Signals, Samples, and Stuff: A DSP Tutorial (Part 2)

As we continue our exploration of DSP techniques, let's look inside an IF-DSP transceiver.

By Doug Smith, KF6DX/7

In Part 1 of this series (*QEX*, Mar/Apr 1998, pp 3-16) we learned about fundamental DSP techniques and algorithms for use in modern transceivers. Here we'll explore an actual IF-DSP transceiver design. Many of the issues relate well to conventional analog equipment, but special emphasis is placed on the unique requirements, advantages and trade-offs in a digital radio.

While the performance of a digital transceiver may exceed that of a traditional design, the basic goals are the same. We wish to fabricate a receiver with good sensitivity and selectivity, the maximum dynamic range and minimum distortion. The transmitter must produce a low-distortion, spectrally pure signal. The frequency stability and tuning resolution should not impose undue limitations on operation.

In this age when DSP hardware capabilities are finally exploiting advances in theory discovered over the last 40 years, we can indeed set our sights quite high! As we begin, let's define the challenges facing us in receivers, so as to assess the influence of IF-DSP technology.

[1]Notes appear on page 37.

PO Box 4074
Sedona, AZ 86340
e-mail **dsmith@sedona.net**

A Receiver: All that Gain and a Whole Lot More

Superheterodyne receivers have been around for awhile, and until DSP hardware can achieve sufficient speed and dynamic range to digitize signals straight from the antenna, we'll all continue to use them. The main advantage of a superhet is that signals are converted to a fixed IF that provides most of the gain and selectivity. To avoid spurious responses, multiple frequency conversions are common. We ought to recognize, however, that minimizing the number of conversions also diminishes the number of oscillators and, therefore, the number of possible internal signals, or "birdies."

At some stage, we'll digitize some signals and perform filtering and other signal processing. We want this point to be as close to the antenna as possible, so we must look at the frequencies, bandwidths and dynamic ranges available in DSP components before choosing an IF. We can eliminate many traditional analog signal-processing stages if we digitize signals ahead of the point that expensive crystal or mechanical filters previously occupied. Our first trade-off is between high-speed analog-to-digital converters (ADCs) and the costly filters they would replace. This decision is driven mainly by cost, although issues of current consumption and processing power definitely come into play.

The final compromise also depends on the performance levels we expect to achieve. For example, many excellent ADCs are quite capable of digitizing signals directly from the antenna: their sampling rates are fast enough for the job, but their dynamic ranges are narrow. As we'll see below, HF receivers must handle a tremendous range of input signal levels without

flinching! So before we can make even this first decision about the receiver's conversion scheme, we must think about dynamic range: What is it, and how much do we want?

Receiver Dynamic Range (DR)

It's every receiver's job to produce a useful replica of the transmitted information and reject all other signals. In today's crowded HF bands, this is an increasingly difficult task! The desired signal might be quite weak, so we need good sensitivity and lots of gain without introducing excess circuit noise.

Sensitivity must be specified as a function of the bandwidth of interest, because we're trying to copy a narrow-bandwidth signal in the presence of noise, which exists at every frequency! In the specified bandwidth, a signal received at the antenna terminals has a certain signal-to-noise ratio (SNR). We fight to preserve this SNR throughout the receiver. Electronic circuits introduce some noise, however. The ratio of the input SNR to the output SNR of a receiver is referred to as its *noise figure* (NF), and is expressed in decibels.

Originally explained by Einstein in 1905, Brownian motion of atoms and free electrons in any conductor produces an available noise power[1,2] (in watts) of:

$$P_{noise} = kTB \qquad \text{(Eq 1)}$$

Where k is Boltzmann's constant, 1.38×10^{-23}, T is the absolute temperature in Kelvins, and B is the bandwidth, in hertz. "Plug and chug" on these numbers, and you'll find that at room temperature (293 K) and in a bandwidth of 3 kHz, this power is –139 dBm or 12.1 *attowatts!* This quantity (12.1×10^{-18} W) represents the minimum discernible signal (MDS) in a perfect receiver using a typical voice bandwidth. Note that as the temperature decreases, the possibilities increase linearly; a receiver operating in a liquid-nitrogen bath is a real gem! Atmospheric and cosmic noise are usually much greater than this theoretical limit, however.

The best HF receivers today have NFs around 7 dB. When noise power equals signal power, the output SNR is 0 dB, and the input signal level is:

$$P_{in_{0dBSNR}} = -139 + 7\,\text{dBm} = -132\,\text{dBm} \qquad \text{(Eq 2)}$$

We define this MDS level as the lower limit of the receiver's dynamic range. (In a 50 Ω resistor this power corresponds to 0.056 μV.—*Ed.*) It's not so easy to find the upper limit of the dynamic range. Because of the manifold ways receivers degrade at high input levels, we'll define several dynamic ranges, one based on each of these.

Receiver Overload: Let Me Count the Ways

Normally, overload phenomena involve large, off-channel signals. Of course, it's also possible to overload on a very strong *desired* signal. For most modern receivers, this level would be so high that radio communication wouldn't be necessary; you could just shout out the window!

Large-signal performance is typically characterized[5] by measuring the following effects:

- Third-order intermodulation distortion (IMD)
- Second-order IMD
- "Blocking," or desensitization
- In-band IMD

Let's examine methods for each of these measurements along with the strengths and weaknesses of current methods.

IMD Dynamic Range and Intercept Point

To measure IMD dynamic range, we inject two off-channel signals of equal amplitude and measure the degradation of receiver performance. Degradation comes in the form of an undesired, on-channel signal produced by the mixing of the off-channel signals. We increase the off-channel signal levels until the on-channel signal power equals the noise power. This is the definition of MDS given above.

We define the IMD dynamic range to be the ratio of this off-channel signal power to the MDS power, expressed in decibels. In the ARRL method for third-order IMD, one interfering signal is placed 20 kHz from the center channel and another 40 kHz from center. The third-order intercept point (IP_3) is calculated by assuming the receiver distortion obeys a perfect cube law: For every decibel increase in interference, the third-order IMD product will increase 3 dB, and the difference will increase by 2 dB. IP_3 is extrapolated, therefore, by adding half of the third-order IMD dynamic range to the interference level obtained in the measurement above:

$$IP_3 = \left(\frac{IMD\,D\,R}{2}\right) + P_{QRM} \qquad \text{(Eq 3)}$$

This is supposed to be the level where the third-order IMD product is equal in amplitude to the interference. Were we to actually inject interference of this level, however, we might find a *real* IP_3 much higher; receivers seldom obey perfect cube laws as they're predicted to do! This normalized procedure is a good basis for comparison, though.

In the second-order test, we inject two non-harmonically related signals and look for the undesired product at the sum or difference of the frequencies. IMD dynamic range is measured as above, and IP_2 is extrapolated by assuming the receiver obeys a perfect *square law*. For every decibel of increase in the interference, the second-order product increases 2 dB, and the difference increases by 1 dB:

$$IP_2 = (IMD\,D\,R) + P_{QRM} \qquad \text{(Eq 4)}$$

How can the receiver obey two apparently conflicting laws at the same time?! In the second-order case, we're mixing the two *fundamentals* of the interference; whereas, in the third-order case, we're mixing the fundamental of one with the internally generated second harmonic of the other.

Note that when we add two fundamental signals, the result is always greater than twice the frequency of one of the signals. For this reason, our second-order performance can be improved by using *half-octave* band-pass filters ahead of the receiver front end. Such filters—when switched or tuned as the receiver changes frequency—always attenuate one of the interfering signals, reducing the deleterious effects.

"Blocking" Dynamic Range

In this measurement, we inject a *single* off-channel source, and look for some degradation in the on-channel performance. In the ARRL method, the output power from a single on-channel input signal is monitored. The interference, 20 kHz away, is increased until the desired output power either increases or decreases by 1 dB.

A decrease is supposed to indicate that some stage or other is saturating, while an increase results in a "noise-limited" measurement. The *blocking dynamic range* (BDR) is calculated as the ratio of the interference power in the measurement above to the MDS power, expressed in decibels.

In reality, saturation seldom occurs in modern receivers before the noise takes over. This noise is the result of *reciprocal mixing*, wherein the interference mixes with the phase-noise sidebands of the LO to produce in-band noise. A state-of-the-art synthesized

LO has phase noise in a 3 kHz BW, and at a 20 kHz offset, of around 100 dB below its injection level. Were the BDR measured using the SNR instead of the average output power, we could call it desensitization or *"desense."* It would be on the order of 100 dB, and would be solely a measure of the synthesizer phase noise. This number is quite a bit lower than that usually obtained with the ARRL method.

The difference becomes evident when trying to measure an IF-DSP receiver with a *digital AGC* system. Such a system holds the peak desired output level constant, and as the SNR degrades, the average output power *decreases!* In a conventional receiver (all other things being equal) the SNR would be identical, but the output power would *increase* because of the added noise. The peak-to-average ratio of noise is high, so monitoring the average or RMS output power wouldn't indicate an increase until much more interference power were added.

To correlate the SNR method with the ARRL method, we might consider using degradation of the output SNR as our criterion, as in the EIA standard. The degradation level could be chosen to equate the new measurements to existing BDR measurements of known receivers. Let's face it, reciprocal mixing gives the most trouble these days. If a blocking measurement is still desired, we ought to use the *peak* output level, not the RMS.

In-band IMD

This is a measure of distortion produced by a receiver when the only signals present are inside the desired passband. Current ARRL methods call for a two-tone input with a frequency separation of 100 Hz. This is excellent. It's roughly the natural impulse frequency of the human voice system. The IMD product levels are examined relative to one of the tones. The AGC speed, if adjustable, is set to its fastest setting.

Digital AGC systems can cause problems here, because they are capable of *very* fast attack and decay times. If the decay time is set fast enough, clearly the two-tone will be subject to extreme distortion; it'll begin "flat-topping." It doesn't make sense to defeat the very system designed to prevent the thing being measured!

Preamplifiers and Dynamic Range

It's obvious that to achieve the best sensitivity, some gain ahead of the first mixer is required. If this gain stage has a low NF, we can improve the sensitivity by almost the amount of the gain. This extends the receiver's dynamic range on the low end.

It's difficult to make up the difference on the high end, though. The large-signal handling will degrade by *at least* the amount of the preamplifier gain, so the dynamic range is generally reduced. Also, notice that dynamic range is just the ratio of maximum and minimum signals that can be handled, and it says almost nothing about actual large-signal handling capability! One receiver might have a greater dynamic range than another, and still have a poor IP. Its sensitivity may be excellent, but it might not be a good large-signal performer.

AGC

We've seen that HF receivers must handle very weak signals (–132 dBm) and strong signals that may approach +20 dBm, near the IP_3. Expressed this way, the dynamic range can exceed 150 dB! As we expect the output level to remain relatively constant and the distortion to stay within limits, a gain-control system is necessary. We must keep analog stages linear, so an analog AGC system is mandatory.

We intend to provide the final selectivity in our receiver using digital filters, as this eliminates the need for expensive crystal or mechanical filters. So it follows that some of the signals we digitize will be undesired —this raises a problem: The digital filters will remove the interference, but the analog AGC will still act on the total bandwidth! A strong interfering signal will reduce the analog gain, as it must, and the level of our desired signal will fall as well. This is where the digital AGC system comes in.

Digital AGC Algorithms

We decide that to keep the desired signal's output level constant, we need a system that measures the ratio of total digitized signal energy to desired signal energy.[1,4] When the interference increases, this system will compensate for the reduction in gain caused by the analog AGC. The effect will be to hold the desired signal's peak level constant. Now we must determine how we're going to measure the critical ratio, and how and when to make adjustments in the gain compensation.

Clearly, the digital gain-compensation algorithm must use two data as inputs: the ratio of total signal level to desired signal level, and the actuation or amount of analog AGC. The ratio of the amplitudes is easily calculated by the DSP system; it need only compare the peak digitized input level with the peak output level after filtering. This isn't quite the whole solution, however, because when the desired signal decreases, the system can't tell if it was because of interference-caused analog gain reduction, or because the other station just stopped transmitting!

So we arrange to monitor the analog AGC voltage in order to find out what it is doing. It turns out we don't need to know the amount of analog gain reduction if we can adjust the digital gain fast enough. We'll examine the analog AGC to detect when the gain is decreasing rapidly and when the amplitude ratio is increasing rapidly, then quickly boost the digital gain until the desired output level is maintained. Notice that both the analog and digital AGC systems maintain a fast-attack characteristic in all situations.

In practice, this system works quite well; the digital AGC decay time can be continuously adjusted as desired. On-channel signals are digitally boosted by the amount necessary to keep the peak output constant. The main drawback is that the dynamic range of the ADC system limits the available digital gain.

ADC Limitations

We learned in Part 1 that the dynamic ranges of ADCs are limited by the bit-resolution, speed and input frequencies of the devices. Looking at current technologies, we see that 16-bit ADCs are available with 96 dB of dynamic range; the input frequency ranges of these are confined to just above audio, however. A low-frequency IF poses a problem only in image rejection because it's difficult to filter out responses close to the desired signal. The trade-off here is between the digital dynamic range at hand for gain boosting and the frequency of our last IF.

Oversampling ADCs can be advantageous in our design, because they spread quantization noise over large bandwidths, then apply digital filters of their own to eliminate most of the noise. Sampling rates and IFs can be chosen so that the ADC filter aids the selectivity of the receiver. We discover that 40 kHz is a good last IF for this reason. The Analog Devices AD7722 is chosen as our ADC.[15]

We further decide that we can apply up to 60 dB of digital gain boost using this device, since its dynamic range is 96 dB; the output SNR cannot degrade to much less than:

$$SNR_{out_{min}} \approx 96 - 60\,dB = 36\,dB \quad \text{(Eq 5)}$$

because of the ADC system. We'll adjust the analog AGC so that it provides a peak input level near the maximum input allowable for the device. Notice that exceeding the maximum input level of the ADC results in instant, catastrophic degeneration of the output signal. ADC overload is the one thing we can't tolerate. So, we allow a few decibels of *headroom* in setting the analog AGC operating point.

The First IF

Now that we've decided on a *last* IF, it's time to figure out how we get there from our RF range of up to 30 MHz. Selection of a first IF depends on the location of spurious responses, as well as availability and cost of components. Spurious and image problems are greatly reduced if the first IF is above the highest RF. *Up-conversion* has been standard in HF receivers for some time now.

Several popular IFs offer cost advantages because of commonality with other radio services.[1,2] These include, but are not limited to:

- 45 MHz—popular in cellular phones
- 70 MHz—the standard UHF and microwave IF
- 75 MHz—aviation service marker beacon

An IF above twice the highest RF is favorable because it eliminates second-harmonic spurious responses. Eg, a strong signal appearing at 22.5 MHz generates a second harmonic at 45 MHz. An IF above 60 MHz is above that harmonic and as a high-side-injection first LO will cover less than half an octave, which simplifies its design.

We can't increase our IF arbitrarily, though; the cost of crystal filters becomes prohibitive above 75 MHz, and the loss and instability of surface-acoustic-wave (SAW) and other VHF filter technologies make them unattractive. In addition, the synthesizer would be forced to higher frequencies, which would increase phase noise and associated reciprocal-mixing problems.

If we want the receiver to cover the range below 500 kHz, we must limit the phase noise. This energy enters our IF directly because of imbalance in the first mixer as the LO approaches the IF. We decide that 75 MHz is a good choice, and now we're ready to draw a block diagram of the receiver.

The Block Diagram

See Fig 1. Beginning at the antenna, we place a bank of half-octave bandpass filters (to extend the second-order IMD dynamic range as described above). These are switched by PIN diodes or relays. Next, a preamplifier with a gain of 15 dB and noise figure of 3.6 dB is used. This amplifier can be switched out using relays, or a 20-dB fixed attenuator can be inserted. Then, we select a high-level, double-balanced mixer to translate our signals to the first IF

This first mixer is critical, because it's likely to determine our IMD dynamic range. We can expect an IP_3 several decibels above the injection level, so we choose a mixer designed for a +17 dBm LO. We anticipate that the extra energy will warrant careful shielding, filtering and isolation to prevent birdies and LO leakage to the antenna.

While filters at the mixer ports mitigate these problems, they can have an unexpected consequence: degraded IMD performance. A mixer is an inherently nonlinear device that generates harmonics of all signals entering it, like it or not! So, although our LO might be spectrally pure, its harmonics and those of the RF and IF will be present. When a mixer port is terminated at its characteristic impedance, this harmonic energy is absorbed by the termination. A filter, however, reflects some of it back into the mixer, where it may add to IMD.

One elegant solution involves "idler" filters. These networks provide a broadband termination impedance *and* the filtering we need. Fig 2 shows the use of idler filters. The filter passing the desired band (eg, a low-pass) is designed to be singly terminated.[6] A similarly designed high-pass filter is parallel connected, and terminated in the characteristic impedance—usually 50 Ω. The mixer then sees a relatively constant broadband load. A trade-off exists between the complexity and cost of these networks and the IMD degradation we'd otherwise suffer.

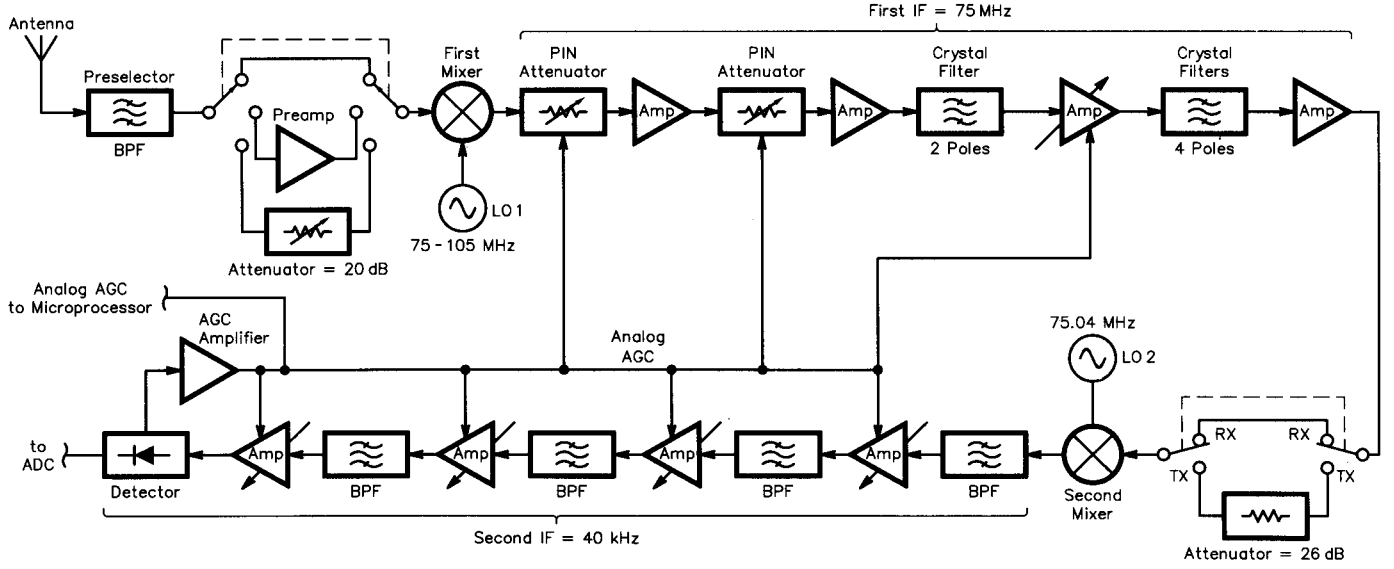

Fig 1—IF-DSP receiver block diagram.

High-Side versus Low-Side

We can elect to put the LO above or below the IF and take either the sum or difference frequency as our 75 MHz signal. High-side injection (taking the difference frequency) is attractive because the required LO range of 75 to 105 MHz is less than half an octave. With low-side injection, the LO range would be 45 to 75 MHz, nearly an octave; this would generate LO harmonics within an octave of the IF. We therefore choose high-side injection.

We'll build ourselves a first LO that covers the range in less than 1 Hz steps by combining direct digital synthesis (DDS) technology with a PLL. The second LO will have a fixed-frequency. We achieve drift cancellation by using difference mixing in both stages. I'll describe the synthesizer design in a future segment.

The First IF Strip

Following the first mixer we must have some gain to compensate for losses in the crystal filters ahead. This stage must have a low noise figure and moderate gain, yet must handle some very large signals to avoid degrading the IMD performance. It also ought to provide a good broadband termination at both its input and output—a pretty tall order so far!

Transistor circuits have been described in the literature,[1,2] which use a combination of voltage and current feedback to achieve simultaneous noise and impedance matching. In combination with FETs of moderate power rating, the performance can be impressive. Fig 3 gives an example. In this circuit, several parallel JFETs increase the current capacity while keeping cost under control. The feedback is taken from a transformer with a turns ratio designed to provide constant 50 Ω input and output impedances. The amount of feedback sets the gain at 6 dB. The measured input impedance is plotted as Fig 4 on a Smith chart normalized to 50 Ω.

The noise figure is 1.5 dB, and output IP_3 is +35 dBm! Clearly, this amplifier fills the bill. It won't affect our receiver's IP_3. Two of these stages are used before any narrow-band filters.

Gain control of transistor stages by varying the bias is impractical because of linearity problems, so PIN diode attenuators are commonly used. Fig 5 shows a typical attenuator circuit. We intersperse two of these attenuators to provide sufficient control range. This circuit must also provide constant impedance, so it uses a combination of series and shunt diodes. Bias voltage and current levels are set so that as the series diode is switched off, the shunt diodes switch on. The bias must be strong enough to avoid excessive IMD. The control range of each attenuator is 48 dB, and the IP_3 is +38 dBm.

Three two-pole monolithic crystal filters provide the "roofing" in our design, with a single MOSFET stage providing 20 dB of gain in their midst. It's advantageous to distribute component gain and loss throughout the receiver with the best noise figure and AGC effectiveness in mind. We must still maintain good SNR when gain control is applied, so concentrating control in one place is unwise.

After the roofing filters, an emitter-follower feeds the second mixer through another PIN diode attenuator. Since we intend to use the first IF in the transmit (TX) mode, this attenuator will provide precise gain reduction to restrict mixer-generated spurious signals while transmitting. It is set for zero attenuation while receiving (RX).

Fixed injection at a frequency of 75.040 MHz translates signals directly to the 40 kHz second IF. Notice that the second mixer can be a lower-level device, because AGC and filtering limit the amplitudes seen at this stage.

The Second IF Strip

Now that we're down to a frequency our ADC can handle, we have to provide enough gain so that the IF output level is near the maximum ADC input. This will allow the greatest dynamic range for digital AGC operation as described above. Getting gain at 40 kHz isn't a problem, but we also want additional analog AGC range here. Finally, we have to further bandwidth-limit the output so that *aliasing* cannot occur.

As described in Part 1, aliasing results when input bandwidth exceeds half the sampling frequency. Once incurred, nothing can be done to alleviate it! We want the sampling frequency as low as possible to minimize

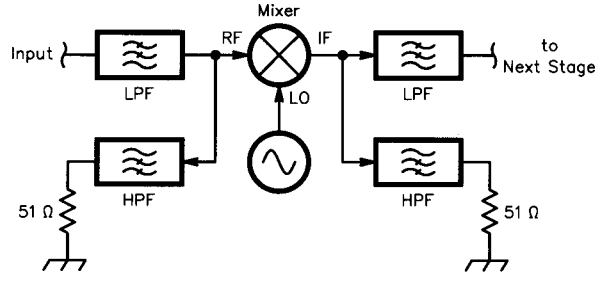

Fig 2—Double-balanced mixer employing 'idler' filters.

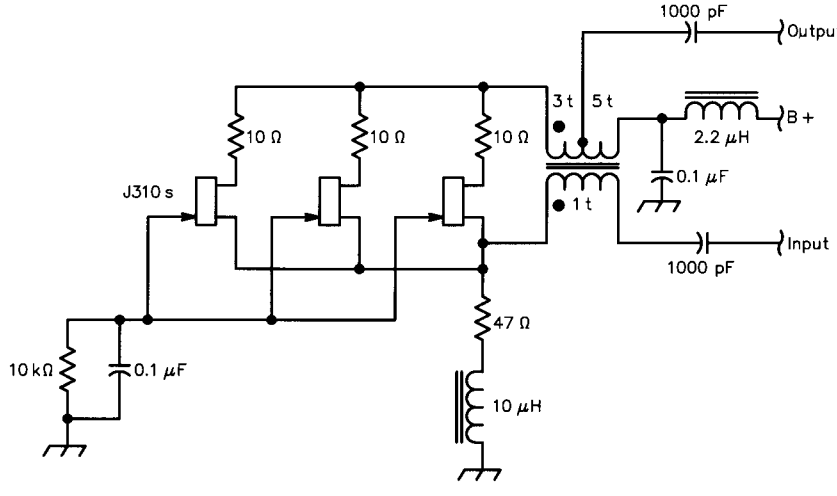

Fig 3—75 MHz IF amplifier stage.

DSP "horsepower" requirements. Narrow-band FM, using bandwidth of 15 kHz, dictates the minimum passband width. This means our sampling frequency must be at least 30 kHz. To ease the design of the filters, we set the sampling frequency somewhat higher than this.

Narrow-band dual-gate MOSFET amplifier stages like that shown in Fig 6 are used in the second IF. These provide uncomplicated gain control and allow us to implement LC filtering in the 15 kHz BW. Since the ADC's maximum input level is 3 V_{P-P}, and its input impedance is high, more than 90 dB of voltage gain is required! Four stages of filtering and amplification get us what we need.

Analog AGC

The analog AGC system is implemented using the traditional detector and amplifier scheme. Gain reduction is implemented first in the later stages, then in the front-end components, so that the SNR can continue to increase with input signal levels. AGC voltage is fed to the DSP so that digital AGC can keep the final output level constant. It's fascinating to watch the second IF output level gyrating with interference while copying a weak signal on an adjacent frequency!

Summary

We've seen that in an IF-DSP transceiver, we had to start the design at the "back end" because of the limitations of available DSP components. This led us to certain decisions about the second IF, but we acknowledge that the rest of the design still resembles that of a conventional transceiver. In the next segment, we'll ex-

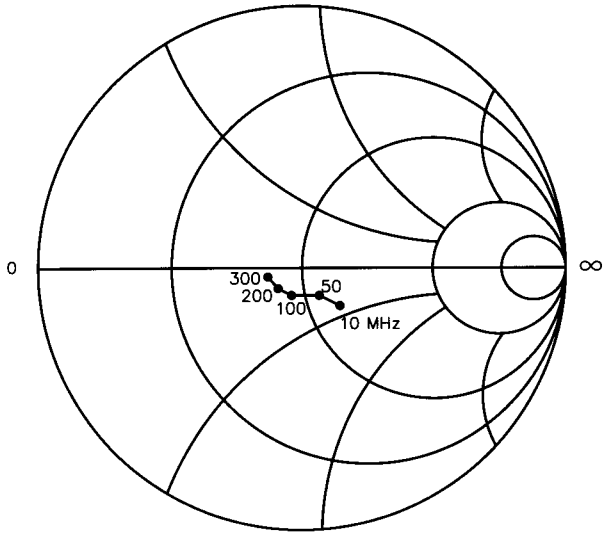

Fig 4—Input impedance of amplifier stage *versus* frequency.

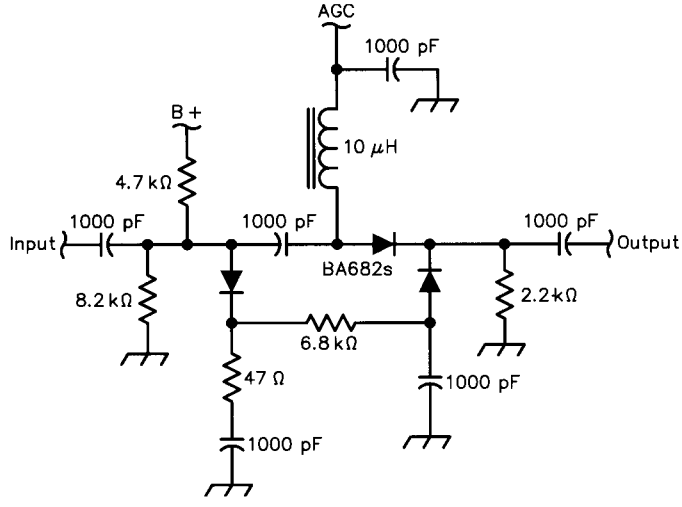

Fig 5—PIN diode attenuator circuit used in 75 MHz IF

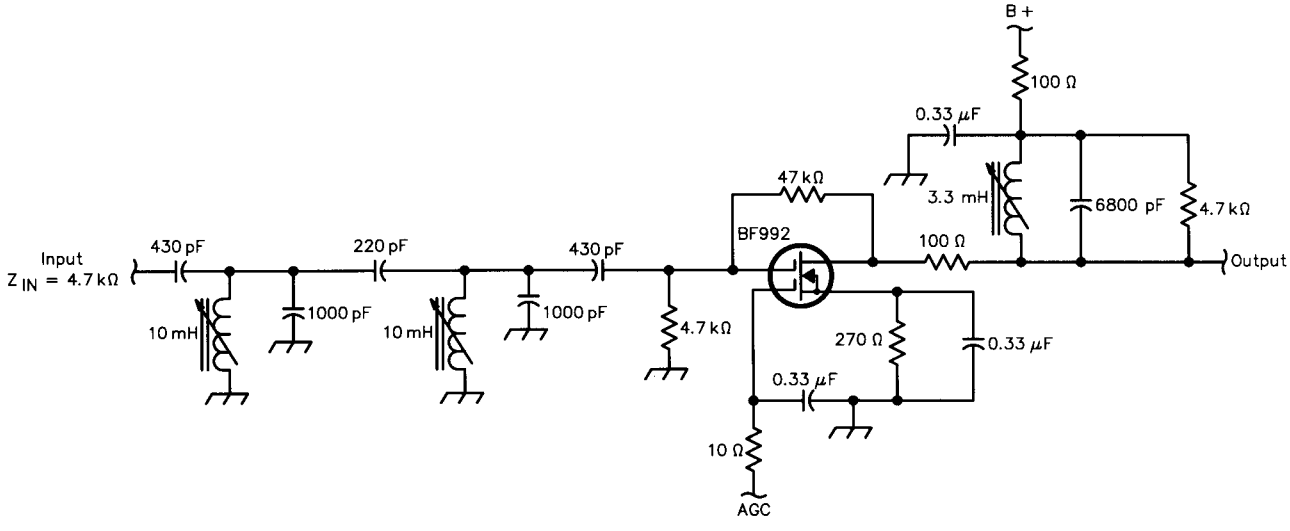

Fig 6—40 kHz IF BPF and amplifier stage.

plore a state-of-the-art synthesizer design to see how DDS/DSP techniques help us get around the bands with the greatest of ease!

The Synthesizer: Excursions in the Frequency Domain

Synthesizers have come a long way since first becoming popular in HF transceivers in the '70s. Availability of components then lagged well behind the development of theory. Now, hardware capabilities have nearly caught up—which is the case for DSP in general—and are driving the very rapid advancement of HF equipment we are now experiencing. Although the gap has narrowed, we're still far from building an entirely digital, direct-conversion transceiver.

Paralleling breakthroughs in the microprocessor and data acquisition fields, progress in direct digital synthesis (DDS) has enabled performance levels only dreamed of a decade ago. Virtually all new designs profit from this technology.

Design Goals

In the previous segment, we defined the frequency ranges to be covered by our synthesizer design, set limits for phase noise and spectral purity and established the output levels. Now let's consider three other critical requirements: frequency stability, lock time and tuning resolution.

Amateurs are free to operate anywhere within large frequency bands, so it might seem that frequency accuracy isn't very critical. Nonetheless, prevalent narrow-bandwidth communication modes require good frequency stability, and operators have come to expect excellent stability from their rigs. It is reasonable to expect ±20 Hz stability over an anticipated temperature range of −10 to +50°C. We'll address the issues of long and short-term stability below.

We wish to attain a tuning speed that doesn't impose limitations on typical use. "Cross-band," or split-frequency operation ought to be considered. We set an upper limit of 20 ms for a frequency shift of ±600 kHz. Lock time is defined as the time required to settle within the above accuracy limits.

The smallest frequency steps should be such that they don't impede performance. A minimum step of 10 Hz used to be good enough, but now certain digital modes benefit from smaller steps. In addition, we'll discover that the narrow digital notch filter described in Part 1 of this series requires tuning within several hertz to achieve the best null! We therefore set 1 Hz steps as our design goal. Table 1 summarizes the work for this major subsystem.

DDS Meets PLL, Object: Matrimony

DDS synthesizers achieve the fastest lock times and—when using a crystal-derived clock—the least phase noise of available methods.[4] It's also common knowledge that their spurious outputs can be excessive, especially as the output frequency approaches one half of the clock frequency. To build a DDS covering 75 to 105 MHz, directly, would be quite a feat!

Reasonably priced devices currently use clocks to about 60 MHz, which limits outputs to well under 30 MHz. Although it's possible to multiply or mix the DDS output up to the proper range, these strategies quickly become complicated and suffer from spurious problems. As a reference input to a standard PLL, however, a DDS can provide the performance we want.

LO_1 Block Diagram

In this design, we operate a VCO at 75 to 105 MHz and phase lock it to a DDS output at $F_{REF} = F_{VCO} / 100$, as shown in Fig 7. F_{REF} is high to get fast lock times, and a DDS clock frequency much higher than F_{REF} reduces spurious content. The ratio of F_{VCO} / F_{REF} is important, because any noise and spurs in F_{REF} are multiplied by the PLL, *within the PLL bandwidth*, by the factor:

$$N = 20\log\left(\frac{F_{VCO}}{F_{REF}}\right) = 40\ \text{dB} \qquad \text{(Eq 6)}$$

Our final output-spurious level is set by the DDS output purity.

DDS Spectral Purity

A DDS is a system that generates digital samples of a sine wave and converts them to an analog signal using a digital-to-analog converter (DAC; see Fig 8). Spurious outputs from such systems are caused by phase and amplitude inaccuracies in the digital and analog circuitry. First, let's look at the errors of the digital portion. These are analogous to the effects of digital word size and truncation of numerical results explained in Part 1.

In a DDS chip, a phase counter increments at each clock pulse, and the phase information is used to look up a

Table 1—Frequency Generation Requirements

LO_1	
Frequency Range	75 to 105 MHz in 1 Hz steps
Output Level	+17 dBm ± 2 dBm
Lock Time (large step)	≤ 20 ms
Phase Noise	≤ −132 dBc/Hz @ F_c ±20 kHz
LO_2	
Frequency	75.040 MHz
Output Level	+17 dBm ±2 dBm
Phase Noise	≤ −145 dBc/Hz @ F_c ±20 kHz
General	
Frequency Stability	≤ 20 Hz over −10 to +50°C ±40 Hz/year (aging)
Harmonics, spurious	≤ −73 dBc

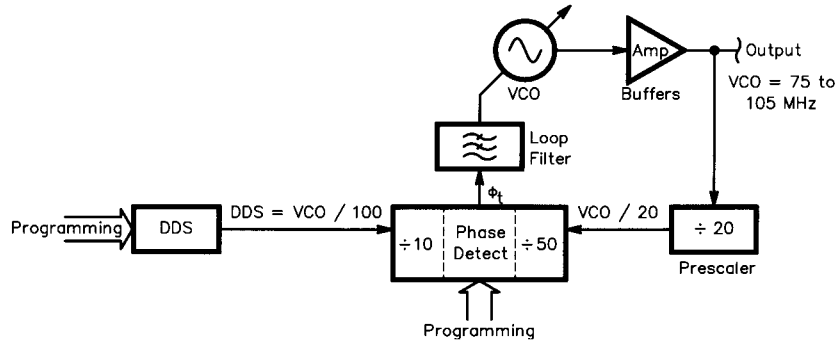

Fig 7—LO_1 block diagram.

sine-wave amplitude from a table. Since the phase is represented by a binary number, with a fixed number of bits, p, errors can develop because resolution beyond p bits isn't possible—the number is *truncated*. The effect is phase-modulated (PM) spurs in the DDS output.

Further errors are related to the output resolution of the look-up table. The table values representing the amplitudes are truncated to some number of bits, a. This mechanism produces amplitude-modulated (AM) spurs in the output.

After Cercas,[7] maximum PM spurs could be:

$$P_{PM\ spurs} = -(6.02p - 5.17)\,\text{dBc} \qquad \text{(Eq 7)}$$

maximum AM spur levels could be:

$$P_{AM\ spurs} = -(6.02a + 1.75)\,\text{dBc} \qquad \text{(Eq 8)}$$

In the analog signal we generate, the DAC introduces more AM spurs, harmonics and IMD because of inherent nonlinearities (see Part 1). Spurs are also likely at the clock frequency and its harmonics. A ninth-order elliptical low-pass filter after the DAC removes harmonics and many of the spurs.

It turns out we can eliminate all AM spurs by squaring the DDS output at the external reference input of the PLL chip! We can do nothing about the remaining PM spurs, so we'd better keep them 40 dB below the desired output spurious level, or:

$$P_{(PM\ spurs)} \le -73 - 40\,\text{dBc} = -113\,\text{dBc} \qquad \text{(Eq 9)}$$

This means a bit-resolution decided by solving Eq 7 for p:

$$p \ge \frac{-113 - 5.17}{-6.02} \qquad \text{(Eq 10)}$$
$$\ge 19.63$$

The Harris HSP45106 has a 32-bit phase accumulator, 20-bit-address sine look-up table and 16-bit output resolution. Since the AM spurs will disappear, a 10-bit DAC is sufficient; we chose the Harris HI5780.[14]

The resulting DDS output feeds the reference input of the Motorola MC145159 PLL synthesizer IC, where the wave is squared and divided by 10 to establish a phase reference of 75 to 105 kHz (F_{VCO}/1000). This should provide very fast lock times!

Frequency Resolution

As stated above, the PLL will multiply the DDS frequency by 100. To get our 1 Hz output steps, therefore, we must tune the DDS in 10 *millihertz* steps! With a 32-bit phase accumulator, DDS step size will be:

$$df_{DDS} = \frac{f_{clk}}{2^{32}} \qquad \text{(Eq 11)}$$

A clock frequency around 10 MHz exceeds the goal by a factor of 4, producing a step size of 2.3 millihertz!

The VCO

A Colpitts design was selected, with a resonant tank circuit switched in eight bands using PIN diodes and three capacitors with binary-weighted values, as shown in Fig 9. The VCO is tuned using back-to-back varactor diodes, in order to achieve maximum voltage across the tank.

The advantage of restricting the frequency range to several bands is one of decreased phase noise. According to

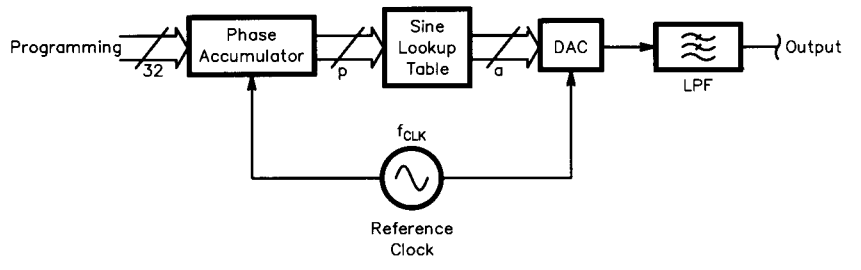

Fig 8—DDS block diagram.

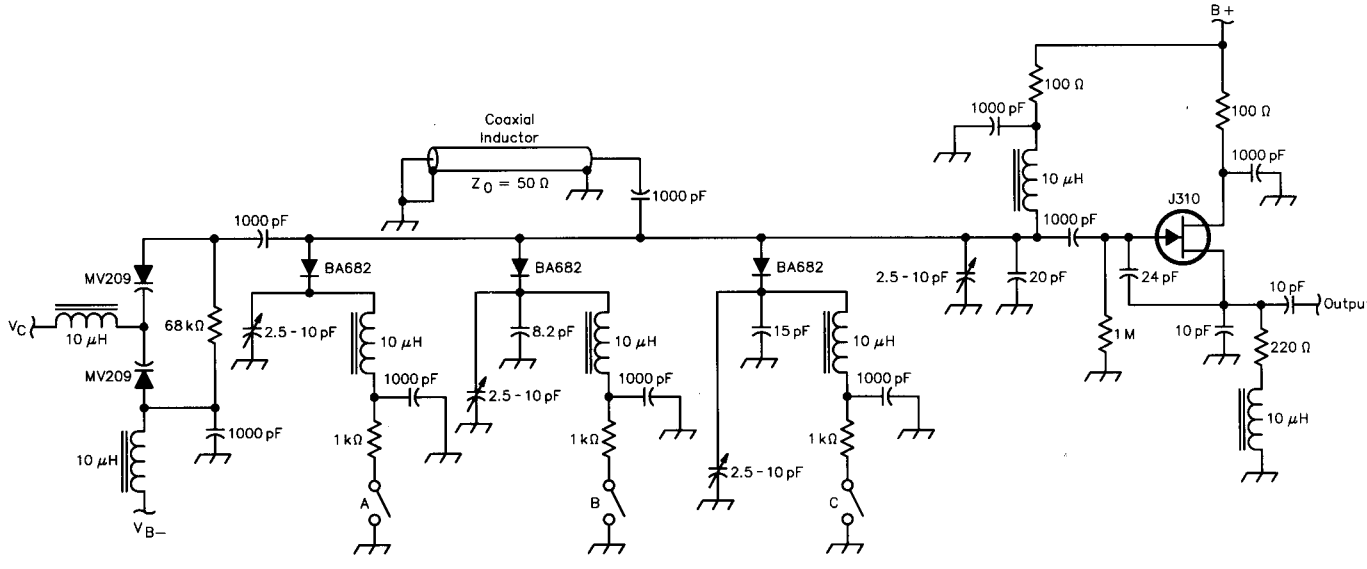

Fig 9—75-105 MHz band-switched VCO for LO_1.

Leeson's model,[8] the phase noise of a VCO may be characterized by the equation:

$$PN_{df} = 10\log\left\{\left(1+\frac{f_c^2}{\left[2Q_{load}df\right]^2}\right)\left(1+\frac{f_f}{df}\right)\left(\frac{nkT}{2P_{avg}}\right)+\frac{2kTR_vK_\phi^2}{df^2}\right\} \quad \text{(Eq 12)}$$

where

PN_{df} = SSB phase-noise power relative to total power, in dBc/Hz

f_c = center frequency

df = frequency offset of noise measurement

Q_{load} = loaded Q of tank circuit

f_f = flicker frequency of active device

n = noise factor of active device

k = Boltzmann's constant, 1.38×10^{-23}

T = temperature in Kelvins

P_{avg} = average power at input to active device

R_v = equivalent noise resistance of varactors

K_ϕ = VCO sensitivity in Hz/V

Great, but what does this incredibly complex equation mean?

First, it means that the phase noise goes up about 6 dB every time we double the center frequency. Second, the phase noise goes *down* 6 dB when we double the loaded Q of the tank circuit. These effects aid each other, because it's more difficult to achieve high Qs at higher frequencies. Third, more power at the input of the oscillator device is better. Finally, the phase noise will approach a lower limit established by the other factors as VCO sensitivity is decreased.

The inductor for the resonant circuit is critical in determining Q. It can also make the VCO "microphonic," (susceptible to vibration, if not solidly mounted). We can borrow from transmission-line theory to address these issues.

A section of transmission line, shorted at the far end, and less than λ/4 long, looks like a shunt inductor[12] having a reactance of:

$$L = Z_0 \bullet \tan\beta L \quad \text{(Eq 13)}$$

where Z_0 is the characteristic impedance of the line, and βL is its electrical length in radians. Eg, a 1 λ line is 2 π radians in electrical length.

When used in this way, the SWR in the line section will be quite high, making it "lossy" and degrading the Q. By keeping it short and using a carefully selected coaxial cable, however, microphonics and any chance of radiation are virtually eliminated! The shield is simply soldered to the circuit board along its entire length; it can even be coiled into a very compact size.

Output Buffering

To maintain our high resonant-circuit voltage, we want to draw as little energy from the oscillator as possible. On the other hand, we must take enough that the noise figure of the buffer amplifiers doesn't become a problem. Two buffer stages are used here. They also provide isolation; changes in load impedance caused by the presence of large receiver input signals would otherwise "pull" the VCO, inducing FM.

The PLL: Closing the Loop

Since the phase reference is at $F_{VCO}/1000$, the VCO output sample passes through an external ÷20 prescaler and a ÷50 *inside* the PLL IC before phase detection. The high comparison frequency produces fast lock times by allowing a large loop bandwidth, while making it easier to avoid "reference spurs," or phase modulation of the VCO at the reference frequency.

The MC145159 PLL chip uses an analog sample-and-hold phase detector, which further eases loop filter design. Details of this technique can be found in the data sheet,[13] but let's note its two major advantages: It has programmable gain, and it outputs an analog level that already resembles the desired loop control voltage. We level-shift and amplify this 0 to 5 V signal until it suits the VCO's tuning range and then filter it to remove noise introduced by the amplifier and sample-and-hold circuits.

The device also outputs a coarse "frequency steering" signal, which is used to move the control voltage by large amounts when phase lock has been lost. When the phase at the reference frequency is within 2 π radians of lock, this output goes high-impedance, and the analog phase detector takes over.

LO_2: A Voltage-Controlled Crystal Oscillator (VCXO) Design

The second LO is fixed in frequency, but it needs to be phase-locked to the master reference so that some "drift cancellation" can be obtained, as explained further below. It should have very low phase noise so as not to add to the total. For these reasons, a VCXO is best.

At 75.040 MHz, an overtone "rock" must be used. A useful property of overtone oscillators is that they produce stability roughly equal to that of the fundamental mode. Because of this, fifth-overtone operation was initially considered. The tuning range must allow the VCXO to track the reference over its limits, though, and a third-overtone design is much more "pullable." The addition of a resistor across the rock aids in achieving sufficient range without lowering the Q enough to affect output noise.

Fig 10 is a modified Pierce circuit. Feedback is taken from a winding on the drain transformer and passed through the series-resonant crystal. Drain capacitors resonate the transformer at the output frequency, and transform the impedance down to 50 Ω.

A PLL using the MC145159 locks LO_2 to the reference oscillator. Two buffer amplifiers are again used.

The Frequency Reference and Drift Cancellation

As both LOs are locked to the same frequency reference, and difference mixing is used at both IFs, some degree of drift cancellation is obtained. Let's look at how this works and calculate the magnitude of the effect.

The DSP sees a 40 kHz IF signal, which has undergone two frequency translations. For a single RF input signal at f_{RF}, the mixing is expressed by the equation:

$$\begin{aligned} f_{IF} &= f_{LO_2} - \left(f_{LO_1} - f_{RF}\right) \\ &= f_{RF} - \left(f_{LO_1} - f_{LO_2}\right) \end{aligned} \quad \text{(Eq 14)}$$

Now each LO varies with a change in reference frequency according to:

$$\frac{df_{LO}}{df_{REF}} = \frac{f_{LO}}{f_{REF}} \quad \text{(Eq 15)}$$

$$df_{LO} = f_{LO}\left(\frac{df_{REF}}{f_{REF}}\right) \quad \text{(Eq 16)}$$

Substituting, we can write:

$$\begin{aligned} df_{IF} &= df_{RF} - \left(df_{LO_1} - df_{LO_2}\right) \\ &= -\left(f_{LO_1} - f_{LO_2}\right)\left(\frac{df_{REF}}{f_{REF}}\right) \end{aligned} \quad \text{(Eq 17)}$$

assuming the RF input isn't changing, only the reference. We see that the total error is proportional to the difference in the LO frequencies, which is approximately equal to the RF:

$$\left(f_{LO_1} - f_{LO_2}\right) \approx f_{RF} \qquad \text{(Eq 18)}$$

At low RFs, the drift cancellation is nearly perfect. As the RF increases, any error increases linearly to the maximum, at 30 MHz, of:

$$df_{IF} \approx 30 \times 10^6 \left(\frac{df_{REF}}{f_{REF}}\right) \qquad \text{(Eq 19)}$$

To get the ±20 Hz accuracy, we now know we must hold the reference to within:

$$\left|\left(df_{REF\,max}\right)\right| = \frac{20 f_{REF}}{30 \times 10^6}\ \text{Hz} \qquad \text{(Eq 20)}$$

$$\frac{\left|\left(df_{REF\,max}\right)\right|}{f_{REF}} = \frac{1}{1.5 \times 10^6}\ \text{Hz} \qquad \text{(Eq 21)}$$

This means keeping the reference, *whatever its frequency*, within $^2/_3$ parts-per-million (ppm) over the range of variables. This is difficult without a crystal oven, but DSP technology once again comes to the rescue!

Microprocessor Compensation

Considering that frequency variation versus temperature is the main factor, we arrange to place a temperature sensor that can be monitored by the DSP microprocessor near the reference oscillator. We specify the AT-cut reference crystal to a tolerance that gives us a reasonable approximation to a straight-line frequency-versus-temperature curve over the range, as shown in Fig 11. We measure the oscillator performance, and use a look-up table and digital-to-analog converter (DAC) to output a tuning voltage that precisely compensates the temperature variations.

Furthermore, the receiver can be tuned on command to an accurate external frequency standard, such as WWV, and the internal reference adjusted to match! In actual practice, this technique results in less than 0.2 ppm error. During this calibration procedure, the internal temperature is noted, and the compensation look-up table is adjusted to reflect whatever curve the oscillator happens to be following. Note that *all* the variables are inside the loop and are, therefore, canceled.

Summary

The marriage of DDS and PLL is an extremely flexible system. Although the PLL programming is normally fixed, its reference frequency can be selected to create different performance characteristics. The DDS can be similarly tuned over a wide range to suit the needs of any particular system.

We've also seen how a microprocessor-compensated crystal oscillator (MPCXO) can exceed the stability of many oven-controlled units—at a small fraction of the power. Over time, repeated calibration to a precise external standard assures the best accuracy on an adaptive basis.

Next, we'll design a transmitter that fits neatly into our plan and exploits many of the advantages obtained in the receiver and synthesizer. The unique benefits of IF-DSP technology will again be highlighted. When reducing cost while improving performance, it's hard to go wrong!

Go Ahead and Transmit!

In this final segment describing IF-DSP transceiver design, we'll define the requirements for the transmitter and look at how they merge with the structures we developed for the receiver. As before, we'll emphasize the significance of DSP-related issues—even to the exclusion of the more mundane aspects, with which most readers are already familiar.

Many receiver sections, such as mixers, filters, amplifiers and IFs, are also required in a typical transmitter design. Considerable cost savings are realized in transceivers by sharing these circuits between modes. To achieve this, we obviously must keep the same frequency conversion scheme, and do a bit of signal switching to toggle between receive (RX) and transmit (TX). First,

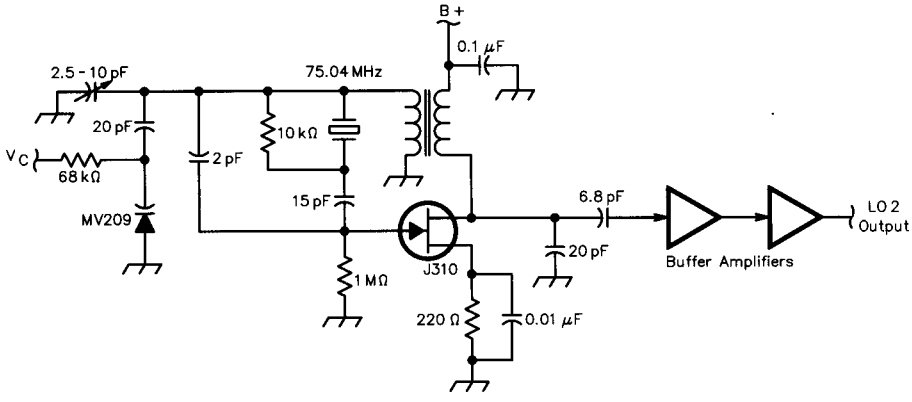

Fig 10—75.04 MHz VCXO for LO_2.

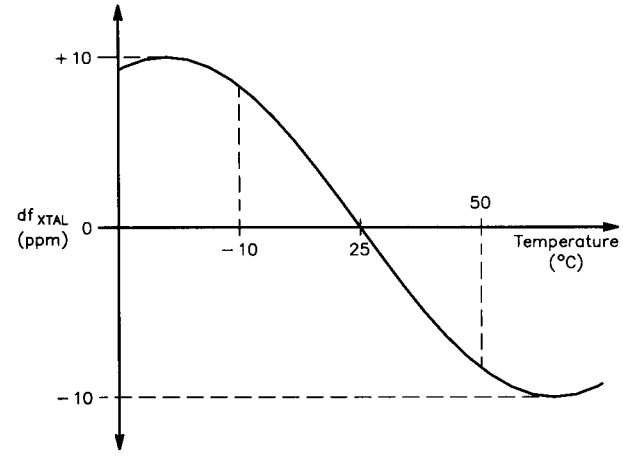

Fig 11—Reference crystal frequency *versus* temperature curve.

let's identify what we expect from the low-level TX, or exciter, function.

Modulation and Drive: The Exciter Design

The goal is to produce a signal—in one of several modulation formats—that can be amplified and transmitted over the air. For an HF transmitter, SSB, AM, FM, CW and various data modes are usually included. Our signal ought to occupy a bandwidth commensurate with good engineering practice and the quality of communication desired—and no more. It shall be, so far as possible, a faithful replica of the input or *baseband* information. We endeavor to design a system that fits simply into the existing architecture and introduces minimum distortion and noise. The receiver's block diagram is presented for review as Fig 12.

With these objectives in mind, it's clear that if we could generate a 40 kHz signal with the desired characteristics, we could translate it up to the 75 MHz IF, then to RF. We wouldn't need narrow second-IF filtering if we could ensure that the 40 kHz transmit signal were already bandwidth-limited. We *would* need the first IF's crystal filters though, to remove the image product and LO bleed-through at 80 kHz and 40 kHz away, respectively.

T/R Switching

The first IF strip can be used in the TX mode by swapping the LOs. This requires a double-pole, double-throw switch, which is implemented using PIN diodes. Good isolation between the injection signals is mandatory.

The input to the strip will be the 40 kHz transmit signal, with the output translated to RF by the second mixer. We must have switches at both ends to select input and output signals, as shown in Fig 13. Good isolation and linearity are critical in these switches. Note that we've avoided switching any low-level, 75 MHz signals in this design.

Many traditional analog signal-processing stages are unneeded, as in the receiver. The balanced modulator, sharp crystal or mechanical filters, speech processor, gain-controlled stages and carrier-null adjustments disappear! In addition, each unit will perform identically to the next, because it's all done in firmware.

Level and Gain Determination

Referring to Fig 14, as the second mixer is a low-level device, we must keep its input level low to avoid objectionable spurious products. For a level-13 mixer, the resulting output level is about –28 dBm. Therefore, 48 dB of gain is necessary to achieve a +20 dBm exciter output. Once we get to about the 0 dBm level, push-pull stages are helpful in reducing second harmonic output and in obtaining sufficient levels. A final 30 dB power amplifier takes us to 100 W. It's best to shoot for at least a 3 dB gain margin, so the exciter output stages are designed to handle up to +23 dBm, or 200 mW.

At the first mix (from 40 kHz up to 75 MHz) we must have a high drive level to overcome the LO bleed through. The LO-to-IF isolation might be as low as 40 dB, so with a +17 dBm injection level, the bleed-through might be as high as:

$$P_{bleed-through} = (17 - 40)\,\text{dBm} = -23\,\text{dBm} \quad \text{(Eq 22)}$$

The crystal filters will attenuate this product by a further 65 dB. So, to keep it less than the design goal of –70 dBc, and considering the mixer's conversion loss is about 6 dB, we must use a drive level of at least:

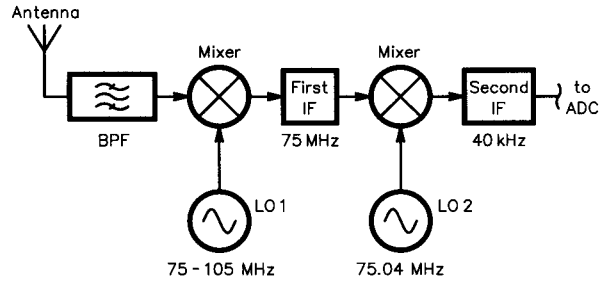

Fig 12—IF-DSP receiver conversion scheme.

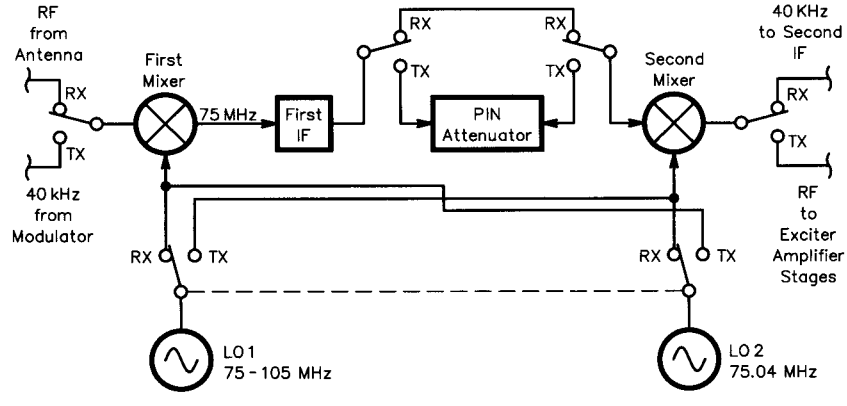

Fig 13—Conversion scheme with T/R switching added.

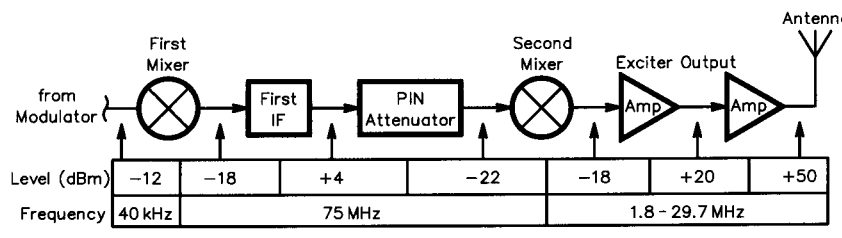

Fig 14—Transmit level diagram.

$$P_{(40\ kHz\ drive)} = (-23 - 65 + 70 + 6)\ \text{dBm} = -12\ \text{dBm} \quad \text{(Eq 23)}$$

Whereas in the receiver, the first IF strip had a gain of about 10 dB, it must now have a loss of:

$$LOSS_{IF_1} = (-12) - (-28)\ \text{dBm} = 16\ \text{dB} \quad \text{(Eq 24)}$$

The PIN diode attenuator just before the second mixer must therefore have attenuation equal to:

$$ATTEN_{PIN} = 10 - (-16)\ \text{dB} = 26\ \text{dB} \quad \text{(Eq 25)}$$

Receive / Transmit Gain Comparisons

Consider the total power gains in the receiver versus the transmitter. The receiver takes as little as –132 dBm from the antenna and amplifies it to around 1 W at the loudspeaker, or +30 dBm; the power gain is:

$$GAIN_{RX} = 30 - (-132)\ \text{dBm} = 162\ \text{dB} \quad \text{(Eq 26)}$$

In the transmitter, a typical dynamic microphone might produce 5 mV (RMS) into 600 Ω, or:

$$P_{MIC} = \frac{\left(5 \times 10^{-3}\right)^2}{600} \approx -44\ \text{dBm} \quad \text{(Eq 27)}$$

The gain is:

$$GAIN_{TX} = 50 - (-44)\ \text{dBm} = 94\ \text{dB} \quad \text{(Eq 28)}$$

The receiver has a far more difficult task, but the transmitter is still doing yeoman's duty! Now think of the maximum path loss of:

$$LOSS_{PATH} = 50 - (-132)\ \text{dBm} = 182\ \text{dB} \quad \text{(Eq 29)}$$

and ruminate on the fact that the total power gain from microphone to loudspeaker must be:

$$GAIN_{TOTAL} = 162 + 94\ \text{dB} = 256\ \text{dB} \quad \text{(Eq 30)}$$

a factor of 4×10^{25}! It's a wondrously large amount of enhancement we get from our electronics!

The SSB Modulator

As described in Part 1, we intend to use the phasing method of SSB generation. We learned it's convenient to choose an output sampling rate four times that of the output frequency, because the injection to the modulator takes on values of only one or zero, simplifying matters. This will be a sampling rate of four times 40 kHz, or 160 kHz. The digital-to-analog converter (DAC) will be humming right along!

We'll bandwidth-limit the baseband information and the resulting SSB signal, but as we saw before, the sampling process will cause the output spectrum to repeat at harmonics of the sampling frequency,[3] as shown in Fig 15. Fortunately, it's easy to build a low-pass filter that will remove these *alias* products.

On inspection, we quickly see it's not practical to sample baseband signals at a rate of 160 kHz. For one thing, we must bandwidth-limit the signals to something like 3 kHz, and building a filter with a fractional bandwidth of:

$$BW_{FRAC} = \frac{3\ \text{kHz}}{160\ \text{kHz}} = 0.01875 \quad \text{(Eq 31)}$$

and decent attenuation characteristics would be a fantasy—even in the DSP world! Additionally, such a high sampling rate is beyond the capabilities of most DSP chips with on-board ADCs. A sampling rate equal to the IF (40 kHz) is more reasonable. This allows us to build a pair of transmit filters having the response shown in Fig 16, with 6 dB points at 180 Hz and 2.9 kHz.[9]

The filtered output is then *interpolated* (see Part 1) up to the 160 kHz rate. An *interpolation filter* removes the alias components due to the lower sampling frequency prior to application to the modulator. The result is a bandwidth-limited SSB signal, ready for translation to RF.

Other Transmitter Modes

Other modulation modes are considerably simpler than the SSB phasing method. In CW, for example, we just output a single, 40 kHz signal. Note that the DSP system makes it easy to shape the rise and fall times of the transmitted CW note. The keyer—which may also incorporate speed and weighting controls—can have adjustable dynamics, from "hard" to "soft," to suit the operator.

FM and AM modes are a little tricky, because we must limit the amplitude of the baseband information to prevent overmodulation. As we'll discover below, the automatic level control (ALC) in AM presents some interesting difficulties.

Distortion and Noise Sources

As discussed previously, numerical truncation, quantization noise and DAC nonlinearity affect the quality of any

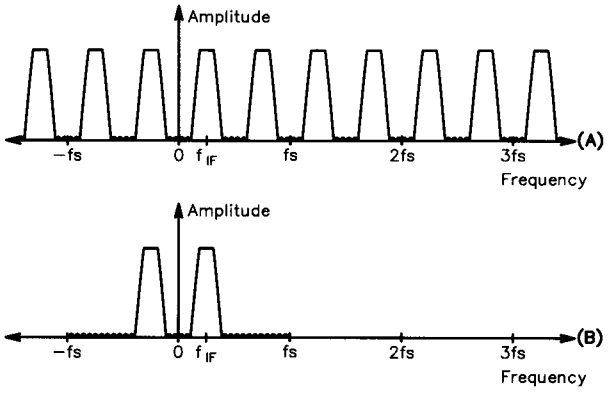

Fig 15—(A) DAC output spectrum showing "aliases. (B) Output spectrum after low-pass filtering.

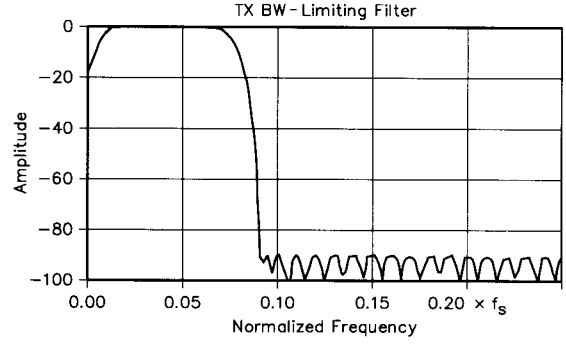

Fig 16—Response of transmit bandwidth-limiting filters.

digitally processed signal. As we revisit these topics, we'll see how they place limits on transmitter performance and why component selection is critical.

First, SSB opposite-sideband rejection and carrier suppression are of concern. Whether analog or digital, amplitude and phase inaccuracies degrade the opposite-sideband suppression in a phasing-method modulator. Maintaining 16-bit data representation throughout the system ensures that computational effects are negligible in a digital implementation.[3] The DAC performance is generally the limiting factor. The best 16-bit DACs produce amplitude and phase accuracy quite adequate for our needs, resulting in typical opposite-sideband suppression of –70 dBc. Note that this level is far lower than the intermodulation distortion (IMD) produced by the final PA.

In a digital modulator of this type, carrier rejection is established mainly by the dc offset present at baseband. This is easy to correct, since in the absence of input audio, the DSP can measure any offset. It's subtracted prior to modulation.

Secondly, noise produced by quantization effects can be significant. A 10-bit ADC has a maximum signal-to-noise ratio (SNR) of:

$$\left(\frac{3}{2}\right)2^{20} \approx 62 \text{ dB} \qquad \text{(Eq 32)}$$

This will also be the SNR of the transmitter output. This level is deemed sufficient for all applications.

Lastly, we must consider mixer performance at an RF-port frequency of 40 kHz. Certainly, the mixer must be designed to handle signals down to this range. The presence of an antialiasing filter will degrade the IMD characteristics (as described in a prior segment) unless idler networks are used.

Digital Automatic Level Control (ALC)

In this design, ALC is realized solely by controlling the amount of 40 kHz drive signal. Information about output power is obtained from a directional coupler, which measures both the forward and reflected power at the output of the low-pass filters that limit harmonic radiation. As it's our intention to place an automatic antenna-tuner unit (ATU) between this point and the antenna, a phase detector is also included.

Refer to Fig 17. As described in *The ARRL Antenna Book*,[12] this directional coupler produces output levels proportional to the square roots of both forward and reflected power, ie, SWR information. These two signals are rectified, filtered and then fed to ADC inputs on the DSP chip. (A small bias current is passed through the detector diodes so they're more sensitive. This allows measurement of as little as 0.3 W.)

The forward output is continually compared to a predetermined threshold. When the threshold is exceeded, drive is reduced. When maximum output levels are not being reached, gain is slowly increased to a preset limit. The whole thing works much like a traditional analog ALC.

When the reflected power exceeds a certain amount, we reduce the forward power to protect the output devices. DSP makes it easy to hold the reflected power to a fixed value, such as 10 W.

ALC in AM

It's long been a problem to hold the carrier level constant in AM transmitters. Because the baseband signal may not have symmetrical positive and negative amplitudes, a suitable analog ALC system would be incredibly complex. In DSP, we can prevent *carrier shift* by using adaptive techniques.

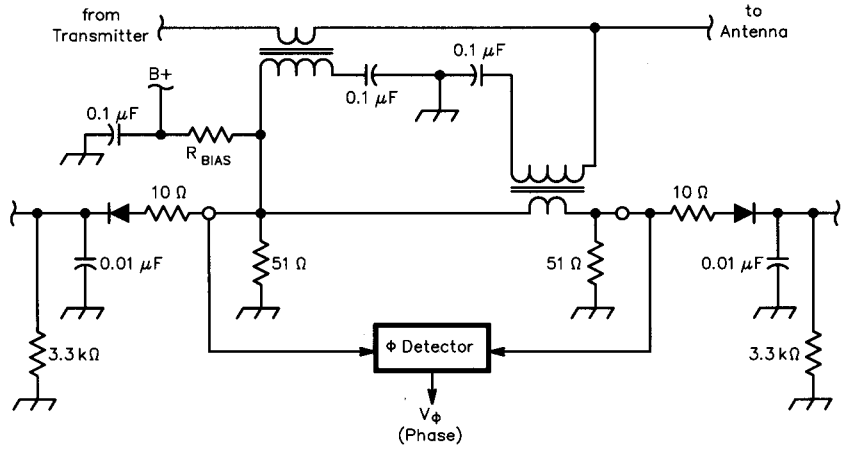

Fig 17—Directional coupler and detectors.

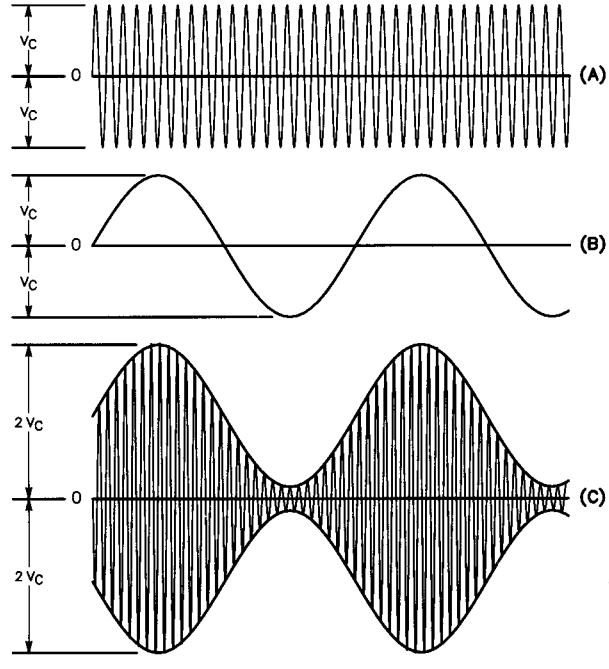

Fig 18—(A) Carrier wave. (B) Baseband input waveform. (C)AM output.

First, the ratio of drive level to output power is easily computed when the transmitter is on, so we can determine the drive level required to achieve a carrier level that is 25% of the peak power setting. Second, the baseband signal must be held to a maximum peak level that equals the carrier drive level. When the carrier and peak-limited baseband levels are added, the result will be a 100%-modulated AM wave,[10] as shown in Fig 18.

This means that *two* ALC servomechanisms operate in our AM ALC: One continuously computes the drive-to-output ratio and maintains the carrier level. The second compresses the peak baseband signal to that same level. The system for this is shown in Figure 19.

Since the baseband peak detector employs a full-wave rectifier in firmware, audio inputs with asymmetrical positive and negative voltage swings can produce unexpected results. That is, either the upward or downward modulation may reach 100% before the other can do so; if the downward modulation limits baseband amplitude first, the peak envelope power cannot reach its set level without introducing a carrier shift! See Fig 20.

SWR Computation

As our detector outputs are proportional to the forward and reflected voltages, the *reflection coefficient* is just the ratio:

$$\rho = \frac{V_{REFL}}{V_{FWD}} \qquad \text{(Eq 33)}$$

and the SWR is calculated using:

$$SWR = \frac{(1+\rho)}{(1-\rho)} \qquad \text{(Eq 34)}$$

To find the actual antenna impedance, we also need to know the *phase* of the reflection coefficient. To get the relative phase of the coupler's two output signals, we'll use a digital phase detector much like those found in PLL chips.

At This Phase of the Game

To determine the phase, we can build a circuit that finds the ratio of the time, t, between the rising edges of the forward and reflected voltages to the total RF period, p. See Fig 21. In signed format, the phase (in degrees) is then:

$$\phi = 360°\left(\frac{1}{2} - \frac{t}{p}\right) \qquad \text{(Eq 35)}$$

A dual-D flip-flop can be configured to output a "1" during time t, and a "0" otherwise, with the forward and reflected voltages as the clock inputs.

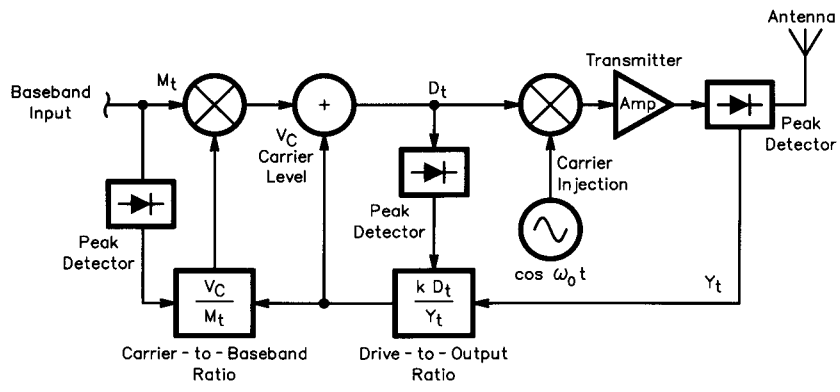

Fig 19—AM ALC block diagram.

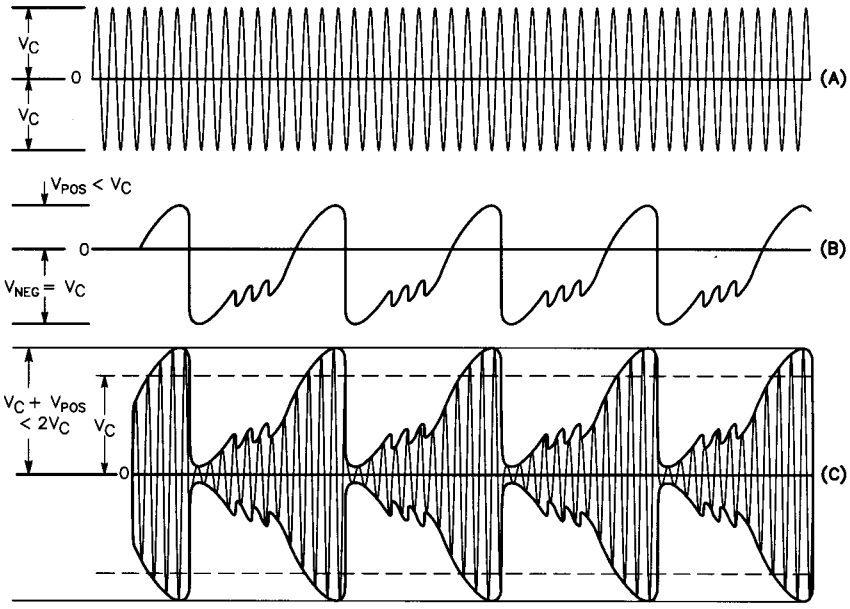

Fig 20—(A) Carrier. (B Baseband input with asymmetrical amplitudes. (C) AM output.

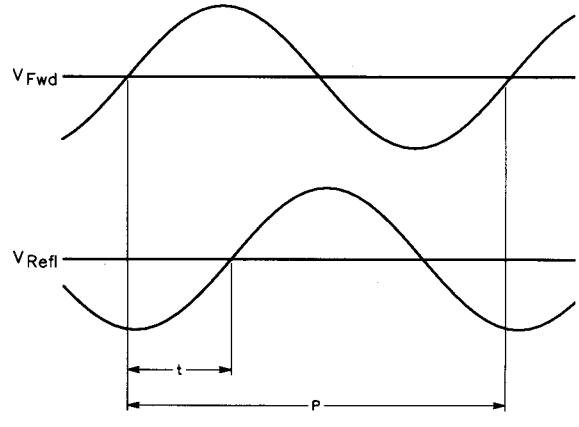

Fig 21—Phase relationship of forward and reflected voltages.

The output is then integrated with a simple RC filter, and the result is a voltage from 0 to 5 V that is directly proportional to relative phase. A circuit to do this is shown in Fig 22.

We can guarantee a constant forward voltage, but the reflected level may be quite low near 50 Ω, so an amplifier is required ahead of this clock input. Even so, the device will fail to clock properly below some value of reflection coefficient. This threshold needs to be less than a SWR of, say, 1.3:1, corresponding to a reflection coefficient of:

$$\rho = \frac{(SWR-1)}{(SWR+1)} \approx 0.13 \qquad \text{(Eq 36)}$$

Almost 18 dB of gain is therefore required to match the forward voltage, and the amplifier can clip at larger input levels to avoid overdriving the chip.

Complex Impedance Transformations

Magnitude and phase information in hand, we can compute the antenna impedance directly! The transforms are easier to present if handled in two steps. The amplitude and phase angle of the reflection coefficient are in *polar form*, so the first conversion is to *Cartesian* coordinates x and y:

$$x = \rho\cos\phi$$
$$y = \rho\sin\phi \qquad \text{(Eq 37)}$$

These numbers, which range from −1 to 1, may then be used to plot the impedance point on a *Smith Chart*[11] of unity radius.

A second step finds the normalized complex impedance $R + jX$:

$$R = \frac{1-x^2-y^2}{(1-x)^2+y^2}$$
$$X = \frac{2y}{(1-x)^2+y^2} \qquad \text{(Eq 38)}$$

To convert to a 50 Ω system, we simply multiply R and X by 50. These data are going to be quite useful as we consider the ATU that must provide a *conjugate match* between our transmitter and an antenna, which may not look like a 50 Ω resistor!

ATU Configuration

In order to hold the internal ATU to a sensible cost, we must limit the range of antenna impedances to be matched. We soon discover that, outside the SWR = 3:1 circle, the voltages and currents in the matching elements rise rapidly. External ATUs must deal with a much larger range of antennas. This places stringent demands on component Q. Physically large inductors and capacitors are employed to reduce losses.

It turns out that we can always achieve a conjugate match to the antenna using an LC network with a series inductance and a shunt capacitance. For antenna impedances with resistance greater than 50 Ω, the capacitance needs to be at the output side; those with resistance less than 50 Ω need the capacitance on the input. We arrange to switch binary-weighted inductance and capacitance values into the circuit using relays (as shown in Fig 23), so we can obtain the range of values we need. Now all we need is an algorithm that drives the network toward a match under microprocessor control.

"Fuzzy-Reasoning" ATU Algorithms

Fuzzy reasoning is a process, like those in the human mind, which assesses a situation in relative terms. For example, if we see the antenna is capacitive, we know inductance must be added; if inductive, capacitance must be used. Further, if the antenna is *very* capacitive, *more* inductance must be inserted. A fuzzy-reasoning system employs *transfer functions* that describe how much adjustment to make based on detector inputs. The transfer functions can represent not only the theoretical requirements of

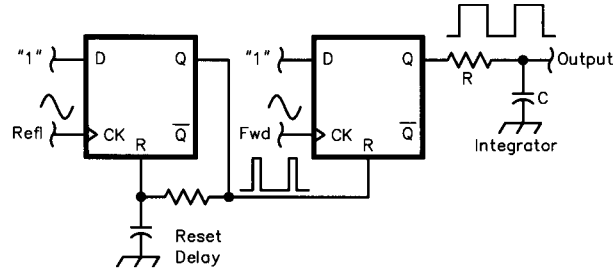

Fig 22—Digital phase detector.

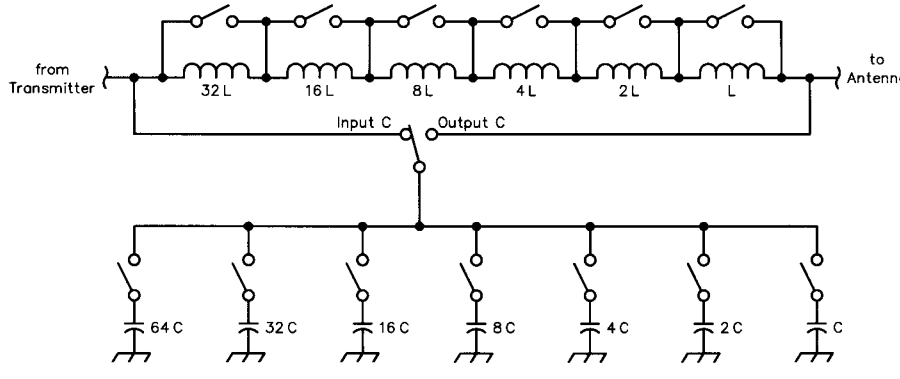

Fig 23—ATU network configuration.

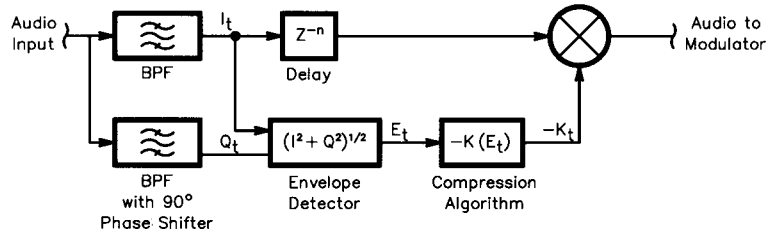

Fig 24—Digital RF compressor block diagram.

the system, but can also incorporate any predictable errors from the detectors and other sources. Fuzzy reasoning tends to overcome errors in systems that can provide only roughly accurate absolute measurements, but produce good relative resolution.

In the case of the ATU, the accuracy of the phase detector degrades rapidly when SWR is less than 1.3:1, and we must rely solely on the reflection coefficient to guide us. A certain amount of "thrashing about" must be employed to find the minimum SWR. Above this level, the phase information is useful in steering toward the goal. To achieve this, we develop transfer functions that embody the matching rules and create a fuzzy-reasoning *engine* that adjusts circuit elements on a step-by-step basis until it reaches minimum SWR.

Step size must be determined by the degree of correlation between the transfer functions and the actual performance of the circuit. Tuning speed is the parameter that suffers because of inaccuracies. In actual practice, measure-and-adjust cycles of around 25 ms yield tuning times well under one second. The use of adaptive, memory-tuning techniques enhances performance.

Adaptive ATU Memory

After several tune cycles on a particular antenna, we begin to get some idea of how it performs. If a frequency is selected near one that has already been matched, a network may be estimated from the previously stored data. The tuning time is therefore greatly reduced. As additional memory points are stored, the number of steps in each tune cycle diminishes until, finally, the antenna system is wholly characterized. If enough points can be stored, tuning time shrinks to that required for switching to the correct network—and no more.

As the network data are available to the control system, the antenna impedance can be plotted across the band. This is extremely useful during initial setup of an antenna, and during continued operation.

Speech Processing: More Bang per Buck

Another distinct advantage in a digital, phasing-method SSB modulator is that the RF envelope can be calculated before the modulation is performed! This allows us to employ RF compression methods on the baseband signal prior to filtering, where they can be effective without adding to "splatter." This scheme is shown in Fig 24.

As explained in Part 1, the envelope of the SSB output is computed as:

$$E_{SSB} = \left(I^2 + Q^2\right)^{1/2} \qquad \text{(Eq 39)}$$

To avoid the time-consuming square-root calculation, we can use an approximation:[1]

$$\left\{ \begin{array}{l} For: |I| > |Q|, \quad \left(I^2 + Q^2\right)^{\frac{1}{2}} \approx |I| + 0.4|Q| \\ \\ \qquad |Q| \geq |I|, \quad \left(I^2 + Q^2\right)^{\frac{1}{2}} \approx |Q| + 0.4|I| \end{array} \right\} \qquad \text{(Eq 40)}$$

This envelope amplitude is used to compress the baseband levels so that the peak-to-average ratio of the transmitted signal is reduced. That is, the average power is increased. The effect is the same as that produced by RF processing. This naturally involves the introduction of distortion, since the transmitter is no longer linear. Nevertheless, this type of distortion enhances the syllabic and formant energy in speech without introducing the "mushy" sound caused by audio clipping.

To elaborate, consider that the human voice has a peak-to-average ratio as high as 15 dB. This doesn't use a peak-limited transmitter very well, and at the 100-W PEP level, the average output power might be as little as 3 W! RF compression enhances the weaker parts of human speech such that intelligibility is improved. As shown by the studies in the reference literature, 15 dB of RF compression can produce up to 6 dB of intelligibility improvement on the receiving end. This is equivalent to quadrupling the output power!

The compressor attack and decay times can be varied to change the amount of processing introduced. As they are made faster, compression approaches the effects of RF clipping. It's widely known that this is the most effective form of speech processing.

Conclusion

We've seen that first, we must design the most linear system possible in a transmitter; then, to improve speech intelligibility, we must destroy the linearity! In Part 3 of this series, we'll examine advanced DSP techniques that further improve communication, in both the transmitter and receiver. We'll introduce adaptive signal-processing methods that wouldn't be possible without DSP technology. They can correct for many traditionally troublesome production variations. Moreover, we'll see how computer control of transceivers makes many interesting features easy to implement!

Doug Smith, KF6DX, is an electrical engineer with 18 years experience designing HF transceivers, control systems and DSP hardware and software. He joined the amateur ranks in 1982 and has been involved in pioneering work for transceiver remote-control and automatic link-establishment (ALE) systems. At Kachina Communications in central Arizona, he is currently exploring the state of the art in digital transceiver design.

Notes

[1]Sabin, W. E. and Schoenike, E. O., editors, *Single Sideband Systems and Circuits*, McGraw-Hill, New York, NY, 1987.
[2]Rohde, U. L. and Bucher, T. T. N., *Communication Receivers: Principles and Design*, McGraw-Hill, New York, NY, 1988.
[3]Oppenheim, A. V. and Schafer, R. W., *Digital Signal Processing*, Prentice-Hall, Englewood Cliffs, NJ, 1975.
[4]Frerking, M. E., *Digital Signal Processing in Communications Systems*, Van Nostrand-Reinhold, New York, NY, 1993.
[5]Gruber, M., *ARRL HF Product Review Test Manual*, ARRL, Newington, CT, 1992.
[6]Zverev, A. I., *Handbook of Filter Synthesis*, John Wiley & Sons, New York, NY, 1967.
[7]Cercas, F. A. B., Tomlinson, M. and Albuquerque, A. A., "Designing with Digital Frequency Synthesizers," *Proceedings of RF Expo East*, 1990.
[8]Leeson, D. B., "A Simple Model of Feedback Oscillator Noise Spectrum," *Proceedings of the IEEE*, volume 54, 1966.
[9]Alkin, O., *PC-DSP*, Prentice-Hall, Englewood Cliffs, NJ, 1990.
[10]Panter, P. F., *Modulation, Noise, and Spectral Analysis*, McGraw-Hill, New York, NY, 1965.
[11]Smith, P. H., *Electronic Applications of the Smith Chart*, McGraw-Hill, New York, NY, 1969.
[12]Straw, R. D., Ed., *The ARRL Antenna Book*, 17th edition, ARRL, Newington, CT, 1994.
[13]Motorola Semiconductor, *AN969: Operation of the MC145159 PLL Frequency Synthesizer with Analog Phase Detector*, Motorola, Incorporated, Austin, TX, 1991.
[14]Harris Semiconductor, *DSP Databook 1994: DB302B*, Harris Corporation, Melbourne, FL, 1994.
[15]Analog Devices, *AD7722 Data Sheet*, Analog Devices, Inc, Norwood, MA, 1997.