offset and drift that must be taken up in the HF transceiver. There is also the issue of spurious responses due to the multiplying process and adequacy of the attendant filters. An alternative would be to use a simplified static synthesizer whose frequency reference is the same as that used in the HF transceiver. This approach, implemented with a low-noise UHF or microwave VCO and phase-lock loop, could yield reasonable phase noise, greater frequency accuracy, fewer spurious responses and smaller size as opposed to the crystal/multiplier chain technique. Again, we can see that the low-noise VCO will be an important element in these designs. These UHF and microwave VCOs would likely employ ceramic or dielectric resonators like those now used extensively in cell phones and TVRO front ends.

It is extremely likely that voltage-controlled oscillators, whether they are crystal, LC, ceramic or dielectric resonator types, will dominate the communications electronics landscape. Very low phase noise will be the cardinal performance specification for these VCOs. High-stability frequency-standard oscillators, while still viable, may give way to GPS as a preferred means of accurate frequency deployment. Mechanically tuned VFOs have all but disappeared from modern equipment and are likely to continue to diminish in application.

Hall, Inc). Now available solely from Compact Software, 201 McLean Blvd, Paterson, NJ 07504, this book contains the textbook-standard mathematical analyses of frequency synthesizers combined with unusually good insight into what makes a better synthesizer and a *lot* of practical circuits to entertain serious constructors. A good place to look for low-noise circuits and techniques.

U. Rohde, "Key Components of Modern Receiver Design," *Part 1, QST*, May 1994, pp 29-32; *Part 2, QST*, Jun 1994, pp 27-31; *Part 3, QST,* Jul 1994, pp 42-45. Includes discussion on phase-noise reduction techniques in synthesizers and oscillators.

Pappenfus and others, *Single Sideband Circuits and Systems* (McGraw-Hill). A book on HF SSB transmitters, receivers and accoutrements by Rockwell-Collins staff. Contains chapters on synthesizers and frequency standards. The frequency synthesizer chapter predates the rise of the DDS and other recent techniques, but good information about the effects of synthesizer performance on communications is spread throughout the book.

Motorola Applications Note AN-551, "Tuning Diode Design Techniques," included in Motorola RF devices databooks. A concise explanation of varactor diodes, their characteristics and use.

I. Keyser, "An Easy to Set Up Amateur Band Synthesizer," *RADCOM* (RSGB), Dec 1993, pp 33-36.

B.E. Pontius, "Measurement of Signal Source Phase Noise with Low-Cost Equipment," *QEX*, May-June 1998, pp 38-49.

F.M. Gardner, *Phase Lock Techniques* (John Wiley and Sons, 1966, 1975; ISBN 0-471-29156-0).