Chapter 31

Switching Converter Power Supplies

The general principles of switching converter power supplies, what they are and how they are used, plus some examples.

General

Chapter 22 covered conventional, powered-from-60 or 50 Hz electric power distribution lines. ***Switchers,*** the short common name for all those supplies who can supply their own frequency on up to 1 MHz, operate a bit differently.[1] They are generally referred to in text as ***converters*** in that they convert one value of DC to another value of DC. Rather than just let their switching rate turn a transformer on and off, they can use small inductors and capacitors to help lower or raise their DC output values. Such switch converter outputs are more efficient, of greater importance in this more-mobile age of electronics. There are two major categories:

A. DC-to-DC Converters. Use as if transformers existed for DC.

B. AC Mains inputs. A good example is multiple outputs in a PC supply running on Mains.

Within each there are four major classifications of switching converters:

1. Buck. DC output is lower than the input. Output is stepped-down from input.

2. Boost. DC output is higher than the input. Output is stepped-up from the input.

3. Flyback. DC output isolated by a transformer, feedback for regulation.

4. Buck-Boost. A hybrid of Buck and Boost, output above or below input.

5. Flying Capacitor. Inductorless supply of with limitations of current.

By eliminating transformers of the old iron kind, there is a great boost in lightness and the packing density of the parts can be made more compact. In the various outputs of specialty switchers based on an IC, Linear Technologies Incorporated may be the king in the USA with ON Semiconductor part of that royal court. That is just on the number of different switch-mode supply ICs both have designed and produced.

Despite their greater efficiency, they have a strong downside: RFI. For a receiver power

[1] Common use also likes acronyms such as ***SMPS*** or Switch-Mode Power Supply or even ***SEPIC (Single-Ended Primary Inductance Converter)*** primarily for Boost switchers.

supply, they produce sharp transients which must be filtered out. RFI or Radio Frequency Interference can be a tough problem to eradicate.

For ***green*** attention to the environment, several IC makers have come out with ***power factor reduction*** techniques to ease drains on AC Mains lines. More on that at the end of this Chapter.

Buck Converter Basic Operation

Figure 31-1 shows a basic block diagram of operating principles. Assume the *Switch* can operate at a high rate. When the switch is closed, the diode does not conduct but current begins to flow in the inductor. That current charges the capacitor and voltage increases across the Load.

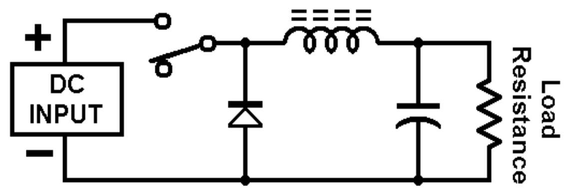


Figure 31-1 Simplified block diagram of a *Buck Converter* supply.

When the switch is open, current from the inductor's built-up magnetic field will discharge through the diode and capacitor charge begins to decay. By choosing the correct inductance and capacitance and the switching frequency, a steady-state voltage can be found across the Load.

Intuitively, *ripple voltage* will be reduced by increasing the switching frequency, given a fixed capacitance and inductance. Likewise, increasing C and L while keeping switching frequency constant will also reduce ripple. Unfortunately, higher inductance will also increase losses due to DC resistance while higher capacitance will also have increased *ESR* (Equivalent Series Resistance).

Regulation is possible by adding a control for the On and Off time of the *Switch.* For an ideal situation (such as Figure 31-1) the output voltage is proportional to the ***On duty-cycle.*** A 5 VDC output from a 12 VDC input would have an ideal *duty-cycle* of 41.667 percent.

Boost Converter Basic Operation

This is shown in Figure 31-2. Again, the *Switch* is operating at a high rate. With the switch off (as shown) current begins flowing through the inductor, through the diode, ***adding*** the input voltage to the Load. Output capacitor is charging up.

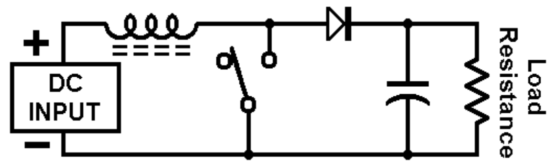


Figure 31-2 Boost converter simplified diagram to show basic operation.

When the switch closes, the magnetic field around the inductor collapses and is dumped to the common connection. Nothing happens at the output because of the diode being reverse-biased. Whatever charge was developed at the capacitor will discharge through the Load.

Again, there is a problem with output ripple voltage. The same cures for that can be done was with the Buck supply. Regulation can be applied by feedback to a duty-cycle controller of the Switch. In this case, the ideal (Figure 31-2) will have a duty-cycle of:

$$D = \frac{(V_{out} - V_{in})}{V_{out}}$$

For a 12 VDC output from a 6 VDC input, the duty-cycle (D) would be 0.5 or 50%. Theoretically, the output voltage could go from input to infinity but practical matters limit parts to much less than that upper region.

Flyback Supply Basics

The word *Flyback* comes from the release of a constant current through an inductor, resulting in a high voltage. Flyback supplies have been used in TV receivers horizontal deflection circuits for a half century, not only forming the horizontal deflection but also providing, via the *Flyback* effect, the very high accelerating potential for the Cathode Ray Tube.

The first widespread use of Flyback circuits other than in TV sets was to replace the old vibrator-induced switching in vehicular radios. Those were essentially transistor oscillators operating at high KHz, unregulated, to power the plate and screen supplies of tube radios. Advent of transistors capable of operating at VHF and UHF and using low voltages spelled an end to that; no high voltages needed. Despite the obvious RFI those continued to be used for a decade, being more reliable than the vibrator used to supply AC for a conventional plate supply.[2]

A modern Flyback supply in block form is deceptively simple. The *power oscillator* may be single-ended or push-pull but always operates at a frequency above audible range. This allows for very little core material and fewer losses. The transformer primary can be constant current driven in a squarewave. When the squarewave quits, the magnetic field collapses and the resulting high voltage is coupled to the transformer secondary. A diode rectifier with ripple voltage lowpass supplies the output.

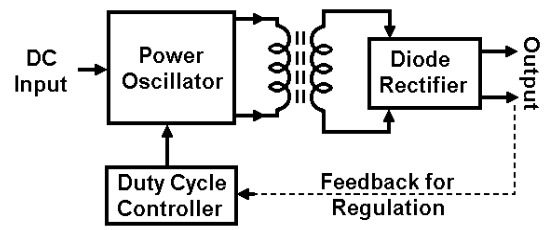


Figure 31-2 Simplified block diagram of a *Flyback* power supply using a high-frequency transformer.

A variable on-time duty cycle controller is usually controlling the power oscillator waveform. Feedback of output voltage may be coupled optically or magnetically back to the controller for regulation. Use of a transformer allows DC isolation between input and output, plus yielding voltages that may be higher than feasible with Boost supplies.

Buck-Boost Converter Basics

These are close-cousin hybrids of conventional *Boost* supplies. The first is the Ċuk Converter shown in Figure 31-4(A) while the SEPIC is shown in Figure 31-4(B).[3] Those are almost the same, only the second inductor and diode changing places. The Cuk converter inverts polarity of its output. The SEPIC does not.

The major purpose of using this type of supply is the large range of input voltages. This suits battery-powered applications where input voltages may range from 0.9 VDC to 2.3 VDC from various single-cell batteries. It could work for vehicular mobile radios with 10 to 30 VDC,

---

[2] A ***vibrator*** referred to a literal vibrating armature of an SPDT switch that took 12 VDC and switched it to a center-tapped AC Mains transformer winding. It resulted in a square wave and compensated for in the transformer design. It was a cheap way to get AC from a DC source. It was also relatively short-lived, requiring replacement often.

[3] The inventor of Figure 31-5(A) is Slobodan Cuk who first publicized it in 1976 while at Cal Tech. The surname Cuk is pronounced like ***chook*** in English. The ***C*** of Cuk is supposed to have a dot umlaut on it. While ***Sepic*** is a proper surname, as used here it is an acronym standing for ***Single-Ended Primary-Inductor Converter.***

depending on vehicle type (12 V for civilian, 24 V for military). It can be used for AC Mains rectified input from 115 to 230 VAC RMS nominal range. For AC Mains input, the DC input is derived from a simple rectifier across the AC Mains and that to the DC input.

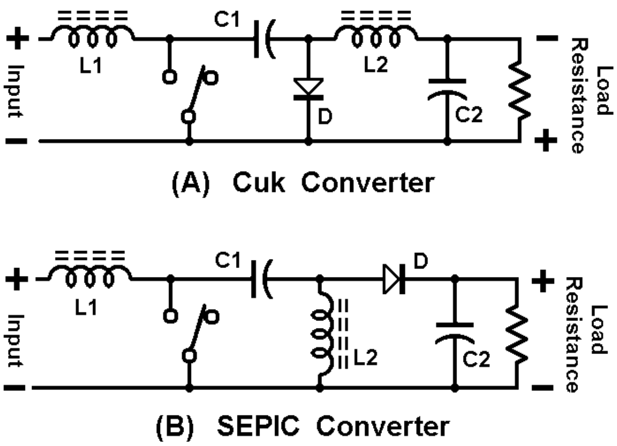


Figure 31-4 Buck-Boost hybrid converters.

Referring to Figure 31-4, the basic operation of both Cuk and SEPIC converters can be derived. In both, C1 will prevent any DC output from feeding back to the input.

In the Cuk converter with switch open, C1 will charge to the input voltage through L1 via a return from L2 and Load resistance. When the switch is closed, C1 discharges to the output through the same path. Diode D does not conduct since it is back-biased. L1's magnetic field collapses but that energy does not go anywhere but to the common. L2 begins to charge its magnetic field while C2 is accumulating an electrostatic charge.

When the switch opens again, L2 will discharge its magnetic field through diode D and that discharge will begin to charge up C2's electrostatic field. When the switch closes a second time, L1 will discharge its magnetic field to common while L2's magnetic field is charging through C2. As the switch continues to open and close, the voltage on the Cuk converter output reaches an equilibrium of slight ripple voltage with a polarity that is inverted relative to the input.

The SEPIC in More Detail

Since the SEPIC arrangement seems to be the preferred one, Figure 31-5 explains the current flow and voltages with ideal parts. Normal schematic is in (A). On-off cycling in shown in (C) and (B).

When the FET is off, C1 gets the start of a charge via L1. L2 has no magnetic field built up and C2 is not charged. Output is zero volts.

When the switch is closed, L1 gets a good magnetic field from input DC but its L1 return line is grounded. C1 electrostatic charge is essentially zero since L2 magnetic field is not built up. C2 is also at zero due to diode D now not conducting.

When the switch opens again, C1 gets an electrostatic charge from L1 and begins to conduct. That conduction is carried through diode D into C2 and Load. L2 begins building its magnetic field.

On closing the switch a second time, the electrostatic change on C1 adds to more magnetic field in L2. C2's charge remains on the output but is dissipated slightly by the Load resistance.

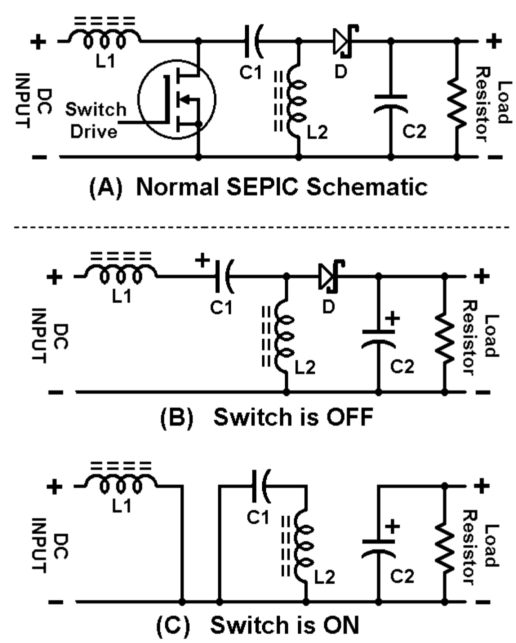


Figure 31-5 SEPIC switch action

As the cycle of switch Off and On continues, C1 continues to charge and discharge and its polarity is not guaranteed. There will be a continuing ripple in the output voltage as C1 determines the positive-going slope and C2 determines the negative-going slope of that ripple. Keeping L1 and L2 fixed values will allow C1 and C2 to determine both the ripple voltage and average DC voltage value of the Load.

Flying Capacitor or Conversion Without Inductors

Figure 31-6 (A) shows the basic switch converter. With switch S to the left, C1 charges to the positive polarity supply. With switch S to the right, C1 transfers its charge to C2. There is no polarity change. Load resistance is approximately equal to:

$$R_{LOAD} = \frac{1}{F \cdot C1} \qquad \text{Where:}$$

$$F = \text{Switch frequency in Hz}$$

$$C1 = \text{Bucket capacity, Farads}$$

Note that the transfer capacitance C1 is commonly called a *bucket.* C2 maintains the load voltage and is normally about equal to C1.

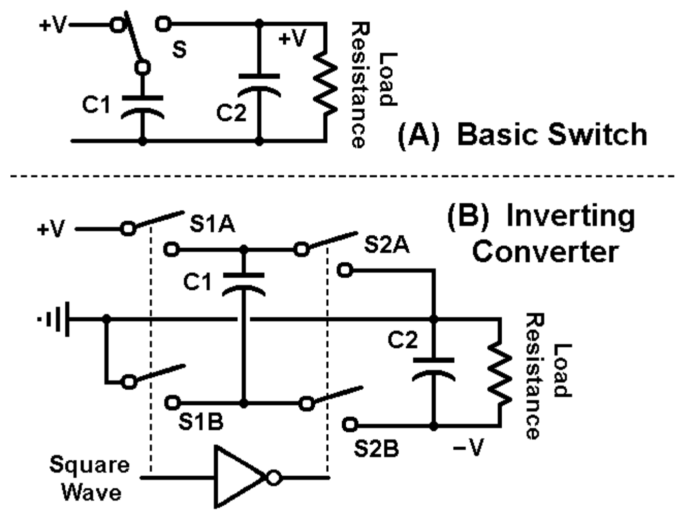


Figure 31-6 Basic *flying capacitor* in (A), more common Inverting Converter in (B).

For a product, three more switch poles and an inverting switch driver are added as in Figure 31-6 (B). This provides a ***negative polarity*** output, albeit unregulated, and can be a low-current source for other circuits.

In (B) when S1 is closed, the bucket capacitance is charged to the positive supply. When S2 is closed (and S1 opened), the bucket charge is transferred to C2 but with polarity inverted. For an older capacitor-only converter such as the Maxim MAX1044, also a second source for Intersil's ICL7660, the light-load voltage is -5.0 VDC but drops to about -4.5 VDC at 10 mA.[4] This was using 10 µFd values for C1 and C2.

An entire light-load negative supply is possible by using a single 8-pin DIP MAX1044, two small electrolytics, plus one positive-supply bypass capacitor. Both the MAX1044 and ICL7660 have internal high audio frequency oscillators and switch drivers. The internal oscillator frequency can be lowered for more current and better efficiency but that drops the frequency into the peak of human hearing.

Design Details of a SEPIC Converter

[4] Reference to Maxim datasheet for their MAX1044/ICL7660 which has several ways to hook up the product, including chaining in series for higher voltages, paralleling for more current, and remote shut-off.

General Details

There are only a few Application Notes available to take one through the steps of designing the values for basic switching converters. All of those few involve specific ICs which contain more circuitry for regulation, shut-down, and various other internal functions. A few converter ICs contain their own MOSFET switches but most have external MOSFETs which must be procured. Scanning the available parts, there are many MOSFETs which can be used. In general, the peak current for initial selection should be (roughly) at a minimum of three times the maximum DC current output. Source-drain current values should be as low as possible to both increase conversion efficiency and reduce values of some other parts.

One thing noticed in regards to MOSFET choices is the ***charge*** in ***coulombs*** between the various MOSFETs. This is often abbreviated to ***nC*** or ***nanoCoulombs*** on datasheets.[5] For IC specific designs there must be enough on the datasheets to include that. As to specific ICs used at the control-heart of converters, they must have enough information to get close to the design, even if interpolated.

For this portion of the Chapter, Application Note AN-1484 from National Semiconductor is used as a baseline along with a National Semiconductor LM3478 control-heart IC. Parts selection follows the Application Note ordering. Specifications for the finished design require an input voltage range of 9 to 15 VDC with an output of 12 VDC at 2 Amperes.

Duty Cycle

Average duty cycle in constant conduction mode is given by:

$$D = \frac{V_{OUT} + V_D}{V_{IN} + V_{OUT} + V_D} \qquad \text{Where:}$$

V_{OUT} is output voltage

V_D is voltage drop of SEPIC diode

V_{IN} is average input voltage = 12 VDC

Assuming the diode drop of 0.5 Volts and output voltage of 12 VDC, D then becomes 0.510. Maximum duty cycle requires only that minimum V_{IN} be replaced by $V_{IN\,(MIN)}$ which becomes:

$$D_{MAX} = 0.581$$

Inductor Selection

A reasonable criterion for ripple current is about 40 percent of maximum input current at minimum input voltage. Given L1 = L2, that yields:

[5] A ***coulomb*** is defined as *a charge transported by a steady current of 1 ampere in 1 second* or *the amount of excess charge of a capacitance of 1 farad charged by a potential difference of 1 volt* according to Wikipedia. All well and good, but it isn't a direct measurable quantity for test instruments in the hobby shop (or even in most professional electronic labs). It has a relationship to switching speed and such but it tends to boggle the mind a bit when it becomes a constant for parts value calculations.

$$\Delta I_L = I_{IN} \cdot 0.40 = I_{OUT}\left(\frac{V_{OUT}}{V_{IN\,(MIN)}}\right) \cdot 0.40 = 2.0\left(\frac{12}{9}\right) 0.40 = 1.07$$

$$L1 = L2 = \left(\frac{V_{IN\,(MIN)}}{\Delta I_L \cdot F_{SW}}\right) D_{MAX} \qquad \text{Where: } F_{SW} \text{ is switching frequency, Hz}$$

Picking 400 KHz as a trial of switching frequency would have inductor values of:[6]

$$L1 = L2 = \left(\frac{9}{1.07 \cdot 4 \cdot 10^5}\right) \cdot 0.581 = 12.2\ \mu\text{Hy}$$

Peak currents in each inductor to avoid saturation of the fields are:

$$I_{L1\,(PEAK)} = I_{OUT}\left(\frac{V_{OUT} + V_D}{V_{IN\,(MIN)}}\right)\left(1 + \frac{40\%}{2}\right) = 2\left(\frac{12.5}{9}\right)(1.2) = 3.33$$

$$I_{L2\,(PEAK)} = I_{OUT}\left(1 + \frac{40\%}{2}\right) = 2 \cdot 1.2 = 2.4$$

By using two windings on the same core, mutual inductance would make each winding:

$$L' = \left(\frac{V_{IN\,(MIN)}}{2 \cdot \Delta I_L \cdot F_{SW}}\right) D_{MAX} = \left(\frac{9}{2 \cdot 1.07 \cdot 4 \cdot 10^4}\right) 0.581 = 6.11\ \mu\text{Hy}$$

MOSFET Selection

Peak switch (MOSFET) source-drain voltage is equal to $V_{IN\,(MAX)} + V_{OUT} = 15 + 12 = 27$ Volts. Peak switch current is the sum of each inductor current or:

$$I_{Q1\,(PEAK)} = I_{L1\,(PEAK)} + I_{L2\,(PEAK)} = 3.33 + 2.40 = 5.73\ \text{A}$$

RMS current through the switch is:

$$I_{Q1\,(RMS)} = I_{OUT}\sqrt{\frac{(V_{OUT} + V_{IN\,(MIN)} + V_D)(V_{OUT} + V_D)}{V^2_{IN\,(MIN)}}} =$$

$$2\sqrt{\frac{(12 + 9 + 0.5)(12 + 0.5)}{9^2}} = 3.64\ \text{A}$$

Note that this RMS current would be the worst-case condition. It would drop to 2.19 A at an input

[6] From the LM3478 datasheet, switching frequency can be set to 400 KHz by about 39 KOhm resistor.

of 15 VDC.

At this point the MOSFET needs to be chosen. A Vishay IRF510 is under consideration. While a bit over-size, it comes in a TO-220 case, has a breakdown voltage of 100, can handle a pulsed drain current of 20A with steady-state current of 5.6A, has a maximum power dissipation of 43 Watts. In addition, it has an $R_{DS\ (ON)}$ of 0.54 Ohms maximum at 3.4 A drain current at 10 V of gate-source voltage. The charge of the gate-drain junction [Q_{GD}] is 3.8 nC. Gate drive current for I_G comes from the LM3478 and is 0.3 A. MOSFET power dissipation is then approximately:

$$P_{Q1} = I^2_{Q1\ (RMS)} \cdot R_{DS\ (ON)} \cdot D_{MAX} + \left(V_{IN\ (MIN)} + V_{OUT}\right) \cdot I_{Q1\ (PEAK)} \cdot \left(\frac{Q_{GD} \cdot F_{SW}}{I_G}\right) =$$

$$(3.64)^2 \cdot 0.54 \cdot 0.581 + (9 + 12) \cdot 5.73 \cdot \left(\frac{3.8 \cdot 10^{-9} \cdot 4 \cdot 10^5}{0.3}\right) =$$

$$13.2 \cdot 0.314 + 21 \cdot 5.73 \cdot 0.00507 = 4.14 \cdot 0.610 = 2.53\ \text{W}$$

There doesn't seem to be a problem there so the IRF510 can be used safely.

Diode Selection

Peak reverse voltage must be higher than $V_{IN\ (MAX)} + V_{OUT} = 15 + 12 = 27$ V. Average diode current is equal to output current maximum or 2 A. An MBRD340, used in other National Semiconductor models, is selected here.

SEPIC Coupling Capacitor C1 Selection

C1 RMS current would be:

$$I_{C!\ (RMS)} = I_{OUT} \sqrt{\frac{V_{OUT} + V_D}{V_{IN\ (MIN)}}} = 2 \sqrt{\frac{12 + 0.5}{9}} = 2 \cdot 1.178 = 2.357\ \text{A}$$

For the C1 value in µFd there is a choice. More microFarads means less ripple voltage. That choice can begin with 20 µFd. Ripple voltage through C1 is then:

$$\Delta V_{C1} = \frac{I_{OUT} \cdot D_{MAX}}{C_1 \cdot F_{SW}} = \frac{2 \cdot 0.581}{22 \cdot 10^{-6} \cdot 4 \cdot 10^5} = \frac{1.162}{8.8} = 0.132\ \text{V}$$

If C1 were selected to be 10 µFd, then the ripple voltage would be 0.264 V. Stay with 10 µFd since this is easier to get as an unpolarized ceramic.

Output Capacitor C2 Selection

Output capacitor C2 current is the same as $I_{C!\ (RMS)}$ or 2.36 A. Assume an output ripple of 2 percent on $V_{OUT.}$ ESR or Effective Series Resistance of C2 can be determined by:

$$ESR \approx \frac{Ripple \cdot 0.5}{I_{L1\,(PEAK)} + I_{L2\,(PEAK)}} = \frac{0.02 \cdot 12 \cdot 0.5}{3.33 + 2.4} = \frac{0.120}{5.73} = 20.9 \cdot 10^{-3} = 21\ mOhms$$

$$C2 = \frac{I_{OUT} \cdot D_{MAX}}{\Delta V_{C1} \cdot 0.5 \cdot F_{SW}} = \frac{2 \cdot 0.581}{0.264 \cdot 4 \cdot 10^5} = \frac{1.162}{105.6 \cdot 10^3} = 11.0 \cdot 10^{-6}\ Farads$$

C2 can be selected as (minimum) 12 µFd although 15 µFd is better due to parts tolerances. A polarized tantalum is fine there.

Input Capacitor C3

AN-1484 recommends a 100 to 200 µFd input capacitor for inputs greater than 8 VDC. Regardless of that, the RMS current of C3 is:

$$I_{C3\,(RMS)} = \frac{\Delta I_L}{\sqrt{12}} = \frac{1.07}{3.46} = 0.309\ A$$

LM3478 Specific Calculations

This requires specific study of the IC datasheet. A comparison circuit within the LM3478 controls the output voltage relative to an internal 1.26 VDC source. R1 (top) and R2 set that voltage with the following equation:

$$R2 = \frac{R1 \cdot V_{REF}}{V_{OUT} - V_{REF}} = \frac{R1 \cdot 1.26}{12 - 1.26} = 0.1173 \cdot R1 \quad \text{[for 12 V output]}$$

A more precise output voltage control can be a potentiometer inserted between R1 and R2 as in the schematic of Figure 31-7.

A switching frequency of 400 KHz can be interpolated from a chart in the datasheet to make the single resistor, R4, equal to about 39 KOhms.[7]

MOSFET Source Current Sensing

This is a very low value resistor which may be difficult to get. It can be made from several paralleled very low film capacitors or from just wire. If done with small wire, it should be wound so that each half of the wire bulk is wound in opposition to the other half. To explain that, take a length of wire, loop it into two pieces; begin winding with the loop first, then wind remaining wire together on each side of the former. Mutual inductance should be minimized leaving only

[7] This, as with several other items, are either difficult to read or to understand. As is the usual case, AN-1484 seems to have been written from earlier work. The latest datasheet on the LM3478 is dated 25 May 2011 while AN-1484 used in here is dated 30 April 2008. It should also be noted that much of AN-1484 is contained in the later parts of the latest LM3478 datasheet. The usual result is a lot more time spent in the lab manually changing things to *improve* performance or, at least, to just get it working.

resistance.

According to AN-1484 there is supposed to be a chart indicating *typical sense voltage versus duty cycle.* Three years after publishing this Application Note, the LM3478 datasheet of May, 2011, has this chart on page 9 and is repeated on Page 16 as *Figure 10*:

$$R_{SENSE} = \frac{0.130}{I_{Q1\,(PEAK)}} = \frac{0.130}{5.73} = 22.7 \cdot 10^{-3} = 23 \text{ mOhms}$$

LM3478 Frequency Compensation

Frequency compensation applies to the inner workings of the LM3478, the part we cannot examine first-hand. Indeed, we are not privileged to have any information whatsoever on that presented; all we can use are the signals at the package pins. In this IC the compensation network is only two capacitors and a resistor. We can take the values as they are presented in AN-1484.

$$F_{RHPZ} = \frac{(1 - D_{MAX})^2 \cdot V_{OUT}}{\pi \cdot D_{MAX} \cdot L_2 \cdot I_{OUT}} = \frac{(1 - 0.581)^2 \cdot 12}{\pi \cdot 0.581 \cdot 12 \cdot 10^{-6} \cdot 2} = \frac{0.176 \cdot 12}{43.8 \cdot 10^{-6}} = 48.1 \text{ KHz}$$

$$F_R = \frac{1}{2\pi\sqrt{L2 \cdot C1}} = \frac{1}{2\pi\sqrt{12 \cdot 10^{-6} \cdot 10 \cdot 10^{-6}}} = \frac{1}{2\pi \cdot 10.95 \cdot 10^{-6}} = 14.5 \text{ KHz}$$

A *cross-over* frequency is taken from the ***lowest value*** of F_R or F_{RHPZ} and divide that by 6 to yield 2.42 KHz which is equal to F_C.

For parts values of the compensation network, some more items must be taken from the datasheet: G_{CS} is the current sense gain or roughly 90 here; G_A is the error amplifier transconductance of about 800 μmho here; V_{REF} is the internal reference voltage of 1.26 VDC. C_{OUT} is the final selected value of C2, in this case 15 μFd.

$$R5 = \frac{2\pi \cdot F_C \cdot C_{OUT} \cdot (V_{OUT})^2 \cdot (1 + D_{MAX})}{G_{CS} \cdot G_A \cdot V_{REF} \cdot V_{IN\,(MIN)} \cdot D_{MAX}} = \frac{2\pi \cdot 2.42 \cdot 10^3 \cdot 15 \cdot 10^{-6} \cdot 12^2 \cdot 1.581}{90 \cdot 800 \cdot 10^{-6} \cdot 1.26 \cdot 9 \cdot 0.581} =$$

$$\frac{51.93}{0.4744} = 109.4 \text{ Ohms}$$

C4 is chosen to set the compensation zero to 1/4 of the crossover frequency:

$$C4 = \frac{1}{2\pi \cdot F_C \cdot R5} = \frac{1}{1.663 \cdot 10^6} = 601 \text{ nFd}$$

C5 is selected to cancel the ESR zero:

$$C5 = \frac{C_{OUT} \cdot ESR}{R_C} = \frac{15 \cdot 10^{-6} \cdot 20.9 \cdot 10^{-3}}{109.4} = 2.87 \cdot 10^{-9} = 2.87 \text{ nFd}$$

It seems likely that 5% tolerance components will work there, so R5 = 120 Ohms, C4 = 560 nFd, and C5 = 2.7 nFd.

Schematic

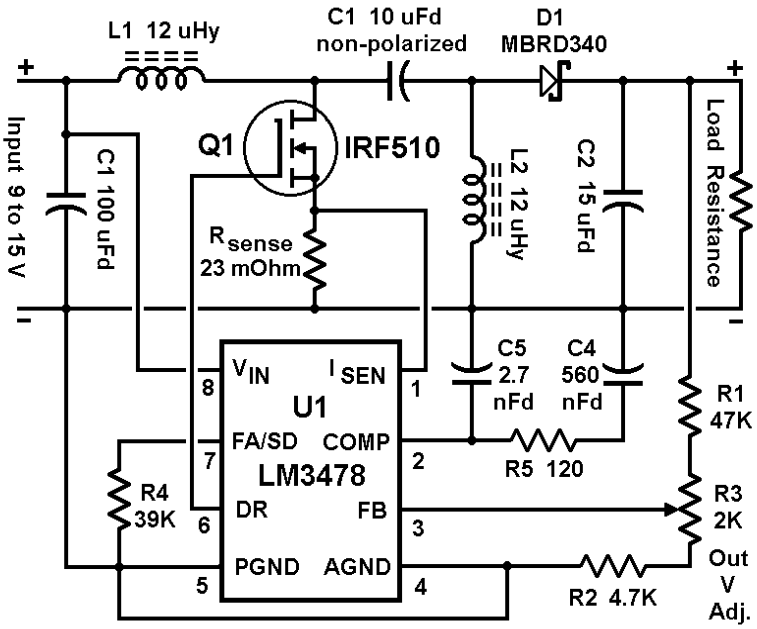


Figure 31-7 Final schematic of SEPIC converter example. U1 is in SO-8 package, Q1 in TO-220, D1 in 0.3 inch square maximum 3-connection case.

Schematic of this long, involved calculation task is in Figure 31-7. Output is 12 VDC at 2 A maximum load.

What is shown is a manual trim adjustment for output voltage. That can be replaced by eliminating R3, joining R2-R1-FB pin, and replacing R2 with 5.6K. Output voltage will be about 12.1 VDC with zero-tolerance resistors.

Using these components, with Q1 off-board, will fit easily on a 2 inch by 3 inch PCB. This converter was never built by the author even though planned for a mobile automotive supply for a VHF application. Efficiency must be measured with a selection of dummy resistive loads.

Another Example, a Boost Converter

General

This example is a ***Boost Converter*** featured on the first page of the LM3478 datasheet, dated May, 2011. It takes in 3.3 VDC ± 10% and outputs 5 VDC at a 2 A maximum current. Steps for calculation of values are found in the last half of the datasheet of May, 2011. This one can create an *old* logic circuit power supply of 5 V from a *newer,* lower-voltage logic circuit supply line such as 3.3 V.

Efficiency of this converter is fairly good. From the datasheet page 7 graph (lower left), it is 52% efficient at 1 mA power demand, 68% at 10 mA, 80% at 45 mA, 84% at 100 mA, 86% at 220 mA, then falls off to 80% at 1.0 A load.

Specific Details of LM3478

This is a fixed-frequency, pulse-width controller IC that varies the on-off times to regulate voltage. The exact frequency is set by a resistor to the following formula:

$$R_{FA} = 4.503 \cdot 10^{11} \cdot \left(Freq^{-1.26}\right)$$

where: $Freq$ is in Hz

This requires a calculation on one number to another number's power. That requires use of a scientific calculator.

Frequency of the LM3478 is between 100 KHz and 1 MHz and resistor values versus frequency are:

1 MHz	12.4 K
800 KHz	16.4 K
600 KHz	23.6 K
400 KHz	39.3 K
300 KHz	56.5 K
200 KHz	94.2 K
100 KHz	226 K

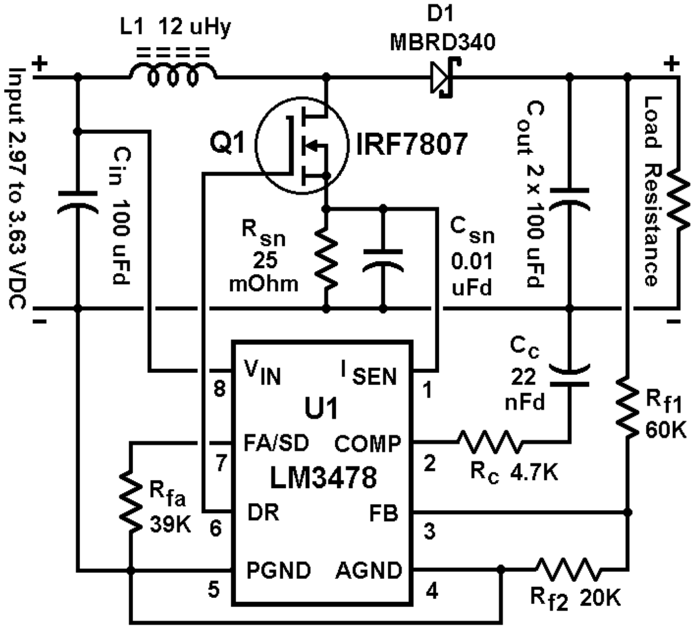


Figure 31-8 Example of a Boost converter for 3.3 V ± 10% input to 5 VDC at 2 A maximum, taken from National Semiconductor data.

Ordinary 10% tolerance values will work for most frequencies. There is no tolerance mentioned in the datasheet for frequency control.

Combining a Cuk and SEPIC Converter

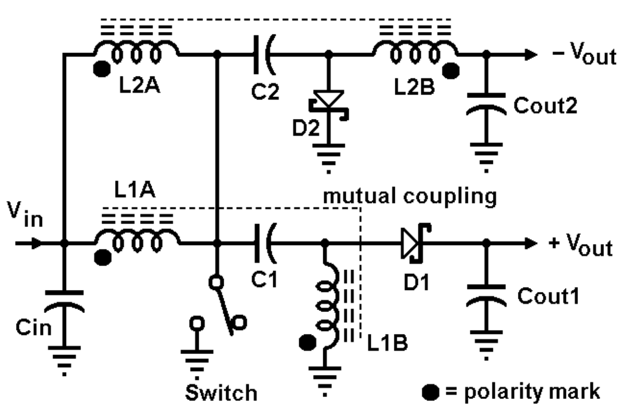


Figure 31-9 Combined Cuk (upper half) and SEPIC (lower half) converter for a ± output voltage from a single source.

Application Note AN-1106 from Analog Devices, Inc., has combination of a Cuk converter for a negative output and a SEPIC for a positive output with both outputs tracking within ±2% and DC input being a unipolar, semi-unregulated source.

The basic circuit is shown in Figure 31-9, a slight variation on Figure 31-4. A single MOSFET switch can handle both converters. Another feature here is the use of mutually-coupled inductors for each converter, glossed-over in this Chapter. The example in this AN uses a pair of Coilcraft LDP4012-153 two-winding inductors for 15 μHy per winding.

Feedback from the $+V_{OUT}$ terminal is used for the Analog Devices ADP1613 combining the MOSFET switch and the controller. This example is for a 5 VDC input although the ADP1613 should be able to work down to 2.5 VDC. The companion ADP1612 should be able to work with input voltages down to 2.0 VDC. Both are in an 8-lead MSOP flat package.[8]

Both the ADP1612 and ADP1613 have selectable switching frequencies controlled by package pin 7. If grounded the switch frequency is 650 KHz, if tied to V_{IN} it is 1.3 MHz; if left unconnected it defaults to 650 KHz. Voltage regulation is by pulse-width modulation. Maximum load current drain is about 100 mA from each output port.[9] Ideally, current loads on outputs should be equal for this combined circuit but tests by Analog Devices show that voltages remain within 5 percent for rather gross variations of current loads from 2 to 100 mA, both supplies.

Design Details

If your computer has an *EXCEL* capability, then you can download the *ADP161x SEPIC-Cuk Designer* from www.analog.com freely, enter the Inputs dialog box contents and get just about everything needed for component identification and specifications. Once on the Analog website, click on *Resources and Tools.* From there you can either search for a part number or click on *Tools, Software and Simulation Models.* Under the latter, you can select the ADP161x Sepic-Cuk Designer or download the entire group of models of Analog switchers.

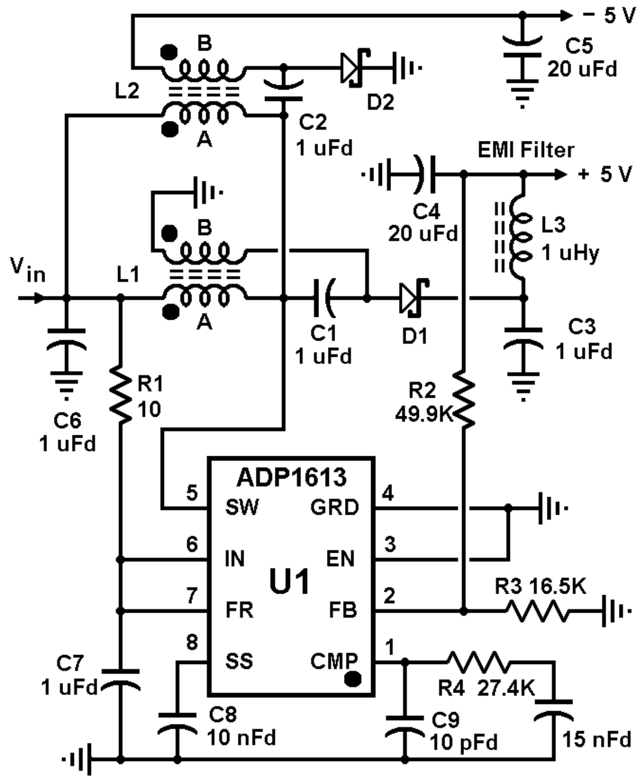


Figure 31-10 SEPIC-Cuk converter (from Analog Designs) given in their AN-1106.

Figure 31-10 is the schematic as given. L1 and L2 are Coilcraft LDP4012-153 and L3 is a Coilcraft ME3220-102MLB. C1, C2, C3 are 1.0 µFd, 16 V, ceramics. C4, C5 are each two 10 µFd, 6.3 V. C6 and C7 can be duplicates of C1 through C3. Resistors R2 and R3 set the positive output voltage level and are ± 0.1 % tolerance devices. Internal error amplifier compensation is done externally by R4, C9, C10. There isn't any data on Schottky diodes D1 and D2 but nearly all that will handle a minimum of 200 mA forward current will work here. C8 is an *inrush limiter*.

Analog Devices does not have an *Evaluation Board* yet, but the circuit is fairly simple to lay out. Warning: Short leads are necessary along with the opposite side foil being nearly intact to take away heat. It will fit on a PCB that is less than 4 square inches in area.

This converter was considered for the WWVB receiver project (in following Chapters) but filed away due to spikes and no need for a battery power source.

---

[8] The 8-lead MSOP is a bit over 1/8-inch square with leads on 0.025 inch centers. Infuriating small size but indicative of nearly all new ICs beginning the second decade in this new millennium.

[9] It should be about 200 mA maximum although the example was not tested fully with that load.

Power Factor Correction

Power factor derives from early in electrical engineering as an attempt at explaining the AC current phase relative to AC voltage phase. Before there were switching converters, most of the *power factor correction* was done by passive elements, generally applied to motors. Power supplies of the earlier times used transformers to change voltages. When silicon rectifiers were used to convert the low frequency AC to DC, there was always a current spike at the leading edge of voltage magnitudes. That came from charging the filter capacitor which maintained a reasonable steady voltage of DC output. With dozens, even hundreds of such electronic equipment in older days, there was a constant spiking of AC voltage, most of it damped by AC distribution transformers and whatever networks were installed by AC power companies.

Consider the time element. Spiking of AC current occurs at 120 Hz with a period of 8 1/3 mSec. A switching regulator can operate with a period of 120 KHz or higher, a thousand times (or more) faster. By adding some active components to a switching converter IC, it is possible to make the AC input current essentially glitch-free. It can do this generally by pulse-width modulation of the AC input, converting the AC voltage into a lower DC voltage which can then be regulated to a finer degree. There would be no 120 Hz current spikes on the AC line. A relatively simple lowpass filter could remove most HF transients from this ***PFC*** or *Power Factor Correction* input circuit.

A PFC circuit is really an outgrowth of a switching converter. Whether it is necessary in a project is up to the hobbyist. With the solid-state era fully established, most electronics uses perhaps a tenth, perhaps less, of what it took in a half century ago. Meanwhile, most appliances in an average dwelling are still drawing roughly the same AC power. The *need* for a PFC circuit would only be reasonable in a high-power transmitter in order to stay within newer AC power regulations.

There are a number of Application Notes and ICs from the leading IC makers available. Most of those are fairly detailed and can be the subject of extra project time in the learning curve of a hobbyist. They aren't included here for the simple reason that they begin to branch away from the usual hobby pursuits.

For a good reference on *Power Factor Correction,* go to www.onsemi.com and download ***Power Factor Correction (PFC) Handbook,*** Revision 4, February 2011, document number HBD853/D, ON Semiconductor.

Warnings

All switching converters generate RFI. One has to be prepared to deal with suppressing such interference both at the input and output. RFI is mostly a problem for analog circuitry. Very little digital circuitry is affected. Switching converters are good for portable equipment to conserve battery drain but are still required to suppress the converter's spikes that can go up high in frequency.

It is possible to be lulled into complacency by things such as PCs at homes. Nearly all PCs use switching converters in one form or another. But, such PCs in wide use are also operating primarily on a digital basis and aren't as affected by switching transients from converters.

Chapter 32

PLLs and DDSs

Phase-Locked Loops (PLLs) and Direct Digital Synthesizers (DDSs). Emphasis on calculation of PLL Loop Filters. Some descriptions of DDS IC types. Peripheral devices used with PLLs. Less abstract theory with simpler solutions.

General

The first Phase-Locked Loop or PLL was invented in 1932, before transistors and digital logic existed as we know it. PLLs would form the heart of many multi-frequency receivers, transmitters, and transceivers during the latter third of the 1900s. As a frequency synthesizer, a PLL can generate hundreds of precise frequencies, each with the accuracy of its single quartz crystal controlled reference oscillator. This Chapter concentrates on the *Integer-N* type of synthesizer although some mention is made of other subsystems.

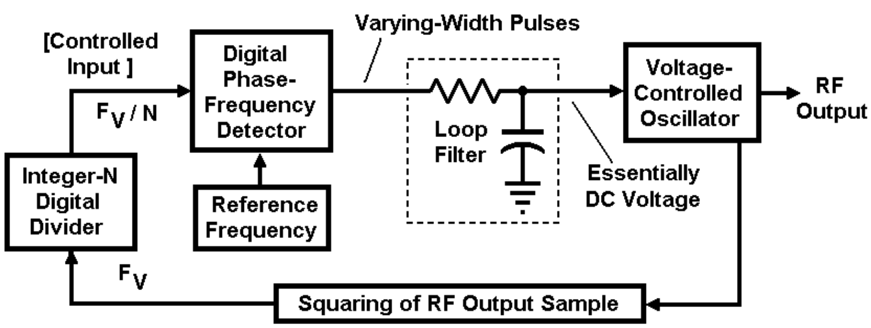


Figure 32-1 Basic PLL block diagram for an Integer-N PLL

The basic PLL subsystem is shown in Figure 32-1. It is a closed system by itself, capable of locking the Voltage Controlled Oscillator to any divided-down frequency from a single-frequency reference.

The *Loop Filter,* shown in generic form, can be configured to be for a voltage-output or current-output PFD (Phase-Frequency Detector). The VCO (Variable Controlled Oscillator) is usually a variable-capacitance diode tuned by a voltage input device covering less than an octave of frequency bandwidth. The Integer-N Divider controls the VCO such that the output frequency is equal to ***N*** times the ***Reference Frequency.*** It should be noted that output frequencies will occur at multiples of the Reference Frequency.

Figure 32-1 shows a voltage-output PFD and a voltage-controlled VCO. That can be re-configured for a *current-output* PFD and a current-controlled VCO. The difference is essentially in the circuitry doing the sensing and controlling; basic equations for the PLL remain the same.

The Closed-Loop System

A PLL is a very independent sub-system. Basic equations involve ***control-system theory.*** As such it is time-dependent to the amount of integer division and its *Reference Frequency.* It can be thought of as a variable motor analogue whose rotational speed can be set to precisely-controlled increments of Revolutions per Minute. But, rather than go through the *control-system theory* basics, the essential component equations can be presented to get the sub-system running.

Both stability and control-ability are dependent on ω_N, the ***loop natural frequency,*** and ζ (Greek symbol Xi). Figure 32-2 is a plot of ζ curves over a given time-duration of ω_N increments. The whole point of Loop Filter calculations is to get ζ within a well-behaved region, usually a range of about 0.6 to 1.0.

If ζ becomes large, say above 4 or so, the system will take many periods of ω_N to close. A very large ζ value may not close in a long time. If ζ is small, say less than 0.05, the system may sit there and bounce back and forth and take a long time to stabilize. A very small value may cause the system to continue oscillating on its own.

An ideal situation is to have a value of ζ within a region of 0.5 to 1.0. Monitoring the frequency correction with an oscilloscope will show the same curvature as shown in Figure 32-2.

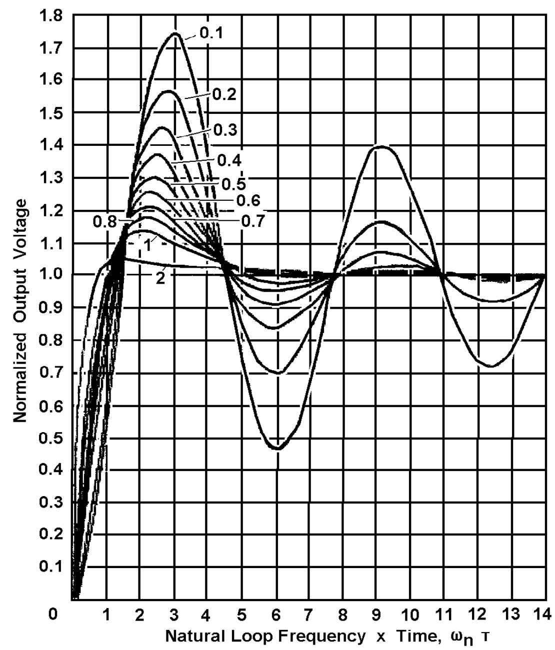


Figure 32-2 Frequency-change response of a Type 2 PFD Loop Filter to an abrupt change from normal.

PLL Loop Filter Calculation, Voltage-Output PFD

Loop Filter

The type of PFD Loop Filter is shown in Figure 32-3 for two types; Passive and Active. Both are considered ***Type 2*** in old control-systems theory. This is considered to be a good compromise for both fixed-frequency and variable-frequency PLLs.

The active filter can use an Op-Amp and has the advantage of allowing an output DC voltage shift as needed by the controlled element.

Note that both Loop Filters allow very low frequencies through. Reactance of C will be very high at low frequencies. At high frequencies it will appear as a near short-circuit so the control

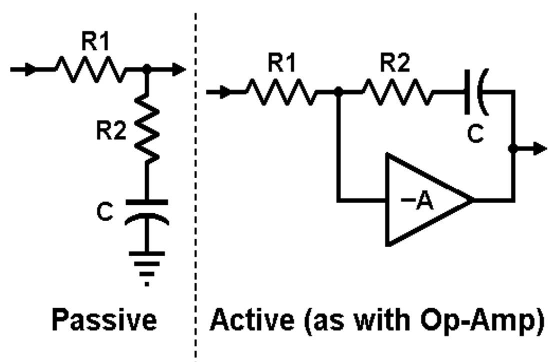


Figure 32-3 Type 2 Loop Filter for a voltage-output PFD, passive and active.

will be limited to an output equal to R1 / (R1+R2).

Constants

Using control-system terminology, the following should be known in terms given:
Kv = Variable Controlled Oscillator in radians/Second/Volt which is equal to (MHz/V) x 2π.[1]
Kn = Inverse of digital divider which is equal to (1 / N-lowest) to (1 / N-highest).
Kp = PFD output in Volts/radian and equal to half the Maximum and Minimum divided by 2π.
Kf = A factor which will determine R1 relative to value of C in the Loop Filter.
N = Integer range of Digital Divider, Lowest value to Highest value.

It is expected that you will know the polarity of the PFD output and polarity of the frequency control voltage. They would normally be opposite to one another. The Active Loop Filter introduces a polarity reversal, necessary for its action as an integrator.

A practical approach is to plan the PLL for an output frequency ratio of about 2:1. This fits the usual frequency-controlling limits. If a variable-capacitance diode is used and the VCO inductor fixed, the capacitance must vary by the square of the frequency ratio, in this case 4:1.

Calculation of Example 1, from AN-535

This used the old Motorola application note AN-535 as an illustration.[2] It featured an all solid-state PLL that tuned 100 KHz increments from 2.0 to 3.0 MHz, using Motorola ICs no longer manufactured.[3] That made N = 20 to 30. The VCO was a voltage-variable astable multivibrator so that Kv = 11.2 10^6 and Kp = 0.111.

Initial choice was to have a lock-in time of 1.0 mSec at a $\zeta = 0.8$ at 3.0 MHz. Interpolated from Figure 32-2, that would make ω_N equal to 4.5 / 1 10^{-3} or 4500 radians/Second. Overshoot (from Figure 32-2) would be less than 20%. From that, C, R1, and R2 can be calculated:

$$R_1C = \frac{0.5\,K_P\,K_V}{N\,\omega_N^{\,2}} = \frac{0.5\cdot 0.111\cdot 11.2\cdot 10^6}{30\cdot 4500^2} = 1.0232\cdot 10^{-3}$$

Choosing C as equal to 0.5 µFd, that made R_1 = 2046.4 Ohms.[4]

---

[1] This is given as both ***Kv*** and ***Ko*** in Garth Nash's AN-535. For simplicity this uses only ***Kv.***

[2] Written by Garth Nash at Motorola prior to July 1994 and available from Freescale Semiconductor under the same AN535 as of February 2006. A good tie-in to control-system theory as well as presenting an example.

[3] It used the old Motorola MC4000 and MC4300 family of TTL-compatible ICs which debuted in the 1970s. Some of those are still available through individual distributors.

[4] If C was selected as 0.47 µFd, then R1 would be 2129.8 Ohms and R2 would be 756.50 Ohms. In staying with 5% tolerance values, C would be 0.47 µFd, R1 = 2.2K and R2 = 750 Ohms..

$$R_2 = \frac{2 \cdot \varsigma}{\omega_N\ C} = \frac{2\ \cdot\ 0.8}{4500\ \cdot\ 0.5 \cdot 10^{-6}} = \ 711.11 \text{ Ohms}$$

The final choice in AN535 was to use 2.0 KOhms for R1 and 680 Ohms for R2.

AN535 shows some oscilloscope photos of the correction voltage which agrees closely to Figure 32-2. Included in AN535 is a computer plot of correction voltage done at divisions of 20 and 30 to prove the point; those are very close to the Figure also.

Calculation of Example 2 For a Project

This is the PLL explained in Chapter 48. It would tune 43.0 to 72.5 MHz in 500 KHz increments with a Mini-Circuit POS-75 as the primary source. Prescaled by a ÷16 the PLL IC would input 2.68750 to 4.53125 MHz in 3.90625 KHz increments. The N divider would be 688 to 1160. Given that the prescaler is fixed, the calculation values would use the lower values except that the tuning N divider would remain the same. The MC145151-2 PLL IC chip would do the work. From that and the IC datasheet, the following constants would be used:

Kp **=** 0.380 Volts/Radian (approximated)
Kv **=** 1.178 10^6 Radians/Second/Volt = [(3.0 MHz/V) / 16] 2π

The choice of ω_N was based on the previous example which used increments of 100 KHz. With increments of about 4 KHz, the times would be about 26 times slower. Note, this is a starter value to see how things fit. With, again, a ζ of about 0.8 and an $\omega_N\ \tau$ of 4.5 as in the first case, assuming a lock-up time of 26 mSec, ω_N would be about 173. The square of that would be 29,956; that can be rounded to 30,000 with only 0.1% error..

Assuming also that the Op-Amp version of Figure 32-3 is used, Loop Filter components can be calculated approximately at mid-band with:

$$R_1\ C\ =\ \frac{K_P\ K_V}{N\ \omega_N^{\ 2}}\ =\ \frac{0.38\ \cdot\ 1.178 \cdot 10^6}{936\ \cdot\ 30 \cdot 10^3}\ =\ \frac{447.64 \cdot 10^3}{28.08 \cdot 10^6}\ =\ 15.94 \cdot 10^{-3}$$

Choosing a value of C as 1.0 µFd, R1 would be equal to about 15.94 KOhms.

$$R_1 = \frac{15.94 \cdot 10^{-3}}{1 \cdot 10^{+6}} = \ 15.94 \cdot 10^3$$

For R2:

$$R_2\ =\ \frac{2\ \varsigma}{C\ \omega_N}\ =\ \frac{2\ \cdot\ 0.8}{1 \cdot 10^{-6}\ \cdot\ 173}\ =\ \frac{1.6}{173 \cdot 10^{-6}}\ =\ 9.249 \cdot 10^3$$

Note that, in both cases, values of R1 and R2 are the inverse of changes in C. If C were chosen as 0.1 µFd or one-tenth, both resistor values would go up by ten times. Using CMOS input Op-Amps, that could be done. Op-Amp input bias current is restricted by C to just through R1.

To keep things practical, R1 can be chosen as 15 KOhms and R2 as 10 KOhms, both 10% tolerance values. The 1.0 µFd capacitor can be a low-working-voltage unit ceramic unit. Using standard values allows a check of both ω_N and ζ with the help of some formulas given in ON Semiconductor application note AN1410D, revision 3 of March 2000, entitled *Configuring and Applying the MC74HC4046A Phase-Locked Loop*.

$$\omega_N = \sqrt{\frac{K_P\, K_V}{N \cdot C \cdot R_1}}$$ Note that only N is a variable here, all others fixed.

$$\varsigma = \frac{\omega_N\, C\, R_2}{2}$$ Formula for ω_N must be done first

These allow checking near-final values of the Loop natural frequency and ζ independently. It also indicates that R2 has influence on setting the final ζ value. Given fixed values of C, R1, and K_P, K_V both can be calculated for two different values of R2:

N	ω_N	ζ for R2 = 8.2K	ζ for R2 = 10K	
688	208.27	0.85391	1.0414	Lowest frequency
936	178.56	0.73210	0.89280	Mid-frequency
1160	158.76	0.65092	0.79380	Highest frequency

It doesn't matter much on the exact value of R2 since the resulting ζ values fall within an acceptable range. There is only a small overshoot indicated on Figure 32-2 at highest frequency with R2 = 8.2K. What does matter is the accuracy of initial calculations that lead to this stage.

Example 3 Examination for a Future Project

This is an examination of a possible PLL for the future SW BC Receiver first described later. It is for a Local Oscillator tuning 26.4 to 37.2 MHz in 1.0 KHz increments. It will be compared with a DDS sub-system that tunes the same range. Because of the tight requirements, this needs to be carefully done to see if the Loop Filter will work.

The LO output is prescaled by a ÷4. The LO source itself is a Mini-Circuits POS-50 (25 to 50 MHz design range). Prescaler output is then 6.6 to 9.3 MHz. The Digital Divider's divisor is 26,400 to 37,200. Original POS-50 tuning range is 2 MHz/Volt. It will work with a discrete 16-bit Divider and the separate Phase-Frequency Detector is part of an 74HD4076 IC. That is a modified HD4046 with a Lock Detector in place of its third PFD internally.[5] The constants:

Kp = 0.30 Volts/Radian as an approximation of the phase detector output (*PC2)*.[6]
Kv = [(2 MHz/V) / 4] x 2π = 3.1415 10^6 Radians/Second/Volt
N = 26400 to 37200 range, 31800 mid-frequency value

---

[5] A standard 4046 could be used with some external circuitry. Having a production model available with that Lock feature incorporated saves PCB space.

[6] That may have been closer to 0.38 Volts/Radian. It may have come from a slight confusion in reading the datasheet.

Tuning will normally be of the near-set-and-forget variety. This is not for the usual radio-oriented user. On the other hand, the user is expected to do some searching in frequency. The 1 KHz increments of tuning is for the possibility of strong adjacent signals; the internal crystal filter of its IF will pass AM signals but most stations are located at 5 KHz increments.

In contrast with Example 2, the Loop Natural Frequency is lowered, but not by much. A trial is done with an ω_N equal to 40. That would make $\omega_N^2 = 1600$. It is faster than that of Example 2 but not a direct numerical change relative to the incremental tuning frequencies. The calculations:

$$R_1C = \frac{K_P\,K_V}{N\,\omega_N^2} = \frac{0.30 \cdot 3.1416\cdot 10^6}{31.8\cdot 10^3 \cdot 1.6\cdot 10^3} = \frac{942.48\cdot 10^3}{50.88\cdot 10^6} = 18.523\cdot 10^{-3}$$

Choosing C = 0.1 µFd would make R1 about 185 KOhms. That is a bit high so choosing C = 1.0 µFd would drop that down to about 18.5 KOhms. To pick R2 with C equal to 1.0 µFd:

$$R_2 = \frac{2\,\varsigma}{C\,\omega_N} = \frac{1.414}{1\cdot 10^{-6} \cdot 40} = 35.350\cdot 10^3$$

To examine this over the tuning range, keep C equal to 1.0 µFd, choose R1 equal to 18K and R2 equal to 33K for 10% tolerance values. For three frequencies this works out to:

N	ω_N	ζ
24600	44.535	0.73483
31800	40.578	0.66954
37200	37.517	0.61903

That is acceptable but changing ω_N to be equal to 80 (roughly twice as fast to lock in) would change R1 x C to be 4.6309 x 10^{-3}. Choosing a C value of 0.22 µFd makes R1 = 21.050K or, choosing a C value of 0.27 µFd makes R1 = 17.151K. Staying with 0.27 µFd would make R2 = 65.468K. For C = 0.27 µFd, R1 = 18K, R2 = 68K, the three frequencies work out to:

N	ω_N	ζ
24600	85.707	0.78679
31800	78.092	0.71688
37200	72.202	0.66281

This seems more desirable and would be tried first in hardware. What can't be found mathematically is the effect of the tuning ***rate*** for small changes in frequencies. It may very well be that an absolutely positive small-increment control of tuning frequency is not practical. This could be replaced with a simpler frequency counter display of a standard variable LO manually-tuned with a variable capacitance or variable inductor. That would save a lot of hardware.

Note that there are more factors in the final decision on hardware than can be found by mathematics alone. On the other hand, a thorough solution, as given above, will prove practical in finding the final choice.

Loop Filter Calculation, Current-Output PFD

General

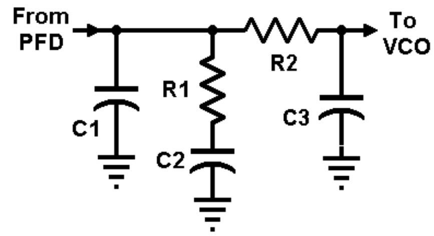


Figure 32-4 Constant-current source Loop Filter.

These are almost the same as the voltage-output except that the PFD output is a ***constant current*** as opposed to the constant-voltage type just discussed. The Loop Filter for a constant current output develops a voltage output from the Filter's reactance times the current source. The constant-current driver is generally called a ***charge pump*** as if *pumping up* electrons on the capacitors of a Loop Filter.

Note that the constant current ***source*** of the PFD has limitations in ***voltage compliance;*** such resulting voltage limits its control capabilities. The PFD Loop Filter output will be a varying voltage which re-tracks the VCO tuning diode into the proper phase lock. The Loop Filter must be a Type 2 as shown in Figure 32-4.

Constant-current source PFDs are generally found in UHF and higher frequency synthesizers which require low control power and small capacitor values. The calculation formulas found here are from a September, 1999 Fujitsu publication in *Microwaves and RF* by Ken Holladay and Dennis Burman. It can also be found in the Fujitsu *Super PLL Application Guide* available separately.

Example 1 from Fujitsu Paper

In Figure 32-4 C1 is the main voltage source. R1 and C2 form the frequency response relative to the Loop control frequency. R2 and C3 is a harmonic-suppression R-C filter to attenuate higher-frequency spurious signals.[7] The constants for calculation:

Fh = Highest frequency in range of frequencies, in Hz.
Fw = Total of all Incremental steps in range of PLL output, Hz.
Fs = Incremental step from one frequency to adjacent frequency, Hz.
Fa = An *accuracy* value of a PLL step frequency, Hz.
Fvs = Controlled VCO tuning sensitivity in MHz/Sec.
Nmax = Digital divider divisor value for ***Fh/Fs.***
t_{LU} = Approximate lock-up time, Seconds, for controlled Loop
Iccp = Maximum constant current out of PFD, Amperes.

For the Fujitsu article, the VCO has a tuning range of 1675 to 1735 MHz, a 60 MHz range. That makes ***Fh*** equal to 1735 MHz, ***Fw*** equal to 60 MHz. VCO tuning sensitivity is 25 MHz/Volt so ***Fvs*** is 25 MHz. N range is 8375 to 8675 and ***Nmax*** is then 8675. Desired Loop lock-up time is 500 µSec. N and ζ (Loop damping factor) have the same meaning as with the constant-voltage PFD source calculations.

---

7 R2 and C3 chosen to be approximate or ***ball-park*** values from the Fujitsu documents. Done for simplicity in calculation. Some of the ***constants*** were renamed by the author to avoid confusion with voltage-source PFDs.

$$\text{natural frequency} = F_N = \left(\frac{-1}{2\pi \cdot t_{LU} \cdot \varsigma}\right) \text{Ln}\left(\frac{F_A}{F_W}\right) \qquad \text{[Note the Natural Log]}$$

$$F_N = \left(\frac{-1}{6.2832 \cdot 500 \cdot 10^{-6} \cdot 0.707}\right) \text{Ln}\left(\frac{1 \cdot 10^3}{60 \cdot 10^6}\right) = (-450.45)(-11.002) = 4955.9$$

$$C_2 = \frac{I_{CCP} \cdot F_{VS}}{N_{MAX}(2\pi \cdot F_H)^2} = \frac{6 \cdot 10^{-3} \cdot 25 \cdot 10^6}{8675 \cdot 965.65 \cdot 10^6} = \frac{150 \cdot 10^3}{8.4117 \cdot 10^{12}} = 17.832 \cdot 10^{-9}$$

$$R_1 = 2 \cdot \varsigma \sqrt{\frac{N_{MAX}}{I_{CCP} \cdot F_H \cdot C_2}} = 1.414 \sqrt{\frac{8675}{6 \cdot 10^{-3} \cdot 25 \cdot 10^6 \cdot 17.833 \cdot 10^{-9}}} =$$

$$1.414 \sqrt{\frac{8675}{2.6749 \cdot 10^{-3}}} = 1.414 \sqrt{3.2432 \cdot 10^6} = 2546 \text{ Ohms}$$

$$C_1 = \frac{C_2}{10} = 1.7832 \cdot 10^{-9} \qquad R_2 = R_1$$

$$\textit{Loop Bandwidth}, \text{Hz} = \left(\frac{2\pi \cdot F_N}{2}\right)\left[\varsigma + \left(\frac{1}{4\varsigma}\right)\right] = \left(\frac{6.2832 \cdot 4956}{2}\right)\left(0.707 + \left(\frac{1}{2.828}\right)\right)$$

$$= 15.570 \cdot 10^3 \cdot 1.0642 = 16.523 \cdot 10^3$$

In realistic values, 5% tolerance values were used. C1 and C3 became 1800 nFd, C2 became 0.018 μFd, R1 and R2 became 2500 Ohms.

Example 2 from *Super PLL Application Guide*

For mobile telephone use, this Loop Filter was designed for a VCO tuning 1005 to 1031 MHz with 35 MHz/V sensitivity and increments of 200 KHz. N would be 5025 to 5155. Constant-current output from the PFD would be 10.0 mA and desired switching time would be 450 μSec. Desired frequency accuracy at each increment would be 1.0 KHz. Damping factor or ζ would be 0.707 for calculation. The constants:

Fh = 1.031×10^9
Fw = 26×10^6
Fs = 200×10^3
Fa = 1×10^3

Fvs = 26×10^6
Nmax = 5.155×10^3
t_{LU} = 450×10^{-6}
Iccp = 10×10^{-3}

$$F_N = \left(\frac{-1}{6.2832 \cdot 450 \cdot 10^{-6} \cdot 0.707}\right) \text{Ln}\left(\frac{1 \cdot 10^3}{26 \cdot 10^6}\right) = (-500.25)(-10.166) = 5085.5$$

$$C_2 = \frac{10 \cdot 10^{-3} \cdot 35 \cdot 10^6}{5155 \cdot (6.2832 \cdot 5085.5)^2} = \frac{350 \cdot 10^3}{5.2632 \cdot 10^{12}} = 66.499 \cdot 10^{-9} \approx 0.0665\ \mu\text{Fd}$$

$$R_1 = 2 \cdot 0.707\sqrt{\frac{5085.5}{10 \cdot 10^{-3} \cdot 35 \cdot 10^{6} \cdot 66.499 \cdot 10^{-9}}} = 1.414\sqrt{\frac{5085.5}{23.275 \cdot 10^{-3}}} = 661.46$$

$$C_1 = \frac{C_2}{10} = 6.6499 \cdot 10^{-9} \qquad \text{and} \qquad R_2 = R_1$$

$$\textit{Loop Bandwidth} = \left(\frac{6.2832 \cdot 5085.5}{2}\right)\left[0.707 \cdot \left(\frac{1}{2.828}\right)\right] = 15.977 \cdot 10^{3} \cdot 1.0606 =$$

$$16.945 \text{ KHz}$$

Final parts were to 10% tolerance values, making C2 equal to 0.068 μFd, C1 = 6.8 nFd, and R1 and R2 = 680 Ohms.

It should be noted that the Loop Bandwidth is the closest to the Loop lock-in time. Since both constant-current PFD Loop Filters were desired to lock-in at, respectively, 500 and 450 nSec, both Loop Bandwidths would be very close. Note also that these more-simplified Loop Filter calculations only figured in ζ at its nominal value, 0.707. In the real world, ζ would be expected to vary, but only slightly from that nominal value.

Some Factors Common to both PFDs

Loop Filter Component Tolerances

Five percent tolerance parts should work out quite fine if the calculations are correct. In many cases a ten percent tolerance will work. What ***is required*** is a good balance (to good calculations) to the ***interactive values.*** It is much like the attenuation versus frequency of passive L-C filters: If there is a good balance of ***ratios*** of values, then the locking to a reference frequency will go smoothly. It can be that there is a slight discrepancy in lock-in time or the time to settle in to a stable frequency but that will not be of much concern.

Using Op-Amps Between Loop Filters and VCOs

The ***amount*** of voltage amplification between a Loop Filter and the VCO must be limited. A VCO input is quite capable of being FMed by small glitches in voltage. Open-loop gain of modern Op-Amps is between 10,000 and 100,000 times so the slightest μV of garbage on the input can result in a large frequency deviation. Figure 32-3 ***cannot be used by itself*** with an Op-Amp**.**

First of all, the Op-Amp will not find an operating point since there is no way for the output to affect the input at DC. Secondly, it will probably need some offset of its output DC voltage to match the VCO. The circuits shown in Figure 32-5 are suggested. In that Figure, R1, R2, and C all refer to the Loop Filter values calculated for a voltage-source PFD.

In Figure 32-5 **A** the VCO control voltage polarity is inverted. That is normal for a positive-going control voltage increasing the VCO frequency and the PFD output is positive-going to decrease the frequency. ***Voltage gain*** of the Loop Filter was not mentioned other than *it should be large.* There was no specification on *how large* the *large* should be...or its effect on the Loop when in correction.

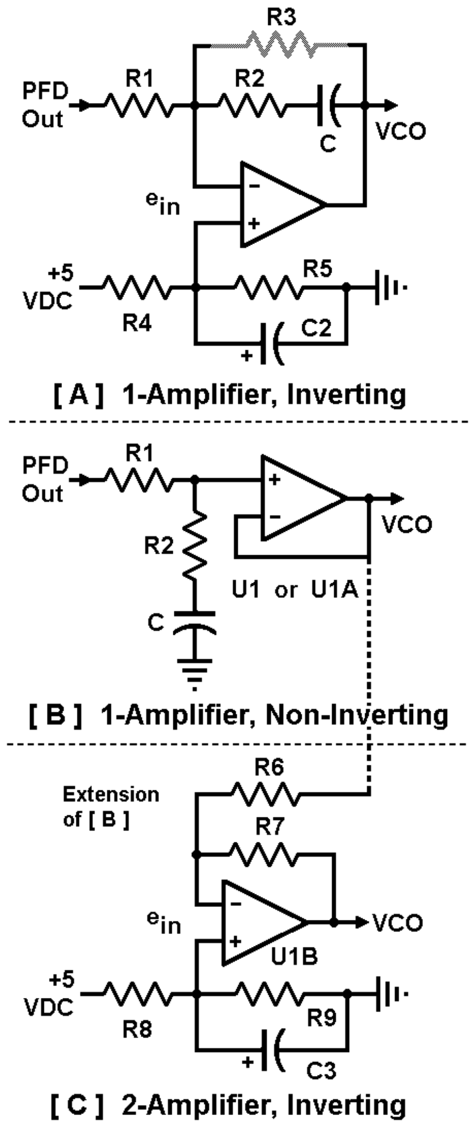


Figure 32-5 Suggested post-PFD amplification and PFD center-voltage adjustments. *Inversion* is relative to PFD & VCO polarity.

In Example 2 for the voltage-source PFD, R1 was calculated as being 15 KOhms and R2 chosen as 8.2 KOhms. The DC gain of an Op-Amp depends on its feedback resistor ratio. But, that feedback resistor, R3 in this case, will effect the series connection of R2 and C. For a DC gain of 10 V output for 1 V input, a 150 KOhm resistance will have little effect on R2 and C in series. Note that, for DC, R2 cannot affect any change in input since it is blocked by C.

To stabilize an Op-Amp requires both inputs to be at the same relative voltage value to ground. Given a VCO voltage range of +2.0 to +8.0 VDC the center of that is about +5 VDC. Approximate center of PFD output would be +2.5 VDC.

If R3 (shown in grey in Figure 32-5 A) were not there, the Op-Amp negative input would assume about 2.5 VDC as a center of control voltage. In that case, R4 would be equal to R5 and both could be 10 KOhms. The inputs would both be +2.5 VDC. There is no telling what the Op-Amp output voltage would be other than at ground or the supply rail.

If R3 were 30 KOhms then the Op-Amp gain would be 3 at DC. Voltage to ground at the negative input would be +3.333 VDC. R4 would have to be 5 KOhms to fit that.

Making R3 as 150 KOhms would have the DC gain at 10. Op-Amp input balance voltage would be about +2.727 VDC. R4 would have to be about 8.333 KOhms.

R4 could be fixed and R5 made variable over a small range. It doesn't matter which. It is a problem in the DC domain and simple arithmetically. C2 is shown as an electrolytic (the + sign) but also needs a ceramic in parallel to reduce any slight disturbances in voltage caused by higher value capacitors, especially old tantalum capacitors from the junk box.

Figure 32-5 **B** is a simple voltage follower with a gain of unity, no inversion. It can effectively isolate the Loop Filter from the VCO. It does not have any offset control for the control center voltage. That is taken care of by Figure 32-5 **C.**

Figure 32-5 **C** is an extension of **B** but keeps separate control of the Loop Filter gain and the calculated R1-R2-C values. As a voltage follower, input and output DC voltages in **B** would be equal. R8 and R9 in **C** handle that the same way as in **A.** The major difference is a more complete division of functions. Loop Filter calculations are not disturbed. Loop Filter gain is left as an isolated topic. Neither are disturbed by the other.

Dual CMOS input Op-Amps are available on the market with ***low noise*** specifications and the ability to operate with just one supply rail.

Lock Detection Pin in the MC14515x Family

Later versions of the MC14515x family had an added ***LD*** pin for Lock Detection.[8] This is useful having a front-panel indicator of ***unlock*** or bad operation of a PLL. For the MC14515x family that occurs with a Logic 0 shift of the *LD* pin from its normal Logic 1 or high-impedance state. A simple circuit to add a driver for the LD pin is shown in Figure 32-6. This one has about a 14 mA LED current across the indicator yet takes only about 120 µA from the PLL IC LD pin when the LD pin is low.

There is a narrow spike in the LD pin output in most applications so the capacitor from transistor base to ground damps that out. The value of the capacitor is not critical. Almost any general-purpose capacitor will work just fine there.

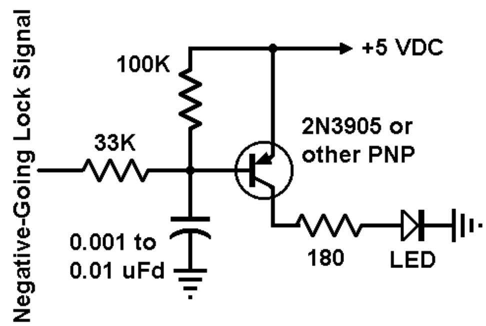


Figure 32-6 Indicator for Out-of-Lock for the MC14515x family.

The single PNP transistor is not critical either. A 2N3905 or other high-β PNP can be used there. Since the PNP conducts only when the input is low, the 33K resistor could be changed down to 27K or 22K if the DC current gain of the PNP is low.[9]

Lock Indicator for the 74HC7046A or 74HCT7046A

That often-used PLL IC has a positive-going Lock signal and can use the same circuitry as with Figure 32-6 except that the negative-glitch damping capacitor is connected to the 7046A pin 15 and the Lock signal comes from Pin 1 for the drive input. Some of the gating circuitry is within the 7046A so it takes two separate connections.

The capacitor value is dependent on the internal PFD reference frequency. The value is about 1.0 µFd for a PFD reference frequency of 100 Hz, about 1 nFd for a PFD reference of 100 KHz, about 100 pFd for 1 MHz. For other PFD reference input frequencies, just scale those points on log-log graph paper. It can also be seen on the '7046A datasheet.

Other PLL Terms and Devices

Single and Dual-Modulus Operation

The main difference there is in the internal divider circuit, particularly with preselector dividers. As an example a ÷65/÷64 means that the fixed divider of a preselector can be firmware

---

[8] That can be checked by the part number suffix. An MC145151 without a suffix has no Lock Detection circuitry. An MC145151**-2** (with that suffix) does have it.

[9] The Lock Display Driver was done by Joseph D. Loritz, N9ZIA, on his Green Bay website page.

controlled for either divisor. That allows both divisors to be running during a PLL operation and can also enable what some call ***Fractional-N*** control.

To explain that fully would take another Chapter or two but is not covered here. It can be found in several application notes in detail.

What Is *Pulse Swallowing?*

This is a euphemism equal to *dual-modulus division.* It is apparently an effort used by some application-note writers to explain the process of using two different divisors in a Prescaler under output of an ***MC*** (Mode Control) pin of a main divider.[10]

Higher-Frequency PLLs

At the close of the first decade of the new millennium, most available one-chip PLLs have become specialized. In essence, the major circuitry within an IC is familiar to what has been discussed but the internal details may be dissimilar. It is up to the hobbyist designer to first ***get to know a PLL IC*** from its datasheet at first. Most of the rest will fall into place based on information in this Chapter.

Fractional-N

The term is another euphemism. It depends on the division of the reference frequency within the PLL subsystem, thus the incremental frequency control. Otherwise, the ***Fractional-N*** PLL is the same as an ***Integer-N*** PLL It all depends on where one wants to put the decimal point in the output frequency plus the fractional increments of frequency control. In some fractional-N explanations, a ***modulus*** term is used, representing the input of a (usually) serial word to represent a binary equality to the output frequency. Using the term *modulus* refers more to the DDS or Direct Digital Synthesis.

Direct Digital Synthesis or DDS

General

A DDS IC internally blends a fast shift-register, a fast digital-to-analog converter, and whatever else the manufacturer wants to insert to provide a quantized sinewave output. None of these can be changed from the outside so any discussion of ***how*** it is done is really meaningless. Input to set the frequency is generally digital.

Control Word

[10] Unfortunately, use of different terms to explain the same thing has become common.

Control Word is the name given to the binary input of a DDS IC by Analog Devices, Inc. In its internal structure, there is a binary digital-to-analog converter that accepts the binary word, thus producing an analog output, the desired frequency.[11] With the ADI model AD9851 DDS IC, this can be represented by:

$$f_{OUT} = \frac{C_W \cdot \text{Ref}}{2^{32}} \quad \text{where: } f_{OUT} = \text{desired frequency in MHz}$$

$$C_W = \text{Control Word in decimal}$$

$$2^{32} = 4{,}294{,}967{,}296 \qquad \text{Ref} = \text{Reference frequency in MHz}$$

When the frequency is known, the Control Word can be found by:

$$C_W = \frac{f_{OUT} \cdot 2^{32}}{\text{Ref}} \qquad \text{where variables are as above}$$

For a hexadecimal representation of the Control Word, and lacking a scientific calculator to show Hex numbers to at least 8 places, a suggestion is to review Chapter 28 and the decimal-to-binary radix conversions.

There are several different powers-of-two in various DDS model numbers. The AD9850 was among some of the first to reach a multiplier of 2 to the power of 32; the AD9851 is a later, improved version of the AD9850. That power allows specifying an output frequency that is within ***milliHz*** of the reference frequency. As such, a few rather accurate audio to HF and to low-VHF signal generators have been devised, principally by or for hobbyists.

Nyquist Response of Output Power

Internally, most of the ADI DDS ICs use a ***sampling*** technique to derive the output. As a result, analog output follows the *Nyquist theorem.* Analog output amplitude is a function of the following:

$$X = \frac{\pi \cdot F}{F_{REF}} \quad \text{where } F = \text{Frequency of Interest}$$

$$F_{REF} = \text{Reference Frequency (same terms as F)}$$

$$\text{Relative Amplitude} = \frac{\sin X}{X} \qquad [\text{ Sin X in Radians}]$$

In taking a number of F values that increase, very low frequencies will have an amplitude close to unity. As frequency increases, amplitudes begin to fall, decreasing until they drop to zero at the Reference frequency. As frequency goes beyond the Reference, amplitude increases to 1.5 times Reference, then falls to zero again at twice the Reference frequency. Amplitude follows the Sin X over X envelope.

[11] Once packaged, it can't be changed. The AD9850 and AD9851 are used here as examples since they have been used by hobbyists for a variety of circuits.

As a percentage of Reference frequency, the following shows output amplitude in db:

Percent	db	Percent	db	Percent	db
5	-0.04	25	-0.91	50	-3.92
10	-0.14	30	-1.32	60	-5.94
15	-0.32	35	-1.83	70	-8.69
20	-0.58	40	-2.42	80	-12.6

For practical applications, a maximum output frequency should be at less than 50 percent of the Reference frequency, usually around 40 percent. For a Reference of 180 MHz, desired output should stop at roughly 75 MHz.

Output Filtering

This is done by a lowpass filter. Cutoff frequency at -3 db point should be at about half of the Reference frequency or just higher than the highest desired frequency. The sampling that is done internally will still apply so that adds to the attenuation of the lowpass. Several DDS ICs have a built-in comparator for squaring up a signal and also to equalize differences due to Nyquist sampling.

There are some cases where a desired output frequency has an integer relationship to the Reference frequency. That can lead to peculiarities in output waveshape close to the integer relationship to the Reference. Not all DDS ICs have that but, if observed, they are mentioned somewhere at the manufacturer. Analog Devices ***Tools*** section has that for their line of DDS ICs.[12]

Reference Clock Versus System Clock

System Clock refers to the internal timing frequency. ***Reference Clock*** is the external standard applied. The AD9850 and AD9851 have selectable ***6-times multipliers.*** With multiplier disengaged, *System Clock = Reference Clock.* With multiplier engaged, internal *System Clock* is the (selected) multiplier of *Reference Clock* frequency.

A multiplier built-in is an advantage. Rather than use a VHF-range precise frequency source, such as 180 MHz, a multiplier allows that external source to be 30 MHz. Internally, operation is at the *System Clock* frequency (180 MHz) but externally the standard *Reference Clock* is lower (30 MHz).

Loading a Control Word

That can be bit-serial or byte-serial. The AD9850 and AD9851 have both. Because of a limited pin count on their packages, those two require ***five*** bytes (of 8 bits each byte) in byte-serial,

[12] ***Design Tools : ADIsimDDS*** for an on-line graphical and numeric calculator. It calculates the Hex, Decimal, and Binary forms of a Control Word, shows all combinations of spurious outputs, includes a waveform viewer for desired frequency output, and an insertable general filter model to see differences in output. Check the ***Remarks*** box. If there is a web name change, use the search feature to find it.

40 bits total in bit-serial. Frequency output, per se, is controlled by four bytes in byte-serial or 32 bits in bit-serial. Seven of the remaining 8 bits are for housekeeping, and one bit is reserved for factory test, always a logic 0.

Housekeeping bits include engagement of the 6-times multiplier when logic 1, a power-down bit, and 5 bits for ***phase modulation*** (32 increments of 11.25 each). For Continuous Wave operation (such as for a fixed but variable Local Oscillator) the 5 phase modulation bits can remain unchanged.

Byte-Serial Loading[13]

A timing diagram is shown in Figure 32-7. Number in parenthesis refer to pin numbers on the AD9850 or AD9851 packages. The positive-going edge of the ***W_CLK*** engages one data byte loading and the positive-going edge of the ***FQ_UD*** signals the internal circuitry that all five data bytes are completely loaded.

Byte contents are organized as follows: Byte 0 contains *housekeeping* data; Bytes 1 through 4 contain the 32 bits of output frequency data. Bits within a byte are ***D7*** (pin 25), the Most Significant Bit (or MSB) to ***D0*** (pin 4), the Least Significant Bit (or LSB). Individual bits within each byte are organized as follows:

	Byte 0	Byte 1	Byte 2	Byte 3	Byte 4
D7 (pin 25)	Phase bit 4	Freq.31	Freq.23	Freq.15	Freq.7
D6 (pin 26)	Phase bit 3	Freq.30	Freq.22	Freq.14	Freq.6
D5 (pin 27)	Phase bit 2	Freq.29	Freq.21	Freq.13	Freq.5
D4 (pin 28)	Phase bit 1	Freq.28	Freq.20	Freq.12	Freq.4
D3 (pin 1)	Phase bit 0	Freq.27	Freq.19	Freq.11	Freq.3
D2 (pin 2)	Power-Down	Freq.26	Freq.18	Freq.10	Freq.2
D1 (pin 3)	LOGIC 0	Freq.25	Freq.17	Freq. 9	Freq.1
D0 (pin 4)	6X Multiplier	Freq.24	Freq.16	Freq. 8	Freq.0 (LSB)

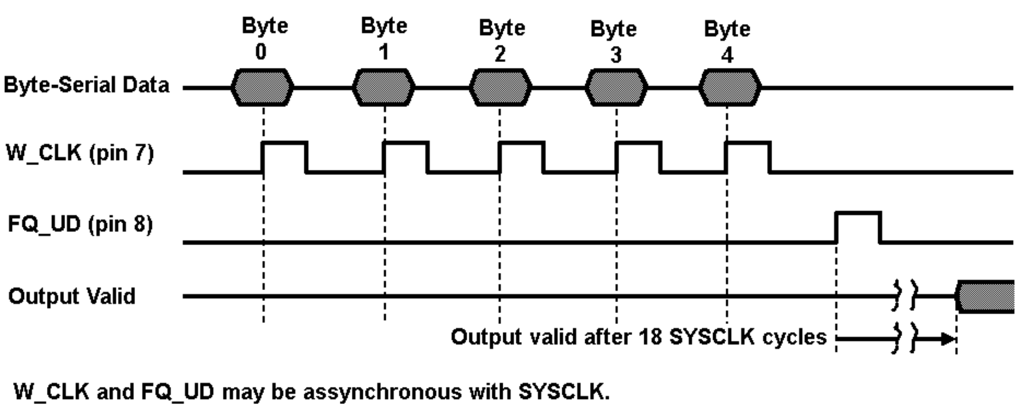


Figure 32-7 Normal Frequency/Phase Loading timing diagram for Byte-Serial mode.

In Power-Down, bit D2 of Byte 0 is a Logic 0 in normal operation. A Logic 1 means the DDS is powered down. Bit D1 is reserved for

[13] Analog Devices datasheets refer to byte-serial data input as ***parallel*** even though there are 40 bits of data but only 8 bits to enter it. This can be confusing to some.

factory test and should always be Logic 0. If D0 in Byte 1 is Logic 0 it means that the *system clock* is the same as the *reference clock*; if it is Logic 1 it means that the internal six-times frequency multiplier is engaged. *Phase bits* 0 through 4 are used for phase modulation in 11.25 increments.

Byte-serial loading mode is the fastest for frequency loading, provided there is an 8-bit data path for input to the AD9851.

Bit-Serial Loading

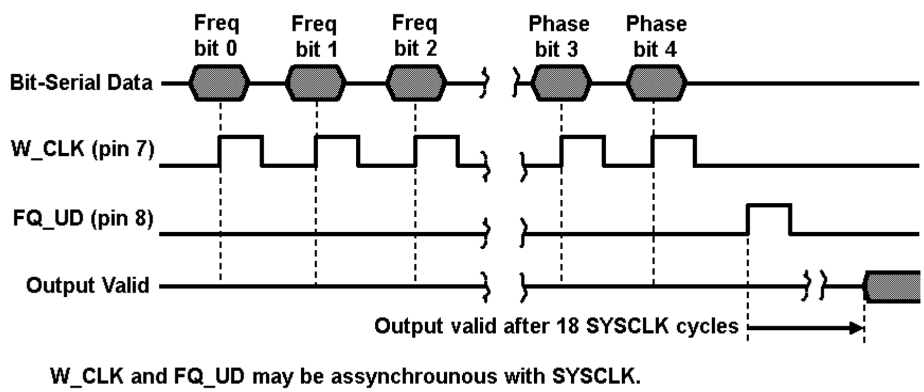


Figure 32-8 Bit-Serial timing with Control Word bit order reversed.

Timing diagram is in Figure 32-8. The order of the Control Word bits is reversed. Input is always to D0 input (pin 4) with D1 through D7 grounded. This is the slowest frequency-loading mode but has the advantage of using only one line for the data. Two control lines are still required, W_CLK and FQ_UD.

Byte-Serial PHASE Modulation

This can be done in 11.25 phase increments. Only the ***first byte*** needs to be input, then the W_CLK, then the FQ_UD. The whole sequence can be done in the shortest span of about 20 nSec followed by a 75 nSec output valid delay. One can figure about 100 nSec total or 0.1 µSec. Longer phase-change times are possible with W_CLK and FQ_UD timed at integer multiples of SYSCLK.

Power-Down and Resetting

Powering down, for whatever reason, is done by setting the *Power-down* bit to Logic 1 and reloading the Control Word. For Byte-Serial mode, that can be is as fast as Byte-Serial Phase Modulation.

Filtering the RF Output

For use as a DDS for an RF output, IOUT pin 21, needs to be filtered. This can be done as in Figure 32-9. Main AD9851 RF output is from IOUT (pin 21). It will have everything including all sampling products. The Lowpass Filter will attenuate most unwanted sampling products.

The asterisk in Figure 32-9 is a DC modifier (two 100 KOhm and 470 pFd) to set the baseline of the Comparator via its negative input. Level into the positive Comparator input is a combination of the Sin(X)/X of the RF output and Lowpass filter characteristic.[14] The 3.9 KOhm resistor connected to R_{SET} (pin 12) sets the RF output level into the Lowpass.

If the Sin(X)/X drop is 4 db and Lowpass insertion loss is 1.5 db, level into VINP would be 5.5 db down or 0.531 times a much-lower frequency. If there is 10 mA RMS of RF output current into the Lowpass and its impedance is 50 Ohms, then input into the Lowpass is 0.500 V RMS. With total attenuation that would be 0.265 V RMS or 0.751 V peak-to-peak. The datasheet graph on Comparator triggering level can be consulted to see if that is enough for the Comparator output at VOUT (pins 14 and 13).

Comparator input impedance is high and its output can drive an optional 74AC14 or 74ACT14 Hex Inverter with Schmitt Trigger.

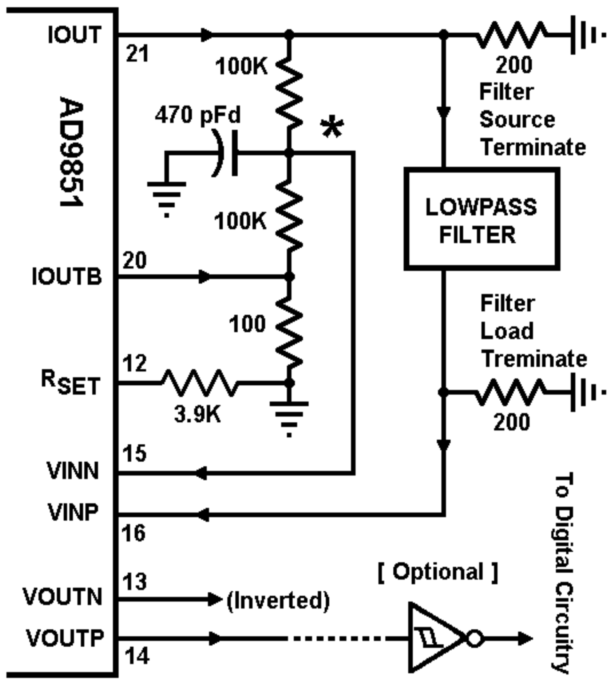


Figure 32-9 Using a lowpass filter and the built-in comparator of the AD9851.

Lowpass Filter Type

It is desirable to have a sharp cut-off, therefore an Elliptic (or Cauer) filter type is usually stated. The AD9851 datasheet shows a 7-section Lowpass with a 70 MHz design cutoff frequency with about 90 db slope in 40 MHz cutoff slope and about -75 db minimum attenuation beyond the first null. Because of the small capacitances of a 200 Ohm terminating impedance would require tight tolerances of parts, a Sensitivity test for component tolerances should be done. As a minimum, five-percent tolerance Lowpass components are also recommended.

Minimum External Reference Input Level

A peculiarity of the AD9851 is the ***minimum frequency*** of the externally-supplied clock input. Without the 6-times multiplier that is 1 MHz. With the 6-times multiplier that is about 5 MHz. This is not true of other DDS ICs. It pays to carefully read everything in a current datasheet.

DC Power Demand

This can vary depending on the external clock frequency. The datasheet carries a rough graph of total power demand. For example, with a +5.0 VDC supply rail and 50 MHz output with an external 120 MHz Reference, current demand is about 90 mA. Total power demand is then 450 mW.

[14] As shown in the AD9851 datasheet, Revision D, published by Analog Devices.

Appendix 28-1

A *Do-It-Yourself* Constant-Voltage PFD for a PLL

A Phase-Frequency Detector provides an output for yielding correction voltages to a voltage-controlled oscillator that is part of a PLL. A PFD will help generate a waveform that represents the *phase lead* or *phase lag* of a controlled frequency relative to a constant reference frequency. As important, the outputs will remain in a low-high or high-low relationship when the controlled input is lower or higher than the reference. That provides a *pull-in* for the control of the VCO if it is too far off in frequency.

With the controlled-signal input lagging the reference frequency phase, the operating waveforms are as in Figure 32-11. All gate designations refer to outputs. Note: The assumption is that both inputs are squared-off around their baselines; see Figure 32-12 waveforms for non-overlapping signal and reference phases.

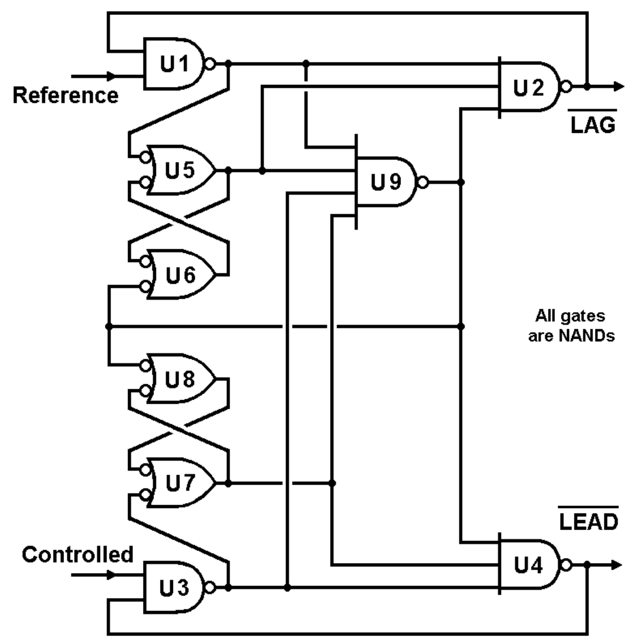


Figure 32-10 An all-gate digital PFD from the original Motorola MC4044 IC.

When the reference input goes low, U1 goes high and that will cause U2 (*LAG-not*) to drop low. Those states remain until the controlled-signal input goes low, causing U3 to go high. That causes a sequence of events to happen, both U4 and U9 going low, U4 low causing U2 to go back high, U9 going low making U4 back high (only two gate delays there), U8 going high, and U6 going high. U6 high will make U5 low which, in turn, makes U9 high again. All states remain as they are until the reference input goes high. That makes U1 low which drives U5 high and, subsequently, U6 back low. When the controlled-signal input goes high it makes U3 go low which then makes U7 high and then U8 low. Note that all gate outputs' states on the right (end of an input cycle) are the same as the ones on the left end (before start of a cycle). The sequence of state changes repeats on next cycle.

What if the controlled-signal phase *leads* the reference? Due to the symmetric arrangement, the state changes essentially change position. U4 will stay low in proportion to the phase lead and U2 will have a very short low state length. Supposing the low states of both inputs do not overlap?

That condition is shown in Figure 32-12, an extreme case being one or both inputs being pulses instead of squared-off sinusoids. Delays are exaggerated to show the sequence of gate state changes.

The major difference in Figure 32-12 versus Figure 32-11 is in the state changes of U5 and

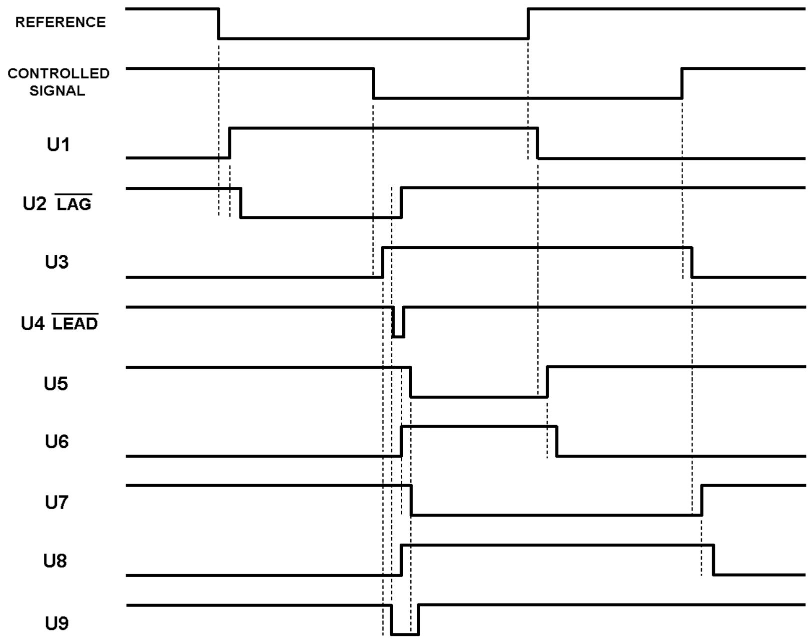


Figure 32-11 Waveforms of Figure 32-10 circuit. NAND gate Phase-Frequency Detector, controlled-signal phase lagging reference. Delays are exaggerated.

U6, one of the two intermediate RS flip-flops in the PFD. U2, *LAG-not*, output will still have a low state proportional to the controlled-signal input phase lag relative to reference frequency input and U4 will still have a very brief low state of about one gate delay.

When the controlled-signal input phase leads the reference frequency input, U2 output will have a very brief low state. That condition, whether lagging or leading, is a fixed low-state length and, if U2 and U4 are made to opposite polarity input of an integrator, would be the same as an essentially fixed offset bias. It would matter little to a PLL control loop following the PFD.

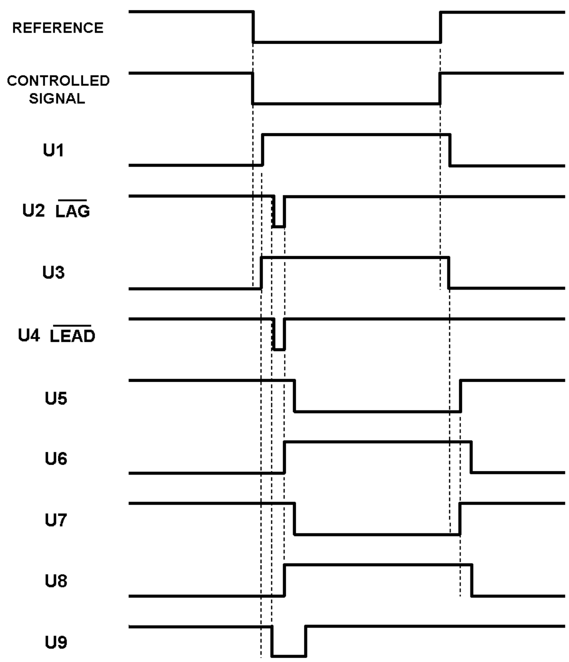


Figure 32-13 Waveforms of reference and controlled-signal inputs overlapping.

If the controlled-signal input phase lagged the reference frequency so much that its frequency was ***less*** than the reference, U2 output would remain low and there would be no brief low-state of U4 output, U4 remaining high. Similarly, if the controlled-signal input led the reference so much that its frequency became ***greater than*** the reference, U4 output would remain low and U2 output would remain high with no brief low state. The low-state time lengths are proportional to the phase lag or lead relative to the reference, up to ±360 before the lag-lead outputs are steady-state.

When the signal and reference are in exact phase of each other, both U2 and U4 will still exhibit the brief one-gate-delay-width low state as can be seen in Figure 32-13. Intermediate RS flip-flops U5-U6 and U7-U8 will follow the input waveforms and U9 will be low for only about 4 gate delay times.

Since U2 and U4 are both low about the same amount of time at synchronization of phases, they effectively cancel each other out. If the signal input lags, U2 output will be low in proportion to the lag in time, up to -360 but U4 will go low for that one gate delay time. Similarly, if the signal input leads the reference, U4 will be low in proportion to the lead time up to +360 but U2 will be low no longer than a gate delay time.

Note there is a ***dead-band*** with the reference and controlled signal inputs at exact or near-exact overlap. This is common and must be accounted-for in circuits used for deriving phase difference operation as a sensor.

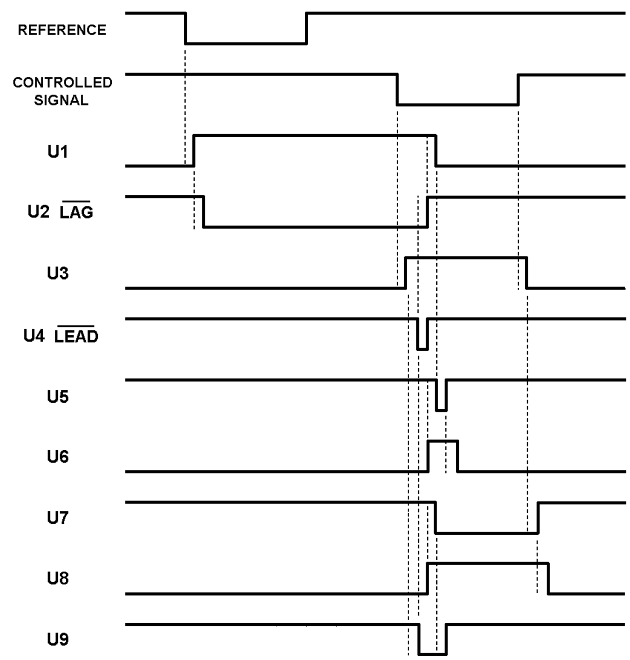


Figure 32-12 Waveforms of Figure 32-10 when there is no overlap.

Personal Reference

This circuit was used at Rocketdyne (then a division of Rockwell International) as the main detector for the phase of 1.0 MHz optical signals in an SDR contract for the USAF. The material here is also adapted from the September 1982 issue of *Ham Radio* magazine under the title *Inside a Phase-Frequency Detector.*

Chapter 33

Modulation and Demodulation

Basic ways to Modulate and Demodulate an RF Carrier wave for AM, SSB, FM, PM, and touching on a few other types for communications intelligence conveyance by radio.

General

Chapter 4 covered the basic mathematics of the three major modulation forms. This Chapter covers *hardware* methods**.** It is possible to combine different modulation forms on the same carrier wave.

In the beginning of radio (1896) technology was quite simple. There were no vacuum tubes. Transmitters, if one can call them that, were either damped-wave machines (***spark gaps)*** or high-frequency rotary alternators outputting on VLF. Receivers were the very simplest, amounting to nothing more than tuned half-wave rectifiers. Useable radio frequencies in the beginning were lower than today's MF. *Modulation* (to stretch a point) were done by turning the transmitter On or Off. Telegraphic codes, already established by a half-century of practice, were used as the *intelligence* medium. It was a time of very high power transmitters and very insensitive receivers, using rather low frequencies and lots of wire for antennas.

By 1915 John R. Carson of AT&T had formulated the first mathematical basis for the three basic means of ***modulation***. The first vacuum tubes were available. Technology had begun. Some of that is explained using semiconductor equivalent circuits following.

First, Amplitude Modulation

This was done as in Figure 33-1. The modulation source is amplified at its normal audio range and that varies the high voltage supplying the carrier amplifier. RF voltage ***envelope*** follows the audio waveform and resulting AM is fairly close to the original audio. This was how it was for the first broadcasting stations...and is still done a century later in MF BC radio.

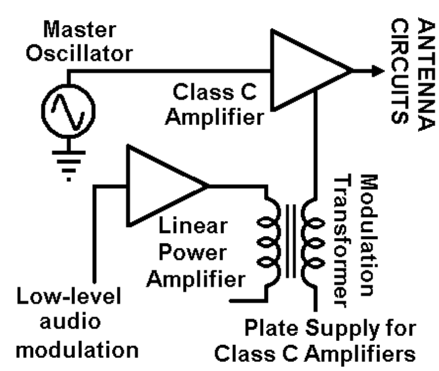


Figure 33-1 First AM transmitter, still used on AM BC band.

For receiving a simple half-wave rectifier, acting much like a power supply AC-to-DC circuit, converted the incoming AM at RF to audio. That audio was amplified to a speaker for all to hear. It was a simple system but there were some things not quite right.

The ***Class C*** RF power amplifier did not quite have the linearity for good audio. It took some time before RF detectors were linear enough to supply negative feedback to the modulator stage for linearization. The audio power amplifier had to be biased ***Class A*** or ***AB*** and have a power rating of roughly ***half*** the Watts of the RF Amplifier.

Receivers needed some amplification ahead of the detector, at first by ***TRF*** or Tuned Radio Frequency stages. Howard Armstrong would be on the scene to improve things. Armstrong's ***regenerative detector***, using ***positive*** feedback, increased the sensitivity of old CW signals in receivers by 20 to 40 db. A major change. By 1918 Armstrong came up with his ***superheterodyne receiver*** patent which would set the stage for nearly all receivers built from then on.[1] A HUGE change in radio technology. Armstrong would later come up with FM broadcasting and the hardware to make that work. Another milestone.

Single-Sideband AM in the Beginning

Telephone companies had been experimenting with ***Carrier*** systems to reduce the cost of installation of hundreds of miles of long wires for telephony.[2] Telcos had taken the basic Carson equations and noticed that AM allowed frequency translating using only one sideband. They could even drop the translated carrier frequency, re-inserting it later to complete the translation back to original audio. Telephonic ***carrier*** systems were the unsung background into the first long-haul radio links using ***Single-Sideband*** beginning in the late 1920s. By the beginning of World War II there were many ***SSB*** networks, each carrying two voice channels and up to 8 teleprinter channels, all independent but sent over the same transmitter. There were only a few single-channel SSB sets.

FM in the Beginning

The last of the basic modulation types to be implemented in radio, it suffered more from lack of circuit technology. It was first tried in some late-1930s police radios with success. It offered almost constant volume, a marked absence of noise, less *wasted power* in transmitters. Trials were stopped due to WWII manufacturing needs but would be picked up by the military during the War.

CIRCUITRY

Amplitude Modulation

This tended to (somewhat) stagnate in the three decades of 1920 through 1950. Many small circuits were tried and used to vary amplitude in transmitters. Receivers tended to hold to the simpler diode detectors. While telegraphy was still a component of radio on HF, the ***beeps*** of Morse Code could be generated by a ***BFO*** or Beat Frequency Oscillator. Since a diode detector is still a

---

[1] There is some dispute on that patent since at least two others claim to have invented the ***superhet*** (as it became colloquially known). It should also be known that civil courts of the 1920s and 1930s were quite busy with all sorts of inventor's claims and counter-suits. Nonetheless, professional radio groups have proclaimed Armstrong to be ***the*** inventor of the superhet. Some claim that naming his device a ***super***-heterodyne was in reference to Aubrey Fessenden's 1906 discovery of ***heterodyning*** with a continuous-wave spark unit seeming to increase the sensitivity of primitive receivers.

[2] ***Carrier*** systems in telephony-speak referred to several voice circuits ***frequency-multiplexed***, typically for four independent voice channels within a 12 KHz bandwidth (type C ***carrier***). Each voice channel took 3 KHz of bandwidth. Given that hundreds of miles of wire were devoted to long-distance lines, this meant that telephone companies could easily quadruple their service use.

form of a mixer, an added BFO, loosely coupled to the diode detector input, would heterodyne an On-Off Carrier code signal at IF into an on-off audio tone. Much easier to listen to, it was introduced into 1930s *communications receivers.*

Figure 33-2 shows a typical diode detector, used from the first days of radio and on into the new millennium. The *Rectifier* in the beginning was a ***coherer*** or a ***Galena crystal*** but much later it was a semiconductor diode of Schottky construction.

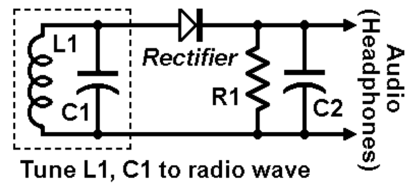


Figure 33-2 Basic AM detector used from 1896 to present day.

SSB with Suppressed Carrier

Success of SSB depended on several things. It needed more-stable Local Oscillators. It needed more quartz crystal devices, for bandpass filters as well as LOs. It needed new circuitry such as balanced mixers. It would need improvements in RF power amplification, particularly in linearity. It would need a more liberal viewpoint in radio use.[3]

Two typical balanced mixers are shown in Figure 33-3. In [A] the input and output transformers and diode quad help isolate Signal Input from the IF Output. In [B] the transistors are all part of the same die and can be made very identical. A ***Gilbert Cell*** depends on such a balanced structure. In both cases the Input frequencies are balanced ***out*** of each others outputs, an important attribute of balanced mixers. It should be noted that such balanced mixers are not confined to SSB use but may be used elsewhere with success.

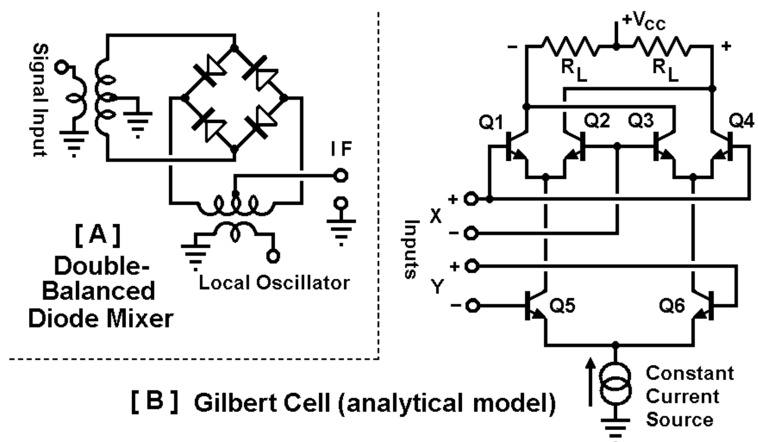


Figure 33-3 Two basic balanced mixers, suitable for SSB suppressed carrier applications.

Detectors for SSB have been renamed as ***Product Detectors.*** While a new name, a *Product Detector* is still a ***Mixer*** and is little different than the basic half-wave diode detector of Figure 33-2. Any circuit allowing such heterodyning can be called a Product Detector.

SSB *Filter* System

Shown in Figure 33-4, this SSB transmitter represents the least-complicated (but usually more expensive) system for transmitting a SSB signal using the ***Filter*** method. A quartz crystal lattice filter has the necessary stopband attenuation for this sort of use.

For reception a similar (or the same) crystal filter is used at the input to the final IF. Only the USB will be received in Figure 33-4. A carrier frequency must be re-introduced into the final detector to complete demodulation. That re-inserted carrier frequency must be below the crystal

[3] USA radio amateurs tended to be rather conservative in demanding that telegraphy remain the forefront of amateur radio skills. Many continued with that viewpoint on into this new millennium. That continued despite all the tangible wonders of modern electronics all around them.

filter's passband. Things would be the same for an LSB signal, transmitting or receiving, if the carrier re-insertion were on the other side of the filter.

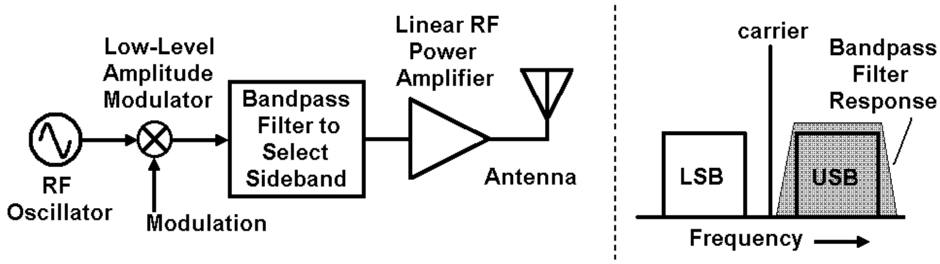


Figure 33-4 A *Filter* method SSB transmitter. The shaded area on the right represents the Upper Sideband passband. Only the USB is allowed to go through the linear power amplifier.

The ***Filter Method*** is the least complicated system. The quartz crystal lattice filter is passive and may be checked out separately in the simpler scalar frequency-versus-amplitude test arrangement.

SSB *Phasing* System

This is more cumbersome to explain since various ***phases*** are involved as well as amplitudes. Both modulation and demodulation require pairs of balanced mixers plus two relatively wideband phase-shift networks. The *Modulation* input filter can be the ***All-Pass*** kind shown in Figure 33-7. The *Carrier* input filter can be that of Figure 33-6.

It should be noted that the phases are ***relative to one another;*** i.e., it is usually drawn as zero phase for one output and 90° shift for the other output but that only means that the 90° shift arm is later in time than the unshifted arm. The phasing difference is ***between the two phases.*** They are not referenced to anything but themselves.

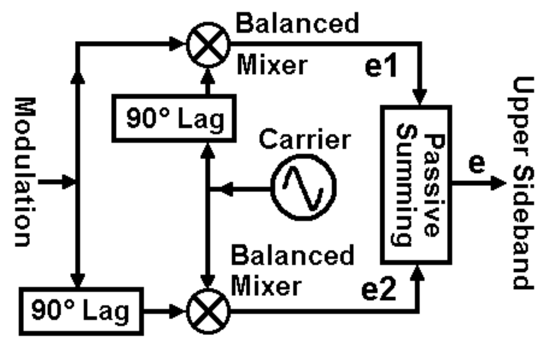


Figure 33-5 An USB Modulator using a Phasing Method of generation.

Figure 33-5 is an example of a phasing method Upper Sideband generator. Both Mixers are balanced types; the Carrier component is missing in both ***e1*** and ***e2***. The ***90° Lag*** at Modulation Input is from either an All-Pass or a Polyphase network. The RF carrier source ***90° Lag*** can be a variety of passive delays or the outputs of a divide-by-4 from a 4X carrier frequency oscillator as depicted in Figure 25-23. The ***Passive Summing*** block is simply a non-reactive summation device such as a resistor pad. It is assumed that both balanced mixers have adjusted amplitude outputs for equal voltages.

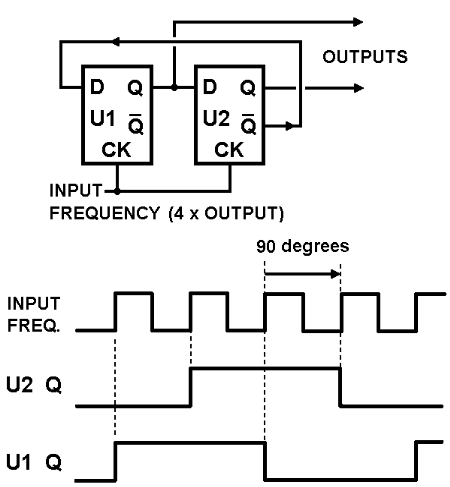


Figure 33-6 Flip-flop method to obtain 0 and 90 degree phasing differences over a wide band.

For a simple explanation of how it works, consider ***e1*** to be a product of (USB+LSB), ***e2*** is a product of (USB-LSB). For a Lower sideband the *Passive Summing* block becomes a ***Passive Subtractive block*** so that ***e = e1 - e2.***

For a trigonometric solution, try the following. Voltage magnitudes become:

$$e_1 = Cos\left[(\omega_C t + 90) - \omega_M t\right] - Cos\left[(\omega_C t + 90) + \omega_M t\right] \quad \text{and}$$

$$e_2 = Cos\left[(\omega_C t) - (\omega_M t + 90)\right] - Cos\left[(\omega_C t) + (\omega_M t + 90)\right] \quad \text{Where:}$$

ω_C = Radian frequency of Carrier = $2\pi f_{CARRIER}$

ω_M = Radian frequency of Modulation = $2\pi f_{MODULATION}$

t = Instantaneous time of amplitude of both $f_{CARRIER}$ and $f_{MODULATION}$

90 = 90 degrees or a quarter - cycle phase lag

It should be noted that the left-hand term group in both formulas represents the Lower sideband and the right-hand term group in both formulas represents the Upper sideband.

After the passive summation the total voltage magnitude is then:

$$e = e_1 + e_2 = Cos\left[\omega_C t - \omega_M t + 90\right] - Cos\left[\omega_C t + \omega_M t + 90\right] +$$
$$Cos\left[\omega_C t - \omega_M t - 90\right] - Cos\left[\omega_C t + \omega_M t + 90\right]$$

By taking an identity of $X = \omega_C t - \omega_M t$ in the above, it simplifies the top-most group to:

$$e = Cos(X + 90) + Cos(X - 90) = -Sin(X) + Sin(X) = 0$$

The bottom-most term group above then becomes:

$$e = -2Cos(\omega_C t + \omega_M t + 90) = 2Sin(\omega_C t + \omega_M t) = 2Sin\left[(\omega_C + \omega_M)t\right]$$

What is left is the Upper sideband of (ω_C+ω_M). The Lower sideband was reduced to zero. Balanced mixers effectively reduced the Carrier to zero.

For the Lower sideband only, just reverse the magnitude of e_2 or make the Passive Summer into a Passive Subtracter. Go through similar trigonometric identities as was done for Upper sideband.

Figure 33-7 shows how to get a 0° and 90° phases over a very large passband using a 74AC74 dual flip-flop. This is a reproduction of Figure 22-23 to make life a bit easier on the reader. Maximum toggle frequency at a +5 VDC supply is 130 MHz. For a typical application this would require (roughly) 5 MHz for a 2° error between the two output phases.

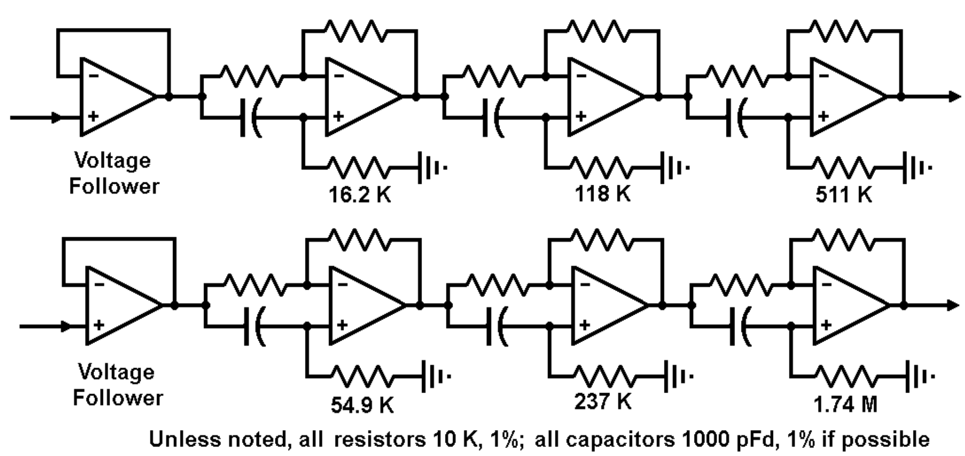


Figure 33-7 All-pass phase-shift network from Philips AN1981

For the audio phase shift, the circuit of Figure 33-7 is a copy of Philips (NXP) Application Note AN1981 that was released in October 1997.

The lower network yields a phase lag of 0.5 to 0.25 degrees from the upper network from 280 Hz to 3.0 KHz. It should also be noted that AN1981 contains a typographical error for the 118 KOhm resistor, mistakenly identifying it as 118 Ohms. A phase error of one degree results in an unwanted sideband that is 41 db down from the desired sideband.

To save design time, an on-line calculator is available at www.werewoolf.org.je/files/apf.exe written in Visual Basic. This may need a dll file from www.werewoolf.org.je/files/msvbvm60.dll to run. The download is available from ***J-TEK All Pass Filter Designer*** available (in 2012) from www.werewoolf.org.je/gj3rax.htm located in the Jersey Channel Islands. That program was based on equations of Dr. (Eng) Ralph Oppell, DB2NP, published in VHF Communications Volume No. 18, Summer, 1987.

A worthy substitute for this *All-Pass* network is the Gingell network shown in Figure 33-8. This was called a ***Sequence Asymmetric Polyphase Network*** in Gingell's PhD thesis. It is entirely passive and produces four phase-shifted outputs It may be fed by a center-tapped transformer or an inverting, gain-of-one op-amp. The four outputs may be isolated by a quad op-amp wired as voltage followers (no inversion, input impedance very high). Gingell received a UK patent on the network in 1969, a US patent on it in 1972.

Again, design solutions for this network require computer calculation to save time. An excellent paper on the subject is ***Understanding and Designing Sequence Asymmetric Polyphase Networks*** by W. J. (Pim) Niessen, PA2PIM, released as Version 4.0 on 25 August 2006. An associated Excel Spreadsheet program quoted in the paper was not found on the Internet in 2012.

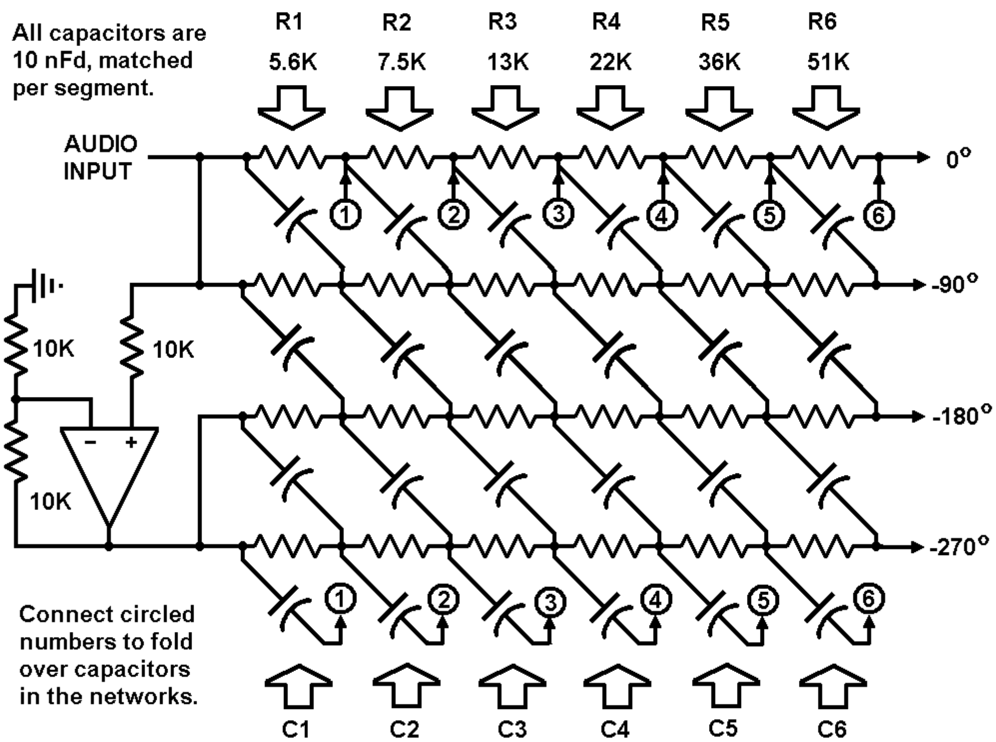


Figure 33-8 The Gingell *Polyphase* network with one set of values.

A later, more detailed analysis, was done on a Niessen Polyphase network at 30 frequencies from 100 Hz to 4 KHz. This showed excellent response, practically identical to Niessen's 2006 paper. Using values given in Figure 33-8, relative phase response was analyzed in 2012 to be less than a half-degree from 250 Hz to 3200 Hz.[4] That is good enough, with a six-segment network and balanced mixers, to have its unwanted sideband reduced to -45 db or better. It is noted that audio output amplitude varies with frequency. To equalize the output amplitude, a shaped lowpass or bandpass filter should be added at the input to compensate for a transmitter. Receiving can use an output filter in recovered audio. Not using an output filter in reception will increase bass tones with slight emphasis on the high end.

As Niessen stated, fixed capacitors will have the widest tolerances. To equalize that, capacitor values can be graded for most-equal values ***per segment.*** A purchase of 50 to 100 lot of 10 nFd capacitors should yield enough to measure them for such per-segment use. Left-overs can be used elsewhere as bypasses or audio coupling. Resistors can be ±1 % tolerance values directly.

SSB Demodulation

Figure 33-9 can be used for demodulation. It takes the same balanced mixers, the same wideband phase-shift circuits and the same passive summer (or subtracter). The Figure shows demodulation of a USB. It can handle LSB equally well.

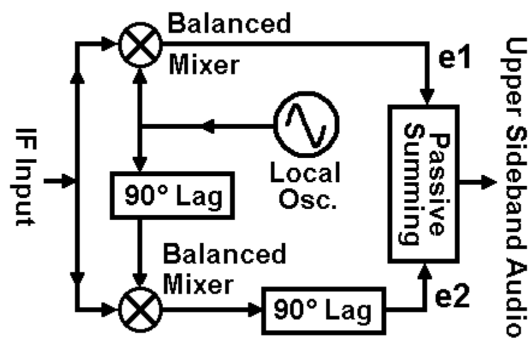


Figure 33-9 SSB USB Demodulation

Generating Frequency and Phase Modulation

Reactance Modulators

The ***reactance modulator,*** shown in Figure 33-10, was the first widely-used audio-input to do FM or PM at a low level. Audio input, isolated from RF by the RF Choke, changes the ***transconductance*** of the FET. In turn, that changes the C_{EQ} of the Source-Drain connection. That is provided that the ***Reactance*** of C is at least 5 to 10 times the value of R. The formula is:

$$C_{EQ} = g_M R\, C$$

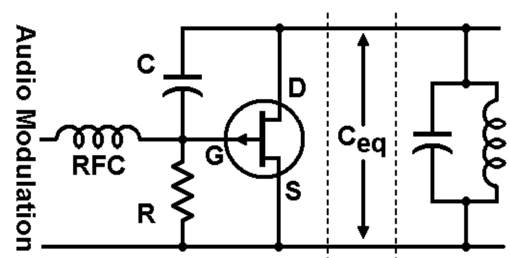


Figure 33-10 A FET Reactance Modulator which creates FM or PM.

If connected across the resonant tank circuit of an oscillator, the changing C_{EQ} does FM. If connected across a resonant circuit, such as for a tuned circuit, it produces PM. There are a couple of problems here.

[4] Done on the author's *LINEA* program, a spin-off from RCA Corporations *LECAP* admittance analysis program, itself a spin-off from IBM's original ***ECAP*** of the later 1960s. The analysis of the Niessen network used 1 MegOhm loading resistors to ground at each quadrature output to emulate a voltage-follower input, done from 100 Hz to 4 KHz in 30 logarithmic steps. There was no appreciable change in relative quadrature output using 22 MegOhm loading nor with source resistances from 10 to 100 Ohms. Added 5 and 10 pFd loading capacitors did not change quadrature errors below 4 KHz. Analysis relative quadrature errors did not exceed ±0.05 degrees.

First, it is somewhat narrowband in that the ***reactance*** of C varies with frequency. Second, the proper DC biasing of the FET is difficult to achieve since the value of R must be quite low. Depending on the active device, a bipolar junction transistor may work better. Original reactance modulators used pentodes with lower transconductance values.

Varicap Modulators

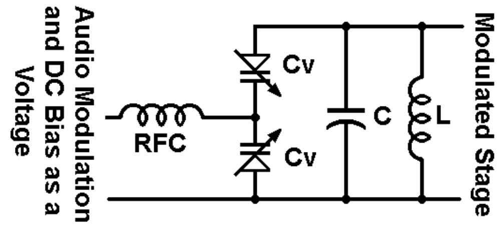


Figure 33-11 Voltage-variable-capacitor-diode modulator for either FM or PM.

Voltage-variable-capacitance diodes, operating with *reverse bias,* are a better choice. These are sometimes called ***varicaps*** colloquially. The diode's capacitance tends to remain constant over frequency, from audio to VHF. As a result they are much more broadband in application.

As shown in Figure 33-11, the two Varicap diodes are connected back-to-back deliberately. This keeps any RF voltages from upsetting the DC bias voltage input. Note also that the upper Varicap is returned to ground through the tank inductance; both diodes *see* the same bias voltage to their cathodes. As a result, Varicap modulators can carry modulation down to the DC level. They can be driven by op-amps.

As with the reactance modulator, varicap modulators do load the external tank circuit (L and C in Figure 33-11) but their Q is rather high. Total tank circuit capacitance is C + diode capacitances.

Frequency Multipliers

These work with FM or PM because the RF amplitude is not varied. Frequency multipliers are generally inefficient insofar as power consumption, but they can work well with FM or PM. A characteristic is an ***increase*** of the FM or PM modulation index by the amount of multiplication.

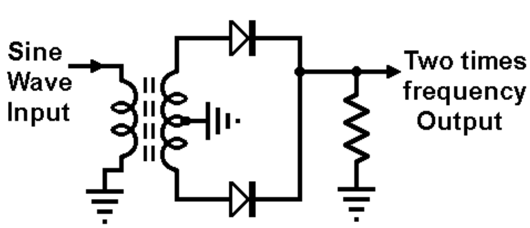
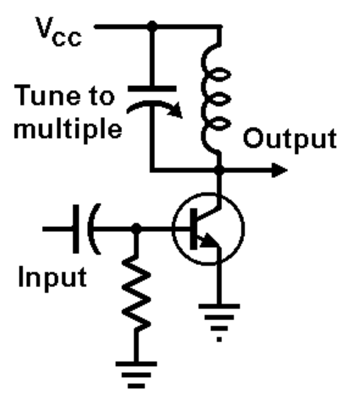


Figure 33-12 A frequency doubler



Fogire 33-13 A simple frequency multiplier.

The circuit of Figure 33-12 is good for doubling a frequency.

Diodes are picked for fast response, are over-driven, with or without biasing, to sum the distorted half-sinewaves out of the transformer. Since this is a passive device, it usually needs an amplifier tuned to twice the input frequency following doubling.

The circuit of Figure 33-13 is more flexible and requires DC biasing to keep the transistor cut off with no input. It also distorts the input sinewave with its output tuned to the frequency multiple. This sort of frequency multiplier becomes impractical about a multiple of 5 times input frequency.

Frequency multipliers need initial modeling to optimize multiplied frequency output with a given input RF power. While frequency multipliers were often used in older radio days, it was about all that was available to multiply frequencies when PLLs and DDSs were not available.

RF Amplifiers for Angle Modulation

Those can be the so-called *high efficiency* (for DC power input relative to RF power output) Class C biased power amplifiers. Since there is no amplitude modulation, linearity is less of a factor for FM or PM. All that is required is attention paid to lowering harmonics generated within Class C bias designs. Harmonics generated in final amplifiers can be attenuated by slightly more elaborate tuned circuits for final amplifier outputs.

Note that a ***linear*** power amplifier can be used with FM or PM. It is not as *efficient*[5] as using Class C bias and the same harmonic-reduction techniques must be used there as well.

The First FM Transmitter by Armstrong

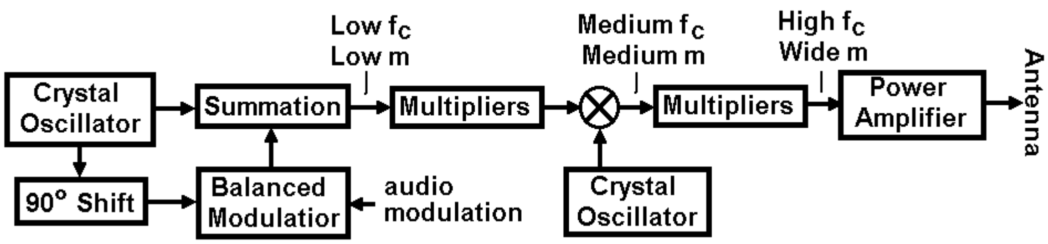


Figure 33-14 The first FM broadcast transmitter in-use in the 1930s.

Figure 33-14 shows the first FM broadcast transmitter by Armstrong that was used in his beginning network of high-fidelity sound transmitters in the old frequency allocation of 42 to 50 MHz (prior to WWII). FM broadcast frequency allocations were changed to 88 to 108 MHz after WWII.

Lacking the reactance modulator forced a slightly-unusual beginning modulation method of splitting a low-frequency crystal oscillator into two phases, modulating one of them with audio, then combining them with passive summation. Modulation index was very low. That was increased with multipliers, then heterodyned with another crystal oscillator to get a sub-multiple of the carrier frequency. With more multiplication, the wider FM spectrum resulted.

This proved out Armstrong's claims by listener acceptance. Fidelity was good and the lack of noise disturbance found most pleasing. A shut-down by the WWII effort resulted in changing the carrier frequencies which was not a great change. Adding another doubler stage in the chain of multipliers would handle that. General technology advances into FM radio, a result of the War effort, would lead to post-War changes in design.

[5] One of the great claims in amateur radio magazines in older times was the Class C ***efficiency*** relative to DC power supplied relative to RF power output. In general, that was a somewhat specious argument. Amateur radio power outputs were generally less than 1 KW at RF and AC Mains electrical power cost less than 10 cents per KiloWatt-Hour consumed. That warning may have originated with higher-power broadcast users. As an example, a 10 KW RF output broadcast transmitter might draw 30 KW total from AC Main supply. At 10 cents per KWHr that was $3.00 per hour or $72.00 per day for 24-hour broadcasting. That was roughly the daily salary of an average broadcasting engineer in the 1950s.

The Crosby FM Transmitter System

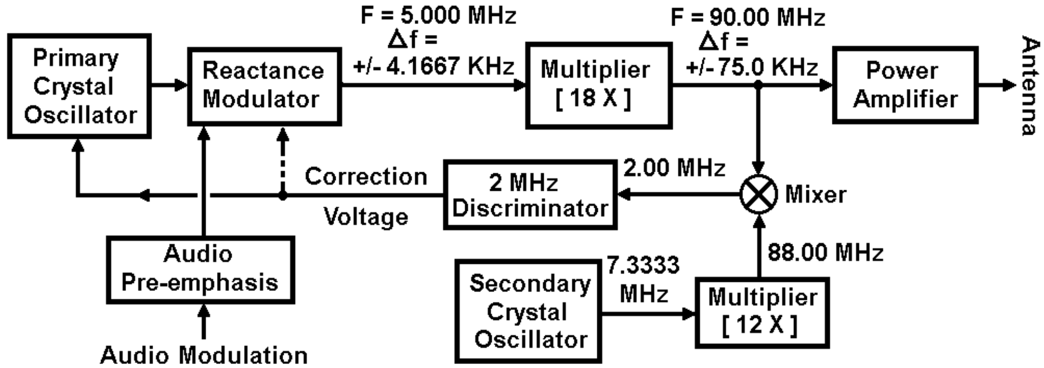


Figure 33-15 The most-used FM broadcast transmitter for 88 to 108 MHz.

Figure 33-15 shows the ***Crosby*** FM transmitter design, fairly standard in the USA. Note the modularized sections. The Primary Crystal Oscillator, Reactance Modulator, and frequency-correction Discriminator with Balanced Mixer were generally interchangeable. The Secondary Crystal Oscillator would set the carrier frequency. Note that increase in modulation index, less than that of the older Armstrong design, made possible by improvements in the Reactance Modulator.

Use of heterodyning and feedback of Correction Voltage to the Primary Crystal Oscillator allowed some feedback of audio to make modulation more linear. That wasn't so accidental. For about two decades after WWII ended there was a large movement for ***high-fidelity*** sound which occupied both the marketplace and hobby pursuits.

FM-PM Reception

The general block diagram of consumer-grade FM broadcast receivers is given in Figure 33-16. The demodulator was called a ***discriminator*** up into the new millennium, despite being replaced in new designs.

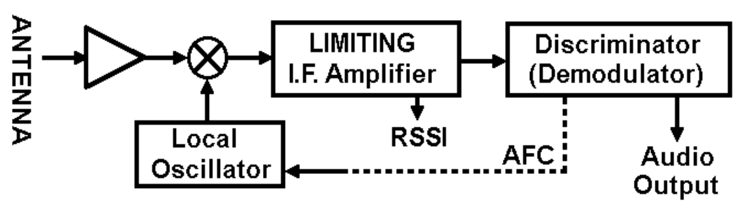


Figure 33-16 Common consumer-grade FM Receiver

The Local Oscillator had AFC (Automatic Frequency Control) voltage applied to keep an FM station in-tune. That was up to more stable crystal-controlled LOs were available.

The most notable block is the ***Limiting IF*** which is an over-driven amplifier used to all but eliminate any vestige of amplitude modulation. Note that FM and PM have a constant carrier level.

The RSSI (Receive Signal Strength Indicator) output is found on many LSI ICs and will function as an old HF Receiver S-Meter. With a lack of amplitude information, there is no way an FM or PM demodulator can output any signal strength information.

A number of LSI ICs have appeared to cover all of Figure 33-16 except for the RF Amplifier and (perhaps) the LO. These are called ***FM ICs*** even if they can also detect PM. The major technological effort had been directed to the demodulator itself.

Round-Travis Discriminators

The first FM-PM discriminator was the ***Round-Travis*** type shown in Figure 33-17. Also known as a ***Balanced Slope Detector,*** it was just two AM-type diode detectors connected back-to-back. In-service in about 1935, it would work fine if attention was paid to tuning L1-C1 to one side of the IF passband with L2-C2 tuned to the opposite side.

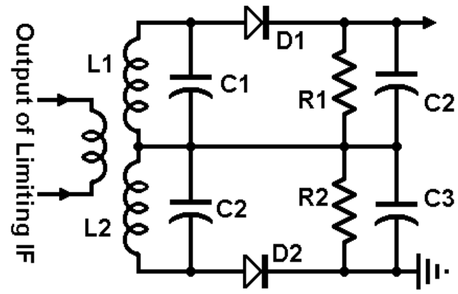


Figure 33-17 A Round-Travis Discriminator

R1 and R2 serve as resistive loads for D1 and D2. C2 and C3 act with R1 and R2 to provide an IF bypassing. Total output voltage is across both R1 and R2. At the IF center frequency the voltages are equal but opposite in phase, the output being zero volts. As the IF varies in frequency, output voltage will vary positive or negative depending on tuning of L1-C1 and L2-C2. AFC voltage is easy to get here.

While this works, linearity over an IF passband has a curvature. That is due to variations in ***Q*** of the two tuned circuits and care used in alignment.[6]

Foster-Seeley Discriminator

First published by Dudley E. Foster and Stuart W. Seeley in 1937, the ***Foster-Seeley Discriminator*** was a better form of FM-PM demodulator.[7] Shown in Figure 33-18 it was easier to align and with better fidelity than a Travis detector.

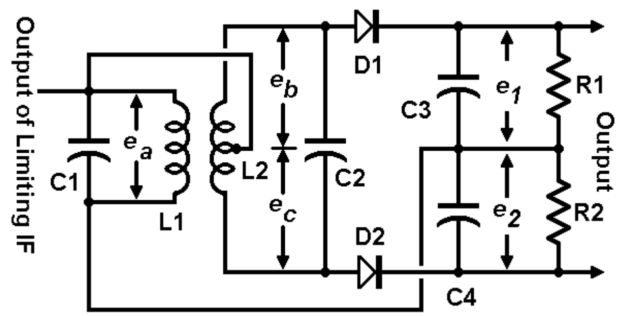


Figure 33-18 Foster-Seeley Discriminator.

Both L1-C1 and L2-C2 are tuned to the IF center frequency and inductively coupled. R1 and R2 are load resistors for diodes D1 and D2, C3 and C4 bypass the IF but allow AF to come through.

Output voltage can be demonstrated better by vector diagrams shown in Figure 33-19. Angles of these vectors are referenced to e_A. Magnitudes represent voltage. In this case, for all three vector diagrams, $e_B = e_C$. Only e_1 and e_2 vary in magnitude depending on the signal frequency, thus the output voltage magnitude varies

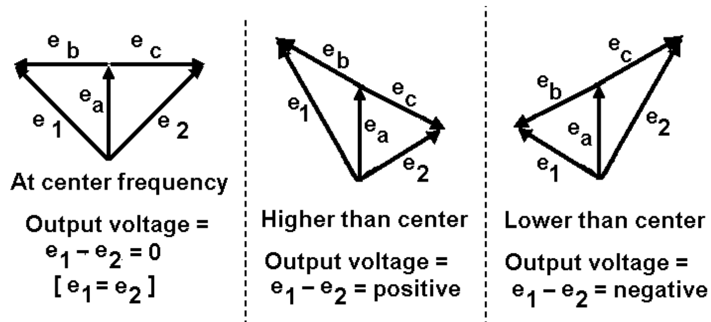


Figure 33-19 Foster-Seeley Vector Diagrams.

---

6 Manufacturing methods were not as exact in the 1930s and L-C ***Q*** values were not held as well. Just the same, the ***Limiting IF*** did promise a constant receiving value output and relative freedom from noise with respect to AM. Considering the general lower-fidelity of received audio, FM appeared to have *better sound.*

7 *Automatic Tuning, Simplified Circuits and Design Practice,* D. E. Foster and W. S. Seeley, Proceedings of the IRE 25, p.289, 1937. The following descriptions of the Seeley Discriminator and Ratio Detector are from RCA Broadcast News, No.42, January 1946, article written by Stuart W. Seeley.

according to frequency.

The major fault with the Foster-Seeley Discriminator was an insensitivity to amplitude variations. It ***required*** an IF Limiter ahead of it so that audio output would be consistent with the modulation index. By 1946 with new frequency bands for FM broadcasting and the beginning of the TV broadcast boom, the ***Ratio Detector*** was born and it could operate with a minimum of IF Limiting.

Ratio Detector

The ***Ratio Detector*** of Figure 33-20 is close to that of a Foster-Seeley, at least up to C2. But D1 is connected in reverse so that diode current is ***additive*** rather than subtractive.

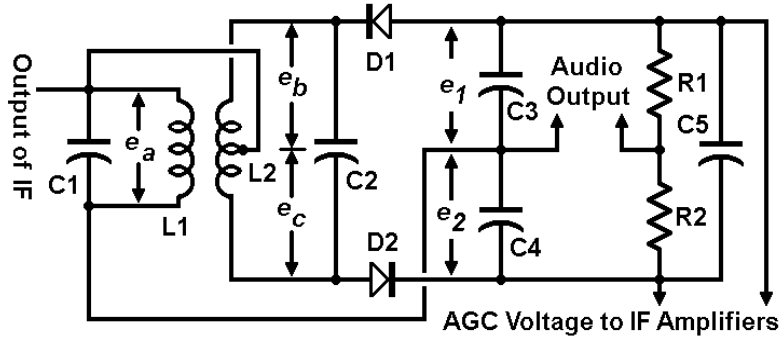


Figure 33-20 A simplified Ratio Detector

A key component is C5, made larger than the twin of C3 and C4. Since the diode current is additive, both currents flow through R1 and R2. Since the time-constant of C5 and (R1+R2) is low, the total voltage can be applied as an AGC voltage to the IF strip.

But, the time-constant of C3 and R1 and its twin, C4 and R2, is faster. Audio modulation will appear across each pair. Values are such that they resemble the time-constants of ordinary AM detectors with AGC output. Audio output is thus suitable for differential input sound amplifiers. The ***ratio*** of the modulation index to the IF input gives rise to its name.

It should be noted that the Ratio Detector is ***not*** immune to all amplitude variations. Under the conditions of holding the IF gain constant and the normal broadcast audio level (and thus the modulation index) the same, it ***appears*** not to need any IF Limiting.[8]

Quadrature Detectors

These surfaced as soon as semiconductor makers began with ***SOC*** or Systems-On-a-Chip for nearly all of an FM receiver. The block diagram is in Figure 33-21.

The IF input ***must*** be Limited, normally the case with SOC ICs of this kind. Relative to ground, the lower balanced modulator input is approximately delayed by 90°. Together, the two mixer inputs act roughly as an Exclusive-OR. Mixer output has a square wave with slight width modulation, coincident with the index of modulation of FM or PM. The

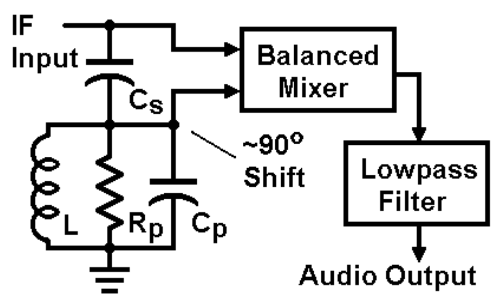


Figure 33-21 Quadrature Detector as used in FM SOC ICs.

---

[8] In the early boom times of the late-1940s through the 1950s, this was generally a good subject for a technical argument. All TV receivers were designed and built to take as few components as possible. AGC was well known and the Ratio Detector would output an AGC voltage, where the original Foster-Seeley could not. It probably saved at least one stage, perhaps two, thus making the cost-to-produce a bit less and allowing some final-sell competitiveness. Design for the consumer marketplace could be severe, not strictly for the love of circuitry.

Lowpass filter removes this width-variable IF, averaging out the voltage to leave just the demodulated audio.[9]

In use with SOC ICs, C_S is usually less than 5 pFd and built into the IC. Only the parallel-resonant circuitry needs to be supplied. The following is from the datasheet for the ML13155, a Lansdale second-source for the Motorola MC13155.

Loaded Q of the parallel-resonant circuit is roughly approximated by (IF center frequency/IF passband). For a 70 MHz IF center frequency and 10.9 MHz passband that would be about 6.42. That can be made slightly less to a value of 5.0. The value of C_P must include a maximum of about 3 pFd for stray and package capacity. If 20 pFd is chosen then the total would be 23 pFd. L is then:

$$L = 1 \;/\; ((2\pi F)^2 \; C_P) = 1 / (193.4 \cdot 10^{15} \cdot 23 \cdot 10^{-12} \approx 225 \text{ nHy}$$

Inductance can be chosen as 220 nHy at center of a trimmable value. Impedance magnitude at resonance is then: $R_T = 2\pi FL \approx 484$ Ohms . There is already about 3200 Ohms inte ML13155 at the quadrature connection so the external resistance should be about 570 Ohms. A 560 Ohm standard value can be used there. Total quadrature component values are then:

C_S = 2.0 pFd (internal to the IC)
C_P = 20.0 pFd (plus 3.0 pFd stray and IC package)
R_T = 560 Ohms and L = 0.22 μHy (trimmable)

This works out well with the standard values given. A 90° shift occurs at about 68 MHz, close to the 70 MHz center frequency. With slightly less stray capacity the center frequency comes closer to 70 MHz. A trimmable L can align that to exactly 70 MHz.

For a narrower FM application, the values can be (somewhat) scaled. The ML3371 and ML3372, both second-sourced by Lansdale Semiconductor from the original, no-longer produced Motorola Semiconductor versions (MC3371 and MC3372 respectively) can be used at 455 KHz. This allows a ± 4 KHz deviation in frequency. Murata has several types of already-trimmed ***ceramic resonators*** for quadratic detector applications, such as their part number CDB455C16.[10]

Modified PLL as an FM-PM Demodulator

The PLL block diagram of Figure 33-22 may be used as a demodulator. The only change is to have the Lowpass Filter cut-off frequency raised to the maximum incoming signal cut-off frequency. In this manner, the *VCO correction voltage* can become the demodulated output.

As a plus, the VCO output frequency can be used as a relatively-low-noise source of the incoming carrier signal (if desired). The VCO will track the incoming signal in frequency, including

[9] ***Averaging*** is not quite technically correct but the resulting ***on to off*** level varies as the Lowpass input changes in width. The ***width*** is a half-cycle of the IF input. At the Lowpass output, this width change comes out as the equivalent of the original modulation.

[10] As of the end of 2012, the full line of Murata ceramic filters and resonators was available in their 2000 catalog publication PO5E-9, along with connection diagrams of several FM SOC ICs. Another Murata publication, C601-I, also published about the same time yields more application notes and theory of design. A problem for hobbyists is that those are all available for purchase only in lots of 100 minimum from Murata. Distributors will have those in smaller quantities but at an increased price.

the frequency or phase modulation.

Note that this was touted in the mid-1970s by Burr-Brown in seminars, using just one, perhaps two ICs.[11] One of the surviving ICs, a 4046 family member, can be used for that purpose now.[12]

For this application the PLL Loop Filter becomes a bit more complex. It is usually DC to the upper modulation frequency bound. If there is sufficient separation of the lowest IF signal and maximum modulation frequency, one can use a simpler R-C filter, almost the same as for a PLL frequency synthesizer.

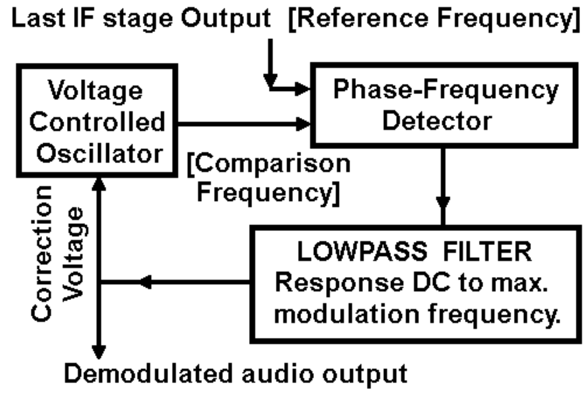


Figure 33-22 A PLL modified slightly to demodulate FM and PM.

Capture Effect

An oddity in consumer-grade FM receivers, this is the apparent insistence of the receiver refusing to leave a strong signal to re-tune to a low-level signal. If each signal is close together, the strong signal tends to ***capture*** the tuning.

Part of that is due to AFC action, insufficient design effort spent on the feedback filter. Part of it is due to the ***Limiting*** action of the IF. About the only way to get around ***Limiting*** is to use a very good bandpass filter at the input to the IF strip. A problem there is to have this new bandpass filter with the lowest possible Group Delay and to have that coincident with a high shape factor. Those two items are generally exclusive.

COMBINATORIAL MODULATIONS

The following are some of the in-use combination modulation-demodulation systems, usually centered on AM (but not always). They are representative, not all such systems.

Analog Television

In the USA, the ***old*** former system of analog TV (now replaced by High Definition Digital TV) was the NTSC (National Television Systems Committee) system. Video modulation was limited to about 4 MHz bandwidth, transmitted in a ***vestigial upper sideband*** method such that, while the majority of modulation content was upper sideband, some of the lower-frequency video was also transmitted on lower sideband. TV sound was sent by FM, much in the way of aural BC methods, with the aural FM carrier located 4.5 MHz up from the visual carrier frequency.

---

[11] Burr-Brown was purchased by Texas Instruments at the start of this new millennium. Some of their one-chip PLLs were still made in 2012. PLLs are described in Chapter 29.

[12] More detailed data on FM demodulation using the '4046 family can be found at the Texas Instruments publication SCHA002A, February 2003. Also the NXP datasheet for the 74HCT9046A, dated 15 September 2009, especially good for the Lowpass filter calculation.

In the 1940s the transmitters were separate, one for video, one for sound. Receivers had separate video and sound IFs. TV channel tuners would pass both video and sound. To save receiver construction costs, the ***intercarrier sound*** system was used. The video detector would pick off the sound FM at 4.5 MHz and send it to a *ratio detector* that did not need any IF limiting. Since this became the standard receiver design, the FCC had to amend its video and aural transmission rules to allow this new *intercarrier sound* system by limiting video black and white levels to avoid aural ***buzz*** from the vertical blanking level interfering with aural ratio detectors.

The old black and white TV industry boomed in the 1940s and 1950s, millions of TV receivers produced in the USA alone. Even TV antenna sub-industries expanded, birthing the existence of cheap, flat 300-Ohm ***twinlead*** balanced feed lines.

Analog Television Gets *Color*

While the TV industry was beginning to prosper, there was a prolonged fight over ***color television standards.*** CBS Laboratories favored the synchronized ***color wheel*** interposed between a single CRT and the screen; RCA Corporation favored a three-CRT projection system with each CRT of a different phosphor color. Neither was ***compatible*** with existing black-and-white TV transmission standards. The NTSC was re-formed after WWII to come up with a *compatible* TV standard that would fit existing TV sets (now numbering in millions) and any new color TV sets.

At the same time, man-made quartz crystal production was beginning to output quartz crystal units cheap enough for TV receivers. This was considered necessary to ***lock*** video color systems into synchronism with TV video transmitters.

The break-through involved ***sideband spectral content*** of video modulation. Energy tended to concentrate, rather like clumps, at multiples of the scanning frequency. By making a color reference frequency (for receiver demodulation) at an odd multiple of the CRT scanning rate, the video energy could be differentiated into black-and-white video and color video clumps. The end result was a color subcarrier reference frequency of 3.579 545 MHz with a slight change in horizontal scanning rate from 15.75 KHz to 15.734 KHz and a vertical scanning rate from 60.00 Hz to 59.94 Hz. Those slight scanning rate changes could be accepted by the millions of existing black and white TV sets and the video modulation sideband clumps would not affect demodulated black and white content. Compatibility was achieved for both monochrome and color TV receivers.

The NTSC compatible color system survived and grew for about six decades. While that was going on, a bigger and longer professional contest was involved to realize ***High Definition Digital Television. DTV,*** as this new system is known, takes much more explanation of how it works. It isn't covered here but is totally superior to analog TV of any kind.

Stereophonic Sound Over Broadcast FM

Binaural sound in the USA began in 1960, also facing a *compatibility* issue with the many existing FM broadcast receivers among the public. This was solved with some interesting audio content differentiation and frequency-multiplexing of the necessary sub-band for stereo sound.

Beginning with ***Left*** and ***Right*** (as in ears) stereo sources are arithmetically mixed to provide ***L+R*** and ***L-R*** channels at the transmitter. The ***L+R*** channel is transmitted in FM the same as in monophonic content. The ***L-R*** channel is up-frequency shifted by 38 KHz. The 38 KHz up-shift frequency is divided by 2 and sent out as 19 KHz. Existing monaural FM receivers would *hear* only the ***L+R*** and disregard the rest. Stereo FM receivers would require a stereo decoder sub-system to do the following:

1. Bandpass filter the ***L+R*** and ***L-R*** channels.
2. Demodulate the FMed ***L-R*** channel to recover the ***L-R*** content.
3. Arithmetically add the ***L+R*** and ***L-R*** contents to recover the ***Left*** audio.
4. Arithmetically subtract the ***L-R*** from ***L+R*** to recover the ***Right*** audio.

Binaural sound recovery is complete. If desired, a stereo FM receiver can add a small circuit to indicate presence of the 19 KHz reference frequency to show that stereo sound is present.

Modifying AM Broadcast Systems

MF BC stations are the longest-running of broadcasters. Since they are AM they would seem to be an easy target for addition of stereo sound, sending one *ear's* worth on one sideband, the other *ear's* worth on the other sideband. Motorola made a big push of their ***C-Quam*** system in the early 1970s and had a few AM stations converted. Unfortunately for them and some few sound lovers, that didn't work out. The reasons are many and varied. It is quite probable that there isn't any advertising market for AM BC in MF to have much of anything new.[13]

DRM (Digital Radio Mondial) has been tried for HF BC stations but has yet to attract enough interest in equipment. In its *test period* from 2005 to 2013, it has proved successful technically, but still hasn't attracted enough new equipment sales. Since HF BC stations have been cutting back on their off-shore service, we will have to see if the DRM market is sustainable.

VOR (Very-high frequency Omnidirectional radio Range)

As a radio-navigation bearing receiver for pilots, VOR has long been a staple of small and large aircraft piloting. Born about 1947, it was conceived to replace the old non-directional beacon stations and so-called ***A-N*** directional beacons. Operating at 108 to 118 MHz, it is basically AM with its antenna electronically rotating for about 30 percent amplitude modulation at a 60 Hz rate. A reference signal is sent at 9960 Hz, FM, to coincide with local magnetic north. Receivers need only to isolate the reference signal and compare it to low-frequency AM to determine ***any*** direction to a ground station.[14] It can be slaved to an auto-pilot to allow it to fly on any magnetic ***radial.***

The normal 300 to 3000 Hz AM capability enables ground controllers to advise pilots for local conditions (such as pre-recorded weather conditions) or to instruct them for any other aero-nautical reason. Voice operation does not interfere with bearing capability. By 1955 this was adopted by the new ICAO or International Civil Aviation Organization as ***the*** directional beacon system for all countries.

High-Rate MODEMs for Computers

While (usually) carried over wire lines, the high-speed (56 KBPS maximum) MODEM can

---

[13] In the author's opinion, there has been a glut of cheap AM BC receivers tuning (primarily) MF since the end of WWII. Known as the ***All-American-Five*** for its five tubes, no power transformer, built-in loop antenna, it cost about $25 US on an average for about four decades. In the USA it carried largely news and entertainment, program content dropping since the 1950s. Spurred on by low-cost stations, many automated, it did not offer much in the way of quality programming nor was it a technical marvel in any way.

[14] In the author's opinion this was a sophisticated (in its simplicity) system for any aircraft. It was simple enough to work with vacuum tube based receiver designs needed for their low weight and low power.

send digital information at a rate higher than the normal line's passband. It does that by ***quantizing*** digital bit ***groups*** into both ***AM*** and ***PM*** components. The *quantization* process requires temporary storage of a digital bit stream, picking out the *groups,* making a composite amplitude-and-phase signal. For reception it is done in reverse. It requires a *modem* to use a special-function IC to handle the quantization process.

There are several different standards of quantization and the subject is enough for a small book on the subject. Suffice to say that the information ***throughput*** is rapid enough to come up to the *Shannon Limit* of about 56 KBPS on a 3 KHz bandwidth telephone line.

Multi-Channel Commercial SSB on HF

This was an out-growth of early *carrier equipment* frequency-multiplexing systems. The first teleprinting *carrier equipment* could carry four 60 WPM teleprinter channels in a single 3 KHz voice bandwidth circuit. The first telephonic *carrier equipment* could carry four 3 KHz bandwidth telephone circuits in a 12 KHz bandwidth. Both used SSB techniques to achieve the frequency multiplexing. Given a total of 12 KHz bandwidth, a SSB radio circuit could carry a number of different, independent communications circuits. A typical SSB radio circuit could handle two voice telephone channels and eight teleprinter channels, all at once.

Of course, each radio link needed separate frequency bandwidths for each path but there was still a compression of total bandwidth versus frequency. Even though SSB began on HF in 1928, there was a considerable cost in radio equipment plus the initial *carrier equipment* cost. There was a scarcity of quartz, both for crystal oscillators and for the large lattice filters in SSB receivers. Plus, there was a doubling of operator salaries since both carrier operators and radio operators were needed. Nonetheless, HF SSB for commercial use prospered until WWII was over.

Appendix 33-1

Eliminating Confusion on the Spectra of Modulation

Most writings on modulation will have a nice picture of AM at full modulation, showing the increase and decrease of RF waveshape amplitude with modulation. Yet, when speaking of the individual spectra, the carrier frequency is said to remain constant. Which is it? Actually, both are correct.

In the first example, whatever is observing the modulation of AM ***sees everything.*** *Everything* is the carrier ***and*** both AM sidebands. What can't be seen readily is the ***phase*** of the sideband RF relative to the carrier. The combination of amplitude and phase of the sideband's RF is mixed by the observing device relative to the constant-amplitude, constant-phase carrier. One can only see the ***individual*** sideband amplitudes with a ***very narrow bandwidth*** spectrum analyzer.

In ***seeing everything,*** the observing device combines all the different phases and amplitudes (such as a conventional detector would do) and the result is what is heard. Using the ***very narrow bandwidth*** spectrum analyzer allows concentration on just a small part of the total spectra to reveal the spectrum amplitude.

A typical observation device is an oscilloscope connected to a receiver's diode detector input. The 'scope will show the amplitude modulation as in the typical modulation drawing because it is displaying everything...carrier and sidebands all together. That demodulated waveform will have amplitude modulation as it is heard. Amplitude will vary from twice the unmodulated peak RF voltage at maximum to near zero for the minimum amplitude modulation.

Yet, for the typical spectra, the sideband is assumed to carry all the modulation and the carrier (if not suppressed) will be of constant amplitude. But, that doesn't give the same intuitive feel of modulation varying in amplitude. The solution is to use a more ***narrow frequency*** observation device such as a narrow-banded spectrum analyzer. Using a SSB receiver with a ***very narrow*** bandpass IF filter will show the same effect. Each will show the constant-amplitude carrier and the varying-amplitude sideband content ***separately.***

With FM or PM the carrier amplitude remains constant and only the ***phase*** changes. Yet the modulation comes through (with varying amplitude) and can be amplified. A good part of the confusion comes about in ***trying to explain that over this medium.***[15] It isn't intuitive such as with AM. The recourse is to use low-harmonic-distortion modulation-frequency tones to simulate the modulation content. Holding the test tone amplitude constant allows observation of the total RF spectrum at any modulation index. A very narrow bandwidth spectrum analyzer will reveal the amplitude of the FM or PM sidebands. That same *sideband* amplitude allows precise calibration of a modulation monitor read-out via the Bessel Nulls method. A modulation monitor is a common monitor equipment in broadcasting stations.

---

[15] It has always been a mind-boggling task to imagine what John Renshaw Carson thought when he came up with the three basic mathematical formulas (shown in Chapter 4) back in 1915. Not even AT&T had the test equipment to examine the modulation effects in full detail back then, no true spectrum analyzers nor much in the way of narrow-band filtering except what was cobbled together for the purpose.. It was as if he came up with ***First Principle*** rules to actually see the concept in his mind, then put it into mathematical form. That might also explain why it took so long for the US Patent Office to finally approve his patent application.

Chapter 34

Miscellaneous Subjects

Among the techniques of analog and digital exist a few curious facts and circuits. Some are used only a few times. Some have lapsed due to other technological changes. Some do not quite fit into other Chapter groupings.

Amplifier Stage *Classes*

This began with early vacuum tubes and referred (mainly) to control grid biasing during an AC input cycle. That scheme is carried on with semiconductors, pertaining to base or gate biasing during a cycle. The Class letters are as follows:

Class A — On during all of an AC input cycle, generally referred to as ***linear,*** common to small-signal circuitry**.**

Class B — On during only ***half*** (or close to it) of an input AC cycle, usually in push-pull so that each active device used alternate AC input halves.

Class C — On during ***less than 180°*** (usually between 90 and 180°). This was common in RF power amplifiers designed prior to 1970.

Class D — Active device was used as a *switch* with a form of pulse-width modulation taking the place of amplitude modulation. This became popular with several audio amplifiers to reduce supply rail drain.

Class E — Same as Class D but the ***switch*** usually employed a fast SCR or Silicon Controlled Rectifier. A few amateur devotees used this in lower-band HF final RF amplifiers.

There was a succession of letters that followed, usually championing a particular form of switch-like active device in a particular choice of circuit.

In the period just after the end of WWII and during the ***high fidelity sound*** boom period, there was a Class AB used in audio power amplifiers, signifying the active device actually drawing a slight amount of power at zero signal. Primarily used with vacuum tubes, this was a way to eliminate the slight distortion from low audio signal levels in Class B operation. That branched into Class AB_1 and Class AB_2 depending on the relative amount of DC current flow. A problem there was a lack of definitive statement of *how much quiescent* current was flowing at no-signal.

Miller Effect

This was first documented in 1920 with vacuum tubes, specifically with triodes.[1] Vacuum tubes always have some capacitance between plate and control grid. Depending on the plate load resistance and amplification, the apparent capacitance can appear to increase beyond the physical manifestation of capacity. This causes a ***lesser bandwidth*** than was expected looking at the circuit by itself on paper.

Figure 34-1 can be used as a rough introduction to Miller Effect on a common-cathode triode and common-emitter NPN, both to an op-amp Integrator circuit. Inter-electrode capacitances C_{PG} and C_{CB} are multiplied by voltage gain and appear as a *Miller capacitance* C_M in parallel with the physical inter-electrode capacitors.[2]

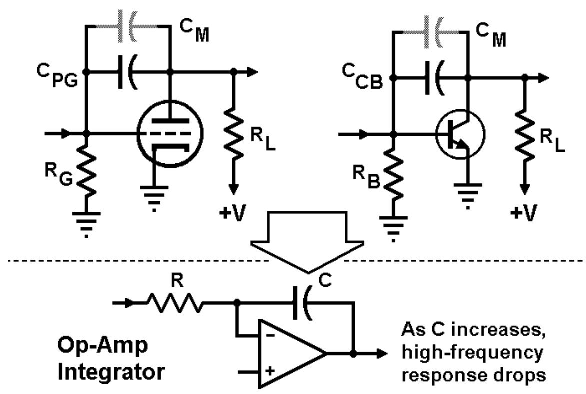


Figure 34-1 Analogous comparison of effect of Miller Effect versus Integrator.

Note that both the triode and common-emitter NPN have inverted voltages at their outputs. So does an Op-Amp Integrator. All three will suffer the effects of Miller capacitance at bandwidth high end. An approximation:

$$C_{MIN} = (1 + g_m R_{TL})\, C_{IEC} \qquad \text{where:}$$

g_m = transconductance in mho

R_{TL} = Total Input Load resistance

C_{IEC} = Interelectrode capacitance

Some notations here. The value of *transconductance* is common in vacuum tubes but seldom found on bipolar transistor junction datasheets. For an approximation, the total Input load resistance can be taken as approximately resistive value at mid-band. Interelectrode capacitance can be found on some datasheets but not all of them. For the sake of illustration, lets assume:

$g_M = 0.012$ mho $R_{TL} = 2.5$ KOhms $C_{IET} = 1.4$ pFd

$$C_{MIN} = (1 + 0.012 \bullet 2500) \bullet 1.4 = 43.4 \text{ pFd}$$

For another approximation, the Miller Capacitance appearing across at the ***output*** load resistance is about:

---

[1] By John Miller, published in the Bureau of Standards publication. *Scientific Papers of the Bureau of Standards,* Volume 15, 1919-1920. Bureau of Standards was the parent to today's NIST, the National Institute of Science and Technology, USA..

[2] Miller capacitance does not appear physically, but, for an analysis model it can exist ***as if*** it did for the purposes of determining the upper end of bandwidth.

$$C_{MOUT} = C_{IET}\left(\frac{1 + g_m R_{TL}}{g_m R_{TL}}\right)$$

Using the above example values:

$$C_{MOUT} = 1.4\left(\frac{1 + 0.012 \bullet 2500}{0.012 \bullet 2500}\right) = 1.4\left(\frac{31}{30}\right) = 1.447$$

Note that there is much less increase of additional Miller Capacitance at the output. Fortunately, that fits the ordinary common-emitter configurion which has a low input resistance but higher output resistance.

Emitter Follower or Common-Base

With the emitter or cathode follower, there is no inversion of the output so there isn't any Miller Effect. With the common Base configuration, the Miller Effect only appears at the output, a much lower value. This applies to Common-Gate in FETs and Common-Grid configurations in vacuum tubes. There isn't much need to calculate the Miller Effect.

Some Conclusions on Miller Effect

In the more modern age, this is a fine project for an engineering classroom exercise, principally on handling equivalent circuits with attendant mathematics. But, it has little to do with applying devices to a circuit other than knowing that there is a slight decrease in bandwidth of ***one*** circuit configuration.

NETWORKS

Bartlett's Bisection Theorem

This is a technique, originally intended for passive filters, of re-ordering and re-arranging parts and values of ***symmetric networks.*** Devised in 1927 by Albert Charles Bartlett (ostensibly for changing from ladder filters to lattice-type filters), it has been expanded by others to allow for changing changing source and load impedances of ladder filters.[3] Considering that passive filters of pre-WWII era were different from the modern filter theory and (at time of writing) was 84 years in the past, it still has some minor use in the design world of today.

The method is shown in Figure 34-2 using a 2-resonator band-pass filter centered on 10 MHz with a bandwidth of 500 KHz.[4] It begins with ideal values at the top. Parts values are based on 1.0

[3] One of the first was Wilhelm Cauer of elliptic filter fame. In reality, passive filters have a usually generous allowance of parts values allowing for variations, changing response only slightly as long as the proportions of passive filter parts values are maintained.

[4] From an example by Richard Harris G3OTK on his website of 2010.

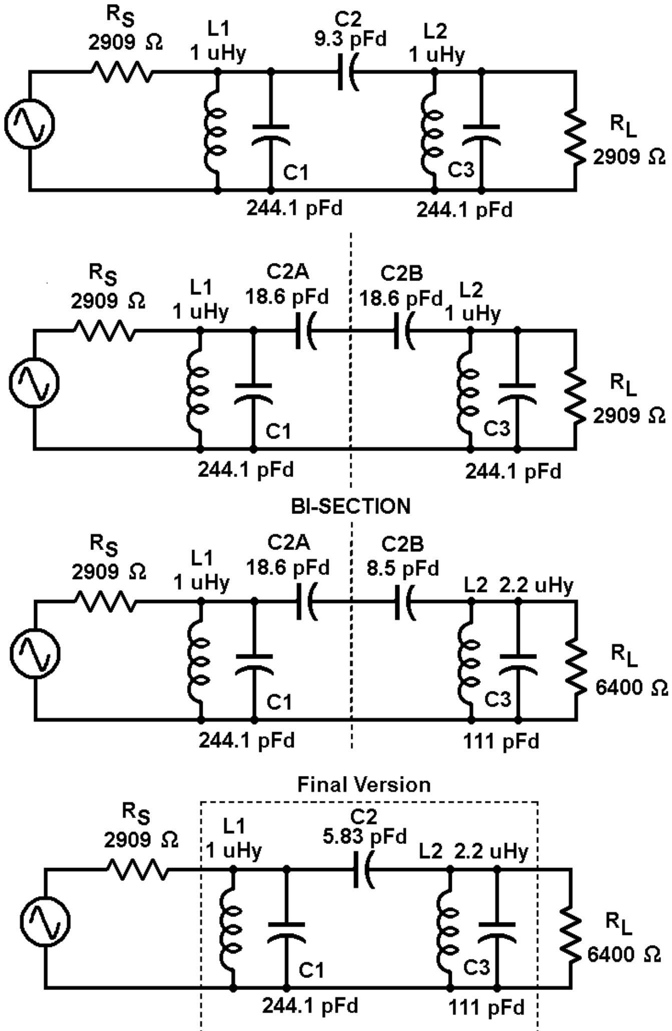


Figure 34-2 The steps leading to changing a load (or source) impedance using Bartlett's Bisection Theorem.

μHy fixed inductors. The example is to raise the load termination using 5% value fixed inductors, in this case to 2.2 μHy.

Bisecting the filter requires that coupling capacitors double to stay in series. If they were in shunt, bisection would be halving their values. Similarly, inductors in series would halve and inductors in shunt would double.

Using a 2.2 μHy inductor requires that termination impedance goes up by 2.2 times but resonating impedance goes down by 2.2 times. C2B changes to 8.5 pFd, C3 to 111 pFd, and termination is almost exactly 6400 Ohms. Combining C2 results in a final value of 5.83 pFd.

Plotting the response by an analysis program before and after modification shows little change in frequency response. One thing that cannot quite be done with bisection-transformed top-coupled bandpass filters is the technique of tuning narrow bandpass filters.

As with all such transformations, it is imperative that an analysis be done *after* all the number crunching is done. That is to check its characteristics before committing to hardware.

Norton Theorem Transformation

This is similar to Bartlett's Bisection method although it goes a bit more complicated. For passive filters it uses some strange concepts on its way to final form. It winds up using ***negative capacity*** sometimes. If that is the case, an inductor can substitute for a negative capacitance, but more than likely there is another positive capacitance in parallel which can be down-valued by the transform's negative capacitance.

We can use the original Richard Harris example here, a two-resonator bandpass filter, and

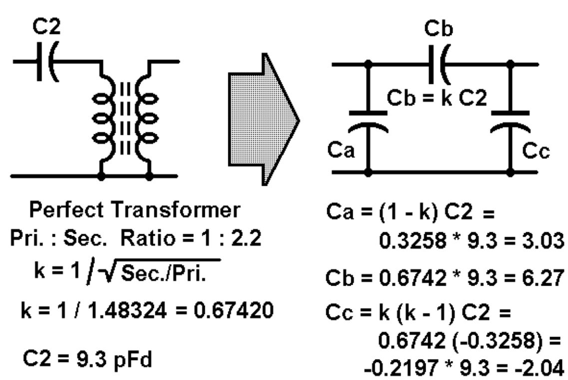


Figure 34-3 Replacement for C2

like that one, increase the second inductor to 2.2 μHy. This must also increase the load resistance and reduce the in-circuit capacitance. A ***bisection*** is not required here, even though the source-end of the filter is the same. What is done is to take a circuit block and revise it in the following steps.

The first step is to replace C2 of Figure 34-2 as shown in Figure 34-3. Note that Cc turns out to be negative. After that, the transform is inserted into the bandpass filter schematic ***in place of*** C2.

Meanwhile L2 is modified to 2.2 * 1 = 2.2 μHy, load-end impedance is multiplied by 2.2 * 2909 = 6400 Ohms, and C3 is divided by 2.2 to equal 111 pFd.

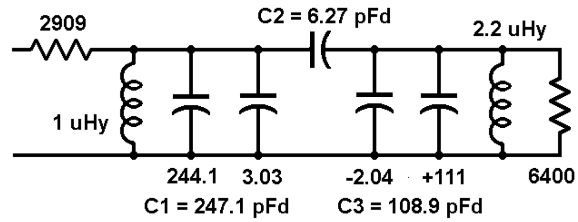


Figure 34-4 Final version of transform.

The final version is in Figure 34-4 and may be compared with Figure 34-2. Values aren't quite the same although an analyzed frequency plot shows them to be nearly equal.

For passive L-C filters this is a lot more work than is normally needed. That's one reason for creating the author's LCie program that includes a Sensitivity (to tolerances) analysis and the ability to change individual values.

Derating

Derating power dissipation of components is seldom considered by too many designers. It should be a ***part*** of design as much as Ohm's Law of Resistance and other basic formulas. The bad part is that some manufacturers of components either do not bother with ***temperature effects*** on parts values or they mention some Military Specification which is either out of date or obscure and hard to find. Nearly ***every electronic part*** considers temperature and (sometimes) air pressure in ***all*** Mil Specs. Military electronics ***requires*** operation at temperature extremes of -55°C to +100°C (-67°F to +212°F). In addition, radios and other electronics for air and spacecraft have to operate from normal air to lower pressures that can include outer space environments.

In so-called *normal room* environments, most parts will perform well within tolerances. In some applications these tolerances are stressed to near limits.

Worst Casing Digital Logic

For TTL, such as with 74xx to 74Fxx devices, one picks the highest ***expected*** ambient temperature and the highest current per-gate or per-package, then adds all of those for the final DC current demand. This results in a higher current demand value than actually occurs when operating. That is probably the best way to avoid a strain on the power supply but is the easiest.

For Advanced CMOS logic, such as with 74AC to 74HC devices, there are two parts to DC current demand. The lowest is the idle or no-signal current, usually given on a datasheet. The highest current depends on the maximum rate of signals. That is the effect of internal IC distributed capacity heating up the IC and is given by the curious datasheet value of $\boldsymbol{C_{PD}}$ or ***power dissipation capacitance*** and always given in picoFarads. From formula (23-1) with capacitance in pFd, all currents in μA, and frequency in MHz: $I_{TOTAL} = C_{PD} \cdot V_{CC} \cdot f + I_{QUIESCENT}$

For a 74AC00 gate running a 5 VDC with a C_{PD} of 35 pFd at 20 MHz and quiescent current demand of 25 μA, total current would be (35•5•20)+25 = 3525 μA = 3.5 mA. Or just calculate the higher demand and add 0.1 mA as a *fudge factor*.

Note that all digital devices acting as drivers need to add in the driven load.

Worst-Casing Active Devices - Small-Signal Analog

The respite for small devices dissipating (usually) less than a half to a full Watt is the datasheet. Most datasheets have information on current demand versus temperature versus collector or drain current. ***Never exceed this wattage!***

Note also the input junction effect from temperature and some from the output junction effect (usually less). In general for BJTs, the base-emitter voltage will *decrease with an increase in temperature.* That may effect a BJT bias network. A lesser effect is the increase in collector-emitter junction with increase in temperature for *saturated operation* as with external loads such as LEDs or relays or small motors used in switching circuits.

Always pay attention to ***SOA*** or Safe Operating Areas of ***all semiconductor*** specifications. This is especially true for high-temperature operation of Class A circuitry.

Derating Resistors at High Temperatures

This is especially true for bias networks. Ordinary carbon-composition or metal-film resistors are rated at 25° C (77° F). These are good to about 70° C (almost 160° F), then derate linearly to zero Watts at about 125° C (almost 260° F). There ***will*** be some changes in resistance, even in safe operating temperatures but that gets rather complicated. ***If in doubt, use a higher Wattage rated resistor to lessen effects.***[5]

It should also be noted that the type of resistor changed should ***not have any AC reactances***. A wirewound resistor at RF would be largely an inductor if substituted for a metal-film resistor.

Heat Sinking

This applies to a limited number of types of circuitry such as power amplifiers, generally over a half-Watt or so, relative to the ***ambient air temperature.*** For nearly all hobby projects, the ambient air temperature will seldom exceed 50° C (approximately 125° F).

Semiconductors have two major thermal effects. Most predominant is the ***Junction temperature***, that area on the die itself where all the little junctions reside. The ***Case temperature*** depends on the type of packaging such as the familiar large diamond-case of a TO-3 to the smaller SMT packages. Datasheets generally give the Junction temperature values relative to the Case types available for a semiconductor.

Having full access to die junctions, manufacturers specify Junction temperature, somewhere on a decreasing linear slope from 25° C down to Zero power at about 200° C (almost 400° F). Then they apply that to the package type to come up with Junction temperatures relative to various

---

[5] Since these are one-of-a-kind projects there is no need to scrimp on Wattage or sizes of resistors. This would be a factor for production lots and total costs, but NOT for hobby projects. There is seldom time to play with circuits for temperature effects in hobby designs. There is always time to play for ***other effects*** at room temperature.

ambient values.[6] There isn't any hard-and-fast rule set to determine that, just the usual skull sweat needed to determine what size and shape of a heat sink is needed.

An electrical analogue is generally presented in texts, a number of quasi-resistances in series to represent the Junction temperature, the Case temperature, the Sink temperature, all relative to the ambient air temperature. All are given as a ratio of temperature (generally in °C) divided by total Wattage. For that, a thorough study of heat sink devices is needed to get the temperature over Wattage to an acceptable level. In general, all of that is rather specialized and not needed unless actual power amplifiers are done as part of a project.

Derating Other Components, Principally Passive

This will require a deep digging into various manufacturer's detailed data to determine exactly ***what*** is changing for a varying environment. Fixed capacitors are graded by part number for a variety of changes, both capacitance and withstanding voltage versus temperature. In some cases they can be obtained with both positive and negative capacitance changes over temperature. Those are good for temperature drift correction to reduce the amount of mis-tuning.

Resistors are generally derated in terms of power-handling characteristics. While wirewound resistors can be used in (essentially) DC circuits, they can get ***hot*** when subjected to high power dissipation. A cure there is to mount them by metal, sinking excess heat to a chassis or metal cabinet. Wirewound resistors are ***not*** to be used in circuits due to internal reactance.

Inductors are generally not specified for holding an inductance value. For RF the best bet is to use toroidal cores specified for a particular RF spectral region. Next best for RF is solenoidal types with small wire held in place by long-lived insulating coating such as varnish. For some power supply applications, ***swinging chokes*** or ***swinging inductors*** will change inductance depending on current through them. Fortunately, their use in most hobby projects is diminishing.

Diodes are generally rated the same as transistors. For rectifier circuits, most rectifier diodes are specified on datasheets for ***peak forward current***, a requirement for initial charging of input filter capacitors at power-on. For diodes, the ***PIV*** or Peak Inverse Voltage is a common value specification. PIV is the sum of reverse voltage ***peaks*** encountered in normal circuits.

LCDs have low temperature limits, depending on their manufacture. At lower than minimum temperatures the display can break up or disappear. Batteries have low temperature limits depending on their electro-chemistry. Depending on type of electro-chemistry, batteries can be stored at low temperatures longer than at room temperature.

EPROM Programming

EPROMs or Electrically-Programmable Read-Only Memories are good for storing digital data for at least about 30 years or more. For most EPROMs, any digital ***word*** can be made available to ***data*** output lines in parallel by setting ***address*** lines in binary states. EPROMs with glass windows can be erased by ultra-violet light and reprogrammed at any time in their life.

In the usual case, a *word* is either 8 bits or 16 bits in length. Addressing can be anything

[6] This can also get complicated, necessitating a thorough study of datasheet values to determine if a heat sink is needed and, if so, the approximate size of the heat sink. That requires input from heat sink catalogs to determine their sinking quality.

from 10 bits for 1K word total size to 18 bits for 256K words. The usual nomenclature has a general part number for ***word size*** followed by ***addressing capability*** with a dash number denoting maximum length from an address change to valid data output, usually specified in nanoSeconds.

As noted before, EPROMs are generally the same price from distributors, actual price increasing only for the faster Address change to Data Output change time. This holds true for old 2716 types (2K x 8, appearing about 1980) to AM27C1024 (64K x 16, appearing about 1998).

To change an EPROM programming data requires ***erasure*** of old data. This is not possible with ***OTP*** or One-Time Program types; those are for new production assemblies, have no UltraViolet window for erasure but are generally the cheapest. ***Erasurable EPROMs have a window.*** Erasure requires a specified time and light density of UltraViolet light. If light sources are not available, exposure of the window to direct sunlight for at least a half day is sufficient. Most EPROMs are made to indicate all digital level 1s when fully erased.[7] Once erasure is done, cover the window with tape. A standard white paper with adhesive is good for this plus ID marking.

Burning-In With a Manufactured Programmer

Most program development software tools can generate the ***Intel Hex*** format. This is read by serial- or parallel-port EPROM programmers for ***burning-in*** data.[8] These are generally referred to as ***Hex files*** now despite several predecessors. Such files can be sent electronically to an EPROM programmer which will do the decoding and appropriate data burning. Users need not read such files, although they can be used for trouble-shooting in rare cases.

To make such files as flexible as possible, each byte is converted to two ***nibbles*** which are represented in ***hexadecimal form.*** This makes such a printed file seem confusing but Assembler and Programmer software can handle those with ease. Note that each byte is represented by ***two*** hexadecimal characters. A Hex File is simply read by software and converted to programming data.

Each line of an *Intel Hex* record contains the data, the beginning data address location, a checksum (for verification) and some control bits in hexadecimal. Data format is as follows, reading from left to right:

Character 1:	An ASCII colon character (":"). Identifies this as an Intel Hex file.[9]
Characters 2, 3:	Record length, the count of the character *pairs* in a data field.
Characters 4, 7:	Address location of first byte of data, limited to 65,535.
Characters 8, 9:	Record ***type,*** either 00 for data or 01 for final line.[10]

---

7 The author has done that. About 40 years ago a very cheap EPROM erasure was purchased having an UltraViolet light source, power supply, all inside an empty 8-track cartridge case. That still works and can erase all EPROMs in about 3 hours of exposure. UV light level for that erasure thing was never measured but it worked to change all outputs to logic 1 levels. It was big enough to take up to 8 EPROMs. The case was good to shield from excess UV radiation to the body.

8 ***Burning-in*** is colloquial for transferring electronic data to actual EPROM bit data. It may have originated in the past with old ***fusible link*** types of digital storage devices just before EPROMs evolved.

9 A colon character can be printed on most printer devices, therefore would appear as a single character. A colon in ASCII would otherwise appear as hexadecimal code of ***3A.***

10 There are more codes for 16- and 32-bit micros plus extended addressing. For details see Intel's own publication ***Hexadecimal Object File Format Specification,*** Revision A, 6 January 1988.

Characters 10+ Data itself as character pairs in hexadecimal, usually given as 32 pairs for 16 bytes or 64 pairs for 32 bytes of data.
Last 2 Characters: Checksum given in 2's complement form.
Line ends in a carriage-return and line-feed, invisible on normal printers.

The ***checksum*** is there for verification (to the programmer device). It is the sum of all digits, concatenated to two hexadecimal digits, then converted to two's complement form. Two's complement is equal to the number in binary subtracted from 100Hex. If the concatenated sum of all digits is 01 then the checksum equals FF; if the sum is 23, then the checksum equals DD.

The very last line would be "***: 00 0000 01 FF***". That ends the Intel Hex copy and stops the EPROM programming.

For microprocessors the program data is not always continuous. In that case a Hex file output would be discontinuous with varying-length data but with the Address location changed for the resumption of continuous data. The Record Type characters would still be 00.

Visualization of a Hex File Record Line

In terms of printed characters, an Intel Hex file line could be as follows:

```
: NN AAAA TT DDDD....DDDD CS
|   |    |    |     |          |
|   |    |    |     |   Checksum, 2's complement of N+A+T+D fields
|   |    |    |  Data as Hexadecimal, always an even character number
|   |    |  Type designator of Record = 00 for Data input, = 01 end of file
|   |  starting Address in Hex, limitation of FFFFh or 65,535 decimal
|  Number of total Data Bytes in D field
Printed Colon character to identify file as an Intel Hex Record
```

As an example, the following two lines represent ASCII characters for **A SAMPLE**, beginning at binary address of **02A6** Hexadecimal:

A _ S A M P L E	<- ASCII contents (printed out)
: 08 02A6 00 412053414D504C45 2D	<- Data record
: 00 0000 01 FF	**<-** End record

Note: There have been several other formats but only the Motorola ***S-Record*** file format seems to survive.[11] Format is somewhat similar but each ***S-Record*** line begins with an upper-case ***S*** instead of the Intel Hex Record colon character.

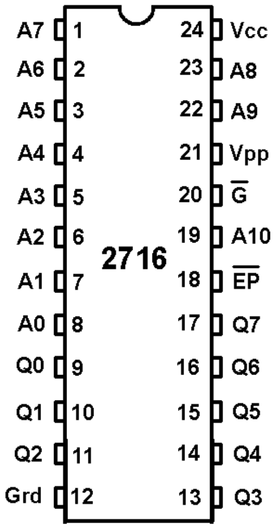
EPROM Control Pin Connections

All EPROMs have several modes of operation. These vary with manufacturer's type numbers. Perusal of datasheets is necessary. Two different types, a 2716 (circa 1980, a 2K x 8) and an AM27C1024 (circa 1998, a 64K x 16) are illustrated. Figure 34-5 shows a 2K x 8 2716 EPROM

[11] Motorola MC68000 Programmers Reference Manual, Rev. 1, 1992, document M68000PM/AD

while Figure 34-6 has the larger 64K x 16 EPROM. In each Figure the control pin states for the functions is tabulated.

The difference between ***Read*** and ***Verify*** modes has to do with the Programmer device. Both mode selections allow normal reading at Data Output pins. *Verify* allows more latitude in reading of burned-in data.



Vcc = +5.0 VDC

Vpp = +25 VDC for Program Input, +5 VDC for all other modes.

An = Address Input

Qn = Data Output (for Read or Verify Modes).

MODE TABLE

	$\overline{EP}$	$\overline{G}$	Q0 to Q7
Read	0	0	Output
Program	Pulse	1	Input
Verify	0	0	Output
Idle	1	X	Hi-Z

Pulse = 45 to 55 mSec at Logic 1 during one byte of data.

X = don't care state.

Figure 34-5 Top view and Mode Table for 2716.

For each EPROM the ***Idle*** or ***High-Z*** mode puts all Data Output pins in a high-impedance state. That allows several EPROMs to have their Data Outputs connected in parallel. That reduces PCB digital components and traces, allows expansion of Addressing to a common Data Output bus.

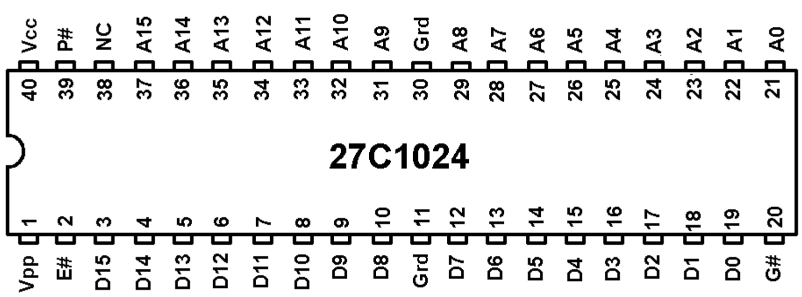
The difference between ***Output*** and ***Input*** of Data Output pins is that ***Output*** is the normal *Reading* mode. Data to be entered is as either ***O*** or ***1*** in digital levels.

There are differences between the old and new EPROM types. The 2716 draws about 100 mA in normal Read mode, 25 mA in its Idle mode. The 27C1024 draws about 35 mA in Read Mode and less than 1.0 mA in Idle.

An overt difference is the Program ***Pulse*** polarity. It is ***positive*** for the 2716 but ***negative-going*** for the 27C1024.

No-PC Burn-In

This is more laborious but is good for a small data group. To do so, consider a circuit such as Figures 34-9 through 34-11, a very ***simple*** type of ***manual*** one-word-at-a-time Programmer. This does not need a PC for data storage. Data to be burned-in must be worked out on paper first.[12]



Vcc = +5.0 VDC

Vpp = +12.5 to +13.0 VDC for Program Input, +5.0 VDC for other modes.

An = Address Input

Qn = Data Output (for Read, Verify).

NC = No connection Grd = Ground

X = don't care state.

MODE SELECT TABLE

	E#	G#	P#	Q0 to Q15
Read	0	0	X	Output
Program	0	1	Pulse	Input
Verify	0	0	X	Output
Idle	1	X	X	Hi-Z

Pulse = 95 to 105 uSec at Logic 0 during one byte of Data.

Figure 34-6 Top View and Mode Select Table for a 64K x 16 EPROM, type 27C1024.

[12] Many old and new Programmers are described on the Internet but nearly all require a PC to put in the code and to run the Programming cycling. As a result they are essentially copies of production units that can (quickly) program an entire EPROM in a few seconds. This one is entirely divorced from a PC.

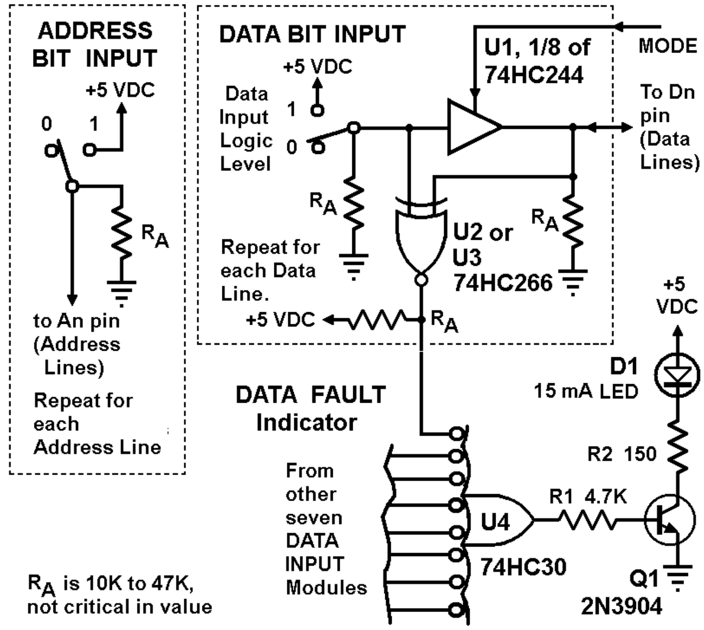


Figure 34-7 Basic MANUAL Programmer

The basic manual programmer is shown in Figure 34-7 and has two sets of ***modules.*** The Address Bit Input module is composed of one switch and one resistor per Address pin. This can be 10 to 16 depending on number of Address lines. All Address Bit modules are equal.

An option is to have the entire Address Bit Input as a 10- to 16-pin resettable counter, counter outputs to the Address Bit pins. If so, there should also be a hexadecimal visual display of the counter state. This is optional and not mandatory.

The Data Bit Input module is similarly expandable but will remain at 8 for most EPROMs (up to 16 for some of the rare wide ones). The major interface is U1, a 74HC244, an octal non-inverting buffer with High-Impedance output switching in groups of four. The two ***Mode*** pins (1 and 19) are connected together and, if low, will allow each input to go direct to each output. If the *Mode* is high or Logic 1, the outputs go into High-Impedance mode and are, effectively, out of the circuit. The way it works can be described as follows:

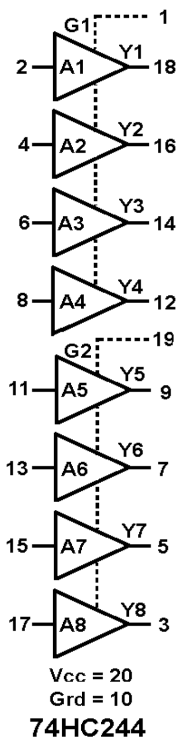


Figure 34-8

Mode line is normally Logic 1. A Word data is entered into the toggle switches. A ***Sequencer*** (Figure 34-9) puts the Mode Line to Logic 0 and writes the toggle-switch Logic states to the EPROM Data pins. The *Sequencer* will write only on a manual command from a push-button. When the *Sequencer* is done, the Mode Line goes back to Logic 1, making all of U1's output lines High-Impedance. What is left is each Data pin compared to the manual toggle-switch setting via an *Exclusive-NOR gate.* The Exclusive-NOR, U2 or U3, is a 74HC266 which normally comes as a quad unit.

If the EPROM Data pin Logic states agree with the toggle-switch settings, then each Exclusive=NOR output is Logic 1. If there is ***no agreement***, then an Exclusive-NOR output will be at Logic 0. Any input to U4, an 8-input NAND gate, that is Logic 0 will send its output high. If that output is high, then base current will flow in Q1 causing emitter-collector current to flow and the LED (D1) will light to indicate an error.

The Sequencer time span is only about 0.12 Seconds to write one byte of Data into the EPROM based on the manual Program button command. If the fault-indicating light remains on, the Program button can be pushed again, perhaps again for several times, until D1 remains extinguished. The operator can than change the EPROM Address pin settings, change the Data pin settings for the next Data Word.

When doing an EPROM Data Word write, the U1 outputs will always be in agreement with the Data Word toggle-switch settings. There will be no Fault indicated then. Only after the Sequencer has completed its one-time Byte-Program pass will be the time when the agreement or non-agreement can occur. In some of the correcting types of commercial programmers, the number of retries can be as high as 20 passes for the same Data Word at the same Address setting. Most of the good EPROMs will write a Data Word just once and then advance the Address line.

Because the 74HC244 is a relatively rare type of gate structure, its pin-out is shown in Figure 34-8 for its 20-pin DIP structure. Pins 1 and 19 are the Mode connections and do not appear on any outputs nor modify any input.

The Sequencer

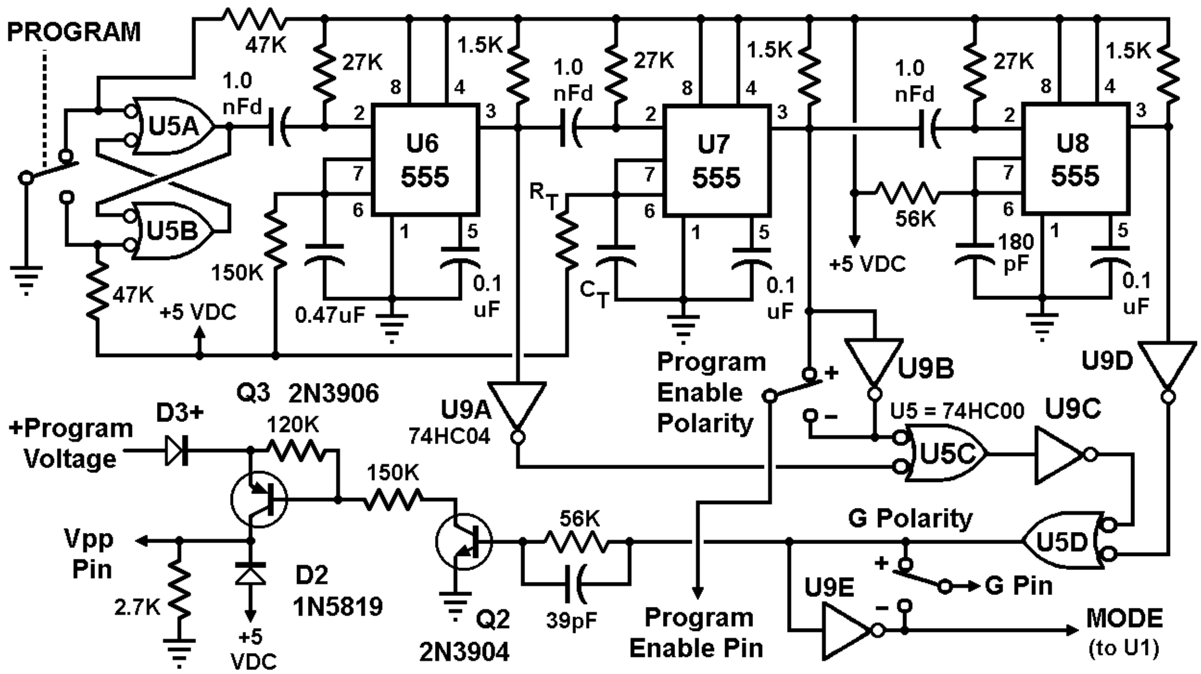


Figure 34-9 Sequencer for Manual EPROM Programmer.

Shown in Figure 34-9, this is basically three one-shots in cascade, triggered by a Program push-button switch through a simple R-S flip-flop made up of U5A and U5B. U6 one-shot is relatively long, about 80 mSec, enough to overcome most switch contact bounces.

The length of pulse out of U7 will depend on the EPROM specification. For the 2716 that is 50 mSec, ± 5 mSec, ***positive-going*** polarity. For a 27C1024 that is 100 µSec, ***negative-going*** polarity. A switch or jumper-wire setup (next to U5C) selects the pulse polarity. See the following table and waveforms for various EPROMs. U8 is about 11 µSec long, enough to complete all EPROM writes.

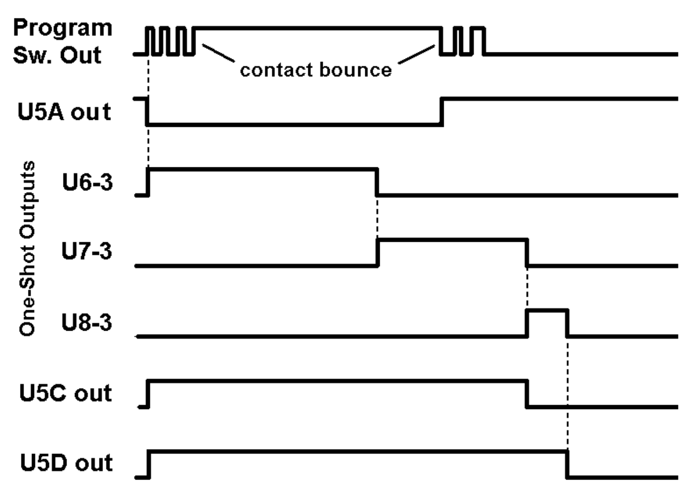


Figure 34-10 Sequencer waveforms. Program push-button switch (top line) is the top stator connection; note the possible contact bounce. Time scale (horizontal) is approximate.

MODE line in lower-right corner goes to U1 in Figure 34-7. Waveforms are shown in Figure 34-10. One-shot circuits using 555 Timers are explained in Chapter 27.

Gates U9 and U5C, U5D form an OR to take all of the one-shot outputs into a single continuous pulse into Q2. Q2, Q3, and D2 make up a variable-voltage electronic switch for the EPROM's Vpp pin. Vpp pin output will be +5 VDC at idle, rising up to just under the +Program Voltage level. The purpose of Schottky diode D2 is to supply the idle voltage to the Vpp pin when both Q2 and Q3 are cut off. When the combined one-shot voltages rise above +5 VDC, D2 is reverse-biased and cut off.

Separation of Data Input Programming Pins

EPROMs generally have ***two*** pins for programming Data Input Bytes. ***Vpp*** is raised to a slightly-higher DC voltage for a Data Byte input. Actual programming of Data within an EPROM comes from one or a combination of pins running between Logic 0 and Logic 1 levels. For the older 2716 this is a ***positive-going*** pulse. For the newer 27C1024 this is a ***negative-going*** pulse with an idle-time normal Read level of either Logic level state. A study of EPROM datasheets is necessary.

In Figure 34-10, time-scale is random. U6 one-shot output is the longest. U7 has its width set for the type of EPROM.

Values of $\boldsymbol{R_T}$ and $\boldsymbol{C_T}$ for U7 can range as follows (formula is found on Figure 27-13):
100 µSec: 33K and 2700 pFd (nominal width of 98 µSec)
50 mSec: 100K and 0.47 µFd
(nominal width of 51.7 mSec)
Note: For larger capacitor values, a polypropylene dielectric capacitor type should be used.

Program Voltage Supply

This must always be more-positive than +5 VDC. D3+ is any conventional silicon small-signal diode such as a 1N4148. Maximum output to the Vpp pin is External Supply minus a diode forward voltage drop (0.7 V) minus 0.2 V for Q3 base-collector saturation drop. D3+ can be omitted or added, depending on External Supply voltage. For a 2716 the External Supply could be +25 VDC with no D3; Vpp pin would be at 24.8 VDC when high. For a 27C1024 and an External Supply of +15 VDC, three D3+ diodes in series would be used; Vpp pin would be about 12.7 VDC when high. Datasheet specifications allow about ± 5 % tolerance on the Program Voltage for Vpp.[13]

Other Pin Connections

This Manual EPROM Programmer operates in ***Program***. Its resting mode is ***Program Verify*** which is equivalent to normal ***Read*** mode. It only sequences through ***Program Write*** *when the* **PROGRAM** switch is operated. In order to fit all EPROMs, at least 3 pins, possibly 4 have to be connected to specific outputs.

For a 2716 (2K x 8) EPROM, the Vpp pin 21 is connected to the Vpp output of the

[13] This tolerance was never verified for correctness on the bench. Instead the Vpp value was set to the specification value in the datasheet with concentration on achieving the correct ***data byte bit values.*** As long as those data bit values were correct, the External Supply could be set by an accurate voltmeter. As an example, a 27C1024 requires a Vpp in Programming of +12.0 to +13.0 VDC according to manufacturer AMD and 12.7 VDC falls within that range.

Programmer. What is called an Ep-not, pin 18, is the ***Program Enable*** output and its ***Polarity*** setting is **Positive.** What is now called a ***G*** pin (pin 20) goes to the ***G Pin*** Programmer output with a ***G Pin Polarity*** that is also **Positive.**

For a 27C1024 (64K x 16) EPROM, the Vpp pin is pin 1. ***Program Enable*** is pin 39 and its ***Polarity*** setting is ***Negative.*** The G Pin, also pin 20, has a ***Negative*** polarity. The ***E-not*** pin, pin 2, is tied to ground (Logic 0).

What is sometimes frustrating is having to study the (sometimes disparate) datasheets to determine the different EPROM operating modes. Most modern EPROMs have about six modes of operation. To simplify this Manual Programmer, the number of modes was reduced to two.

BODE Plots

General

Bode plots are a very simple form of frequency or phase response, either relative to frequency. Done on ***logarithmic*** paper, they are the ***asymptote*** of approximate response. That is, their straight lines ***do follow*** response, but only for ***most*** of it. As shown in Figure 34-11 for a lowpass R-C network, magnitude response is down 6 db in voltage at the ***corner frequency*** (intersection of unfiltered response and the filtered response).

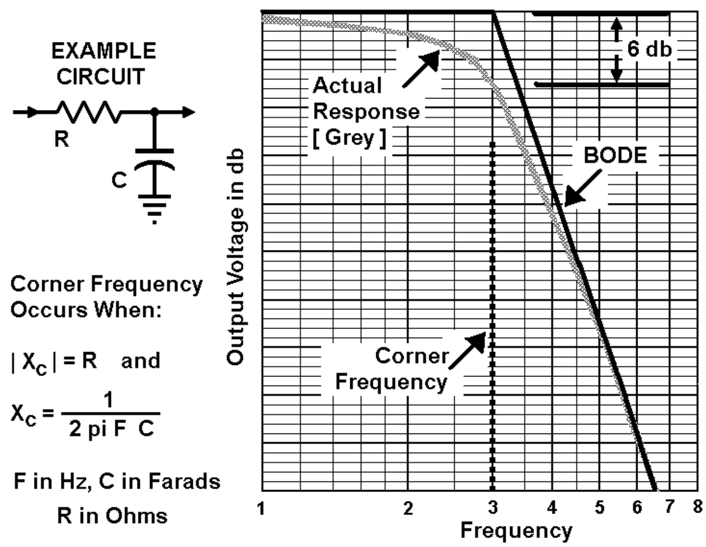


Figure 34-11 A sample Bode plot of an R-C lowpass filter. Actual response is shown in grey.

This is a convenient method of showing ***approximate*** response, done first by Hendryk Wade Bode at Bell Labs in the 1930s. But that was over 75 years ago, before there were PCs and easy ways to show ***actual*** response.

Nyquist Plots

Not to be confused with *Nyquist criteria* (as in sampled waveforms), a ***Nyquist Plot*** is a ***circular*** type of Bode Plot that yields both Magnitude and Phase information on one piece of paper. Think of it as a precursor to the ***Smith Chart*** used in RF design. Nyquist was attempting to clarify feedback design, also at Bell Labs at a time just after the Bode Plot was devised. The first *Smith Chart* came into being in the late 1930s just before the USA was into WWII; it was improved later and the Smith Chart became a de factor standard for RF work and measurement.

Both Bode and Nyquist Plots were good in their day but that was over 75 years ago when there were no PCs and frequency-magnitude-phase plotting was done by very expensive specialized equipment. Both of these plots are good electronics ***classroom*** subjects but this author thinks that they are rather far back in history now that personal computers can do the same thing at far less cost and directly from circuit data. Several of older texts will mention both of them, but be aware that the writers of such texts were schooled on them and they do not leave the mind very well.

Early Effect

James M. Early (1922 - 2004) worked at Bell Laboratories with the Bardeen-Brattain-Shockley team that developed the transistor. He came up with a number of things all lumped under the ***Early Effect*** name. One was that the common BJT Collector voltage-current Base current curves all have an origin well outside the Collector voltage on curves displayed on so many datasheets. This is shown in Figure 34-12 following. That origin applies to most FET Drain voltage-current curves as well.

This *Effect* does not show up in vacuum tubes. Tube curves are much more influenced by the many electrodes' shape and physical position which can alter tube characteristics. Transistors are simpler in their structure, therefore they relate much more directly to simpler physical laws.

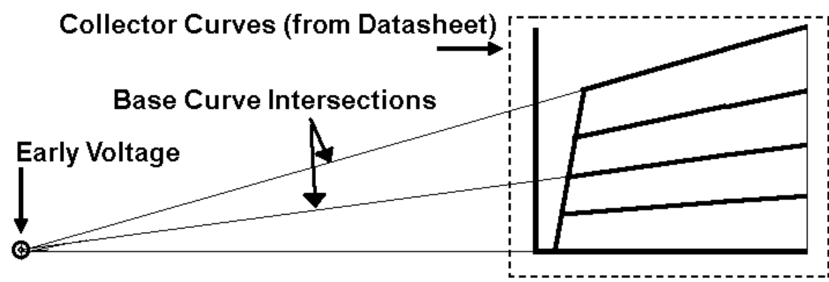


Figure 34-12 ***Early Voltage*** **point intersection of Base current curves of Emitter-Collector voltage-current curves (Shown within the dotted lines) found on datasheets.**

This data has no real effect on transistor circuit design but is shown as something rather unique to active semiconductor devices.

Radiation Resistance of Antennas

For antennas, this is an old term that is supposed to indicate the amount of ***power*** of the RF wavefront actually radiated. It has been picked up in other technologies and the exact meaning has been both confused and mis-stated. It is principally applied to physical antenna design and the relationship of RF current and voltage in antenna elements. As a term, it is still carried as meaningful to ***antenna efficiency;*** i.e., it is an old figure of merit which was once, in the 1920s to 1930s, to explain how *good* an antenna was for both transmission and reception of RF signals.

In the age of personal computers, the ***Method of Moments***[14] analysis, working on a fairly standard Personal Computer, is able to combine the electrical properties of feeding an antenna and the created RF wavefront, able to establish the spatial pattern method of RF wavefront radiation. Since antennas operate essentially bi-laterally, the same principles apply to reception of an RF

[14] ***Method of Moments*** or, colloquially, ***MoM*** is a general mathematical tool applicable to many different technologies. In particular this was applied to the ***Numerical Electromagnetic Code*** or ***NEC*** developed by G. J. Burke and A. J. Poggio at the US Navy Postgraduate School in Monterrey, California in 1981 as ***NEC-2*** and made available over the ***National Technical Information Service*** or ***NTIS*** for the cost of reproduction. There are four versions of the ***NEC: NEC-1*** (1979), ***NEC-2*** (1981), ***NEC-3*** (1985), and ***NEC-4*** (1992). As a USA government work it is copyright-free and NEC-2 has been ported to several commercial versions on the market.

wavefront into its feedpoint (and thus to a receiver input).

Efficiency of a transmission line to an antenna is part of the problem and is easily solved by an Antenna Tuner to deliver the most RF power to an antenna. The other part of the problem is coming up with the various antenna elements to transfer this wired RF power to free space. That is a much harder problem which has been nearly solved by the ***NEC - Method of Moments*** analysis. Such analysis also provides the RF wavefront power in three-dimensional data for simpler numeric conversion to RF signal level into the far-distant receiver, after free space path loss computation.

The end result for *Radiation Resistance* has, essentially, been rendered obsolete. However, for those who wish to appear all-knowing, specific formulas may derived from a few Internet sources for those who wish to appear more knowledgeable.

Chapter 35

PIC Microcontrollers

A primer on microcontrollers and their 8-bit devices from Microchip Technology for those not yet fully into such one-package devices. Some previous Assembler programming experience is helpful to fully appreciate the contents in this Chapter.

General

This Chapter is a primer on ***MPUs*** (Microcontroller Processing Units) as single-package ICs containing a microprocessor, memory, and dedicated peripheral circuitry. It concentrates on the ***CPU*** (Central Processor Unit) in its simplest form, the instructions for it and what the instructions do. Emphasis is on the Microchip PIC 8-bit ICs since those were judged to be among the simplest on the market.

Microchip began its product line with a ***PIC*** or Peripheral Interface Controller, developed the general scheme of it, kept its ***RISC*** (Reduced Instruction Set) as its basic language for all following 8-bit MPUs. The word ***microprocessor*** got into the field of small digital computing in the 1970s, probably to distinguish the 8-bit word length from the 16-bit ***minicomputers*** and the 32-bit to 64-bit bigger computers (later called *mainframe* types).

When the ***ASCII*** (American Standard Character Information Interchange) standard was adopted in the early 1960s, the basic word length needed was just 7 bits. The 8th bit was considered a *parity check* on information transfer. That evolved to a final version, sometimes called the *Extended ASCII* set. That could handle all the umlauts and special characters of alphabetic languages. Only the languages using ideographic symbols needed two bytes per character. All of that could still be handled by 8-bit MPUs.

The Basic MPU

As envisioned by the small but growing digital industry, the Microprocessor was basically an Arithmetic and Logic Register with at least one byte of temporary memory, an Instruction Set decoder, all with an external memory storage input-output that could hold 65,535 bytes. Limitations of hardware production kept the memory separate but the Instruction Sets were kept for specific models.

When Microchip began their PIC they did two major things: The old ***Harvard Memory***[1] storage was used with separate data and instruction memory sections; the MPU was based on the old ALU (Arithmetic and Logic Unit) but it was pared down to a single temporary data register.

---

[1] Memory storage is done two ways: The von Neuman method with combined Instructions and Data; the Harvard method with separate Instruction and Data locations. Most PCs use the von Neuman method for the simple reason of so many of these units used it. Both methods were known and used since post-WWII times in very early digital computation, before the solid-state era began.

Both of those meant they could have their ***RISC*** and also cut down the execution time to (nearly) a single ***instruction cycle*** for most of their reduced instruction set of 35 different functions, later to 49/51 instructions.

As semiconductor technology increased, the instruction cycle time could be dropped until reaching 0.2 µSec per cycle (5 MHz rate) by 2005. ***Flash*** memory locations allowed the specific program instructions to remain for at least 40 years. Adding more ***special function registers*** allowed a ***memory banking*** selection of four times the original memory storage to a maximum of 8192 steps.

All things in consideration, the 8-bit PIC line is now fast enough to do full frequency measurement of a Local Oscillator to ±1 Hz resolution (along with compensation for an IF offset) several times per second.

Simplest MPU

Shown in Figure 35-1, the basic CPU is an ALU plus a single register called **W** by Microchip. Internal gating handles the 8-bit pathways into and out of the ALU and W register.

Not shown fully is half of the ***STATUS*** Special Function Register (SFR) which holds the **Z** (Zero) bit 2 indicating the ALU output is all-zeroes if a Logic 1 and non-zero if a Logic 0; **C** (Carry) bit 0 which is a Logic 1 if a Carry-out has occurred, Logic 0 if no Carry-out. There is also a **DC** (Direct or Decimal Carry) bit 1 which is a Logic 1 if there was a Carry-out from the 4th bit of the ALU, a Logic 0 if no Carry-out. This ***DC*** bit is hardly ever used in examples and may have been intended for 4-bit arithmetic conversion.

Input and output is directed internally depending on the Instruction type. Inputs would be ***from*** the data memory (dubbed ***F*** for file), the ***W*** register, or a ***Literal*** binary value from the Instruction word (not the data memory). Those could go to either the ALU itself or the W register. Outputs would be ***to*** the data memory or W register outputs.

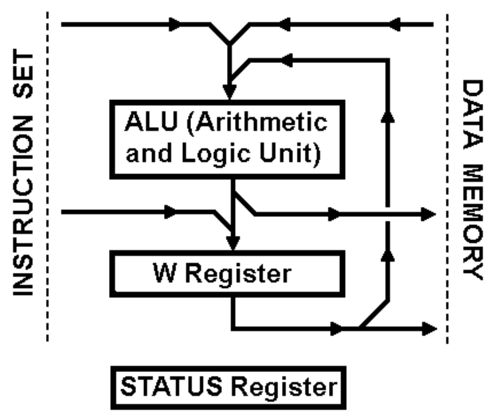


Figure 35-1 Simple CPU in the Microchip PIC 8-bit family.

Making Some Sense of Instructions' Suffixes

Fourteen of the 35 instructions in the (original) Instruction Set have the suffix ***F, D*** following an operand. Four others have the suffix ***F, B***. Two have only ***F*** for a suffix.[2]

The ***F*** suffix denotes the instruction operates with a data memory ***file word*** given following the instruction operand. That is usually a stated programmer-named variable or (doing it the hard way) an actual binary address for the variable's location in Data memory.

The ***D*** second suffix stands for *Destination* (of output) of the ALU and is one bit. Bit state is a Logic 1 for storing ALU output in the ***F*** Data memory, Logic 0 for storing ALU output in ***W***.

The ***B*** second suffix refers to the ***Bit*** of the operand and would be within the range of 0 to

[2] Capitalization is solely that of the author's choice to avoid conflicts in reading. Microchip literature shows those as lower-case.

7, corresponding to least-significant to most-significant within the byte. That second suffix occurs in only 4 of the 35 in the Instruction Set, specifically the Bit-Oriented operations.

The ***K*** suffix following 9 operands of the remaining 13 instructions actually refers to the operand, a binary value of 8 bits. This is a ***Literal*** input, usually given as a hexadecimal value following an ***H'*** and itself suffixed with a closing single quote. It can also be given as a binary number or a direct numerical value in decimal. Some examples of the same value are:

Decimal: D'123'
Hexadecimal: H'7B' -or- 0x7B
Binary: B'01111011'

Some General Terms

An ***Opcode*** is the name of the instruction. It is named for the purpose of identification in an Assembler listing but has an internal binary number value. An ***Operand*** is the named part following an *Opcode* to further instruct what the instruction should do with the named *Opcode.* A few *Opcodes* have no *Operands* at all.

A ***Word*** is a single binary collection of the same data. For most instructions a *Word* is simply an 8-bit data byte. In the original *computerspeak* terms, a *Word* could be anything in terms of bit length, usually from 8 to 64 bits long. In the Microchip terms, a ***Mid-Range*** MPU series (such as the 16Fnnnn family) has *Words* of 8-bit byte length. A few special cases will have *Words* slightly longer, up to 13 bits in the ***Mid-Range*** line from Microchip.

Destination, the **D** in an Operand, refers to the final location of the modified variable. For those instructions indicated as having *Destinations* of **F** or **W**, that is given by the character following the ***Operand*** by a comma. Only one bit is used for which one. A **0** or **W** names the *Destination* as the **W** register, a **1** or **F** names the *Destination* as the File or memory. Most Assembler development programs accept either number or letter. The author prefers the letter to indicate where the Destination will be.

Ranges of Named Files, Literal Values, Locations, and Bits

Files are limited to **127** by direct Assembler Operand**.** Instruction words are limited to 2047 total and the **F** suffix is allotted only seven bits to name a file. This may seem like a terribly small limitation if thinking in the PC mind-set of megawords of memory. The number of actual variables is usually less than 128 for typical applications. Memory paging allows ***banks*** of words switchable from software.

Literal values can be up to **256.** That is the size of the maximum value in an 8-bit byte. It fits directly into handling ASCII character data.

Bits can be any of the 8 bits in a microcontroller word. Bit 0 refers to the least-significant bit while Bit 7 is the most-significant bit. Note the *computerese wording* for numbers where, digitally, Bit 0 exists, but decimal non-logical wording would have least-significant as 1.

Location in memory is limited to a range of 0 to 2047. Eleven bits in an instruction word are allocated to specific locations. That can be increased by **memory banking.**

Memory Banking or Paging

In the beginning the PIC memory was limited. A memory range of 2K of RAM was thought a maximum. As solid-state technology advanced, much more memory could be added. Since the instruction set and general form were fixed at the beginning, a way to increase memory location was (sort of) solved by **banking**, that of creating a way to add **pages** of memory. Memory **pages** are selected by the three most-significant bits of the **STATUS** special-function register. *Banking* would change by 2010 resulting in a new scheme with up to 32 Banks. The old way is kept but the new way lends more flexibility.

The Original RISC

General

Terms for digital instruction here follow the general rules of ***binary*** Logic operation. You should be familiar with such basics. It is suggested that the digital logic beginner review Chapter 23 to make sure such binary terms are understood well enough to continue. The following copies the original RISC set, the additional 14 to 16 instructions beginning in 2010 shown later. Instructions are differentiated by the author into specific types.

Arithmetic Operations

ADDWF	Add **W** with named file **F,** result to destination **W** or **F**; affects STATUS **C, Z** bits
ADDLW	Add **W** with Literal Operand value, result to **W**; affects STATUS **C, Z** bits
SUBWF	Subtract **W** from named file **F**, result to destination **W** or **F**; affects **C, Z** bits
SUBLW	Subtract **W** from Literal Operand value, result to **W**; affects STATUS **C, Z** bits
INCF	Increment named file **F** by one, result to destination **W** or **F**; affects STATUS **Z**
DECF	Decrement named file **F** by one, result to destination **W** or **F**; affects STATUS **Z**

Both SUBWF and SUBLW use two's complement arithmetic so the Borrow and Zero bits of STATUS register are reversed.

Binary Operations

ANDWF	AND named file **F** and **W,** result to destination **W** or **F**; affects STATUS **Z** bit
ANDLW	AND **W** with Literal Operand value, result to **W**; affects STATUS **Z** bit
IORWF	OR named file **F** with **W,** result to destination **W** or **F**; affects STATUS **Z** bit
IORLW	OR **W** with Literal Operand value, result to **W**; affects STATUS **Z** bit
XORWF	Exclusive-OR named file **F** and **W,** result to destination **W** or **F**; affects **Z** bit
XORLW	Exclusive-OR **W** with Literal Operand value, result to **W**; affects STATUS **Z** bit
CLRW	Clear **W** register, no Operand, result to **W**; affects STATUS **Z** bit
CLRF	Clear named file **F,** destination file **F**; affects STATUS **Z** bit
SWAPF	Swap upper and lower nibbles in named file **F**; no change in STATUS register
COMF	Complement logical bit value in named file **F**; result to **W** or **F**; affects **Z** bit

A COMF operation exchanges Logical 1 and 0 values in the Register, resulting in a ***one's complement*** final value.

Shift Operations

RLF Rotate named file **F** to Left one bit through **C** bit, result to **W** or **F**
RRF Rotate named file **F** to right one bit through **C** bit, result to **W** or **F**

Rotational Shift instructions are shown in Figure 35-2. No original bits are lost since the bit of the named Register moving *out* is put into the STATUS Register **C** bit.

Both instructions are good for simple *binary* multiplication, RLF the same as doubling original Register value and RRF halving original Register value.

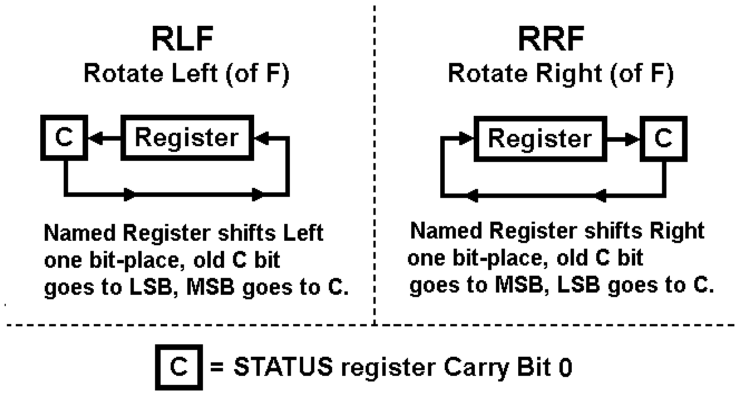


Figure 35-2 Visual representation of RLF, RRF

Word Movement Operation

MOVF Move named file **F** to destination **W** or **F**; affects STATUS **Z** bit
MOVWF Move **W** to named file **F**, destination is **F**; no STATUS bits are affected
MOVLW Move Literal value in Operand to **W**; destination is **W**; no STATUS bits affected

A MOVF Opcode to the same place in memory might seem redundant. Since the **Z** bit is affected, such an instruction can be used to check the file if its contents are all zero; it will not change the data contained in file **F**.

Unconditional Program Flow Movement

GOTO Jumps to the named LABEL given in Operand; no STATUS bits affected. Note: the LABEL location must be within 2048 program steps, allowed Operand has 11 bits.
CALL Jumps to the named Subroutine Label given in Operand, keeps the old CALL location in the STACK memory. Same 2048 program step limitation.
RETURN Last instruction of a normal Subroutine, returns to CALL program location plus 2 via data in the STACK memory; no Operand.
RETLW Same as RETURN except Operand contains a given Literal 8-bit value
RETFIE Same as RETURN but used only in Interrupt-Servicing routines

RETLW can be used in certain ***table-read*** routines to read constants stored in EEPROM. The RETFIE is different in that it changes INTCON bit 7, the Global Interrupt Enable bit.

All of these Unconditional Program Flow Movement instructions take ***two instruction cycles*** to complete. All others given before take only ***one instruction cycle.***

Conditional Program Flow Movement Operations

BTFSC Bit Test named file **F** and Skip immediate-next instruction if named bit is Clear or zero. If bit is Not Clear, then execute the immediate-next instruction.
BTFSS Bit Test named file **F** and Skip immediate-next instruction if named bit is Set or Not zero. If bit is Clear, then execute the immediate-next instruction.
DECFSZ Decrement named file **F**, Skip immediate-next instruction if value is Zero. If bit is Not Zero, then execute the immediate-next instruction.
INCFSZ Increment named file **F**, Skip immediate-next instruction if value is Zero. If bit is Not Zero, then execute the immediate-next instruction.

These four take either ***one*** or ***two*** instruction cycles to execute. If the Test is met, then they require two instruction cycles. If the Test is Not met, they take just one instruction cycle. None of them affect the STATUS register bit settings.

Individual Bit Operations

BCF Clear a given bit of a named file **F**; no change of STATUS bits
BSF Set a given bit of a named file **F**; no change of STATUS bits

Miscellaneous Operations

NOP No Operation. Good for timing loops to use one instruction cycle without doing any specific operations. No Operand.
CLRWDT Clear Watchdog Timer. Sets STATUS bits TO-not and PD-not. No Operand.
SLEEP Enter Sleep mode. STATUS bits TO-not is Set, PD-not is Cleared. No Operand.

A NOP instruction is good for software adjustment of timing loops by exactly the time of one instruction cycle. More NOPs can be used for slightly longer delays.

Both CLRWDT and SLEEP will probably not be used in any radio projects. Microchip Technology makes microprocessors for non-radio uses, particularly portable devices which must be powered-down to a minimum before being made operational. That is done to conserve battery power source drain. For most radio uses, a radio will usually be on when the listener is present. When the listener leaves, the listener turns off all power.

Some Basic Routines

Moving Data Variables

A *variable* is any particular data storage unit. It has become synonymous with *word.* Much of a program is concerned with moving a variable (or a word) from one place to another. Often that includes *doing something to a variable.* The following is a set of instructions on moving such data:

```
MOVF     WORDAA,W    ; Get a word out of memory into W register
MOVWF    BWORD       ; Send word from W to NEW place in memory.
```

The first instruction uses ***MOVF*** which can have its destination depending on the final suffix. Rather than use a 0 (for W) or a 1 (for a file location), the author uses a specific letter allowed by the Assembler program; this is a mnemonic to realize which direction the word (or variable) is moving. In this simple case, variable WORDAA is simply duplicated in variable BWORD. Remember that MOVWF has no descriptor, just goes directly to file memory.

To actually ***do*** something, an instruction can be put between the two as in:

```
MOVF        WORDAA,W    ; Move WORDAA into W register
ANDLW       H'0F'       ; AND literal (hex) 0F to W (WORDAA)
MOVWF       BWORD       ; Send this modified word to file
```

This technique is called ***masking.*** It masks off the upper nibble, allowing only the lower nibble to exist in the W register. Bits 4 through 7 of W are made Logic 0. W is then sent to the file under the new name of BWORD.

The new second instruction can be most anything that operates on the W register. It can have additional instructions such as:

```
MOVF        WORDAA,W    ; Move WORDAA into W register
ANDLW       H'0F'       ; AND to allow only lower 4 bits in W
IORLW       H'30'       ; INCLUSIVE OR to make bits 4 through 7
                        ; into 0011 ---- for an ASCII capital
                        ; chracter.
MOVWF       CWORD       ; W register now sent to file in CWORD.
```

Examples of Conditional Branching

A ***conditional branch*** has been common in program languages since the beginning. It means that the ***condition*** of a result determines the flow of instructions (or ***branching to***). If a *condition* is a so-called normal result, flow continues in one direction. If a *condition* results in an opposite state, then program flow goes a different direction. As an example, assume a repeated program is executed and a *counter* is decremented. If the *count* reaches all-zeroes, then program flow exits.

```
            MOVLW       H'14'       ; Put a decimal 20 in register W.
            MOVWF       COUNTC      ; Hold this counter state in file COUNTC
REPEAT       ...
             ...                    ; Do a particular program action here...
             ...
            DECFSZ      COUNTC,F    ; DECrement variable COUNTC by one, SKIP
                                    ; over next instruction if all ZEROES.
            GOTO        REPEAT      ; Jump to label REPEAT and do everything
                                    ; of a particular action again, or
NEXTONE      ...                    ; drop out and do the following action.
```

This is good for actions such as making a timing loop or just operating on a block of data. The action taken (marked by ellipses) is not directly related to variable COUNTC. COUNTC is merely a counter and begins with a logic state of binary 14. At the DECFSZ instruction it will decrement by one and check logic status of bits in COUNTC. If any bits contain a Logic 1 then the variable is ***not zero*** and program flow moves to the GOTO instruction. Label REPEAT is the beginning of the same program action (marked by ellipses).

Variable COUNTC began in a state of decimal 20. On the first pass through, it decremented to decimal 19. After the second pass through, it decremented to decimal 18. The same pass throughs will occur for COUNTC going to 17, 16, ... 2, 1, and finally 0. On COUNTC reaching decimal 0 program flow skips the GOTO instruction, in effect ***branching out*** of this loop and continues with the rest of the program at label NEXTONE.

The DECFSZ instruction makes COUNTC (and the intermediate loop) go through the ***same*** number of loop iterations as it was originally formed in the opening, decimal 20. It is a ***count-down*** type. A ***count-up*** loop can be formed by changing the initial MOVLW to a state of H'EC' or decimal 236, then changing the DECFSZ to an INCFSZ. COUNTC would begin with decimal 236, then increment to 237, 238, 239, 240, ... 254, 255, and finally 0. Branching is the same.

Count-up loops are a bit more difficult to set. They begin with a value of (decimal) ***256*** minus the number of repeated loops (in this case 20).[3] For 20 iterations, initial value of COUNTC would be (decimal) 256 - 20 = 236.

Besides DECFSZ and INCFSZ, there are only two other conditional branching instructions: BTFSC and BTFSS. Instruction names are a contraction of ***Bit Test F Skip on Clear*** and ***Bit Test F Skip on Set.*** Those two allow testing of *individual* **bits** within a variable. Branching is still the same: Test the bit, skip the next instruction if the condition is met, otherwise execute the next instruction.

Addition and Subtraction

There is no Carry or Borrow acted upon after either an addition or subtraction. A Carry or Borrow is stored in the STATUS register. A Carry is ***not added*** nor is a Borrow ***subtracted*** directly. It must be taken out of the STATUS register and added/subtracted separately. As an example, if ADONE and ADTWO are going to be added in sequence, sum to be stored in ADTWO, then the code could look like this:

```
MOVF      ADONE,W     ; Get ADONE into W register
ADDWF     ADTWO,F     ; Add W + F (ADTWO), result in ADTWO
BTFSC     STATUS,C    ; Test CARRY bit, skip over it if CLEAR
INCF      ADTWO,F     ; Not Clear so add Carry into ADTWO.
```

Of course, if the sum came out to 1111 1111 and a Carry was done separately, it would have to be added in separately afterwards. The same would result with subtraction except the INCF would be changed to DECF. Both conditions would be solved by changes to the Instruction Set forthcoming.

Reading of Constants

Constants, in programming, are numeric values that do not change as a result of program manipulation. A *variable* may be a constant, in this instance. To add to this semantic upset, any variable is always considered a binary numeric *thing* within the program and microcontroller. That is, it appears to have a *name* but is really just a collection of 1s and 0s within the micro.

Constants may be declared (in Assembler) before a program begins. This is done by an ***EQU*** (equate) statement. That variable may be called at any time by the program, read-in as the constant. Such variables can be over-written by the program, thus resulting in new values for that named variable.

---

[3] Maximum value of 8 bits is decimal 255. The state of decimal 256 would be binary 1 0000 0000, requiring nine bits.

Instruction Enhancements

General

In the beginning of this new millennium, Microchip made some additions to the Instruction Set, raising the number of Instructions from 35 to 49.[4] A few other improvements were added, such as increasing the internal ***stack*** size, used to keep track of program locations with subroutine calls. In the spirit of merchandising, this was named the ***enhanced 8-bit MCU*** feature and appeared with the 16F1517, 16F1784 and a few others.[5] These new devices will run older instructions developed for older devices; older devices may ***not*** operate (or perhaps fail) trying to run these new instructions.

Byte Arithmetic Operations

ADDWFC	Add W and F (named in instruction) and Carry, result in File or W
SUBWFB	Subtract F - W - Borrow, F named in instruction, result in File or W

What to do with an arithmetic Carry or Borrow was solved by adding the ADDWFC and SUBWFB instructions. Presumably the last Carry or Borrow will remain in the STATUS register for those new instructions. For multi-byte, chained additions or subtractions, it must be determined ahead of time as to the total number of bytes in the expected answer. That will apply to using both the old and enhanced instruction sets.

Logical Shift Operations

ASRF	Arithmetic shift Right of instruction variable, result in File or W
LSLF	Logical shift Left of instruction variable, result in File or W
LSRF	Logical shift Right of instruction variable, result in File or W

In ASRF the original bit contents are shifted right once (LSB going to C) but original Bit 7 is duplicated in the new Bit 7 position.

LSLF and LSRF are same as in old microprocessors..

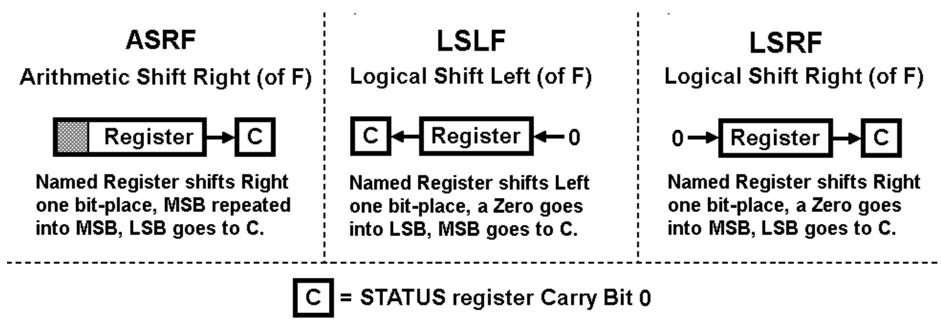


Figure 35-3 Illustration of added Enhanced Shift Instructions.

---

[4] There is some confusion here. According to the Microchip extensive documentation, the ***Enhanced 8-bit*** line has, by count, a total of ***fifty-one*** instructions, not *forty-nine.* Every datasheet has the same ***Instruction Set*** so anyone can count all by themselves.

[5] The 16F1517 and 16F1784 are very close to the older 16F884, both in package size and general internal features, although the *enhanced* versions have a few more things inside.

Byte Movement Operations

These four instructions should be used with great care since they are rather complex and require more information of the FSR (File Select Register), INDF (Indirect File Register), BSR or PCLATH Registers.

MOVIW Move ***indirect FSR*** to W (see following greater description)
MOVWI Move W to ***indirect FSR*** (see following greater description)

Both MOVIW and MOVWI have a large set of Assembler suffixes, too complex to state here; see Datasheet Instruction Set Summary for exact syntax and ordering.

MOVLB Move given 5-bit Literal value to Bank Select Register (BSR)
MOVLP Move given 7-bit Literal value to PCLATH Register

Conditional and/or Branching Operations

BRA Add the signed 9-bit Literal value to the PC (Program Counter), branch to the Label following the opcode.
BRW Add the un-signed W Register value to the PC (Program Counter).

For both BRA and BRW the PC will be incremented by one so the destination will be (PC) +1+(k) or (PC)+1+(W).

CALLW Calls subroutine by value in W, Operand is the instruction Literal value

CALLW has no operand but requires W Register already set to the Address for Bits 0 to 7; operation results in (W) loaded into Bits 0 to 7 of the PC, and PCLATH into PC Bits 8 to 14.

Other Operations

ADDFSR Signed 6-bit Literal is added to the FSR register pair; FSR result is limited to between 0000 and FFFF in hex; a roll-over to same FSR if exceeded.
OPTION Load OPTION Register with contents of W
RESET Executes a RESET of hardware from a command in software**;** this is risky to use unless there is a way to restart the microcontroller, so be warned.
TRIS Loads W to TRIS register, TRIS named by numeric value in instruction; that value would be 5 for TRISA, 6 for TRISB, 7 for TRISC.

Some Thoughts on the Enhanced Instruction Set

Microchip is marketing their PICs to a wide market. As such, the Enhanced Instruction Set can be a bit bothersome in *interpretation and implementation.* For relatively-ordinary program sets such as frequency meters, the standard set of 35 will do nicely. There are perhaps only six of the Enhanced instructions which might be an improvement in terms of reducing the overall program size.

The ADDWFC and SUBWFB instructions will reduce the bit-handling for arithmetic operations. Since they handle the Carry bit directly, they can make greater-than-two-byte additions or subtractions without the extra steps of getting the Carry-or-Borrow bit to add or subtract as extra steps following a conditional-branch check.

The LSLF an LSRF are good for shifting Left or Right when a logical operation can do without either the MSB or LSB. This is good primarily for certain algorithms where there is an advantage to *Logical shifts* as opposed to *Rotational shifts***.**

The TRIS instruction can save the extra steps needed to change the A, B, or C Port settings for Input/Output data exchanges. This is good for low pin-count packages requiring many and different Input and Output interfacing.

The OPTION instruction is, perhaps, limited. It would apply to many changes of TIMER0 and TIMER2 setting plus some algorithms with analog-to-digital or pullse-width-modulation.

With the other Enhanced Instructions there is little to remark upon. Most of those get into reducing some of the Assembler code in certain, specialized functions. Those are not the subject here unless one's hobby pursuits are largely devoted to microcontroller programming.[6]

Planning and Organization of a PIC Program

General

For the majority of microcontroller applications to hobby style projects, nearly all can be relegated to the original 2047 bytes of program memory. Bank Switching need be done only to change certain FSR contents on-the-fly. The same is true of RAM addressing where most applications will be satisfied with, at most, 127 bytes of variables. On-board EEPROM memory can be (usually) just left there, unused, in the package.

Some Minor Philosophy

As stated, Microchip markets to a very wide user audience. That audience of potential buyers may love to have all sorts of peripherals handy with a large program and RAM variable space at its disposal. For most needs of the hobbyist, the simpler microcontroller will work out fine.

It is an interesting observation that, once a die mask has been proved by test, it can be produced in quantity for approximately the same cost as an older, simpler die. The major change there (as it affects the unit price) is the package variation. There is little difference between one device of a family having one RAM memory size and another one with four times the RAM for the same purpose. The same is true for Program memory sizes. The number of I/O pins of the package will make the larger difference in price and that makes the great variation among family members.

Avoiding *Bells and Whistles*

It is generally a practice of the beginning microcontroller application designer to *use everything within a package...because they* ***bought it all***...and, perhaps, *because it is there.* A case

---

[6] That subject can be fascinating in itself and there are many texts available on the market just for the PICs. Some of those can be obtained via Microchip's own website.

in point is making routines with lots of alphanumeric output, as in messages of various kinds. In the various Internet user groups, alphanumeric I/O and small LED displays are a big thing. They are *big* because such is more eye-catching to the Internet page reader and the page looks *interesting.* That holds true when considering the medium. It does not hold true necessarily for the ***project.***

Consider a large EPROM. They are mostly binary memory. What is held in memory can be quite large. There is very little desire or application to ***converse with an EPROM***. The same should hold true with a microcontroller, even a fancy one for many bands. There should be a limitation on the number of warning outputs or fancy alphabetic messages sent to a display. For most of its use, the microcontroller is simply one of several different parts within a radio. It has a definite function but its function is generally just *output to the project's user.* Such a user is ***not*** trying to converse with it. The user is just relying on getting an output.

Bells and whistles is a colloquial expression of having lots of extras. The major task is to do the (rather) plebeian task of ***performing a certain function.*** Once that has been proved, the hobbyist ***might*** add bells and whistles. If there is program or variable space. A microcontroller can be re-programmed many times. Or the hobbyist may be satisfied to leave it as-is.

Being Organized...and Creative

The main thing in writing a microcontroller program is to be ***organized.*** Limit the number of subroutine and conditional branches to elsewhere in program memory. This keeps a program from getting too confused, hard to read later. Within a block of code one can get ***creative.***

Chapter 36

Elemental Metrology

Metrology is the science of measurement and measurement of characteristics of a design are absolutely necessary to ensure it operates as intended. Measurement is required in the hobby workshop environment in everything from testing parts to setting up laboratory power supplies. Test instruments can be grouped into two main types: ***Stimulus*** and ***Response***. Stimulus instruments are represented by signal generators and power sources. Response instruments would include voltmeters and oscilloscopes.

The First *Meters*

The first response instruments were little more than electrostatic potential indicators of *static electricity*, that of attraction of charged conductors to low-mass things like dust and lint. The first meters (if one can call them that) consisted of metal foil strips in slight vacuum of a glass bulb. If one foil, connected to an outside electrostatic potential, moved away from another foil that was grounded, it gave an *indication* that an opposite charge was there. Hardly quantifiable and of little use as a meter, it was more of a scientific curiosity or novelty. Once experimenters in electricity got to know more about the subject, that there was both voltage and current present in any circuit, the ***galvanometer*** was born

Galvanometers

The name derives from *galvanic current*, one of the first names for what we now term *electricity*. The basic design of the analog (moving-needle) meter was invented by Jacques-Arsene d'Arsonval (1851-1940). It is a coil of fine wire in the field of a permanent magnet, the coil held in rotational tension by a small stiff wire. As electrical (DC) current flows in the coil the magnetic field set up by it reacts with the permanent magnet field to physically move the coil. The first galvanometers used a small mirror attached to the coil to reflect light onto some nearby screen. While quite sensitive (full scale a few µA) there was no precision as we know it today.

Galvanometers, just as all d'Arsonval analog meters of today, measure ***both*** current and voltage. The moving coil has a finite resistance. Current through the coil will create a voltage potential across that coil. While the *meter motor*[1] is normally rated in Amperes, it always creates a voltage drop from the coil resistance which is seldom specified but can be readily measured.

Few electronics technologists of today bother with galvanometers although such still seem

[1] Term used by analog meter makers in the latter 1900s for the basic coil-magnet-needle movement.

to be worthy of mention in academic circles.[2] Electronics and solid-state technology enables us to measure *picoAmperes* with accuracy, and at lower voltage drops than is possible by non-active-device galvanometers.

Voltmeters

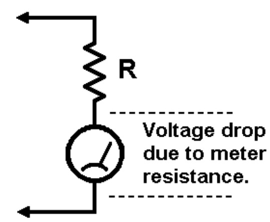


Figure 36-1 Basic analog voltmeter

An analog voltmeter basic circuit is shown at right in Figure 36-1. It is nothing more than a milliammeter or microammeter with a series resistance. The series resistance **R** is calculated from the desired full-scale voltage indication divided by full-scale meter current, ***minus*** the meter motor resistance. In older times, such voltmeters were called out by their *Ohms per Volt* after the value of the series resistance. A 1.0 mA full-scale meter would be a *Thousand Ohms per Volt* and a 50 μA full-scale meter would be a *20 Thousand Ohms per Volt* kind. Note one thing: Such simple analog meter circuits will ***always take some current from the circuit being measured.*** The amount of current drawn is termed *loading* after its effect on the circuit being measured.

Multimeters were introduced to the market prior to World War II. These had several full-scale voltage ranges, selectable by rotary switches or connections to a string of series resistors. Multimeters also had some current measuring capability plus marginal-accuracy resistance measurement ranges. The better ones used low-full-scale meters along with mirror scales for accuracy in reading the meter face markings.[3]

During and after World War II the ***VTVM*** or Vacuum Tube Volt Meter became available and it had a standard input resistance of 10 or 11 MOhms at all voltage ranges (depending on model). Those loaded the measured circuit the least and were thus preferred. A VTVM used a dual-triode basic circuit in a differential configuration, essentially two cathode followers with the display meter between the cathodes. One side's grid went to the voltage range selection resistors while the other side's grid went to an adjustable bias voltage *zeroing* manual control. Powered from AC mains, the three-tube VTVM consumed little power and remained in calibration as long as the non-active-device multimeters.

The ***DVM*** or Digital Volt Meter would replace all its predecessors. Those use an internal ***ADC*** (Analog to Digital Converter) which is referenced to a very accurate internal ***band-gap*** voltage source. Input impedance is still high, equal or better than a VTVM. While nearly all require batteries for operation, the common ***DMM*** (Digital Multi-Meter, combining voltmeter, ammeter, DC or AC, plus ohmmeter and, sometimes, a frequency counter) can *float* without regard to AC mains connection effects.

---

[2] This is a curiosity but possibly a result of academic practice of referencing older works in textbooks. The whole science of electricity has yet to see a bicentennial (as of this writing) and radio, as a communications medium, is just past its centennial. It takes time to produce textbooks.

[3] Meter needles have to be spaced out from the marked meter face. That space can result in a reading error if the reader is not viewing it perpendicular to the meter face. The mirror on the face allows the observer to line up the needle exactly on the needle reflection and thus be in the perfect perpendicular position. Such can be found on an old 1950s to 1960s' *Simpson model 260,* a 20 Thousand Ohms per Volt multimeter.

The analog meter coil resistance can be measured via the Figure 36-1 circuit. Apply a variable DC voltage (with a suitable value of **R**) until full-scale indication is made. Analog meter motors have a rough average of 50 mV drop at full scale. A DVM or VTVM can measure that voltage drop and the coil resistance would be the drop voltage divided by the full-scale current.

Voltage Standards

Until the arrival of the solid-state era around 1960, the local standard for voltage was the *Standard Cell*. Made in a tubular glass figure-H shape, the first Standard Cell was devised in 1878 by Latimer Clark using zinc and mercury electrodes in a sulphate electrolyte. In 1893 Edward Weston replaced zinc with cadmium and that version became the standard Standard Cell for decades. The voltage was quite temperature stable at 1.0183 Volts (very little current loading) and can still be found in Calibration Laboratories as a backup to various forms of the Band-Gap IC voltage reference.

The Effect of Meter Loading

Figure 36-2 is an example of error measurements as might be encountered with higher-resistance vacuum tube circuits. Figure 36-1(A) is the circuit as-is and it is desired to check the voltage divider output. In Figure 36-2(B) an ancient 1,000-Ohm-per-Volt voltmeter, set to 50 V full-scale, is put between divider output and ground. The resulting error is an astonishing -50 percent! In Figure 36-2(C) a less-ancient 20,000-Ohm-per-Volt voltmeter is used. Even though the loading with it is 1 MOhm, the indicated divider voltage will be low by about 4.76%. In Figure 36-2(D) a 10 MOhm input resistance VTVM is used, resulting in an error of only -0.48%. Always keep in mind the potential effect of error by the measurement device, any measurement device introduced into the circuit.

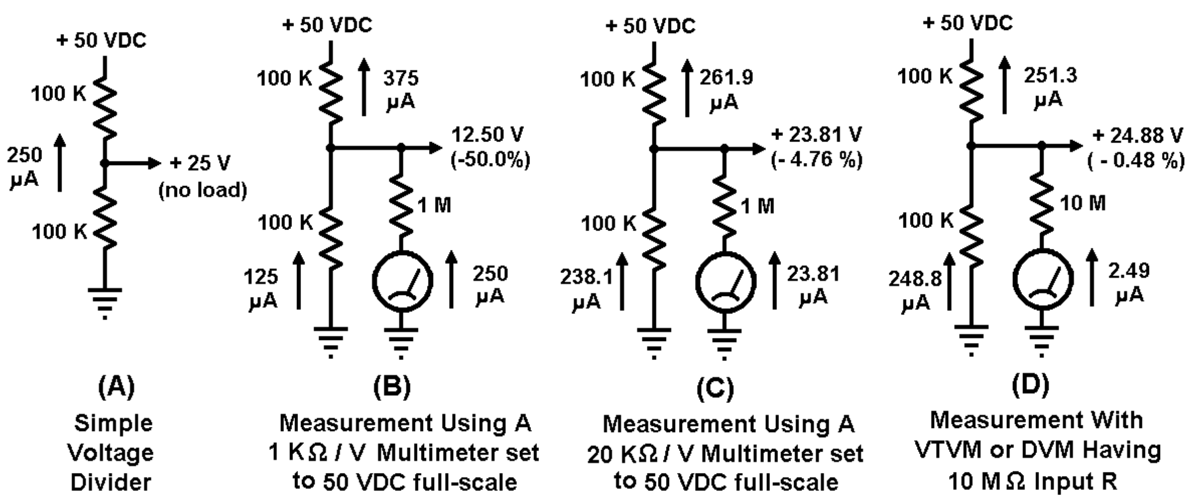


Figure 36-2 Effects of circuit loading of various voltmeters' input resistance.

AC Voltage Meters

A symmetric sinewave into an analog meter circuit will not indicate any voltage; asymmetric waveforms will indicate something not Quantifiable. For AC voltage measurements a simple rectifier is connected across the meter motor. The older 1000-Ohms-Per-Volt meters used a cheap (but effective) copper-oxide rectifier bridge; the most inexpensive models omitted filter capacitance. Later models used germanium, then silicon diode rectifiers. In some multimeters the AC ranges had a half to a quarter of the *Ohms-Per-Volt* input resistance. It should be noted that most analog AC voltmeters will ***also indicate the DC component*** in a circuit; depending on the circuit that DC component may be adding or subtracting from the AC component.

VTVMs do not need any input circuit changes to pass both DC and AC components (AC to upper end of audio frequencies) courtesy of the differential triode internal circuit. Isolation of input versus display circuitry allows switching in an AC rectifier plus the ability to sense only AC or only DC components (depending on type and model). Digital voltmeters have similar input versus display circuit isolation

Root-Mean-Square

AC voltage and current is generally referred to as ***RMS*** or Root Mean Square. What that means is that any AC voltage or current is defined as *the square-root of the sums of the squares of (many points) along the AC waveform period.* That sort of measurement allows great latitude in the shape of an AC waveform, anything from a pure sinusoid to a distorted periodic shape. See Appendix 37-1 for more detail.

With pure sinusoids, the RMS voltage or current is the peak divided by the square-root of 2 or about 0.70711 times the peak value. Most, but not all, simple rectifier circuits will result in an *average* value of voltage or current. With pure sinusoids the average value is extremely close to the RMS. With distorted waveforms the rectified voltage or current tends towards the peak value. Most multimeters have average-reading AC measurements even though they are specified with RMS. In general, noting or writing an AC voltage without a qualifier implies that it is RMS.

RMS became the standard arbitrary-waveform measurement system due largely to Charles Proteus Steinmetz (1865 - 1923) while at the Edison Company.

DC Current Meters

These ***always*** involve a voltage drop across the current meter whether they are analog or digital types. For DVMs the drop depends on the meter specifications. For analog meters the meter motor must be known for both resistance and voltage.

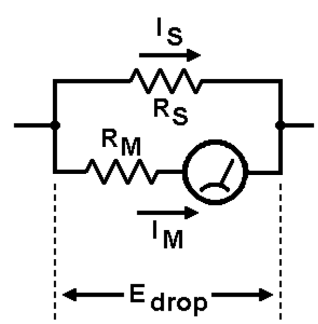


Figure 36-3 Typical analog current meter and shunt resistor.

Let: I_M = Meter full-scale current

R_M = Meter motor resistance

I_{FS} = Full-scale current of final circuit

Then:

$$R_S = \frac{I_M - R_M}{(I_{FS} - I_M)} \qquad \text{and} \qquad I_S R_S = I_M R_M$$

An easy way to check on the calculated value of shunt resistance is to use the second equality for the equal voltage drops across the two branches of the circuit.

AC Current Meters

The most-used type of AC current measurement is the ***clamp-on*** ammeter. Good for mains frequencies into the audio range, the clamp is part of the core of a high turns ratio transformer. The transformer's primary winding is approximately a quarter- to half-turn equivalent of the line to be measured and the secondary winding is inside the clamp-on's housing. The secondary winding is connected the rectifier and calibration components, display either in the clamp-on's housing or made to external voltmeter connections. Any voltage drop in the line to be measured is quite small due to the impedance ratio being the square of the turns ratio, thus reflecting a low impedance from the rectifier-display circuit to the equivalent primary winding.

Inaccuracies come about by improper placement of the clamp jaws' plane relative to the to-be-measured line's run. Note: Manufacturers' instructions generally tell how to place the clamp jaws for best accuracy. Such is directly visible by applying a clamp-on and changing the jaw planes' angle to the wire carrying a load. While of less precision than a fixed transformer, the clamp-on ammeter can measure in areas of high voltages with safe isolation and on any one of several current-carrying lines. No DC components will come through a magnetic transformer coupling.

A toroidal core transformer is perhaps the most accurate for fixed applications. A high-current line can go through the core's center hole, acting as an approximately quarter-turn primary winding of the transformer. The secondary circuit can have either an AC voltmeter or AC current meter depending on the step-up of voltage or current. Frequency range is dependent on the constancy of the core material's permeability over frequency. *Ferrite* or *tape-wound* cores are good for AC mains through audio frequencies; *powdered-iron* core material is best for RF from MF on into VHF.[4]

RF Current Meters

In early radio transmitters the antenna RF current was measured *indirectly* using various forms of ***thermionic*** devices. These were very low-resistance heating elements whose heat was physically transferred to a thermocouple-like dissimilar metal junction. That junction, in turn, generated a current at a low potential which could be measured by a D'Arsonval DC analog meter. These RF current meters were extremely broadband, the heating element being nearly entirely

---

[4] *Ferrite* is heat sintered iron alloy particles in a solid mass, good as transformer cores below MF. *Powdered-iron* core material is iron alloy powders held in suspension in a plastic binder, best for transformers at MF and above. *Tape-wound cores* are made from a high-permeability iron alloy tape wound into a roll or semi-toroidal form, best at audio and below. Note that EMI-suppression ferrites used on PCs and peripherals do not generally make good HF through VHF transformers due to varying permeability over their high frequency response characteristics.

resistive, hardly any inductive or capacitive reactance. The limitation was needing a relatively high antenna line impedance and relatively high RF current. Typical uses at HF bands were with 10 KW or greater transmitters with 600 Ohm balanced line outputs. A 10 KW transmitter into a 600 Ohm line produces a 2.45 KV RF voltage with 4.08 A RF current; a 0.1 Ohm series resistance produces 0.4 Watts of heat for the thermionic junction. The 100 mOhm series resistance upsets the 600 Ohm line balance only a slight amount if used singly.[5]

Thermionic, or properly, ***calorimetric***, RF power meters have been used at low powers as well. A notable example is the Boonton 260A Q Meter (see Appendix 1) where a very low resistance is both the load for an L-C Q measurement RF input and for indication of the RF voltage applied to the L-C test circuit.[6] Other than Q Meter RF excitation applications, the thermionic indirect current meter is relegated to high-power transmitter output current measurement below VHF.

The fixed toroidal core transformer is probably best for hobbyist use at RF powers below a KW on a 50 Ohm characteristic impedance system. Current in a 50 Ohm system at 1.0 KW would be 4.47 A and voltage at 224 V. With a 1:100 turns ratio (assuming primary being a quarter turn), the impedance reflected to the primary would be 1/10,000 of the secondary circuit. A 200 Ohm secondary circuit would then appear as a series resistance magnitude of 20 mOhms in the primary winding, creating a VSWR of no more than 1.0004. The transformer structure can be measured at low power by itself, exclusive of the secondary's metering circuit. The secondary's metering circuit can be characterized and calibrated at a higher voltage, lower current.

Parallax Errors With Analog Meters

Analog meter dials are made and factory-calibrated assuming the viewer's line-of-sight is exactly perpendicular to the meter face. If the viewing line-of-sight is at a different angle the meter's needle may appear to be at a different position compared to the perpendicular or 90 view angle. The reason for using *mirror scales* (just the mirror, no markings) is for the meter reader to line up the meter needle exactly over its reflection. That insures an exact perpendicularity.

Power Measurement, DC to Low-Frequency AC

Power *output* at DC to LF can be measured simply by attaching a ***resistive*** load to an output and measuring the voltage drop across the *dummy load.*[7] AC to LF might be affected by the inductive reactance of wire-wound power resistors, the series inductance increasing the magnitude of the load impedance relative to measured DC resistance. A number of such dummy resistances

---

[5] Two meters were generally installed, one on each side of the balanced line, less for symmetry and more for quick determination if one side of the feedline was open or grounded.

[6] The line blurs between voltage and current in some applications. The point with the 260A being that such an applied power load is extremely broadband, going from LF up to the low end of VHF. A similar load and applied RF voltage meter circuit was used in the HF to UHF Boonton 190 Q Meter.

[7] Colloquial term, appearing to have been borrowed from RF use-measurement, so called because the resistors are a *dummy* or substitute for the actual loads.

may be required to characterize a power supply or audio amplifier output.

There are *electronic loads* which can be manually set to be almost any dummy load within such an instrument's range. While useful in a laboratory environment those are generally too expensive to consider for hobbyist design work.

Power *input* measurement (for some required output) is more tedious. This requires knowing the impedance at each AC frequency to be measured. It also requires that phase angle of the applied input voltage and current be known, then complex number arithmetic applied for impedance or admittance to find the actual input power. As a gross measurement of input power, a known power is applied through a variable attenuator (at least 20 db) and adjusted until the output power is of a required value. The approximate input power can then be calculated by comparing the effect on the attenuator's characteristic impedance versus the magnitude of the input impedance relative to what it should have been to fit the attenuator impedance.

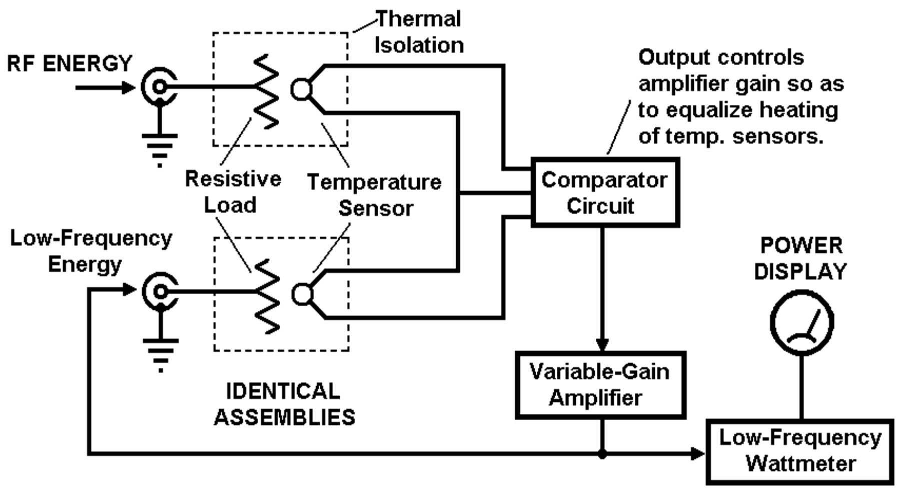


Figure 36-4 ***Substitution*** **measurement system of measuring RF power using identical DC-to-RF terminations, equalizing heating.**

Calorimetric Power Measurement at RF

If an RF power is terminated in a load that absorbs all of it, that power is converted to ***heat.*** An impedance or admittance measuring device can determine how much the load is actually absorbing by measuring the deviation from the system's characteristic impedance. The heat absorbed by the load can be measured by a variety of very sensitive resistance versus temperature devices: Thermistors; barretters, bolometers, and (amazingly enough) 1/100 A cartridge fuses.[8]

[8] Thermistors on the market in the new Millennium are designed for different purposes than RF power measurement and there are a great variety of them, available in either positive- or negative-going resistance coefficient versus temperature. Barretters are now rather rare but operate about the same. Bolometers are now mainly used for IR (Infra-Red, thus heat) measurement and imaging but are extremely sensitive in changing resistance with applied heat. 1/100th Ampere instrument fuses, in the proper RF-matching holder, function as less sensitive bolometers-barretters. It should be noted here that one does ***not*** check out 1/100th Ampere fuses in any common multimeter's Ohms function. The limit current of those is already 10 mA and some of the early analog multimeters would have more current than that on low Ohms ranges. Current in any set-up to check such low value fuses should be much less than their rated blow value.

Figure 36-4 shows a generic *substitution* method RF power measurement system. Two identical terminations are required, each with identical temperature sensors. Termination-loads are resistive with a bandwidth from DC to a frequency higher than that to be measured.

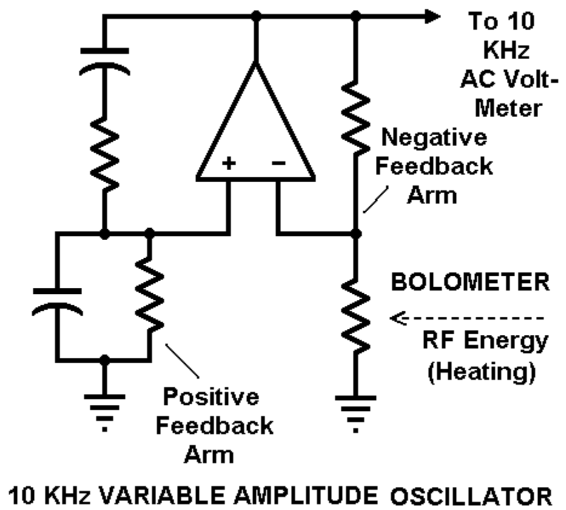


Figure 36-5 General form of self-balancing bolometer bridge in a HP 430A power meter. Same system used in model 432.

The comparator circuit output controls a variable gain amplifier to provide a low-frequency energy input to the bottom load assembly that results in the same heating of the load as does the RF energy into the top load assembly. Amplifier output may be anything from DC to some audio frequency, anything that can be readily converted to a display of power for the user.

The system *self-balances* such that each of the identical load assemblies are heated the same at balance. Using DC-to-UHF terminations, the heating from a low-frequency source will be that of an RF source.[9] Since the load assemblies are identical, they can be measured separately for both VSWR and low-frequency heating of the internal temperature sensor. The type of sensor is not critical so long as both are the ***same*** kind with as much equality as possible in resistance versus temperature characteristics. Those characteristics can be determined prior to assembly in a load and a matched pair selected from such separate testing.

Combining Self-Balancing With A Single Termination

The 1950 Hewlett-Packard 430A RF Power Meter set a standard for relatively easy-to-use RF power measurement in the 300 *Mc* to 4000 *Mc* frequency region.[10] With the proper frequency range bolometer mounts the 430A could directly measure power in five full-scale ranges: 100 µW, 300 µW, 1 mW, 3 mW, and 10 mW. The measurement system used a *self-balancing bridge amplifier* very similar to the Wein Bridge audio oscillator (model 200, the first large-selling instrument from HP).

---

[9] 50 Ohm terminations in the *DC-to-GHz* category generally use a tapered absorber with equal DC to RF characteristics, in contact between center and outer conductors of a coaxial assembly. The limit on the high frequency end of characteristics is due primarily to structural differences, center and/or outer conductor dimensional differences (from design) and the amount of taper (which influences distribution of heat in the absorber material). Upper-end limits vary from 1 to 4 GHz depending on model and maker.

[10] 300 MHz to 4 GHz. Scientific value multipliers had not been adopted in 1950. Information on the 430A is taken from Agilent (the test-and-measurement division of Hewlett-Packard) *HP Journal*, Volume 1, Number 9, May 1950. The more-common HP432A and HP478 thermistor mount full information can be found in military technical manual TM 9-6625-2469-15, dated December 1969.

The R-C values in the positive feedback arm cause an oscillation at about 10 KHz. The negative feedback arm values are selected to continually keep the bolometer resistance at 200 Ohms. When RF Energy is applied to the bolometer, it heats up and changes resistance. The oscillation amplitude then reduces in order to maintain the bolometer resistance at 200 Ohms. The oscillator-amplifier circuit thus outputs a 10 KHz amplitude that is inversely proportional to the applied RF energy going into the bolometer mount. The circuit *balances itself* to keep the bolometer resistance constant, with or without applied RF energy.

The end result of self-balancing is to have the 10 KHz amplitude decrease with an increase in RF energy input to the bolometer mount. This is easily compensated in the power input AC voltmeter following the bridge oscillator by having the voltmeter output read in reverse. With the bolometer always in circuit the bridge remains balanced even with no RF energy input to the bolometer.

Note: Applied RF is bypassed in the bolometer's output leads and does not affect the 10 KHz oscillator operation; the 10 KHz of the negative feedback arm does not interfere with applied RF energy to the bolometer mount. NB: Not shown in Figure 4 is a DC bias current applied to the bolometer to compensate for power input range switch changes; that bias is bolometer resistance dependent.

The only drawback to the model 430A system was the requirement to manually tune the bolometer mounts for maximum power input at any frequency change. Hewlett-Packard changed that after 1950, coming up with broadband, ***untuned*** bolometer and thermistor mounts. The self-balancing bridge was essentially the same. By the 1960s the improved, restyled-case model 430C power meter and model 477 thermistor mount were standard fixtures in most RF laboratories in the USA for VHF through microwaves power measurement.[11]

Calorimetric methods of RF power measurement have been in use for over a half century. The elegance of the method is that entirely separate measurements can confirm the amount of power actually absorbed by a power measuring sensor mount and that all the RF energy absorbed is converted to heat. Heat itself can both be measured and supplied as a stimulus to sensor mounts for mount and instrument calibration.

Indirect RF Power Output Measurement

This method uses a broadband resistive dummy load (almost always at a 50 Ohm characteristic impedance) with an attenuated voltage probe and diode rectifier to indicate the RF power equivalent absorbed by the load. A good example is the Bird model 43 *Thruline Wattmeter* popular with radio amateurs. However, neither the Bird model 43, nor similar models from Coaxial Dynamics or Telewave are wattmeters per se. They should be called *Directional Power Samplers.*

When used with a good 50 Ohm termination, all three will indicate RF power within ± 5 %.

[11] Hewlett-Packard undertook a complete modernization of their instrument enclosures after 1950 along with improvements to circuitry within those enclosures. Most-visible changes used cast-aluminum enclosure pieces replacing the old-style all-sheet-metal style of the 1940s. The 430C made its debut in 1955 along with its same-size enclosure companion the 415B VSWR meter (for use with slotted lines), 476A universal bolometer mount, 477A untuned thermistor mount plus some other popular instruments such as the same-width 200CD audio oscillator and huge, heavy 524B 10 MHz frequency counter. In another decade H-P would go through another enclosure look-and-feel change suited for rack installation of many instruments in one convenient place.

A cut-away view is shown in Figure 36-6.[12]

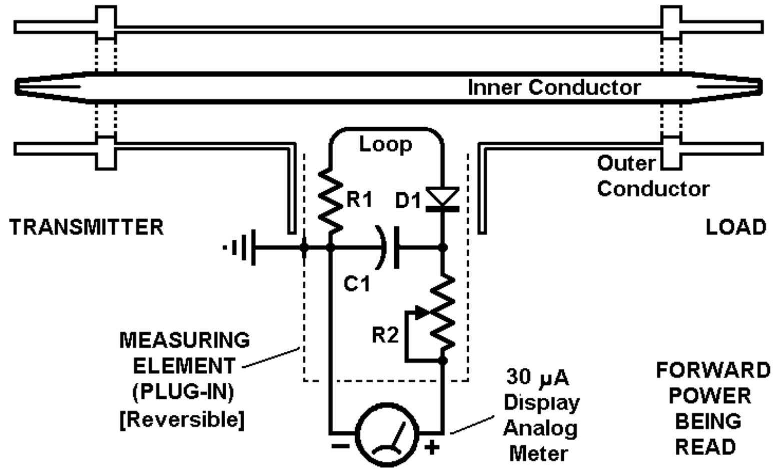


Figure 36-6 Generalized arrangement of a coaxial line power sampling wattmeter and reversible element.

The heart of the instrument is the plug-in sampling element colloquially called a *slug*. Each element has a small, approximately quarter-turn loop mounted at the end of the precision-machined element housing. The loop has both capacitive coupling to the inner conductor and inductive coupling to the magnetic field between inner and outer conductors. Without R1 (typically 68 Ohms or so), the loop current would have only the magnetic field induced current into the diode detector. With R1 the electrostatic field is also a contributor to diode input. In combination each element offers *directivity* of about 25 db between forward and reverse powers.[13] All such plug-in elements may be inserted for forward (incident) or reverse (reflected) power indication and each one is marked with a big arrow symbol for the power-reading direction. These sampling power detectors can be used to yield an indication of any load's VSWR.

D1 is a high-frequency diode in a proprietary, very-low-inductance mounting. C1 is a low-inductance bypass capacitor. R2 is a small potentiometer used to calibrate each element for the standard 30 µA analog meter in every main housing. It is not possible to access R2 without some disassembly of an element. Each element is designed to connect to the display meter whether inserted for forward or reverse power.

Since the grant of the Bird patent in 1958, a number of adaptations of the forward-reverse samplers have appeared, nearly all in amateur radio literature.[14] Most are used for VSWR indication of antennas, usually in conjunction with manual or automatic antenna tuners.

Direct RF Power Measurement Through Logarithmic Detector ICs

---

[12] Illustration adapted from Bird Electronics Corporation Model 43 instruction book, part number 920-43, Revision E, 2004, page 6. Bird Electronics US Patent 2,852,741 was granted 16 September 1958.

[13] *Directivity* is the power difference between coupling in one direction versus the opposite direction.

[14] The *Bruene detector-sampler* appears to be the first widely-used electrostatic and magnetic sampler, used in the T-195 transmitter of the AN/GRC-19 vehicular-transportable receiver-transmitter designed and made by Collins Radio for the USMC circa 1955. That detector-sampler was the heart of one of the first automatic antenna tuning systems for field radios anywhere. See also *Single Sideband Principles and Circuits* by Pappenfus, Bruene, and Schoenike, McGraw-Hill Book Co., 1964, Chapter 10.

Logarithmic detectors were first adapted to radar receivers in the 1950s, easing their requirements of providing a more level a radar return signal for display. A conventional AM detector will output a DC voltage that is proportional to the amplitude of the RF carrier. A logarithmic detector outputs a DC voltage proportional to the *logarithm* of the RF carrier. The *dynamic range* of the logarithmic proportionality may be in the 30 to 120 db span, maximum to minimum RF carrier input power.

While Texas Instruments was first to offer both logarithmic amplifiers and detectors with bandwidths up to low VHF, *Analog Devices* later produced a series of *log amps and detectors* with bandwidths from DC to 2.7 GHz. The Analog Devices ***AD8307*** has been popular with radio hobbyists as the heart of a *milliwatt meter* that can accurately read RF power levels in a 50 Ohm system from +17 dbm (50 mW) down to -75 dbm (32 pW) in either dbm or Watts directly, no scale changes required.[15] The AD8307, available in both DIP and SOIC packages, contains a succession of limiting detectors all summed to a single output pin. Detector characteristics are tailored such that the overall input versus output follows a logarithmic function of voltage within a typical ± 0.3 db tolerance to a true logarithmic function, ± 1.0 db worst-case tolerance. A microcontroller, Microchip PIC16F876 in Scherrer's project, does myriad tasks to convert the AD8307 DC output to dbm, control an LCD display panel, and even display Watts in parallel with dbm (through an internal routine of conversion), plus show a coarse bar graph of power input on the display.

Any one of several PIC microcontrollers can be used. A few other projects have done away with the microcontroller, using a digital voltmeter assembly with adjustable end-limits of output voltage such that the millivolts of log detector output correspond to tenths of dbms. The versatility of this type of RF power measurement is its 90 db dynamic range, considerably greater than an older calorimetric power meter limited to, at most, 30 db dynamic range.

Calibration of a milliwatt meter needs to be done by a known RF level attenuated by known attenuation values. The dynamic range can be calibrated by using a signal source with known *waveguide-below-cutoff*[16] adjustable attenuator in a substitution-method set-up with a standard waveguide-below-cutoff attenuator into a common receiver.[17]

---

[15] One of the best examples is that of Danish radio amateur Thomas Scherrer, OZ2CPU, and well illustrated on his website http://www.webx.dk/oz2cpu/radios/milliwatt.htm with photographs, drawings, and microcontroller software listings. Also on that website is a series of photographs of other hobbyists' duplications of Scherer's milliwatt meter. [web URL as of mid-2006] Several other project versions using the AD8307 can be found on the Internet.

[16] *Waveguide-below-cutoff* attenuators are common in high-end signal generators for linear-in-db/dbm output level adjustments. Frequency-independent, they are circular waveguides with the guide's lowest frequency at least ten times that of the upper frequency of the signal generator (hence the *below-cutoff* part of their name). These have the unique property of the physical spacing between coupling loops being exactly proportional to the attenuation in decibels. Such attenuators have a minimum attenuation of 10 to 20 db but have a dynamic range of 100 db. They are not recommended as projects for hobbyists due to construction requiring fine machining and a design necessary to insure minimum RF leakage (both in and out) at maximum attenuation.

[17] *Weinschel Engineering* has made several below-cutoff attenuator standards, such as the model PA-2 calibrated at 30 MHz with accuracy of ± 0.001 db / 10 db, ± 0.0005 db, over a 100 db range of attenuation. Narda, PRD, and Ailtech have made similar below-cutoff attenuators. The statement *calibrated at 30 MHz* refers more to the NIST calibration method, NIST using an uncommonly-sensitive 30 MHz receiver to resolve very tiny power input differences.

Once the calibration tools are available it should be possible to include correction tables in an internal microcontroller's flash memory to compensate for inaccuracies of the DC output voltage logarithmic function versus known RF input levels. Scherrer's project does something like that in some compensation of input frequency versus level relative to input VSWR of his meter.

Waveguide-Below-Cutoff-Attenuators as a Standard RF Attenuation

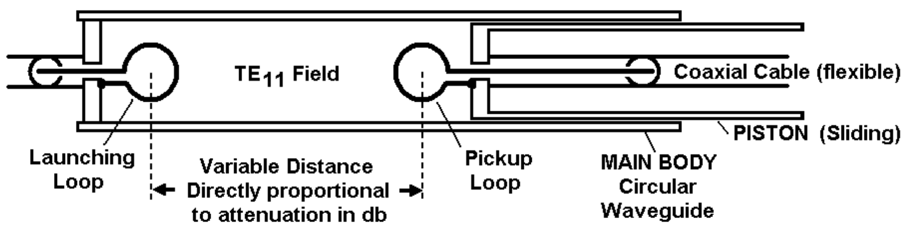


Figure 36-7 Idealized cut-away of a waveguide-below-cutoff (*piston*) attenuator. Loops must be on the same plane.

One of the vexing problems in early radio was establishing the ***low-power*** accuracy of any RF source, especially for quantifying the sensitivity of receivers. In 1935 D. E. Harnet and N. P. Case first published (in the Proceedings of the IRE, later to become IEEE) the first paper on using a waveguide well below its low-frequency limit at a broadband (at low frequencies) adjustable attenuator. What was remarkable was that the *linear position* of the *piston* (adjustable part of attenuator) was directly proportional to the attenuation in decibels! Adapted for all high-end signal sources, RF output of the sources could be calibrated at a high output level with lower output powers dependent only on the manufacturing-machining tolerances of the sliding piston drive.

A general cut-away diagram of a piston attenuator is shown at left. The piston assembly does not have to be in contact with the circular waveguide body but should have a close, sliding fit to it. The main body should be circular and hold its inside diameter as close as possible. One formula for attenuation is:[18]

$$\alpha_{(\text{db / cm})} = \frac{15.99 \bullet \sqrt{1 - (f \,/\, f_C)^2}}{r_{(CM)}} \qquad \text{Where:}$$

α = Attenuation constant in decibels per cm of distance

f = Frequency of attenuation

f_C = Cutoff frequency of circular waveguide (same units as f)

r = Radius of inside diameter of waveguide in cm.

[18] From *On 30 MHz TE$_{11}$ Mode Piston Attenuator,* by R. Swarup, J. R. Anand, P. S. Negi, Review of Scientific Instruments, Volume 72, #3, March 2001. This formula presented because it is the simplest compared to other formulas in older textbooks.

It should be noted that the numerical value of terms inside the square-root sign are very close to unity. If the cutoff frequency of the circular waveguide is 3 GHz and the attenuation frequency is 30 MHz, that numeric value is 0.01; squaring that and subtracting it from 1.0 yields 0.9999 and the square-root of that is 0.999 950. Such has no appreciable effect on holding a ± 0.01 db accuracy from audio frequencies on up to 100 MHz using a large-radius (low cutoff frequency) circular waveguide.

The TE_{11} *Field* as indicated in Figure 36-7 is the *dominant mode* in circular waveguide and describes the electromagnetic wavefront inside the guide. Voltage lines of force will be *transverse* (going perpendicular to) the inside walls; magnetic lines of force will be at right angles to voltage lines. For that reason the input and output coupling loops must be oriented the same way, remaining so for the entire travel of the piston. There doesn't seem to be much explicit information on the construction of the loops relative to some attenuation frequency range.[19] That and impedance matching to the input-output coaxial lines is something that has to be worked out by hobbyists on a case-by-case basis. The author has seen that some *WBCO* (Waveguide Below Cut-Off) attenuators have discs of absorber material between loops and their mounting plates; that may be more as a device to reduce coaxial cable coupling discontinuities over the attenuation frequency range than any protection from other modes.

Waveguide modes are many and depend on the shape and position of the launching device (as well as field container) for their existence. Fortunately for WBCO attenuators, only one other mode can reduce accuracy, the TM_{10} characterized by magnetic lines of force in circles on the same plane as a cross-section of the guide, voltage lines of force going into the interior circumference from the guide center-line. TM_{10} mode propagation is much less than TE_{11} but affects large attenuation spacings the most. In the referenced paper, a *mode filter* was installed close to the launching loop. This mode filter consisted of several thin parallel, separated wires, the plane of this pseudo-Faraday screen perpendicular to the guide's axis.

All WBCO attenuators have a minimum attenuation of 10 to 15 db due to closeness of the loops. Maximum (accurate) attenuation is a function of signal leakage around the input and output connectors, including the woven outer conductor of the flexible coaxial cable. The semi-armored type of coaxial cable should be used, the kind having an additional copper weaving over the normal outer conductor. If used, coaxial connectors should be of the screw-together type for minimum leakage. The piston assembly does not have to be in direct contact with the circular waveguide but a waveguide length at least twice that of the piston travel will reduce leakage through the narrow space between piston and inside of the guide. In practice, on laboratory-quality signal generators, there is usually a large spacing between the output attenuator RF source and the front panel output connector. That spacing helps minimize any RF leakage to devices under test.

Calibration of Large Attenuation Values

High-end RF signal sources now use *step attenuators* for larger values of attenuation (1.0 db and greater, usually in a binary progression of 1-2-4-8- etc). Fixed attenuators with

[19] It is certain to exist, but as very proprietary information within manufacturer's design labs. Part of that is for protection from competition and part is from a lot of man-hours spent in achieving proper loops along with their wideband matching networks.

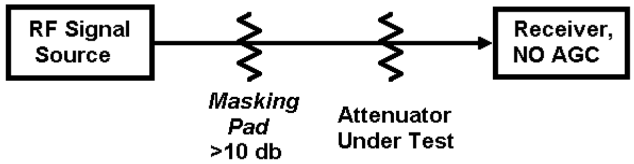


Figure 36-8 Coarse-to-medium accuracy setup for measuring fixed attenuators.

electromagnetic or PIN diode switching under microprocessor control allow front panel manual control or under a predefined stored-program control. Such fixed attenuators are of the precision variety, sometimes laser-trimmed to exact values via an automated system. Electronic attenuation of the RF source generally handles attenuation under 1 db.

A reason for precise trimming of fixed attenuators is the error build-up possible with off-the-shelf coaxial attenuators. Accuracy of individual fixed coaxial attenuators is on the order of ±1 db per 20 db.[20]

To get 120 db total, six 20 db fixed attenuators would have to be connected together. However, that seriesing also results in an aggregate ± 6 db uncertainty!. Individual attenuator errors also add. One can get by using the old square-root of the sums of the squares of individual errors but that would still result in an error of ± 3 db. Quantifiable receiver testing for sensitivity that way would not be accurate with a ± 3 db error.

A simple comparison setup for *pad*[21] measurements is shown at left. The receiver must not have any automatic gain control in order to show the greatest differences in input signal level. The attenuator under test can be replaced with of known accuracy in the beginning, then those to be tested compared in level with that or, if the signal source has a calibrated WBCO attenuator, used to make the receiver level the same; the attenuation under test is then the difference between the WBCO attenuator readings. The *Masking Pad* is used solely to keep the VSWR at a minimum in a 50 Ohm system. Another can be used between the attenuator under test and the receiver to make sure the receiver input impedance does not change the characteristic impedance of the system.

Once some fixed pads are known, they can be used in series with a waveguide below cutoff attenuator to ascertain where the linear-with-db begins and continues. Note: All WBCO attenuators will have non-linearity region between 10 to 20 db total attenuation (10 db insertion loss is common); that can be calibrated also but the linear-in-db-versus-position, once found, can be mechanically marked using a machinist's rule or micrometer or fine-position caliper in order to calibrate out to the maximum attenuation.

Why NIST and Other Calibration Agencies Use 30 MHz in Testing WBCOs

That is primarily for the ultra-sensitive to small signal level differences required to be accurate enough for federal agency standards. Those can resolve an amazing 0.001 db difference at 130 db down! That requires elaborate frequency-locked receivers fixed-tuned to a single

[20] In the 1960s the general rule was *one db in ten* accuracy. While that has improved over the years, the fine print in most fixed attenuator specifications has dropped to *one in twenty* for the DC-to-GHz variety. While those can be calibrated at audio frequencies quite accurately, internal reactance can result is more error at higher RF. Expensive fixed attenuators with better accuracy are available but seldom within a hobbyist's budget.

[21] *Pad* is colloquial for a fixed attenuator, usually prefaced by the attenuation in db.

frequency source. The accuracy over a wider frequency range can be determined from physical laws governing the TE_{11} mode propagation versus attenuation frequency at any given attenuation setting. At NIST the calibration service is better than ± 0.003 db per 10 db attenuation.[22]

Resistance Measurement Using Multimeters

The first multimeters measured and external resistance by means of a fixed voltage from a 1.5 VDC dry cell with the analog meter indicating the current. Such resistance readings were non-linear and depended on good markings on the meter face. Accuracy was poor with external resistors' value towards the high end of the scale. More modern digital multimeters use a constant-current source excitation and read the voltage drop across the external resistor. That results in a DC voltage directly proportional to the external resistor value. Such accuracies are as good as the constant current and the internal millivoltmeter calibrations.

With the advent of transistors and the 0.6 V base-emitter junction potential, the old multimeter's Ohms function excitation by a single dry cell was hazardous to the transistor as well as inaccurate due to the junction characteristics. DVMs have limited-compliance[23] constant-current sources with a much lower voltage potential. The actual voltage used depends on the make and model of the DVM but can be measured with another DVM.

A few dealers and distributors sell resistors with ± 1.0 %, ± 0.5 % and even ± 0.1 % tolerances. These are more costly than the common ± 5 % and ± 10 % variety but a few are useful to have on hand to recheck DVM resistance ranges. Those can also be used in ***bridges*** for more-accurate measurement, the main subject of the next chapter which includes measurements at AC and low frequencies.

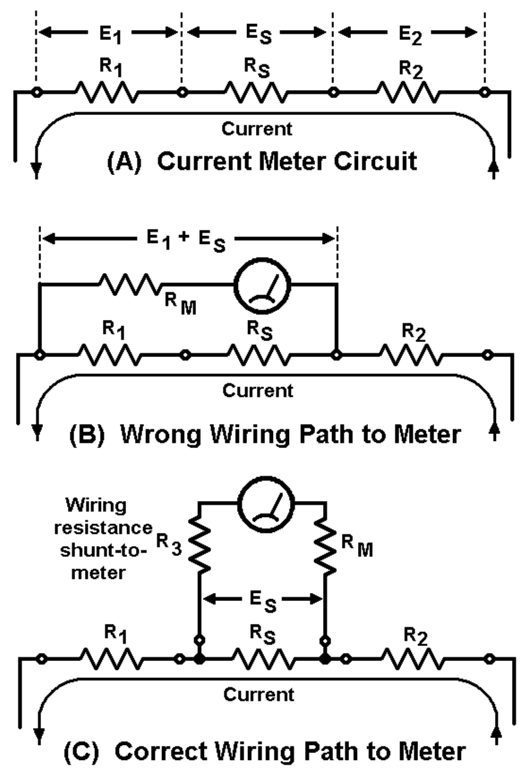


Figure 36-9 Current metering circuit showing wiring resistance effects.

Kelvin-Varley Connections

Kelvin-Varley connection is a misnomer. Just the same, it refers to measurement connections of a low resistance as close to, or the same, as calibration connections as shown in the figure below. This is particularly true for shunts across an analog meter of small full-scale current to make it read at larger currents. In Figure 36-9 R_S would be the shunt and the

---

[22] Exact details at NIST are found in http://ts.nist.gov/ts/htdocs/230/233/calibrations/Electromagnetic/RF-microwave.htm#61300C or follow the links at http://www.nist.gov.

[23] *Compliance* refers to the voltage limits of a constant-current source, minimum to maximum.

other two resistances being that of the connecting wires. A calibration measurement of R_S ***cannot include*** all or parts of either R_1 or R_2 without being in serious error.

As an example, assume a 1 mA meter with a 200 Ohm meter motor (R_M) will be used to indicate a full-scale current of 1.0 A. From the formulas on page 37-4 the shunt resistance would have to be 200.2002 mOhm for a 200 mV drop at 0.999 A. It should be intuitive that just 20 mOhms of wire resistance (R_1 in Figure 36-9B) would cause a full-scale current reading to be off by 10 percent.

A *Kelvin-Varley* or, sometimes just *Kelvin connection* refers to the desired resistance having the low-current connections very close to the main current flow in the circuit. If the low-current path has 20 mOhms extra resistance (R_3 in Figure 9C), but the connections to the example shunt were right next to R_S, the error would be only 0.1 %. The meter circuit itself takes only 0.1 percent of the total current. Most high-precision resistors have ***two*** connections on each end or an extra wire on each end of pigtailed resistors. The larger at each end would be for the main current path. The smaller would be for the voltage drop measured by the displaying circuit.

OSCILLOGRAPHY

General

Oscilloscopes are of enormous value from design breadboarding through design evaluation through troubleshooting. Their *vertical deflection* can show DC to video range voltages from mV on up. Their *horizontal deflection* has a precision internal sweep circuit selectable from Seconds per graticule division to nanoseconds per division, triggered either internally from vertical deflection or externally from some signal source. The bandwidth of vertical deflection extends from DC on up to 20, 40, 60, or 100 MHz depending on the make, model, and price. Vertical inputs of two independent channels may be overlapped, on alternate sweeps, or chopped by an internal switcher. Oscilloscope ***probes*** may be used for 10:1 attenuation and less loading effect of the 'scope vertical input impedance on the circuit being probed. Horizontal deflection *sweep* is produced by a precision *time base* generator. More up-scale models may include an extra sweep called *delaying sweep* that allows both a time delay of the trigger plus some magnification.

Analog oscilloscopes have direct voltage amplification into a CRT or Cathode Ray Tube, both deflection directions. Digital oscilloscopes, sometimes called *DSOs* (Digital Sampling Oscilloscope) do sampling of the horizontal input voltage waveform, convert the voltage samples to a binary value at specific points in time, and store the binary value words for slower-speed display on an LCD panel.

DSOs are more expensive than analog 'scopes but convenient for direct digital printer output or personal computer storage, plus including vertical and horizontal range settings. Analog 'scopes require film photography to record oscilloscope *traces*.[24]

---

[24] The author and several others have had success in recording analog oscilloscope images using an ordinary digital camera capable of focusing at shorter distances. The digital image can be stored on a PC. This is a bit better than the older analog 'scope recording using expensive *Polaroid* film camera. The LCD display of the image on the digital camera is big enough to show the proper distance of magnifying glass in front of the camera lens.

Vertical Deflection Bandwidth

One of the major specifications of oscilloscopes, the vertical deflection can be defined either by frequency (to an approximate response roll-off) or by ***rise-time***, the ability to display the transition of a step waveform. For the more-common analog oscilloscope rise-time and bandwidth are related by:[25]

$$\text{Rise - Time}_{(\text{SECONDS})} = \frac{0.35}{\text{Bandwidth}_{(\text{HERTZ})}}$$

The rise-time is measured between 10% and 90% of displayed voltage of an infinitely-fast transition step applied to the vertical input. A *60 MHz Bandwidth* analog oscilloscope could resolve step transitions to 5.8 nSec. Note: *Rise-time* also applies to *fall-time*, opposite transition direction.

DSOs have slightly faster rise-time resolution (the 0.35 in the analog formula replaced by 0.4 to 0.5) but the time position of the step suffers from slight *aliasing*[26] in the sampling process. This is generally not of great concern. A waveform recording will show this aliasing as if the step display was slightly fuzzy in appearance.

Bandwidth of oscilloscopes is now stated for DC up to the specified bandwidth frequency.[27] Nearly all oscilloscopes now have an *AC* input selection by front panel switch, that selection introducing a DC-blocking capacitor at the input. This resulting low-frequency cut-off varies among different makes and models of ‘scopes but seems to average out at about 10 to 20 Hz.

Many of the moderate-cost modern oscilloscopes include a stable reference voltage (usually a squarewave) on front panels. This is most useful in periodic checking of gain settings of vertical input as well as periodic checking of vertical calibration.

For hobby work the author recommends a bandwidth of 40 to 60 MHz (depending on the hobbyist budget). This allows more-accurate viewing of digital waveform rise times and ability to resolve very short spikes in digital circuitry, spikes generally due to digital time-delay problems in gating and multiplex circuits. A delaying sweep is convenient but not absolutely necessary. Cost is proportional to bandwidth and the higher vertical bandwidth is best.

Why Modern Oscilloscopes Have *Calibrated* Input Capacitance

Modern oscilloscopes have calibrated vertical deflection per graticule range settings. The

---

[25] Agilent application note 1420, 1 December 2002.

[26] *Aliasing* is a complex subject in itself and is concerned only with sampling systems. In oscillography it is the inability to correctly display the original from time samples or the creation of artifacts of the display that were not in the original waveform.

[27] The introduction of the *Heathkit O-1* oscilloscope (one of Heath Corporations’ first electronics kits) in the early 1950s was typical of low-cost, AC-coupled ‘scopes with a bandwidth of only 100 KHz! A decade-plus later there were relatively low-cost oscilloscope kits which could reach 5 MHz in bandwidth. *Tektronix*, formed just after the end of World War II, improved oscillography industry-wide with their 511-series of precision time-base sweep models, followed by the modularized vertical input option series of 530s and 540s, the latter bringing vertical bandwidth up to 35 MHz.

input resistance remains fixed at 1.0 MOhm, in parallel with a small capacitance, regardless of the range setting. That small capacitance is important for its vertical deflection bandwidth.

The *small capacitance* is better shown in Figure 36-10 as ***KC.*** In parallel with the DC resistance of ***R***, it can form a voltage divider with ***KR*** and ***C*** (in the test probe). Provided there is little or no inductance in the circuit, the multiplier of ***K*** forms the usual 10-times divider with ***K = 10 - 1 = 9.***

Achieving a constant divider over the whole vertical deflection bandwidth requires the matching given in Figure 36-10. The squarewave test waveforms to the right show gross mis-values of too little or too much trimmer capacitance. Most oscilloscopes of medium-to-high cost now have such squarewave sources on the front panel.

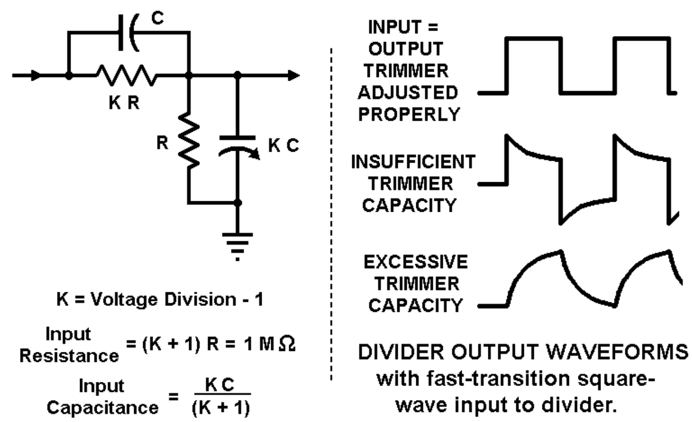


Figure 36-10 Wideband vertical deflection range voltage dividers compensating for stray capacity of vertical amplifier inputs.

The major reason for standardizing input capacitance was to allow interchangeability of *attenuating probes,* at least those of the oscilloscope manufacturer. Parallel input resistance and capacitance varies with different makers. In the 1956-debut Tektronix 530 series oscilloscopes the input resistance was 1.0 MOhm with 47 pFd parallel input capacitance. With the 2004-debut Protek 6500 series oscilloscopes the input resistance was held at 1.0 MOhm, ± 2 %, with a lower parallel input capacitance of 25 ± 3 pFd.

The *Ten-to-One* 'Scope Probe

A 1 MOhm input resistance is high but not high enough for some higher-frequency circuits. It would be better if that were ***10 MOhms*** and the parallel capacitors lowered as much as possible. Doing that required a 10:1 voltage division. The same general circuit of Figure 36-10 was used for the probe except that the capacitance of probe-to-oscilloscope cable was part of the trimmable capacity and ***R*** is 10 MOhms, not 1 MOhm.

Oscilloscope Probe Loading (10 MOhms || 12.5 pFd)

Frequency	Impedance, Ohms, Polar Form
100 Hz	9.9693 Meg / -4.4908
1 KHz	7.8644 Meg / -38.146
10 KHz	1.2630 Meg / -82.744
100 KHz	127.31 Kilo / -89.271
1 MHz	12.732 Kilo / -89.927
10 MHz	1.2732 Kilo / -89.993
100 MHz	0.1273 Kilo / -89.999

As the table indicates, the impedance of a parallel R and C rapidly becomes capacitive at the high audio frequencies and above. It should always be kept in mind that the probe's input capacity is finite. As little as 12 pFd will definitely detune high-impedance resonant circuits.

Oscilloscope Accuracy

You cannot expect more accuracy than stated in manufacturer's specifications. At best the vertical deflection accuracy will be roughly the same as a 2 ½ digit digital voltmeter. Horizontal deflection accuracy in time per graticule division is about the same. For more-accurate measurements of a waveform period or frequency an external frequency or period counter is necessary. Some 'scopes have an extra connector and cathode-/emitter-follower circuit for that, usually on the back panel. For home workshop calibration, that can be done via procedures specified by the manufacturer. If not included in the operating manual those are available as extra documents; contact the manufacturers if you can't find it in catalog listings.

Miscellaneous Equipment

Dippers

The name refers to a handheld variable oscillator with a meter or other display device to show disturbance of the oscillator when it is coupled to a resonant circuit. As with the original ***grid-dip*** meters, the oscillator meter would show a decrease in reading at resonance. Dippers have been made with vacuum tubes, transistors, and tunnel diodes as the active elements.

While dippers can cover (variously) 0.5 MHz to 200 MHz with their plug-in coils, their frequency calibration and stability under load are worse than ± 10 %. The worst observed offense in use is placing the dipper's coil too close to an external resonant circuit. *Loose coupling* is mandatory to hold frequency accuracy. Too close or *tight coupling* causes frequency detuning.

In the author's opinion, Dippers of any kind are poor investments for a home workshop.

Ordinary Thermometry

This refers to very ordinary thermometers found nearly everywhere: From supermarkets to Do-It-Yourself stores to Hardware stores. Given their limited read-out, those generally ***are*** a good investment.[28] The main purpose is to find *differential* temperature, the change between room temperature and powered down to the active state and producing heat. Such differential temperature readings are good for determining oscillator frequency drift when applied to parts (such as capacitors) which have been carefully measured at room temperature. If the dielectric constant of capacitors is known, the change can be found from the difference between *at-rest, room-temperature,* to *hot, heated* frequency which is then squared to find the capacitance change.

Different in both cost and use, heat-sensitive *paint* such as the products of ***Tempil*** are good for measuring hot spots around power transistors and similar. All that is required is to know the

[28] Simple thermometers can be from $5 at Supermarkets to a couple hundred dollars for non-contact infra-red sensing units.

temperature at which a color change occurs, then use that. It might take at least three different temperature grades of paint to use them even near-accurately but they are excellent for being in small areas.

Appendix 36-1

RMS - Root Mean Square

RMS is appended to many AC voltage and current values, yet its meaning is hardly ever defined. It is a statistical technique for describing an AC waveform characteristic such that the resulting power dissipation from AC is ***the same as if the voltage or current was DC.*** Since such dissipation results in ***heat*** for either AC or DC, the ***heat can be measured and compared to DC***. The name comes from taking the ***root*** of the ***mean*** value of ***squared*** periodic samples, thus abbreviated ***RMS.*** In simplified arithmetic form it is:[29]

$$e_{[RMS]} = \sqrt{\frac{e_1^{\,2} + e_2^{\,2} + e_3^{\,2} + e_4^{\,2} + e_5^{\,2} + \ \dots\ e_N^{\,2}}{N}} \qquad \text{and}$$

$$i_{[RMS]} = \sqrt{\frac{i_1^{\,2} + i_2^{\,2} + i_3^{\,2} + i_4^{\,2} + i_5^{\,2} + \ \dots\ i_N^{\,2}}{N}}$$

Where: e_n = Instantaneous voltage at point n along a period.
i_n = Instantaneous current at point n along a period
N = Number of points of instantaneous voltage or current in one period, all points being at equal time increments in that period.

Note: The numerator under the square-root sign could be of any length in time provided that the points of time are evenly-spaced. As an example for a sinewave whose peak value is exactly 1.0, one can take 5 increments to five places from a Sine table between 0 and 90 , square each one, add the squares, then divide the sum by 19 (number of increments of angle) and get exactly 0.50000. Taking the square-root of that yields 0.70711 which, not surprisingly, is the RMS value of a sinewave having ± 1.0 Volt peaks.[30]

If several different waveforms, all symmetrical, are examined mathematically, there will be another AC definition: ***Crest Factor.*** Crest Factor is the peak voltage divided by the RMS voltage (or peak current divided by RMS current). The following table shows Crest Factors for various waveforms:

[29] The *correct* mathematical format is not used to show the summation as it really is without the math shorthand and to illustrate better how the formula could be applied in programming instructions.

[30] Squaring a negative number results in a positive quantity. Since a sinewave is symmetrical, it will be numerically accurate to use just one-quarter of a period. That is considerably easier than repeating the 90 calculation three more times. Note: That is accurate only for a perfect, symmetric waveform such as a sinusoid.

Waveform	Crest Factor
Sinusoid or Full-Wave Rectifier without filter*	1.41421
Symmetric Square-Wave	1.00000
Symmetric Triangle Wave	1.73205
Half-wave rectifier output without filter*	2.00000
Pulse of 10% duty-cycle	3.0
Pulse of 1% duty-cycle	10
Pulse of 0.1% duty-cycle	30

* Sinewave input to rectifiers

It should be obvious that different waveforms will have rather different RMS values if all have the same peak values. With random or even pseudo-random noise, the waveform is impossible to state numerically (it has no period, is always varying randomly) yet it is capable of being quantified by a ***True RMS voltmeter***. Achieving such a *True RMS* condition over a reasonably-wide bandwidth rules out conventional rectifier circuits. A single diode or half-wave rectifier requires a low-frequency (large capacitance value) output filter and it will quite likely output a DC voltage equal to the ***peak*** of the input. But, the 2:1 difference in Crest Factors of a half-wave rectifier versus a symmetric squarewave would result in the apparent RMS (from rectified input voltage) varying the same amount. Substituting a full-wave rectifier is of no help, resulting in a shift of the DC output equivalent to RMS but still having an almost equal error depending on input waveshape.[31]

Calorimetric Measurement Method

Going to *first principles*, of using dissipated energy (as heat) to equate AC RMS voltage or current is still the simplest as well as most accurate means to measure AC RMS. The principle is almost identical to RF power measurement described in this chapter's main body, that of using a self-balancing comparison circuit. Figure 36-11 shows the basic block diagram of this RMS-to-DC converter.

The input signal is amplified by a wideband video stage which can have optional range-setting attenuators. Output to R1 of the left-hand heater and sensor assembly results in the amplified

[31] US Patent Number 6,856,185, granted 15 February 2005 to Henry Herbert Sully claims that a conventional full-wave rectifier with two series output resistors and a capacitor across only one of the resistors, output taken from the resistive voltage divider and capacitor connection, will handle *any waveform* input and still yield an approximate RMS equivalent of input as DC output. The author did some slight simulation of the circuit given in this patent grant but did not achieve a reasonable DC output with all waveshapes over a bandwidth greater than an octave. The author's analysis was not complete and others might spend more time on it to ascertain its validity. In the author's opinion this might be applicable to narrowband RMS-to-DC converters but does not have the accuracy required for calibration purposes.

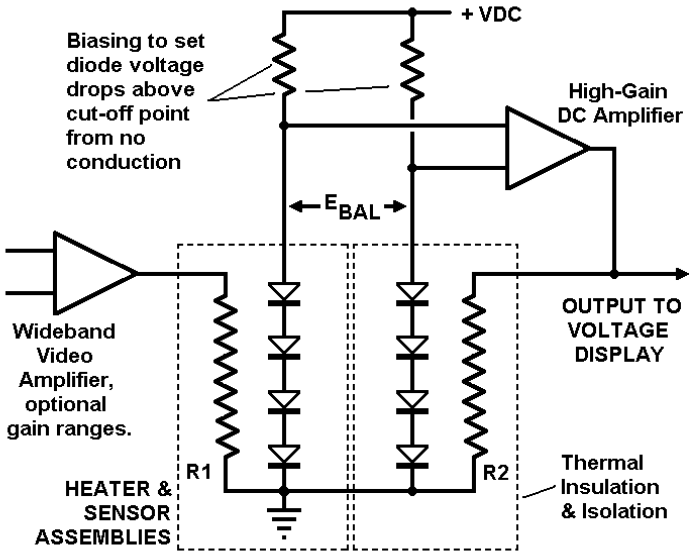


Figure 36-11 Thermal RMS-to-DC converter, self-balancing when heater-sensors are identical.

voltage dissipating all output as heat.[32]

The diode string is forward-biased so as to make the DC voltage (to the left of E_{BAL}) positive and on the low side of the diodes' forward-conduction voltage curve. Both diode strings have the same value of bias.

The high-gain DC amplifier outputs a voltage sufficient to heat R2 when there is an imbalance at E_{BAL}. If both heater-sensor assemblies are identical, the energy dissipated in R2 will equal that of R1 and the DC amplifier output voltage will be proportional to the RMS value of the AC input waveform.

The problem in designing such a relatively-simple circuit lies in the heater-and-sensor assemblies. Interior thermal masses must be light to provide some quickness of response. Thermal insulation needs a balance between heating-up and cooling-down the sensors; too much insulation and the heat from high input levels will not dissipate rapidly when input level drops. The assemblies must be thermally isolated from each other but also be in close proximity to avoid differences in their external thermal environment. Linear Technology Corporation once made such a dual heater-sensor in their part number LT1088, good from DC to slightly higher than 100 MHz.[33]

The use of multiple diodes as sensors is based on a diode junction temperature characteristic of about -2 mV per degree Celsius at about 0.6 VDC forward voltage. Using four diodes yields about four times that change.[34] There is no worry about diode AC characteristics since the diodes can be AC-shielded and do not pass any AC; in this configuration they react only to heat. Sensitivity in achieving a balance depends on the DC amplifier gain and the characteristic curves of forward

[32] From EDN magazine 11 May 2000, pp 55-57 and Linear Technology application note AN-22, September 1987, revised slightly by the author. Note the theoretical identity to Figure 36-4 in chapter main part. Also Linear Technology application note AN-61, August 1994, pp 16-18 and 28-38.

[33] From about 1986 to at least 2000 (footnote 5) but it was a discontinued item by mid-2006. Three other bench multimeters, HP3400A, HP3403C, and Fluke 8920A all use thermocouple junctions as sensors with resistive heaters. Those three are also discontinued instruments as of mid-2006. Note: The trend in bench multimeters has been towards combination digital-logarithmic conversion methods; Sampling and ADC techniques allow rapid true RMS calculation a la the mathematical form on previous page.

[34] The number of diodes was made 4 solely for the illustration. The only limit on the number of diodes is the DC bias voltage and the fact that each diode will drop about 0.6 Volts. Ten of them would drop about 6 Volts and a +9 to +12 VDC supply rail could handle that very well.

conduction of each diode string being as near-equal as possible.[35]

A video amplifier is shown in Figure 36-11 but is not essential to basic operation. Its use is to provide gain for lower-level AC signal input well below the heating levels or R1 and R2. That amplifier should be reasonably amplitude-linear. Non-linearity will introduce errors in the converted RMS-to-DC output. Overall bandwidth will be limited by reactances in R1 reducing its HF magnitude relative to DC.[36]

As to calibration, that can begin at DC, power applied directly into R1. Since DC can be measured quite accurately, the dissipated heat's effect on the sensor diode's forward conduction can be seen and characterized accurately. If both heater-sensor assemblies are identical, each can be measured separately and the pair's characteristics compared, DC R1 and R2 input versus DC voltage into the DC amplifier.

A perfectly symmetrical squarewave has a Crest Factor of unity so the RMS value would equal peak amplitude. That could be one way to check calibration dynamically. There should be no overshoot or undershoot and the transition edges should be as fast as possible.

Reactances in R1 can be checked separately by any one of several impedance-measuring instruments or a bridge measuring both R and X, in parallel or series equivalent. The video amplifier bandwidth can be checked independently by both oscilloscope observation and a peak-reading AC voltmeter. Depending on the output impedance of the video amplifier, its gain can be checked with R1 present as its load..

Crest Factors Versus Peak Values

From the definition, peak voltage or peak current is the RMS value times the Crest Factor. This will define the needs for dynamic range in any AC-RF input amplifier. The now-discontinued Linear Technology LT1088 has a specified maximum Crest Factor of 50, but that is only the dual heater-sensor; any signal-conditioning circuit ahead of it would be required to have a dynamic range at least 14 db greater than the output RMS value for a Crest Factor of 5. Note: Random noise is generally presumed to have a maximum Crest Factor of 4 to 5; any amplifier handling both noise and signals, for measurement or not, should have a dynamic range of 14 db to handle peaks.

Some Clarity on *Effective Values* of Low Duty-Cycle Pulses

The Crest Factors on the second page of this Appendix for pulses assume AC coupling. That is, the baseline is slightly below the zero horizontal axis. For DC coupling with very low duty-cycles, the *average* values can be taken to get the effective heating effect. *Average* voltage or current assumes DC coupling, baseline between pulses being zero. Just sum the time intervals on the pulse itself and divide by the total number of time intervals in a period.

---

[35] Thermocouples with resistive heaters were used in the HP3400A (an analog meter readout instrument), HP3403C, and Fluke 8920A bench multimeter. A comparison chart of accuracies with a random noise source as generators is included in the EDN article mentioned.

[36] This is the same limitation of the HF-VHF-microwave power meter, providing a non-reactive DC load for the incoming RF energy to preserve bandwidth at a specified accuracy of power measurement.

Chapter 37

Metrology - Advanced

A continuation of the preceding Chapter, this one covers more advanced measurement apparatus, including Bridges, Q-Meters, Signal Generators as well as a few others.

Bridges

One of the early, innovative devices, principally for metrology, was the *Wheatstone Bridge*.[1] Given that Georg Simon Ohm's *Law* was published just 16 years prior (and communications at that time was rather slow), it is a remarkable feat.[2] The basic Wheatstone Bridge of Figure 1 is a ***null-balance*** instrument. One manually adjusts the arm resistances until the detector indicates nothing.

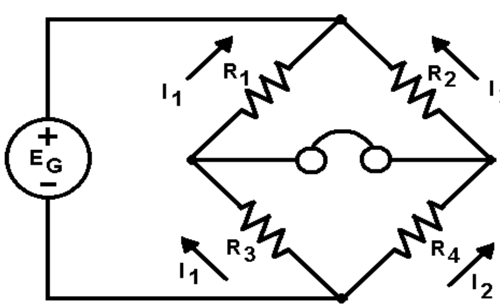


Figure 37-1 A basic Wheatstone Bridge for measuring resistances. The stylized headphones at the center denote a *detector* whether headphones, galvanometer, DC µVoltmeter. A DC or AC source (at left) provides excitation.

If the voltages across each resistor have sub-scripts equal to the resistor subscripts, then they can be in simple formulas:

$$E_1 = I_1 R_1 \qquad E_2 = I_2 R_2$$
$$E_3 = I_1 R_3 \qquad E_4 = I_2 R_4$$

Note the two currents in Figure 1. Those presume that the detector has no resistance. At null the DC voltages at each end of the detector will be the same magnitude and polarity so there is no current through the detector. At null:

$$E_1 = E_2 = I_1 R_1 = I_2 R_2 \quad \text{and}$$
$$E_3 = E_4 = I_1 R_3 = I_2 R_4$$

By some algebraic manipulation of the null formulas we can get:

---

[1] It wasn't invented by Charles Wheatstone but by S. Hunter Christie, some time on or before 1843. Wheatstone graciously gave Christie credit for it in an 1843 presentation. Those who heard the presentation apparently thought more of Wheatstone that Christie and, in spreading the information about it, christened it the *Wheatstone Bridge.* The diamond shape of the bridge drawing was Wheatstone's idea and that may have contributed to identifying his name with Christie's innovation.

[2] Ohm published his *Law* in an 1826 paper but it's inclusion in his larger 1827 book would probably be more widely disseminated.

$$I_1 = \frac{I_2 R_2}{R_1} = \frac{I_2 R_4}{R_3} \quad \text{or} \quad \frac{R_2}{R_1} = \frac{R_4}{R_3} \quad \text{and}$$

$$I_2 = \frac{I_1 R_1}{R_2} = \frac{I_1 R_3}{R_4} \quad \text{or} \quad \frac{R_1}{R_2} = \frac{R_3}{R_4}$$

From either there result:

$$R_4 = \frac{R_2 R_3}{R_1} \quad R_3 = \frac{R_1 R_4}{R_2} \quad R_2 = \frac{R_1 R_4}{R_3} \quad R_1 = \frac{R_2 R_3}{R_4}$$

The remarkable thing of this bridge is that if three arm resistors are known, the fourth arm resistance can be found solely from calculation. There is no dependency on a voltage source value. The sensitivity to a null condition is much greater than to any peak condition.

One of the known resistors can be made variable and an operator can simply adjust it for the null. Its marked resistance can then be applied to the final formula to determine the unknown.[3] In actual practice of the 1800s to early 1900s, a low audio-frequency buzzer could have been substituted for the DC source and sensitive earphones used as a detector (hence the commonly-used *detector* symbol).

Resolution of the null condition is quite good. Using a 10 Volt audio buzzer and 1000 Ohm headphones and three arms at 1000 Ohms each, a ± 0.1 % change in a fourth resistor would result in about 2.5 mV across the headphones. At ± 0.01 % deviation of the 4^{th} resistor, the headphone would have about 0.25 mV.[4] If 1.0 mW across headphones would be almost excessive to hearing, then the errors would be about -62 and -72 db down from that 1 mW level, still audible in headphones of a century ago.[5]

Expansion of the Basic Wheatstone Bridge to Complex Quantities

If the basic Bridge is redrawn as in Figure 37-2, there is ample opportunity to measure capacitance and inductance using standard resistors as the main reference point. Ohm's Law works

---

[3] This may be the origin of the word *potentiometer* as we know it today, that of adjusting the *potential* of the detector. Such a potentiometer could be both arms of a bridge with the rotating contact being the common point. In very old bridge instruments a variable resistance was often called a *slidewire*, the description of its construction, a sliding contact moved along a resistance wire.

[4] Accuracy of standard resistors developed over time and that subject is wholly different from using them and measuring with them. The *resolution of error* using a null-balance measurement was - potenitally - as great in 1850 as it was in 1950.

[5] Such sensitivity via headphones is due to the low perception of any sound of the normal human ear and to the very close acoustic coupling of headphones to the ear. Amplification is not necessary to have *perception of sound* as low as -75 db below 1 mW. Note: Harmonics of buzzers' fundamental frequency would be rather high but the ear-mind can also discriminate by sensing the *tone* through the headphone detector; that allowed some reactances among the standard resistor values to exist without unduly disturbing the measurement accuracy.

equally well with impedances as with resistances. However, measuring impedances requires knowledge of the expected phase angle resulting from Bridge unbalance by reactance and that dictates *where* to place unknowns and standards. The mathematical relationships of the Bridge arms now becomes:

$$Z_4 = \frac{Z_2\, Z_3}{Z_1}$$

...and that is the same as for all-resistive-arm bridges.

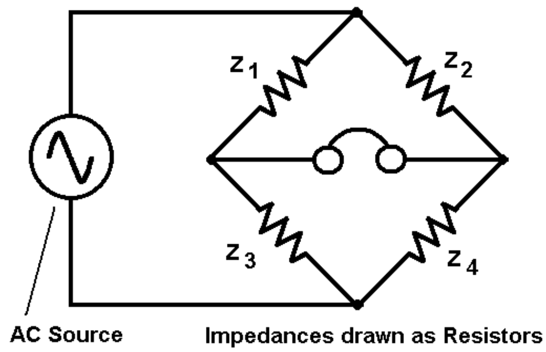


Figure 37-2 Basic Wheatstone Bridge adapted for measuring impedances.

Series-Resistance-Capacitance Bridge

The Bridge of Figure 37-3 is non-resonant; i.e., it does not require a specific AC frequency for balance. Both Rs and Cs could have calibrated dials to use in finding unknowns Cx and Rx. At null:

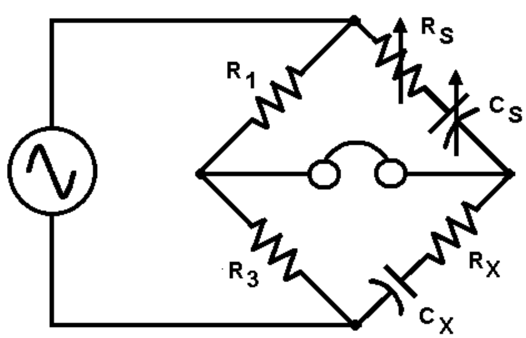


Figure 37-3 Series-Resistance-Capacitance Bridge, unknowns at Cx and Rx.

$$C_X = \frac{C_S\, R_1}{R_3} \qquad R_X = \frac{R_S\, R_3}{R_1}$$

This same Bridge can also measure inductances by substituting them for the capacitors. There is some difficulty in finding variable inductors so the Owen or Maxwell Bridge would be more suitable; those find unknown inductances via known variable capacitors.

NOISE Bridges

This is a more modern variation on the basic null-balance bridge, that of using ***random noise*** for the generator and having a tunable wide-band receiver as the frequency-selective detector. The principle is the same although the method may seem confusing. A version of a Noise Bridge is shown in Figure 37-4 for measuring impedances along transmission lines. Two of the four arms of the basic Bridge are represented by close-coupled, identical windings of the coupling transformer.

The Unknown series impedance is balanced out by the series of Rs and Cs. The *detector* allows sensing the imbalance and can be as sensitive as the receiver itself, down to the µV level.

The noise source is a wideband-amplified noise from any noise source, usually a zener diode with a slight current flow. Since true random noise covers such a wide frequency range, it can be used at any frequency within tuning range of the receiver. While this creates an RF signal into, say,

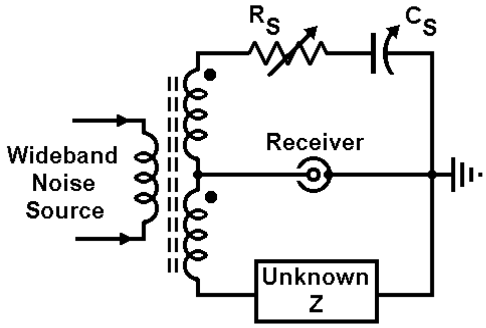


Figure 37-4 Basic Noise Bridge

an antenna, the amount of actual noise power is too low to be measurable to any distant station.

At first glance it would seem that the Noise Bridge (as shown in Figure 37-4) is unsuitable for measuring an unknown having an *inductive* component. That is true. But, revising the Noise Bridge to that of Figure 37-5 will solve that problem.

In Figure 37-5 the *standard* part (Rs and Cs) is made a ***parallel*** R-C comparison. It really doesn't matter if it is parallel or series since the final value of the *unknown* has to be calculated. The matter of having just a series R-C solution is made simpler: Simple add a ***fixed, high-quality*** capacitance across the *unknown* connection and make the change of Cs have ***twice*** the expected value of the unknown ***parallel*** connection.

Calibration of Cs will have to be at zero susceptance at center travel with a positive value on one side (capacitive susceptance of unknown) and negative value on the other side (inductive susceptance of unknown). Note that the ***unknown*** is now an ***admittance*** because of the parallel connection.

Calibration is necessary beforehand. The transformer shown should have its secondaries as alike as possible. One way to make that so is to use twisted magnet wire, winding this *bifilar* wire together. The primary is either wound as close to the center-tap or a *trifilar* arrangement is used. Keeping structure lead lengths equal as much as possible, allow connections separately to Rs and Cs and the fixed value equal to Cs/2 at the unknown connection. That Cs/2 should be ***locked in place*** if made from an air-dielectric variable; it will directly affect calibration of Cs if it is allowed to change.

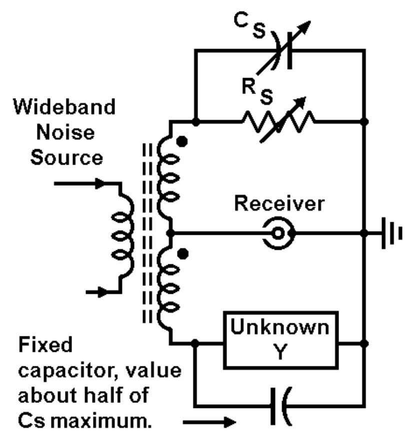


Figure 37-5 Noise Bridge modified for Admittance.

Rs can be calibrated-measured with a DC ohmmeter directly in Ohms. Cs can be calibrated-measured with another, known variable capacitance and labeled in pFd. Cs is best measured at a single frequency using the Noise Bridge, the standard capacitance connected across the the unknown-admittance terminals. Choice of Cs marking polarity away from center is up to the builder.

Formulas:

If: G_U = conductance of Unknown, B_U = susceptance of Unknown

$$Y = \left(\frac{1}{R_S}\right) + j\left(\omega C_S\right) \quad \text{[in mhos]}$$

with ω = radian frequency of $(2\pi F)$, F being frequency in Hz

$$\text{Series Connection in polar form:} \quad Z = \left(\frac{1}{\sqrt{G_U^{\ 2} + B_U^{\ 2}}} \right) \angle \left[\text{ArcTan} \left(\frac{B_U}{G_U} \right) \right]$$

On some *scientific* pocket calculators the conversion to series connection can be done with a ***polar to rectangular*** function. Note that sign of the Arctangent function is ***negative.***

As an example, suppose the frequency is 5 MHz and the Rs reading is 100 Ohms and Cs reading is 100 pFd. Conductance is then 0.01 mho and susceptance is 3.14159×10^{-3} mho. The series equivalent connection is then $95.4028 \angle -0.314159°$ in polar form, and, from Chapter 6, the rectangular form in series connection is equal to 91.0170 Ohms in series with 1113.21 pFd.

With a ***programmable scientific calculator*** that can be automatically calculated so that the only entries are the Rs and Cs readings with a result equal to what is wanted.[6] Note that this is a combination of techniques: Simplicity of the hardware instrument with the sophistication of precision calculation possible with at pocket programmable calculator.[7]

Instruments for Combined L-C-R Measurement

General

One can get by with very simple meters and also have very simple accuracy. One can splurge on more sophisticated instruments and get calibration-laboratory accuracy. There is very little in-between. One of the in-between instruments the author uses is the AADE ***L/C Meter II*** for capacitance and inductance.[8] Another, older, is the no-longer produced Data Precision model 938 for capacitance, having 8 full-scale ranges of 3-decade read-out from 200 pFd to 2000 µFd. Such in-between models depend on microcontrollers (AADE) or specific digital logic (Data Precision) for their operation.

More Accuracy With Added Phase Measurement

Getting better than 1% tolerance measurement of inductance and capacitance requires more complex internal circuitry. Frequency of measurement must be known accurately and there must

---

[6] This was covered in the May, 1978 issue of *Ham Radio* magazine by two different Andersons: the author (Leonard H.) for the Hewlett-Packard Model 25 and T. J. Anderson, then WD4GRI, for the Texas Instruments models 58 or 59. No family relationship. The author's article included using the RPN programs for Smith Chart plot points and finding the electrical length of transmission lines. Both pocket calculators have ceased production but the basic programming steps are applicable to following versions for both H-P and TI.

[7] A number of later magazine articles past May 1978 have appeared that cover calculation with a programmable pocket calculator, too many to list here. The principle is the same, to make a calibratable hardware instrument and combine it with the sophistication of a numeric calculator.

[8] Almost All Digital Electronics in Washington state. They have a website at www.aade.com and are located near the city of Seattle.

be a way of measuring the ***phase*** across the device under test. That last requirement is needed to separate the equivalent of resistance as well as capacitance or inductance of the device being tested.

For this more-sophisticated measurement, some references tell the complicated needs and are listed following:

1. ***Impedance Measurement Handbook,*** December 2003, 114 pages, Agilent Technologies.

2. ***LCR Measurement Primer,*** 4th Edition, February 2005, 55 pages, IET Laboratories, Inc., split from QuadTech in early 2012, it carries on some of the measurement of its origin, General Radio Company. Website: www.ietlabs.com

3. Hewlett-Packard Journal edition of December 1977 with a technical insight to their model 4263 LCR Meter. Hewlett-Packard Journal was a semi-monthly brochure-magazine published from September 1949 to November 1998 and on-line copies (PDF) are available from www.hpl.hp.com/hpjournal/PDFs/hpindex.html

4. H-P Technical instruction manual on the HP 4342 LCR Meter, printed March 1983, obtained over the Internet.

Frequency Measurement

Sources

The United States' NIST operates Time-Frequency standards at Fort Collins, Colorado on HF at 2.5, 5, 10, 15 and 20 MHz over WWV with a synchronized transmission from Kauai, Hawaii on 2.5, 5, 10, and 15 MHz over WWVH. Those are double-sideband AM. In addition to HF at Fort Collins, WWVB at 60 KHz, with a pulse-width coding of 60 bits per minute. Both sites reach most of the United States 24 hours a day on VLF, and are on continuously over HF, subject to ionospheric propagation.

RF power output on HF is 10 KW at 5, 10, 15 MHz in Colorado and Hawaii. In Colorado the 2.5 and 20 MHz power output is 2.5 KW. Hawaii's 2.5 KHz output is 5 KW. RF power output from WWVB is about 70 KW. Slow rate of modulation allows automatic synchronization of radio wrist watches and radio clocks every 24 hours, even with very limited physical-space antennas.[9]

RF carrier accuracy as transmitted is better than ± 1 part in 10^{11} or better than ± 1 Second every 76,000 years. That applies to both HF and VLF. Propagation variations will reduce that accuracy. Radio spectrum allocations have created both HF and VLF as *standard time-frequency* portions of the spectrum.

Local frequency standards are limited to better than 50 PPM for quartz crystal oscillators at

---

[9] Receiving bandwidth of about 20 Hz allows greater sensitivity over WWVB on VLF, much better than the 1.0 KHz time-ticks over WWV or the 1.2 KHz time-ticks over WWVH. Full details of RF characteristics is available in NIST Special Publication 432, 2nd Edition, ***NIST Time and Frequency Services***, written by Michael A. Lombardi, available free over www.nist.gov/pml/div688 (home page of Time and Frequency Services).

low relative prices to better than 5 Parts per Billion for Cesium Beam oscillators at much higher prices. An intermediate local frequency standard is the controlled TCXO or Temperature Compensated Crystal Oscillator having a small frequency trimming adjustment.

Using Simple Crystal Oscillators With WWV

This is simple heterodyning a local frequency oscillator with WWV at one of its HF carrier frequencies. With a weak WWV or WWVH signal, a 10 MHz local oscillator can be set to within ± 5 Hz, equal to a tolerance of about 2 PPM. With a strong HF signal the carrier heterodyne can be set down to about a half Hertz or a tolerance of roughly 200 PPB.

Synchronizing a 10 MHz VCXO to WWVB

The Figure 37-6 borrows an old trick from the Hewlett-Packard 117A Comparator instrument, synchronizing a local 10 MHz standard to the WWVB carrier. If the 60 KHz signal can be received clearly, the carrier can be extracted and divided by 3 to get 20 KHz. That is applied to a phase detector as the Reference.

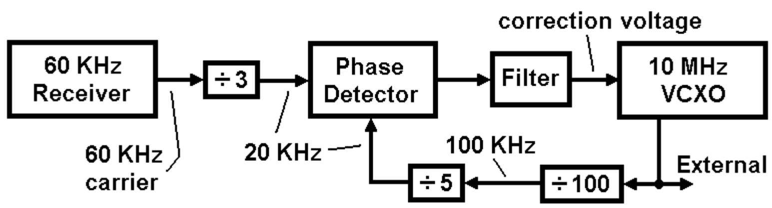


Figure 37-6 A simple synchronizer, 10 MHz to WWVB

A local frequency standard, a temperature-stabilized VCXO or Voltage Controlled Crystal Oscillator is then divided by 500 to get 20 KHz as a Comparison signal to the phase detector. Phase detector output is lowpass filtered to about 1.0 Second and used to control the 10 MHz VCXO.[10]

The phase detector can be a pair of audio transformers or a digital-gate type of phase-frequency detector. The main thing is to have it relatively noise-free. The Filter must have a low gain and a corner frequency of about 1 Hz to allow it to control the VCXO.

World Time-Frequency Stations on VLF[11]

BPC	68.5 KHz	China
DCF77	72.5 KHz	Germany
JJY	40, 60 KHz	Japan
MSF	60 KHz	United Kingdom
WWVB	60 KHz	United States of America

[10] A Crystek model VCT32-10.000 was used here with about a 2.0 VDC control voltage. This comes in an SMT and operates from 3.0 VDC, tucked in a temperature-stabilizing mini-oven of about 140° F. A wideband buffer amplifier follows to the external output and to the ÷100 (a single 74H390 dual decade). In the HP 117A the phase detector operated at 60 KHz but there doesn't seem to be a good technical reason for doing so. In the 1960s most in-house local frequency standards were 100 KHz.

[11] From Horological Journal, March 2010, article by Michael A. Lombardi. All have varying bit pattern arrangements for time, date, and other data.

Making a Frequency/Time-Interval Counter

This takes advantage of existing digital logic devices or uses a microcontroller to be both the counter, gate, display driver, and local frequency reference. A block diagram is shown in Figure 37-7 that allows a choice of either Frequency or Time-Interval, both governed by the local frequency standard.

For frequency measurement, the local frequency standard is divided down digitally to produce a ***gate*** that allows the signal input to be accumulated in a chain of digital counters. At the end of a gate time, the accumulated count is transferred to a numeric display such as numeric LEDs or an LCD unit.

Resolution of the count, the LSD (Least Significant Digit) is determined by the gate time and is always ± 1 count statistically due to non-synchronicity of the timing oscillator frequency. This is tabulated as follows:

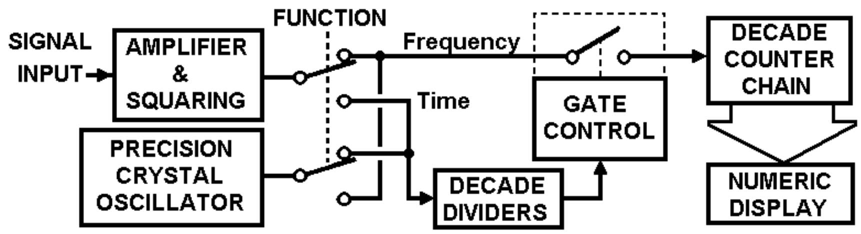


Figure 37-7 Frequency/Time-Interval Counter block diagram.

Gate Time	Resolution	Decimal Digits for 10 MHz Count
10 mSec	100 Hz	6
100 mSec	10 Hz	7
1 Second	1.0 Hz	8
10 Second	0.1 Hz	9

For time-interval measurement, the signal input goes to the gate time dividers and the internal precision oscillator frequency is measured. See the *Function* switch in Figure 37-7. Resolution is now the period of the internal oscillator and the gate time is the period of the signal input. For a 10 MHz precision internal oscillator the tabulation is:

Resolution	10 MHz Divider	Decimal Digits for 1 KHz Period
100 nSec	1	5
1 µSec	10	4
10 µSec	100	3
100 µSec	1000	2
1 mSec	10,000	1

Time-interval measurement is good for relatively low-frequency signal inputs. The time-interval displayed can be inverted to get the exact frequency on a pocket calculator.

Time-interval measurement also requires a ***comparator*** at signal input to set the logic levels accurately. That allows the signal input to be positive-going or negative-going. For a single input, the measurement is really ***period.*** For exact ***time intervals*** the gate time must be turned on at the start and turned off at the ending of the time interval. That usually requires ***two comparators.***

Increasing Period or Time-Interval Accuracy

A technique called ***time-interval averaging*** has been in use since 1970 in commercial test equipment. The reason is the same non-synchronous behavior of the measurement meter's internal oscillator and the AC signal being measured. Its first symptom was the ± 1 count ambiguity of the least-significant digit.

By extending the gate time by the ***square*** of the desired minimum resolution, statistics prove that displayed accuracy is increased by the ***square-root*** of the gate time increase. For example, to increase the reading 10 times, the gate time is extended by 100 times to get that 10-times-increase. To get a 100 times display accuracy, the gate time is extended by 10,000 times for that 100-times-increase in accuracy.

This works fine for lower-frequencies such as a power line frequency of 60 Hz which has a period of 16,666.667 µSec. The ± 1 count reading ambiguity would still be there but that now becomes ± 0.001 µSec or ± 1 nSec. The more-significant digits can be divided into unity to get an exact 60.000 0 Hz.

The same ***averaging*** can be applied to frequency measurement but at a great cost of ***time.*** To get a precise frequency reading of 1.0 MHz to within ± 1 milliHertz with an internal precision 10 MHz oscillator (100 nSec period) would require 10,000 count cycles at 10 Seconds per cycle or 1667 hours. That isn't feasible. Even a reading to within ± 10 mHz would take most of a 24-hour day.

Increasing Frequency Measurement Accuracy

Heterodyning and unknown frequency with a ***known, precise*** frequency has been known since before the World War II period.[12] All it requires is a mixer with a low-pass filter to pass the difference between the known, precise frequency and the unknown frequency source. ***Prescaling*** techniques such as with presettable accumulator counters are good for handling LO measurements that require the IF offset to read the tuned frequency but do ***not*** offer any advantages equal to the heterodyning technique for more precise frequency readings.

Instrument Errors

Since frequency counters and time-interval counters involve on-off conditions, their construction can be satisfied by digital logic devices. The only real uncertainty comes from the local precision frequency oscillator and, if used, signal input comparators' settings.

Local precision oscillators can be tested separately against a very good source such as VLF or HF time-frequency broadcast carriers. As an example, dividing a 10 MHz master time-base frequency by 20 gets 500 KHz. Harmonics of that can be used to beat against any HF broadcast of WWV.

Comparator settings require oscillographic confirmation of settings, such as for very low

---

[12] This technique was used for the tremendous quantity crystal unit production in the USA since World War II. It does require the ***known*** frequency source to be at least 10 times better than production standard. This heterodyning process was also used in the BC-221 frequency calibrator in use at that time.

frequency sources. As an alternative, precise knowledge of trigger settings is good. A saving grace is that repetitions of such low-frequency signals are usually precise enough that they will remain stable over a measurement period.

Comparator circuits at signal inputs should have a Schmitt trigger function to convert slow rise and fall signal slopes into sharp digital on-off states for the remaining counter circuits.

Signal Generators

General

These can be a very large variation in accuracy, both for carrier frequency and for very low power outputs.[13] On top of that the harmonic content of the carrier and linearity of any included AM or other modulation source are not considered by many. Everything matters if measuring things like receiver sensitivity.

Very Low Carrier Level Attenuators

Waveguide below cutoff attenuators or double-shielded strings of switched attenuator pads are necessary for setting very low carrier levels. Such attenuators must be ***calibrated*** at least once. For most hobbyists this means calibrating them in substitution mode against a ***known*** attenuator. The attenuators of a cheap model signal source is always suspect. Those are usually made of resistors arranged as a fixed pad attenuator.

One can approach some calibration with a few fixed attenuator pads that have been measured on commercial equipment and known at various frequencies. Those can be strung in most any order with in-between-calibration-measurements interpolated.

Shielding

This is mandatory for very low signal levels. A full surrounding shield must be in place on things such as switched attenuator pads. Two things on those switched attenuators: First, make the arrangement of switches and pads in-line within a structure that resembles a waveguide, roughly an inch high by two inches wide. The whole attenuator structure then becomes a waveguide below cutoff with the output shielded from input by at least 120 db or so. Second, assume a full foil or thin conductive plate to attenuate about 60 to 80 db if an odd shape. Use two shields to get maximum shield attenuation.

On the bench, check for leakage around large attenuation settings by noting any changes due to hand positions or changes in tool placement. If those are seen, there is insufficient shielding. Trying to get 120 db or more shielding requires a ***full surround*** of conductive structures, no slots or openings.

---

[13] Variations run from a Heathkit originally costing about $30 to more expensive test equipment costing $4000 or more new from the factory.

Output Levels

This is also mandatory and can be done with a simple wideband detector for the frequency range. Using a medium- to high-value load on a diode detector for the HF to low-VHF range (such as using a 1N4148 or equal), that can be calibrated at low frequencies by reading the peak voltage. That should hold up to about 50 MHz. A Schottky diode such as a 1N5819 is a better choice for a linear output-indicating meter scale.

For VHF and upward in frequency, calibration almost requires the aid of commercial equipment and may require a thermoelectric sensor. Most thermoelectric units are broad band enough to handle frequencies up to VHF-UHF. What is important is that only a sample of the output power is measured, such as using diode detectors with a couple thousand Ohms of load resistance or thermoelectric sensors using a small part of the output current.

Output power ***must be at a standard impedance*** such as 50 Ohms.[14] This allows interchangeability of equipment on the bench. If the output impedance of a signal generator is unknown, then it should have a wideband ***masking pad*** of 10 to 20 db between output and a variable attenuator. While that seems to *waste* RF power, consider that full output power is seldom used. If a higher level of RF output is required, a small amplifier can be inserted for such purposes. Most testing of receivers for sensitivity requires considerable, calibrated attenuation in the output.

Frequency

Don't expect cheaper signal generators to have accuracy, especially if the maximum-to-minimum frequency ranges are in the 2:1 to 3:1 ratios. They *might* be stable when warmed up but a sample of the RF to a wideband frequency counter is a much better indication.

Lacking a full-blown synthesizer, an equivalent can be made from a DDS IC such as the Analog Devices 9851. With a known reference frequency such as a tripled 10 MHz standard, and the internal 6-times multiplier switched in, it can output very stable signals anywhere from audio to 60 MHz or so with digital control. No mechanical coupling to tuned circuits is required and the 9851 can be assembled on a small PCB that can be shielded.

A DDS-based signal generator can do very small frequency changes anywhere in its output range. That is enough to test the response of narrow quartz-crystal filtered receivers down to a few Hertz per point.

Harmonic Distortion and Modulation

Harmonic content of cheaper signal generators can be high, on the order of less than 20 db down from the carrier. Higher cost signal sources are much reduced. One has to consult the specification sheets for the full picture.

Modulation in AM can be relegated to an *identifier* category, just enough to tell the signal generator from any other stray signals that wander in. Do not expect any sort of expectation of modulation, AM or FM or PM, unless one pays good money for such quality.

---

[14] Over a half century past, the electronics industry went to a 50 Ohm input and output impedance for interchangeability, of connecting anything to anything else on the bench..

Principle of Operation and Calibration of a Q Meter

The legacy Boonton 260 Q-Meter was a standard inductance-testing device for a number of years. Its simplified measurement section is given in Figure 37-8. There are three binding posts on its top with the external inductor to be measured put between the ***L*** and ***C*** posts.[15] That inductor will have a loss component equal to inductive reactance divided by Q. It also has some distributed capacity indicated by the C_0 (shown shaded). Frequency range is 50 KHz to 50 MHz and the calibrated variable capacitor is about 405.3 pFd to 40.53 pFd.[16] When the RF signal source is set to 79 KHz, 250 KHz, 790 KHz, 2.5 MHz, 7.9 MHz, or 25 MHz, the inductance markings on the variable capacitor dial can be read directly in ranges of 10-100 mHy, 1-10 mHy, 0.1-1.0 mHy, 10-100 µHy, 1-10 µHy, and 0.1-1.0 µHy, respectively. A single variable capacitor dial scale suffices for all frequencies in the 79 and 250 multiples; a second dial scale is marked directly in picoFarads.

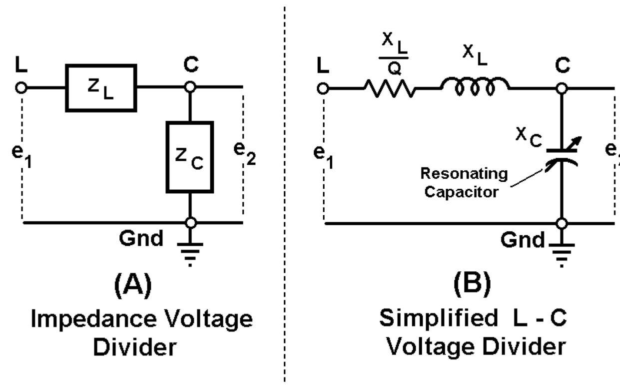


Figure 37-8 Frequency-independent relation of output RF voltage relative to constant input.

At resonance of L and C, the set frequencies for decade measurement ranges correspond to increments of multiples of the square root of ten or 3.16228. If 250 KHz is the exact choice, the next lower inductance decade range with the same variable capacitor requires (250 / 3.16228) or 79.0569. Picking an exact 79.0000 results in a frequency error of only -0.072%. The inductance ranges and their corresponding frequency settings are marked on the Boonton 260A front panel. Q at other frequencies can be read out directly but the inductance would have to be calculated from the resonance formula using the capacitance dial scale markings.

How the Q is Read Directly and the Same at All Frequencies

The internal RF source supplies current to the 20 mOhm R1. The current is monitored by a thermal milliammeter (second meter on the front panel). For direct reading of Q the adjustable RF source current is set to the same marking on the meter at all frequencies. R1 and the thermal meter are one assembly in the 260Z and are very broadband. The result of that manual operation will yield an RF voltage drop across R1, given as e_1 in Figure 37-8. What follows shows that the magnitude of e_2 will be exactly equal to $Q \bullet e_1$ at resonance!

The impedance voltage divider of Figure 37-8 (A) operates the same way as a resistive voltage divider but does so on the *magnitudes* of the impedance blocks shown. Z_L represents the inductor to be measured and its equivalent series resistance representing Q loss. Z_C represents the calibrated

[15] Markings on drawings are for explanation clarity, will not be found exactly like that on equipment.

[16] It is slightly wider than this. Variable capacitor range end values correspond to the 1:10 inductance range ends at any particular *79* or *250* increment of frequency.

variable capacitor. Assume for the moment that the variable capacitor is perfect, having no losses.

At Resonance:

$$e_2 = e_1 \left| \frac{Z_C}{Z_L + Z_C} \right| \qquad \text{Where } e_1 = \text{ input RF Voltage, } e_2 = \text{ output RF Voltage}$$

$$Z_L = \left(\frac{X_L}{Q}\right) + j\,X_L \qquad Z_C = 0 - j\,X_C \qquad Z_L + Z_C = \left(\frac{X_L}{Q}\right) + j\,0$$

$$|X_L| = |X_C| \text{ at resonance so imaginary parts cancel out}$$

$$\frac{Z_C}{Z_L + Z_C} = \frac{0 - j\,X_C}{\left(\frac{X_L}{Q}\right) + j\,0} = 0 - j\left[\left(\frac{X_L\,X_C}{Q}\right)\left(\frac{Q^2}{X_L^2}\right)\right] = 0 - j\,\frac{Q\,X_C}{X_L} = 0 - j\,Q$$

$$e_2 = e_1\,Q \qquad [\text{ magnitude of } (0 - j\,Q) = Q\,]$$

This bit of seeming legerdemain only happens at resonance.[17] It is frequency independent insofar as the resonance frequency is concerned! This was a major innovation in electronic measurement when it was introduced in the 260A's predecessor, the Boonton 160 Q-Meter. All one needs now is a very high impedance input wideband amplifier with its output made to an amplitude detector for meter (analog or digital) read-out of Q.

High Impedance RF Voltmeter to Read Q

The 260A used a special vacuum tube in its RF voltmeter, one with such low grid current that the grid return resistor could be as high as 100 Megohms. The parallel resistance of the voltmeter input across the calibrated variable capacitor will cause a lower indicated Q at the lower frequencies. The parallel R-C transforms to a series R-C with the series resistance component causing the error[18]. The change in the transformed C value is too small and would cause only a very slight shift in the variable capacitor's dial rotation. To gauge the extent of this frequency-dependent error, Q of exactly 100 was calculated for three voltmeter input resistances at the four lowest test frequencies, inductance as the maximum value, variable capacitor at minimum value; note also, exact *79* frequency of 79.056941 and minimum variable capacity of 40.528473 pFd values were used in calculation to avoid numerical errors.

Calculated Q Readings of Test Coil With Exact Q of 100 versus Voltmeter Resistance

[17] Those who aren't convinced by the mathematics is invited to enter this simple impedance voltage divider into a circuit analysis program to prove the point. In any series L-C circuit the peak-to-peak voltage across either L or C will be higher than across the whole series circuit (if the Q is greater than 1).

[18] Chapter 6, equation set (6-12).

Test Frequency	100 MOhm	22 MOhm	10 MOhm
2.5 MHz	99.8	99.3	98.5
790 KHz	99.5	97.8	95.3
250 KHz	98.5	93.3	86.4
79 KHz	95.3	81.6	66.8

Errors decreased inversely proportional to frequency above 2.5 MHz. Its rather obvious that the lower frequencies **need** the higher voltmeter input resistance.

The General Radio 260A manual gave correction values for readings at lower frequencies. There is no need to slavishly follow their method of a one-tube RF voltmeter conception; the real requirement is to make the voltmeter input impedance resistive and the highest practical value. Figure 37-9 shows one way to that, copying the *compensated voltage divider* common to every 10:1 oscilloscope probe. A 100:1 divider is used here with the input resistance of the 100 MOhm value. Any parallel capacitance of the input will appear across the variable capacitor and become a part of its calibration.

The video amplifier can be one or two of several gain block ICs available. Two stages would provide a voltage gain of 200 to 300 times with a 1 V peak-to-peak output across that wide frequency range. *Rectification efficiency* of the diodes at higher RF would enter in here, equivalent overall as a reduction in gain.[19]

The entire video-amplifier-detector chain ***can be tested separately*** to quantify its broadband characteristics or flatness insofar as the Q indication is concerned. If necessary, the response versus frequency could be shaped as needed to compensate for other non-flat-response characteristics. The input voltage divider's frequency compensation trimmer capacity would be set by injecting a 10 to100 KHz nominal squarewave and adjusting it while observing the source follower's output on an oscilloscope. The trimmer's nominal capacitance would be roughly 100 times the capacitance across the high value resistors. The 33 MOhm resistors were shown because those were the highest value carbon-composition resistors available locally. Do ***not*** use the spiral-band high-MOhm types since those will introduce some series inductance that drops the higher frequency response. Some *film* resistors have more internal inductance than others so one should be wary of those in this kind of circuit. In the 260A the RF voltage across the variable, and thus the RF voltmeter input, is 5.0 V RMS for a Q of 250. That level into the Figure 37-9 circuit would result in 50 mV RMS at the source follower input. It does not need to be that high and could be cut down to a tenth of that. The

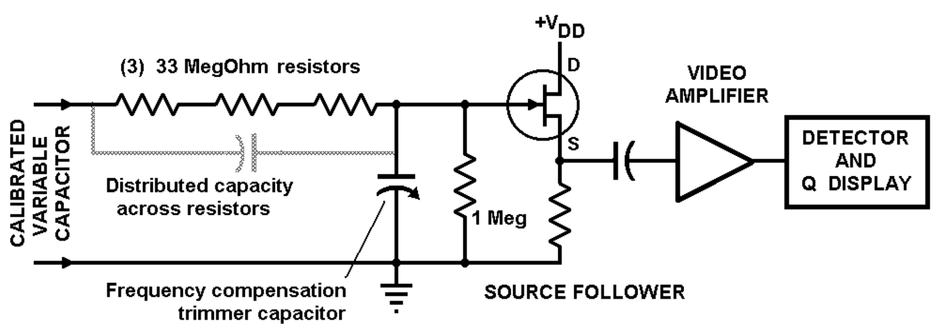


Figure 37-9 Conceptual version RF voltmeter of high-R input.

[19] Rectification efficiency can vary among common diodes, silicon or germanium, at higher RF since it is not normally a specified parameter.

on-resonance voltage equivalent to a Q of 250 could be dropped to 0.5 V RMS, relieving the RF source power injection demands.

RF Source and a Low Point

In the 260A there is an average of 1 A RMS into the 20 mOhm resistance-thermal-milliammeter. It was a simple choice for 1930s, being broadband for the entire frequency span, but it results in a very low injection voltage, both a good point and a bad one.[20] At resonance, the resistive component of the inductor arm (reactance divided by Q) is in series with the resistive part of the capacitive arm, both in parallel with the 20 mOhms of the thermal milliammeter. There is little effect of that on the lower bands but requires more source power on the higher bands. Operators of the 260A are continually resetting the *Q Multiplier* control (RF source level) after changing frequencies. Part of that is due to the internal RF source output varying with frequency. The other part is the variation of the inductance's resistive part changing with frequency. The following table shows the impedance magnitudes of the L and C by themselves, then in shunt with 20 mOhms.

		Magnitude of Test Network, Ohms		Total Circuit Magnitude, mOhms	
Inductance Range	**Test Frequency**	**Max. Variable**	**Min. Variable**	**Max. Variable**	**Min. Variable**
10 mHy - 100 mHy	79 KHz	20.12 -	223.4	19.98 -	20.00
1 mHy - 10 mHy	250 KHz	6.308 -	65.30	19.94 -	19.99
100 µHy - 1 mHy	790 KHz	1.989 -	20.12	19.80 -	19.98
10 µHy - 100 µHy	2.5 MHz	0.6286 -	6.308	19.35 -	19.94
1 µHy - 10 µHy	7.9 MHz	0.1987 -	1.989	18.17 -	19.80
0.1 µHy - 1 µHy	25 MHz	0.06283 -	0.6286	15.17 -	19.38

Total Circuit Magnitude columns are the parallel of resistive parts of L and C arms in parallel with the 20 mOhm thermal milliammeter resistance. The greatest change comes from the lowest inductances within a range and the higher Qs. Since the resistive part of an inductor arm is (X_L/Q) that resistance is inversely proportional to Q.

If the internal RF source can couple to milliOhm loads, then the source current need vary no more than about 1:2 in power, including 50 MHz used with 0.1 to 1.0 µHy inductors. What would be highly desirable is a *leveled* RF source; i.e., where operators didn't have to re-tweak that *XQ* control for the index mark on the lower (260A) meter. Operators need only to set the test frequency and concern themselves with the variable capacitor and peaking the Q reading.

Leveling the Playing Field

[20] Boonton Radio's first product was a Q Meter, circa 1934. Model 160, predecessor to the 260, came out in 1946 using a 40 mOhm thermal milliammeter. Some texts claim that all of this type inject a *current* into the L-C divider but it is really just a voltage drop so that output to the Q voltmeter will be proportional to Q. Hewlett-Packard acquired Boonton in 1959.

Figure 37-10 block diagram is a revision for somewhat leveling of the RF. R1 is made 1.0 Ohm resistive, made from 10 each 10 Ohm carbon resistors. R2 is still 100 MOhms (as input impedance of the 100:1 voltage divider). Both video amplifiers and their RF detectors are identical in frequency response and overall gain versus meter RF input-voltage read-out. The gain-controlled RF amplifier following the source represents a variable-gain stage with RF *current* output capability. At low frequencies, 10 mA RMS into R1 will produce 10 mV RMS. If the inductor has a Q of 250, the input voltage to the 100:1 divider will be 2.5 V RMS at resonance. The video amplifier following the divider will have an input of 25 mV RMS. With nominal Qs in the range of 50 to 200 the voltage after the divider and into the video amplifier will be 5 to 20 mV. Voltage amplification in the video amplifiers is about 100x so that output into the RF detector is 0 to 3.0 V RMS. What is important for minimum error is that both video amplifiers have identical frequency response and gain. Reasons for that are found in the tabulation following:

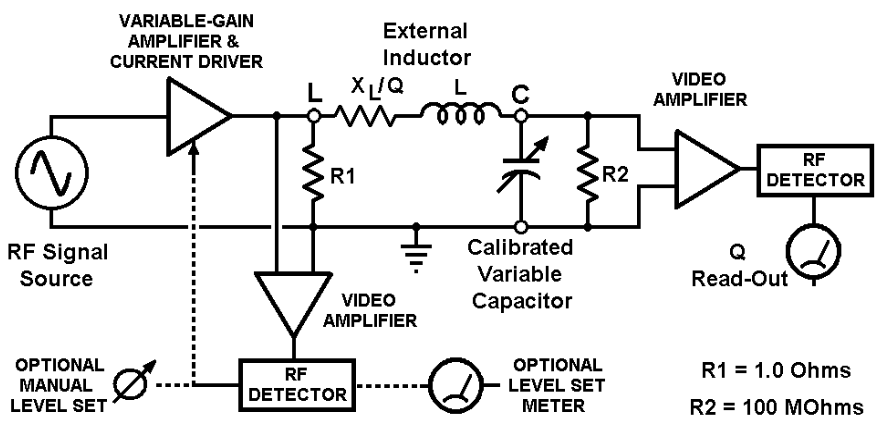


Figure 37-10 Concept of a Q Meter using two identical video amplifier-detectors with possibility of automatic Level Set.

		RF current, mA RMS For 10 mV Reference			**Q Reading Error In Percent**	
Inductance Range	**Test Freq.**	**Max. Variable**		**Min. Variable**	**Max. Variable**	**Min. Variable**
25 mHy - 250 mHy	50 KHz	10.31	-	10.03	-1.92	-16.4
10 mHy - 100 mHy	79 KHz	10.50	-	10.05	-1.23	-11.1
1 mHy - 10 mHy	250 KHz	11.59	-	10.15	-0.38	-3.78
100 µHy - 1 mHy	790 KHz	15.03	-	10.50	-0.13	-1.23
10 µHy - 100 µHy	2.5 MHz	25.91	-	11.59	-0.02	-0.38
1 µHy - 10 µHy	7.9 MHz	60.32	-	15.03	-0.02	-0.13
0.1 µHy - 1 µHy	25 MHz	169.2	-	25.91	-0	-0.04
25 nHy - 250 nHy	50 MHz	328.3	-	41.83	-0	-0.01

Note: ***-0*** refers to a negative percentage smaller than 0.1%

Q of the inductor is a result of the effective RF voltage magnitude magnification by the L-C ell-arrangement. If both video amplifier and RF detectors are identical, then the indicated Q will be at a minimum possible error condition, that caused mainly by the 100 MOhm parallel resistance across the variable capacitor at the lower test frequencies.

If the RF source has to provide a much higher RF current at high frequencies, then the voltage drop across R1 will still be the 10.0 mV needed for a 1x Q reference setting. The remainder of the RF current in excess of 10 mA will be flowing through the resistive part of the inductor's equivalent resistance due to losses. From the previous page's tabulation, that resistance will be below 1 Ohm. In the comparative system of Figure 17, that doesn't matter. All that is required is a 10 mV RMS drop to ground at inductor connect point ***L***. Q of the inductor will result in an equivalent RF voltage multiplication from 10 mV RMS as the input to the 100:1 voltage divider. At a Q of 100 at resonance the divider's voltage input will be 1.0 V RMS; the divider's output will then be 10 mV.

RF current source at the higher frequencies will be a problem in circuit design and high-f_T transistor choices (to reach 50 MHz). The RF source current variation is about 9:1, 10 mA at low frequencies but requiring almost 90 mA at 50 MHz with a 100 nHy inductor having a Q of 250. There's a way out of not having enough RF source current called the *Q Multiplier.*

Q Multiplication, Compensation for Lack of RF Source Voltage

Since Q is essentially a voltage magnitude multiplier of the RF source voltage (refer to Figure 37-8 explanation), it follows that halving the RF source voltage will require doubling the indicated Q readings. Most of the older Q Meter models have index marks on the Level Set meter for such *X2* and *X3.*[21] Indicated Q readings must be multiplied by 2 or 3 respectively when the RF source level is set to those marks.

Using the *times Q* reference levels for lower RF source voltage will ease the high RF currents necessary at high frequencies with low inductance, high Q inductors. The downside is that operation will require always multiplying the indicated Q value to get the true Q.

Calibration

The variable capacitor dial scale can be calibrated by a precision digital capacitance meter or bridge when the Q Meter is powered down. Connect between ***C*** and ***Gnd*** terminals, no inductor on the ***L*** terminal. Mark both capacitance and inductance on the dial (for the test frequencies). Since some Q tests need to be done on frequencies other than the 79-250 test variety, a capacity scale is needed to allow manual calculation of inductance. Inductance values at the 79-250 test frequencies are equal to:

$$C = \frac{405.2847346}{L} \qquad \text{Where: C is in pFd}$$

L is in mHy at 250 KHz, in μHy at 7.9 MHz

This yields a decade scale at exactly 7.90569415 MHz (1.0 to 10.0 µHy) or 250.000 KHz (1.0 to 10 mHy) with variable capacitors of 405.28473 to 40.528473 pFd.

The two video amplifiers-detectors should have *exact* amplification and frequency response from 0 to 30 mV RMS, 50 KHz to 50 MHz. The *Level Set* amplifier-detector of 10 mV RMS input is the reference of L-C network input voltage. That would correspond to a Q of 100 in the *Q Display* amplifier-detector output or 10 mV RMS input directly or 1.0 V RMS input to the 100:1 divider. Both can be initially calibrated at high audio or supersonic frequencies, then checked with a good

[21] Front panel markings of **XQ** on the 260s and 160s refers to ***times Q*** not *ecks-Q* as if it were a reactance.

RF signal generator for frequency response up to 50 MHz.

The *XQ=1* index mark on the *Level Set* display should have *XQ=2* and *XQ=3* (even *XQ=4)* additional indices for lack of RF source voltage at high HF and low VHF. If both video amplifier and detector voltage responses are equal, their higher frequency gains can be down. IF so, all that is required is to boost the RF signal source output so that one of the index marks is reached. Appropriately multiplied, the Q readings will still be accurate. Additional RF source voltage should not affect passive L or C components.

Standard inductors, those of specially-measured inductance and Q, aren't necessary.[22] Measured Q is a function of the relative voltages of the RF source and the variable capacitor's voltage at resonance. Note: Do not use a 'scope probe and oscilloscope for RF level calibration across the variable capacitor; the probe will introduce at least a 5 to 10 pFd additional capacitance and its 10 MOhm resistance to ground will upset the lowest frequency range Q readings.

The following table is for calibrating the variable capacitor's dial for inductance, using the 79-250 test frequencies.

Capacitor Dial Inductance Markings Versus Capacitance at Test Frequencies

Inductance Scale	Capacitance pFd	Inductance Scale	Capacitance pFd
1	405.285	3.2	126.651
1.1	368.441	3.4	119.201
1.2	337.737	3.6	112.579
1.3	311.757	3.8	106.654
1.4	289.489	4	101.321
1.5	270.190	4.5	90.063
1.6	253.303	5	81.057
1.7	238.403	5.5	73.688
1.8	225.158	6	67.547
1.9	213.308	6.5	62.351
2	202.642	7	57.898
2.2	184.220	7.5	54.038
2.4	168.869	8	50.661
2.6	155.879	9	45.031
2.8	144.745	10	40.528
3	135.095		

[22] Coil standards were once a staple set of accessories for Q Meters. Those were sold from the 1930s onwards, probably due to difficulty of most metrology labs had in measuring video-frequency-range voltages. Precision LCR Meters and Q Meters of today can be calibrated independently of such coil standard. A set of *working coils* for checking capacitors and other things is handy.

Spectrum Analysis

General

As a conventional oscilloscope measures amplitude (vertical deflection) versus time (horizontal), the spectrum analyzer measures ***amplitude*** (vertical) versus ***frequency*** (horizontal deflection). Shown in Figure 37-11 is a general block diagram of an ***analog*** spectrum analyzer. Modern variations use FFT (Fast Fourier Transforms) with digital filtering for best performance.[23]

If Figure 37-11 looks like a *receiver,* that is what it is basically. The difference is that the output is made to an oscilloscope, not a speaker or TV screen.

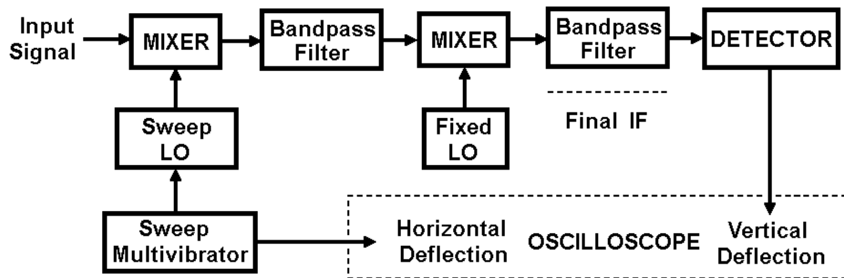


Figure 37-11 Generalized block diagram, spectrum analyzer.

Block Diagram

In the block diagram, the signal to be analyzed is first up-converted and passed through a VHF to UHF bandpass filter. It is then down-converted and passed through a narrow HF bandpass filter. The shape of this last filter will determine the spectral resolution. Detector output will go to the vertical deflection of an oscilloscope.

The *sweep local oscillator* will determine the width in terms of frequency of an analysis. This is generally a form of VCO which covers a particular band of interest. That sweep voltage is generated by an astable multivibrator whose output shape matches the VCO sweep LO so that the frequency is approximately linear across its band. The sweep multivibrator also goes to the horizontal deflection of the oscilloscope so that the display is complete.

Detection is best done with a ***logarithmic detector*** having a dynamic range of 60 to 80 db. This provides a vertical display more in line with other spectral plots and specifications. That also eliminates some amplification necessary to overcome conversion and filter insertion losses.

The middle section of Figure 37-11 can be several stages of mixers and filters to allow switching in different bandwidths. The ***rate of repetition*** of the multivibrator is determined by both the ***bandwidth*** of the detector and the total ***span*** of frequencies displayed. ***Minimum*** sweep rate follows this approximate formula:

$$\text{Sweep Time, Seconds} = \frac{2\ \text{Span, Hz}}{(\text{BW, Hz})^2}$$

For a span of 200 KHz and a bandwidth of 3 KHz, minimum sweep time would be 44 mSec. Sweep

[23] Digital filtering and FFTs are beyond the scope of a single Chapter. See references.

time can be slower than the formula indicates, limited only by persistence of the eye and persistence of the oscilloscope display.

Detector Filtering

For an analog spectrum analyzer, the IF (equivalent) should be as narrow as possible, such as a 4 to 5 resonator synchronously-tuned bandpass filter and approximately Gaussian in shape for frequency versus amplitude. Analog filter shape should be no more than 12 times wider at -60 db than at -3 db. Output of the detector can also be filtered if desired, but no more than the expected sweep time and bandwidth/resolution allow.

References

There are several good references on the subject of spectrum analysis available over the Internet. Some of the best ones are:

[70] ***Fundamentals of Spectrum Analysis,*** Rhode and Schwarz, 2001, 213 pages, document number PW 0002.6635, www.rohde-schwarz.com.

[71] ***Agilent Spectrum Analysis Basics, Application Note 150,*** Agilent Technologies, 2006, 119 pages, document number 5952-0292, www.agilent.com.

[72] ***Agilent Spectrum Analysis Amplitude and Frequency Modulation,*** Application Note 150-1, January 1980, document number 3954-9130, www.agilent.com.

[73] ***Spectrum Analyzer Measurements and Noise,*** Application Note 1303, Agilent Technologies, 2000, document number 5966-4008E, www.agilent.com.

Those documents all complete the necessities of various precisions of spectrum analysis besides giving way to some slight salesmanship for Agilent and Rhode and Schwarz. The first one is heavy on mathematical formulas while the Agilent documents (done with smaller font) go for more examples in the practical world. The Application Note 150-1 covers the basic modulation types which was presented slightly differently in Chapter 4.

Chapter 38

Military Radio Systems and History

Receivers, transmitters, transceivers, and test equipments can each be called stand-alone *systems* that perform a specific communications function. Their components (circuits and sub-systems) and structure vary depending on the available state of the art at the design time. A review of the state of the radio art with some selected military radio sets is useful; those illustrate the systems concepts, innovation and how they came to exist.

In the Beginning

Radio was born in 1896, then described basically as primitive.[1] A receiver was merely a tuned detector feeding a pair of earphones. A transmitter was invariably a damped-wave tuned arc (*spark)* turned on or off by a telegraph key. In some commercial radio installations the RF was created by high-speed alternators generating in the VLF part of the EM spectrum. Radio frequencies were considered high at MF, common at LF; large wire antennas were the rule, necessary for the typical 20 to 150 miles of any radio link. Amplifying vacuum tubes would not be in production until after 1910. Voice modulation had been done in 1906 but as a one-way experiment, not viable for any sort of broadcasting. All communicated intelligence via radio was done by on-off keying telegraphy codes.

Early radio worked. It could reach beyond the visible horizon and allow near-instant wireless communication to relatively distant locations, particularly to ships at sea. Unfortunately, large antennas were required and the radio equipment was fussy in maintenance with relatively poor performance. The amplifying vacuum tube changed all that. Detected audio from *crystal set* diode detectors could be amplified for more sensitivity. An *RF amplifier* could be added ahead of the diode detector to yield even more sensitivity. In transmitters a tube *oscillator* circuit could produce relatively clean RF energy at a selectable frequency, perhaps amplified to higher output by another tube or tubes.[2] The *MOPA* or Master Oscillator Power Amplifier two-stage basic transmitter structure was born.

Edwin Armstrong's *regenerative detector* allowed one-tube receivers to approach 100 μV

---

[1] The first public demonstrations of radio communications was done in Italy by Marconi and Popov in Russia, both in the year 1896. Most radio histories credit Marconi as first in that year.

[2] Spark transmitters usually had a sort of amplitude modulation from rapidly moving arc contacts, usually a low frequency growling sort of note. This helped identify the signal in the diode receiver, basically just an AM detector. Sideband and harmonic content of spark transmitters was great.

sensitivity levels through controlled positive feedback. With careful adjustment of that feedback, a regenerative detector can produce the *beep* of an on-off keyed RF signal. Such *regens* can detect both AM and on-off keyed CW signals. Unfortunately, regenerative detectors also acted as small transmitters so an *RF amplifier* stage was added between the antenna and detector. That shielded much of the detector's regeneration from being re-radiated out the receiving antenna.

Armstrong was granted the first patent for the *superheterodyne* receiver in 1920, the basic architecture form for nearly all radio receivers to follow.[3] Heterodyne mixing or frequency translation had already been accomplished in the beginning of wireline *carrier* equipment for multiple voice channels on a single pair of longer-distance telephone lines. That also marked the beginnings of *single sideband* AM (SSB) techniques using FDM or Frequency Division Multiplexing in telephony. SSB would be applied to HF radio by the beginning of the 1930s, using Type C carrier equipment that allowed four voice bandwidth (3 KHz) channels in a 12 KHz bandwidth.[4]

By the mid-1930s AM broadcasting on MF had become an established fact and single-channel FSK modulated HF transmitters were carrying the bulk of long-distance messages, on-off keyed CW primarily for medium-distance maritime or land wireless communications. Teleprinter telegraphy had taken over landline manual telegraphy services. A few municipal police departments in the USA had installed one-way and two-way AM voice communications with patrol cars on a band just above the MF broadcasting band. The fledgling air transport services had begun using AM voice radio through airlines or private communications services and had minimal radionavigation through *A-N beacons* on LF.[5] The international distress frequency was 500 KHz with on-off keyed international morse code as the required mode.

Most-used (per transmitter) modes in 1930 were on-off keyed CW, frequency-shift-keying CW, and AM voice/music. Radio frequencies were mainly from LF to HF; VHF and above was largely considered strange, unexplored territory. Receivers were nearly all single-conversion superheterodynes, manually tuned. Transmitters had a typical RF stage chain of oscillator, frequency multiplier (if required), final amplifier. If AM was used an audio power amplifier was

---

[3] It was invented when Armstrong was on US Army duty in Paris, France, in 1918 after the Armistice, but patents take a while to go through the Patent Office mill. It took several years for other countries to recognize that patent after 1920. The name with the *super* prefix apparently came about due to earlier receiver design by Reginald Fessenden and his *heterodyne* detector. That receiver used a low-power *spark* generator tuned to a frequency near the desired signal and the heterodyne action yielded greater sensitivity. That can be considered the precursor to the later *BFO* or Beat Frequency Oscillator used for creating the beep in received radiotelegraphy.

[4] *SSB* would denote commercial or military four-voice-channel radio "C Carrier" internationally for the next two decades until single-voice-channel single sideband radios began in the 1950s. Commercial-military SSB, typically used teleprinter telegraphy carrier sub-FDM of four to six channels in one 3 KHz bandwidth or 8 to 12 teleprinter channels in 6 KHz. The remaining 6 KHz was split between one voice channel as an *order-wire* or control-command circuit for the radio link, the other voice channel for radiotelephone communication through a wire switchboard. Type C carrier equipment is described as having a 30 KHz bandwidth but it was most often installed as a 12 KHz four-channel system per long-distance wire pair.

[5] Modern ATC or Air Traffic Control did not begin in the USA until after WW2 although the Army Air Corps had established a primarily-military airways system during the war. A-N Beacons used a two-lobed antenna pattern on LF with one lobe transmitting a repeated morse code *A* (dot-dash) while the other lobe transmitted an *N* (dash-dot); an airway direction azimuth was at the centerline between lobes where the signal was constant (dot-dash blended with time-synchronized dash-dot for a steady tone).

added to vary the final amplifier plate voltage in accord with audio, that causing an amplitude modulation of the RF output. Transmitter oscillators were quartz crystal controlled for broadcasting and fixed locations while mobile, vehicular, and many HF transmitter oscillators were manually tunable. Most of the powerful VLF and LF on-off keyed CW transmitters retained their Alexanderson alternators for international communications. SSB on HF was gaining popularity and, while traffic was considerable, was known to only a small percentage of all radio users. Radio equipment design and general knowledge was temporarily stratified for at least the next decade. Change was on the way.

FM radio was undergoing experimented both for broadcasting and in vehicular applications such as police cars. However, FM needed wider bandwidth channels for broadcasting to fulfill its promise of less noise and better sound fidelity. Edwin Armstrong once again pioneered in radio with experimental FM broadcasts at carrier frequencies just above the top end of HF. In vehicles, the FM receiver's *limiter* stages removed impulse noise and insured a constant audio output without needing any automatic gain control.[6] Mobile antennas were more efficient at higher frequencies and shorter wavelengths. All-electronic television was proved feasible, its black-and-white pictures better than the mechanical-scanning systems proposed in the 1920s. TV bandwidths would require operation at VHF and higher so the VHF region of both antennas and tube circuits was a new focus of experimentation in laboratories.

Radar was a hot but rather secret area of research, particularly for military purposes but also tried for more peaceful maritime uses in inclement weather.[7] Radar would also require wide bandwidths, preferably at UHF or higher for narrow-azimuth antenna patterns.[8]

Almost Everything Military a *Crystal Set*

In the exciting new research area of the 1930s there was little mention of quartz crystals in frequency control of tuning. For fixed or semi-permanent frequencies, quartz crystal oscillators are ideal in portable military radios for holding frequency under severe environments.

Bell Telephone Laboratories and the General Electric Company established quartz crystal laboratories in 1923 but the remaining USA companies in the quartz frequency control business were perhaps three dozen and all with just a few employees each. Raw quartz was obtained mainly from Brazil; man-made synthetic quartz growth processes would not appear until 1958. AM

---

[6] Pre-WW2 vehicle ignition systems had considerable broadband RF impulse noise from various arcs (spark plugs, DC generators, ignition coil interrupter and distributor contacts). This was alleviated in post-WW2 years by designing-in resistance-lead sparkplug harnesses, AC generators with rectifiers for DC, electronic ignition, and, in some vehicles, shielded wiring.

[7] Harbor control in France, trials of ice location at sea on British ocean liners.

[8] Microwave frequencies are preferred for 1 to 3 degree antenna widths, also very high peak pulse power that would be possible only with the *magnetron*, a magnetically-biased diode vacuum tube in a tuned cavity, capable of MegaWatt peak powers but using an average power of only a few hundred Watts. Radars would also require single-conversion superhet receivers with a *klystron* tube as the LO and an IF in the neighborhood of 60 MHz center frequency. Radars would retain that essentially simple radio structure for the next three decades until new military applications required fast radar frequency hopping through new techniques to avoid jamming. The first radars of the late 1930s were in the 100 MHz and up frequency region using elaborate triode tube transmitters.
Radar development during WW2 is well documented elsewhere.

broadcast stations had gone to quartz frequency control before the new USA Federal Communications Commission was established in 1934.

The US Army decided in 1939 that quartz crystal control was essential to Army communications and set up a civilian-military advisory group to insure a supply of quartz crystals. In 1939 the total USA production of such crystals was about a hundred thousand annually. Millions would be needed if the USA ever got into a widespread world war.[9]

The USA was already supplying some crystals to Great Britain for use in their airborne VHF AM radio, the predecessor to the US SCR-511. That radio used four crystals for transmit control and four for receive. The US Army had started in on the *SCR-508* series of *tank radios*, a family of high-HF, low-VHF FM voice radios having ten preset channels selected by robust push-buttons in each receiver and each transmitter.[10] About six different SCRs existed in that family and each set had a case of 72 quartz crystals mounted in FT-241 holders, marked by channel number (for that radio family). Crystals were in the 350 to 550 KHz fundamental frequency range although the RF output frequency range was 30 to 40 MHz range. Modulation was actually PM but the necessary multiplication resulted in an acceptable modulation index equivalent to FM.. Thousands of sets were built and fielded.[11]

The same type of FT-241 holder crystals were produced for the AN/TRC-1, -3, -4 VHF radio relay sets operating in the 70 to 90 MHz range. Those sets had separate receivers and transmitters, each having crystal control, again with cases of crystals per set. Operational by 1943, the radio relay sets proved useful to quickly establish gaps in landline communications. They could handle 12 KHz bandwidth of Type C carrier equipment to carry four simultaneous voice channels.[12] As with tank radios, transmitter crystal oscillators were phase modulated and multiplied many times before reaching the final amplifier.

An innovation in WW2 military field communications was the SCR-536 (BC-611), the first Handheld Transceiver or *HT*, more familiarly called the *Handie-Talkie* during and after WW2. It used five miniature tubes, operating solely from an internal dry battery pack good for about 4 hours of continuous operation. It used a pair of FT-243 holder crystals and stayed fixed until changed by

---

[9] The US Army made a fortuitous decision about two years before December 7, 1941, that brought the USA into WW2. The QCS (Quartz Crystal Section) was formalized at the Pentagon a month before that attack. At that time the US Air Force was a part of the Army as the Air Corps.

[10] *SCR* is an acronym thought to stand for *Set, Complete, Radio* but may have been an invented acronym to mean *Signal Corps Radio* since the Signal Corps was the only establishing and procuring branch in the US land forces. Individual units within the SCR nomenclature (receivers, transmitters, special control boxes) were designated by *BC-nnn* with the *BC* meaning *Basic Component*. The SCR and BC designators were dropped later in WW2 in favor of a more logical and universal *AN/* (Army-Navy) prefix and a three-character set designation followed by a number. The WW2 tank radios are easily spotted by their chrome-plated push-buttons.

[11] SCR numbers are 508, 528, 538, 608, 628, and 638, each differing only slightly from the other.

[12] Typically that carrier was a CF-1, also made by Western Electric. Each voice channel could take a CF-2 carrier bay that enabled four simultaneous teleprinter loops. Each carrier bay was mounted in its transit case that could stand upright on one end and each weighed in excess of 300 pounds. Two CF-2s and a CF-1 yielded eight TTY circuits and two voice circuits with a single AN/TRC-1. Radio range was a maximum of about 30 miles in level terrain over a three-element Yagi antenna and a 30 foot mast. Those were used throughout the active phase of the Korean War, 1950 through 1953, saw fixed-station operation past 1956.

a battalion-level signalman from a case of many crystals.[13] Many other WW2 radio sets used at least one quartz crystal each.

The total military quartz crystal procurement and production between 1941 through 1945 was about 40 million units peaking at about 1 million per month. The number of quartz crystal manufacturers in the USA had grown to about 130 at the end of WW2. USA defense priority during WW2 had quartz crystal manufacturing as number 2, overridden only by the Manhattan Project (making of the atom bomb). Quartz crystal frequency control contributed greatly to maintaining communications in all military networks through the worst operating environments in the field and at sea.

Figure 38-1 SCR-194 pack radio carried by a signalman.

Arrival of the Talkies

The first *handie-talkie* came about after some Motorola executives observed US Army maneuvers in 1939 in regard to communications.[14] At that time the Army had a *sort* of portable in the SCR-194 pack radio. Using only two tubes, one as a modulated oscillator for transmit and a super-regenerative detector for receive, it operated 27 to 56 MHz with a transmit power output of 2 W; a companion unit was the SCR-195, same characteristics but a frequency range of 53 to 66 MHz.[15]

Needless to say, those field radios of the 1930s were less than spectacular. They were fussy to tune and difficult to stay in calibration. Added weight (see photo) meant a signalman could not be as mobile on the ground as a regular soldier. By 1939 the all-glass envelope miniature 7-pin vacuum tube had arrived in the electronics industry. Some of those had 1.5 and 3.0 VDC filaments that reverted to the original vacuum tube directly-heated cathode design. Those were ideal for portable radios powered by dry batteries. Motorola engineers came up with a long square box containing a 5-tube receiver-transmitter and a dry battery pack with a telescoping whip antenna. Both receiver and transmitter were crystal-controlled, microphone and earphone mounted on the case with a simple push-to-talk switch on the side. Extension of the bottom section of the antenna was the power switch.

---

[13] Changing crystals was an easy task. Alignment was aided by a small, specialized test set, just a few adjustments. According to some veterans, the mechanical disassembly and reassembly took longer than channel changing.

[14] The etymology of the name is unclear but it appeared in Motorola advertisements during the early 1940s. Presumably it is a more familiar form of *HT* or Handheld Transceiver.

[15] The technical manual or TM for the SCR-194 and SCR-195 is TM 11-238. The receiver-transmitter unit is BC-222 and BC-322 respectively. Both could operate from dry batteries or a vehicle battery source using a separate power supply.

Figure 38-2 SCR-536 Handie-Talkie.

Shown in final production form in Figure 38-2, the unit was a bit larger than a common telephone handset and could be held in one hand while communicating. The familiar name of *handie-talkie* came easily to anyone using it. AM was used on a frequency range of 3.5 to 6.0 MHz, crystals spaced for 40 KHz *channels*. While transmitter power output was about 40 mW, that was sufficient for a typical one mile range, quite good enough for small-unit field operations. A battery pack was good for four hours of continuous use, was easily changed. That battery pack took up almost the bottom 2/3 of the case interior. Its weight was just over 5 pounds and any soldier could use one. The Army was impressed and Motorola got an order for a small quantity after some changes were done to the prototype.[16] Production began in July, 1941, on what was now designated SCR-536 and BC-611.[17]

The *handie-talkie* had a maximum useable range of about 3 miles in level terrain. Vehicular radios could enable communications in the field over larger distances but a portable *manpack* radio was needed for intermediate distances. The US Army had a concept of such a radio in 1938 and had done some prototyping of an HF AM set. In 1942 Dan Noble of Motorola convinced the US Army that they needed a different approach: FM for clarity of communications and a VHF range to avoid frequency congestion.[18] Noble urged a redesign and the concept was a plateau-leap above all previous designs in the military, portable or fixed.

This new *walkie-talkie* would use FM and a 100 mW transmitter output. The advantage of FM was in keeping constant receiver output level over its 5 mile range and almost elimination of vehicle ignition noise and lightning discharge noise. Such noise had always been a problem at low HF. A hundred milliWatt output is a compromise between range and battery life, but at the 40 to 48 MHz tuning range of this portable radio, the antenna would radiate closer to the horizon angle than at low HF and thus be more efficient.

What was extraordinary in design was a manually variable tuning range requiring a minimum of quartz crystals and double conversion in the receiver. The latter was thought inconceivable for

---

[16] The push-to-talk switch was given a rubber cover to protect the interior from external moisture. The telescoping whip antenna required a special non-reflecting metal plating; then as now, such telescoping whips were chromium-plated and enemy soldiers could spot the reflections in daylight. Chromium plating is also a good resistance to abrasion in repeated extension-retraction. Some special effort was required to get a non-reflective metal coating that would not abrade. That was successful. A small cloth sling was added so that it could be carried easily over one shoulder.

[17] See also footnote 9. In the case of the *handie-talkie* being such an integral unit, the *BC* designation also stands for the full *SCR*. The BC-611 had revisions A through F from 1941 to 1945. Motorola would produce over 40 thousand *handie-talkies* in the four years of WW2.

[18] Daniel E. Noble had already engineered the first vehicular FM two-way radio system for a Connecticut city police department. Paul Galvin, president of Motorola, persuaded him to quit full-time college teaching and join Motorola in 1940. The reason was principally to help Motorola's efforts in civilian vehicular FM radios.

a dry battery power portable set. Adding to the revolutionary concept was 18 all-glass-envelope miniature tubes crammed into a small chassis roughly 5 inches deep by 10 inches wide. Going to VHF meant very careful design and selection of the tunable oscillator circuits to avoid frequency drift. An advantage in manual frequency selection was flexibility in the field, especially so in fluid mobility of ground units in action. Field units could change frequencies quickly to talk with adjacent units. Such flexibility meant another compromise in added size and weight versus time lost while the operator trudged back to headquarters for new crystals.

To understand the problem in holding to an assigned frequency, consider that the WW2

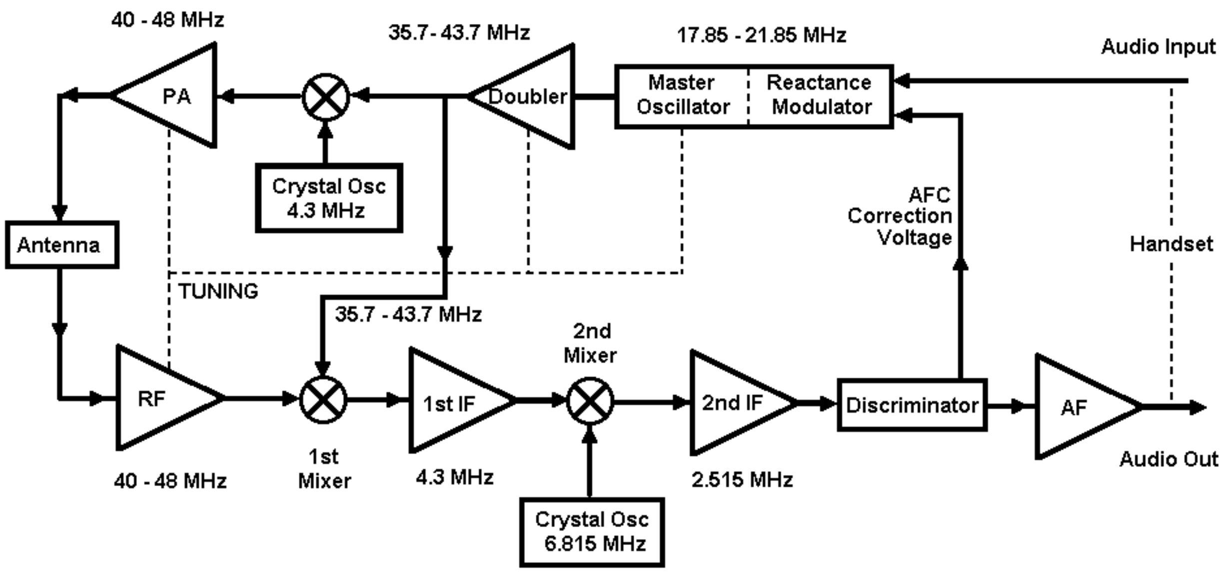


Figure 38-3 Block diagram of SCR-300 with emphasis on frequency control.

quartz crystals could hold about 200 PPM accuracy over a harsh field temperature environment.[19] That is equal to 0.02 percent. At 40 MHz this corresponds to 8 KHz in potential drift. Frequency control architecture of the SCR-300 walkie-talkie is shown in Figure 100-3.

The transmitter section master oscillator is the local channel tuning reference. After doubling it is 4.3 MHz below the required dial frequency. The transmit mixer up-converts to the 40 to 48 MHz range. On receive, the transmit doubler output serves as the 1st LO while the 2nd Mixer down-converts the 1st IF to 2.515 MHz in the 2nd IF. The discriminator circuit produced a DC output equal to any deviation of the other transmitter and that DC would become a *correction voltage* to bring the transmit master oscillator to the correct frequency. This AFC or Automatic Frequency Control handled transmitter frequency differences of up to 50 KHz. The transmit master oscillator would have to stay in calibration within about 0.1 percent of any frequency tuning setting. Operation channels were picked to 200 KHz increments for a total of 41 channels on the tuning.

---

[19] While 50 PPM is considered common for crystal drift over extreme temperatures of the 1970s and beyond, quartz technology was not fully optimized in the 1940s nor did oscillator circuits develop with any temperature compensation. 200 PPM is the worst-case tolerance of one 100 PPM crystal with another adjacent-channel 100 PPM crystal. 50 and even 25 PPM quartz crystals could be made then but at much increased cost.

To achieve FM in a transmitter of the mid-1900s there were two basic ways: Phase modulate a low frequency crystal oscillator output and multiply that frequency many times; reactance modulate an L-C oscillator (for wider deviation than PM) at a frequency near the desired one. The BC-1000 transceiver of the SCR-300 had the latter method. Reactance modulators could be DC biased to achieve the AFC action.

Figure 38-4 SCR-300 on back of soldier. Whip has semi-rigid base to allow vertical positioning. Bottom 3/5 of case is batteries.

Figure 38-5 Operator view of SCR-300 controls. Tuning by crank knob, left middle, channel dial right middle, Whip antenna base at lower right.

Receiver squelch had been brought over from vehicular FM radio use to this new portable. Radio operators had only three basic controls: Tuning, volume, and squelch. Different front panel jacks allowed different microphone-earphone combinations; the SCR-300 was shipped with a push-switch-to-talk telephone handset. With 18 7-pin miniature tubes the chassis required all components to be small (Figure 38-6).

Figure 38-6 Rear view of the BC-1000 R/T chassis. Tuning capacitor shield is long box structure at middle-left. All tubes have shields holding them in place.

Total radio set weight was about 40 pounds. A soldier's *combat pack* was about the same weight and field radio operators had a choice of other unit members carrying most pack items or adding the combat pack to the set bottom by means of straps. Despite the weight, the true walkie-talkie (SCR-300) was reliable. Some 50 thousand SCR-300s were built in the four years of WW2. The set was essentially copied by the British Army after WW2 as the Wireless Set No. 31, the Swedish Army as Model RA100, and by several other armies. The SCR-300 was in operational use in the US Army until 1965, seeing service in the Korean War and later in Vietnam.

The SCR-300 was modified as the AN/VRC-3 for vehicular mounting in WW2. This was basically a vehicular battery input voltage power supply that replaced the internal dry battery pack.

The BC-1000 R/T component was identical in both sets.[20] The case was the same size as the portable set.

One vehicular problem during WW2 was that vehicle batteries ranged from 6 (older) to 12 to 24 (newer) VDC. Auto batteries had standardized on 3 lead-acid cells for 6 VDC ignition, however aircraft had standardized on 12 cells or 24 VDC to cut down on electrical wire weight.[21] The US military standardized on 24 VDC vehicle power systems after WW2.

Since vacuum tubes generally required higher voltages for plates and screens, vehicle system power had to be changed by small motor-generators called *dynamotors* or use *vibrator* supplies that created square-wave AC. In a few designs this was not necessary such as with the BC-1206 LF radio range receiver. The 1206 used six 14 V filament tubes and those were wired in series-parallel for a 24 to 28 VDC supply input from an aircraft. Plates and screens of this single-conversion superheterodyne were made directly to the same supply. Tube stage efficiency dropped with lower *B+* but that was a compromise with much-reduced weight.

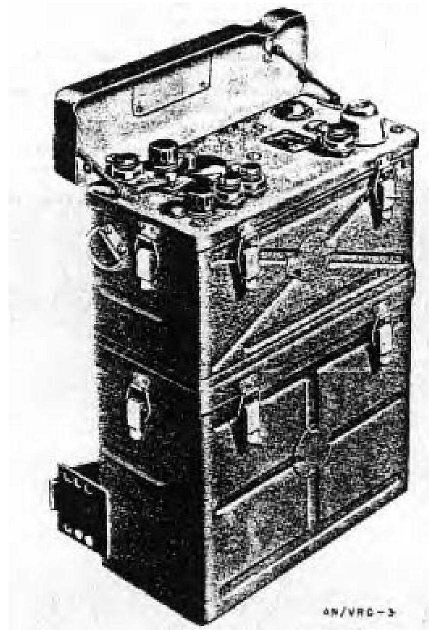


Figure 38-7 SCR-300 modified to be the AN/VRC-3 vehicular radio.

Post-WW2 Planning and New Talkies

The World War 2 period was one of intense development and production. For the USA this effort was squeezed into a relatively brief four-year span. In many designs there were compromises that could be overcome with new technology or original concepts changed to accommodate new experiences of use in a battlefield environment. In some designs radical new concepts were proven effective and retained.

The US Army, drawing on the effectiveness of FM in short-distance battlefield conditions, adopted a three-band spectrum plan over the 20 to 60 MHz range. The three bands were ostensibly allocated to artillery (20 to 27.9 MHz), armor (27.0 to38.9 MHz), and infantry (38.0 to 58.4 MHz) units. The purpose served to separate spectrum use for field needs but also provided some overlap for inter-branch liaison communications.[22] HF radio from 4 to 20 MHz would still exist in the field but intended for large (regiment and higher) unit communications over medium to long distances.

---

[20] The *SCR* designation as well as all other old military branch designations were dropped by the US military in later years of WW2. The *AN/XXX-nnn* designation was then adopted, the *AN* meaning combined branch use by Army or Navy. The SCR-300 was later redesignated the AN/PRC-2 or -4, the *PRC* meaning Portable Radio, Communication. *VRC* stood for Vehicular Radio, Communication. *GRC* was a catch-all designation for all Ground-use, standing for Ground Radio, Communication.

[21] For the same amount of load power, wiring in a 24 V system needs to carry only 1/4 the current of wiring in a 6 V system...which means lighter wiring weight.

[22] While good planning at that time, the narrow band overlaps and greater field mobility experienced during the Korean War and Vietnam following showed that a general, non-branch-specific band plan would be more effective.

Radio relay would be 70 MHz to higher frequencies.

For small-unit communications came the biggest change: The AN/PRC-6 handheld transceiver replaced the SCR-536. The AN/PRC-6 had about the same weight as the SCR-536 at 6 pounds with battery pack. Its case was contoured rather than boxy and the antenna was now a flexible flat spring. At 200 mW output on a crystal controlled frequency between 47 and 55.4 MHz it could reach the same 1 mile range but had FM clarity.

Figure 38-8 AN/PRC-6 with H-33 handset (extra).

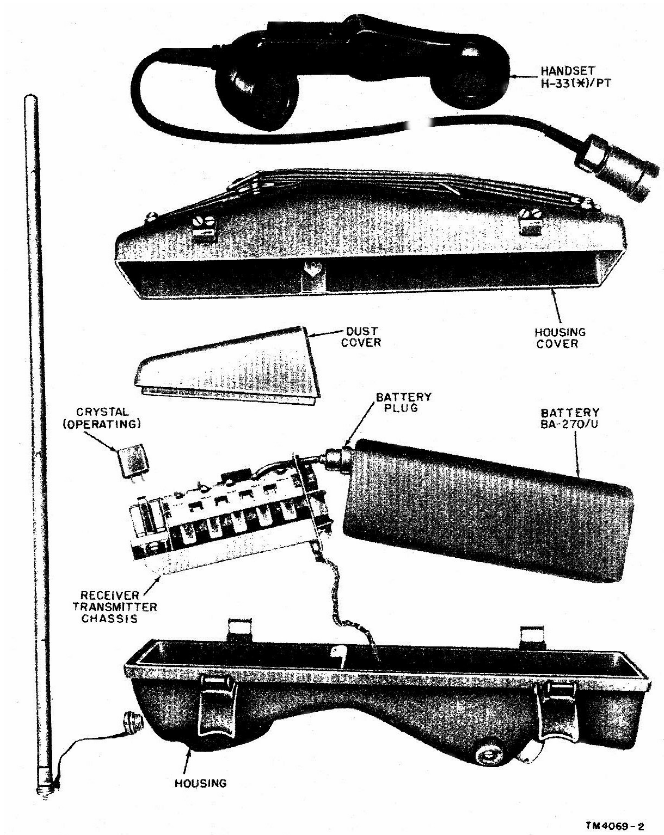


Figure 38-9 AN/PRC-6 Components.

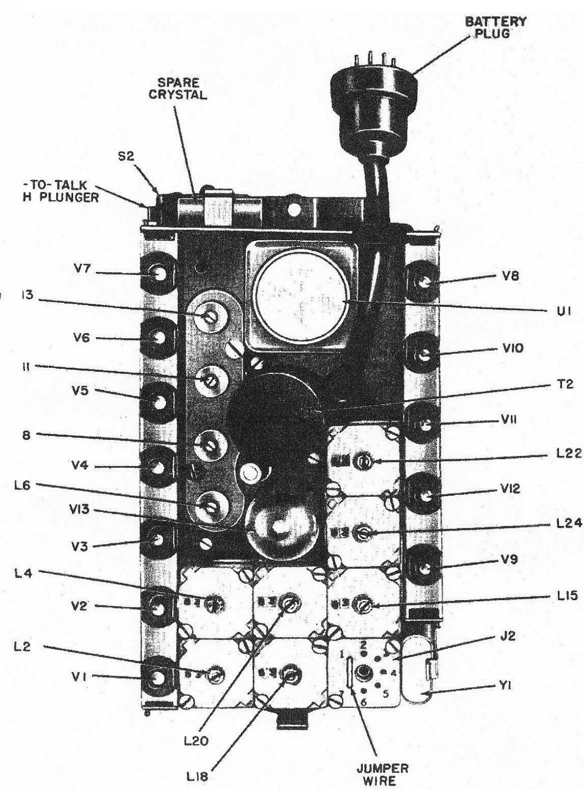


Figure 38-10 PRC-6 transceiver chassis. Very compact.

It could use the now-standard H-33 handset having a push-to-talk switch. The H-33 had a much flatter earphone end than the old SCR-300 handset and could fit under a helmet edge (see Figure 100-8). The handset connector was planned for *external control* such that it could possibly be used for remote operation or with another PRC-6 as a radio relay system. That radio relay adaptation was covered in its Technical Manual.

A total of 13 vacuum tubes were packed into an approximately 2 ½ by 5 inch chassis, 12 of them being a new subminiature *pencil tube* envelope design. Transmitter final amplifier was a 3S4 miniature, the *giant* of the tube lineup. With fixed frequency operation controlled by a single plug-in crystal (200 KHz channel increments), crystal changing also required re-alignment of 7 inductors. Each of those (see Figure 100-10) had a small 2-digit counter to aid preliminary settings. The 43-crystal set was accompanied by a small alignment test set, the case lid marked with preliminary settings for each channel.

The PRC-6 may have been the first radio design to use semiconductors for other than rectifiers, the design probably finalized about 1950.[23] The FM discriminator used two diodes in the Foster-Seely arrangement. The FM modulator was a diode, not strictly a variable-capacitance type, but part of an R-C arrangement across the transmit oscillator's L-C resonant circuit. What is surprising there is that the receiver was *always on*, even when transmitting. AFC voltage from the receiver discriminator was fed back to the frequency modulator when transmitting to keep it on frequency. Only the single-conversion receiver's LO was crystal-controlled. No separate transmitter crystal was needed. IF was 4.3 MHz, similar to the SCR-300's 1st IF.

While on in receive the PRC-6 drew 1.25 W total from the battery. A conventional 2-cell hand flashlight draws about 1.5 W. For transmit the audio output amplifier filament was turned off and the transmitter tubes' (four) filaments were turned on. Transmit battery drain was an added 4.6 W. The case push-to-talk switch consisted of a rubber boot on the case itself with a push assembly to the switch proper on the transceiver chassis. The push-to-talk switch could still be operated manually when the chassis was removed from the case for servicing.

The PRC-6 power switch was a rotary with three positions: Internal microphone and earphone and push-to-talk switch (in case), off, external microphone and earphone and push-to-talk switch (handset connected through bottom connector on case). This yielded greater field flexibility potential. While the older SCR-536 had its power switch combined with the telescoping antenna, the flexible tape antenna of the PRC-6 did not allow that feature. This was made up by the antenna being able to fold back into the small sling. The sling allowed the PRC-6 to be carried over the shoulder even with the handset connected. The H-33 handset was interchangeable with all other new field radios and weighed only 14 ounces.

Second-Generation Walkie-Talkies

The AN/PRC-8, -9, -10 manpack radios had an immediate ancestor in the SCR-300. They had essentially the same circuit structure, differing only in the receiver-transmitter units' frequency ranges. They were manually tuned (and lockable) on 200 KHz channel increments, transmitters relying on part of the receiver for AFC, had a squelch control. They differed from their ancestor in having three possible antenna selections (two kinds of whip antennas supplied or an external antenna), remote control capability (via wire up to two miles), higher transmit power (1.2 to 0.9 W depending on band), greater receive sensitivity (0.5 µV for 12.5 KHz FM deviation), single receiver conversion to an IF of 4.3 MHz and a different transmitter AFC system. Modular construction was done with about half the stages of the 16-tube set (15 pencil triodes, 1 miniature 7-pin tube as transmitter oscillator). Receiver calibration was done by zero-beating harmonics of a 1 MHz quartz

[23] Fielding of the PRC-6 took place about late 1952, the last half of the active period of the Korean War of June 1950 to a Truce established in July 1953.

crystal calibrate oscillator at the input with another crystal oscillator at 4.3 MHz feeding the FM discriminator.

While approximately the same dimensions as the SCR-300, the PRC-8, -9, and -10 were nearly half as thin, less than 3 ½ inches and had a total radio plus battery plus harness plus H-33 handset weight of 20 ½ pounds with the 3-foot metal tape antenna. The longer whip antenna with a semi-rigid adjustable base was about 2 pounds heavier. This was about half the weight and a smaller profile than the WW2 walkie-talkie, a definite plus for radio operators on foot in the field. External antennas available at the time was a loop antenna for homing (seldom used) and a drooping ground-plane on a mast (rarely used). The ground-plane had adjustable telescoping sections to match the 20 to 59 MHz range. The AN/TRC-1, -3, -4 radio relay sets introduced in WW2 had 3-element Yagi antennas adjustable in length and element spacing by markings to cover a 70 to 90 MHz band.

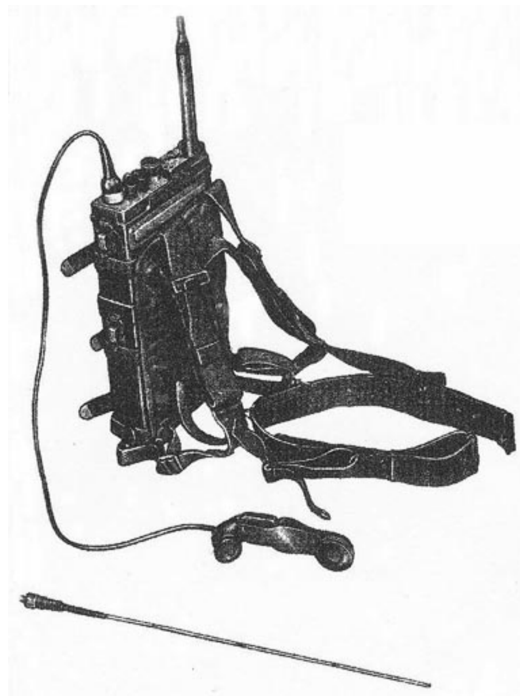


Figure 38-11 AN/PRC-8 series manpack with 2nd antenna.

Receiver RF, Mixer, and LO stages are kept on during transmit. As with the PRC-6 there is no switching of antenna to either transmitter output or receiver input. Transmitter output from the 5A6 oscillator-output stage is sufficient to cause limiting in the 4.3 MHz AFC driver-limiter (V1) of the transmit discriminator assembly (U1). DC correction voltage from the discriminator biases reactance tube modulator V2 to keep the transmit oscillator V3 on-channel. The receiver's LO is the reference frequency for both receive and transmit. It should be noted that the transmitter AFC assembly, all 5 IF amplifier stages, and the receiver discriminator are all plug-in units containing both individual tubes and stage L-C tuning. This preserves the IF tuning at 4.3 MHz from getting misaligned in the field; each stage is fix-tuned at the manufacturer. Both discriminators were Foster-Seely types using two germanium diodes each.

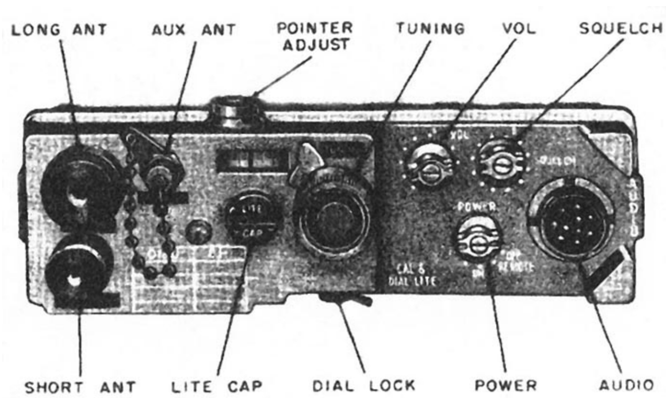


Figure 38-12 PRC-8 series front panel, channel tuning knob in moddle, three antenna jacks at left. *Audio* connector is to HS-33 handset or external audio amplifier (in vehicle mounting).

Receive-transmit switching is done by a three-pole relay driven by the push-to-talk switch circuit. This is a compromise to the ability of remote controlling the set. Squelch switching is done by a pentode fed by grid current from the 5th IF amplifier stage, its plate in series with a small relay. Those relay contacts short the input grid of the audio output amplifier when squelch is on. Opposite contacts on that relay are made to an external audio control connector for re-transmission control. Squelch is operator-adjustable and far CCW position switches it off by opening the squelch tube filaments and saving a little bit of battery power.

The battery pack supplies 1.5 VDC for most tube filaments, 67 VDC for receiver B+, and also 6 VDC for C- or *DC bias* and the transmit oscillator filament, and 135 VDC for the transmit B+.[24] During transmit the AFC driver-limiter, reactance modulator, and transmit oscillator are turned on while the receiver 1st RF stage, all five IF stages, and the audio amplifier are *turned off*. This saves battery power as much as possible during transmissions. Transmitter AFC is still possible through capacitive coupling in the receiver 1st RF stage feeding the 2nd RF stage; 2nd RF stage, Mixer, and LO remain on during transmissions. The physical layout of the PRC-8 series is more compact than the SCR-300, relying on *modular* assemblies both for size and manufacturing ability. Point-to-point wiring was still used; printed circuit boards had yet to be invented.

Front panel controls were simple and direct, befitting hasty field operations and minimally-radio-trained operators. The *power* switch had four positions: Off, external (for external control through *audio* connector), internal (normal operation), and calibrate. The calibrate position turned on both the 1 MHz and 4.3 MHz crystal oscillators plus a small dial light for night operations.

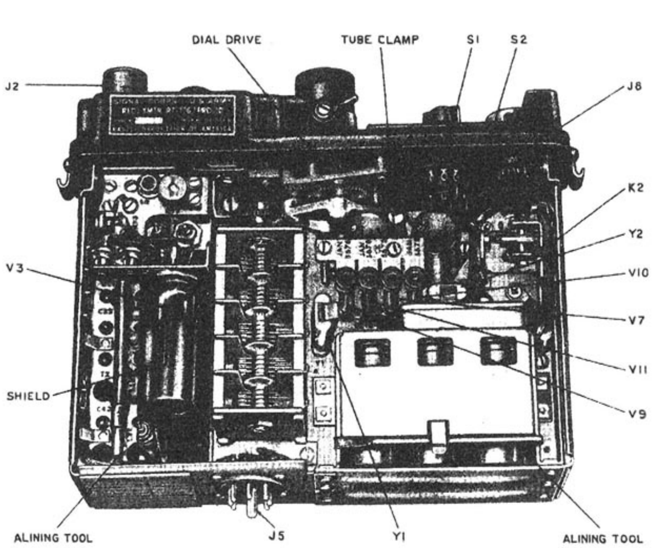


Figure 38-13 PRC-8 series chassis removed from top half of case. Tuning capacitor at middle left. Transmit output is V3.

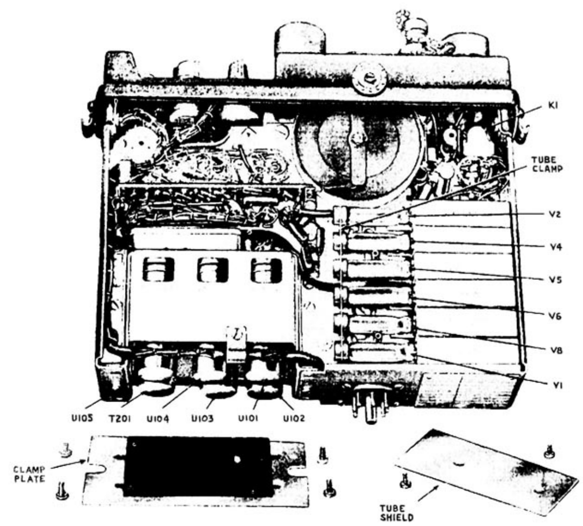


Figure 38-14 PRC-8 series chassis, opposite side view. Tuning dial drum at top middle right. Plug-in IF modules at left. RF modules at lower right.

Harmonics of the 1 MHz oscillator were good for setting the dial pointer at every 5th channel. Manual tuning accuracy need only stay within about 100 KHz or about 0.15% for the PRC-10; the internal calibrator could handle more precise dial settings. Manual channel selection provided more field operation flexibility.

The transceiver chassis itself, exclusive of front panel assembly, is approximately 8 inches long by 10.5 inches wide with a thickness of 3 inches. Sixteen tubes and a five-gang tuning

[24] This is a throwback to old vacuum tube sets were *A batteries* were for filaments, *B batteries* for plate supplies, and *C batteries* (always opposite polarity to B supplies) for negative DC biasing.

capacitor plus part of its dial mechanism had to fit into this space with point-to-point wiring. The tuning capacitor had an additional variable capacitor for impedance matching of the longer whip antenna plus a variable inductor for matching the shorter whip antenna.[25] Those are mechanically ganged with the 5-section variable, tucked part-way inside the cast magnesium front panel housing.

Designers were faced with several problems on the receiver IF and the transmit AFC system. First of all, it was the prime determinant in receiver bandwidth. A 5678 pentode (typical tuned amplifier in the PRC-8 series) had only 1.1 mmho of transconductance at a 67 VDC plate and screen supply so any tuned circuit had to have a high loaded impedance or low capacitance value for a good voltage gain. With a low tuned circuit capacitance, necessary tube replacements might require re-adjustment due to differing tube internal capacitance. Further, designers wanted to keep alignment adjustments to a minimum to simplify overall radio maintenance.

A solution was to make an *entire* IF amplifier as a total plug-in module, tube and all tuned circuits together. A tuned output circuit for the pentode could be a transformer type with a low-impedance secondary to the following stage. That minimized changes in output circuit capacitance between modules. Unfortunately, that resulted in a relatively low voltage at the module output and vacuum tubes have very high impedance inputs, being essentially voltage-driven circuits. The solution was a simple 2-section series-L, shunt-C tuned circuit normally used in low-to-high impedance matching networks. That is represented by *L1* and *C1* in Figure 38-15.
Assuming T1 primary has a loaded Q of about 40 and the total plate-to-ground capacitance is 52 pFd (47 pFd fixed plus 3.8 pFd tube plate capacity plus strays), the plate load will be about 22.5 K Ohms (120 K fixed shunt resistance plus plate resistance of 1 M Ohm). At 1.1 mmho transconductance, the voltage gain, grid to plate, will be about 24.7. However, to achieve a low-impedance module output to a (typical) 72 Ohm impedance, T1 turns ratio primary to secondary, would be about 18:1. Without any other circuit at its input, the voltage gain of an IF module would be only about 1.37, way too low for any practical amplification.

Assuming the 5678 input impedance is purely capacitive and L1 has a Q of about 40, the voltage across C1 with module *grid* pin grounded will be much higher than the module input voltage at 4.3 MHz resonance.[26] The general effect was akin to a conventional double-tuned IF transformer used plate-of-preceding-stage to grid-of-following-stage with 1:1 coupling, high impedance. The end result of the module design has two benefits: More flexibility in placement in the chassis due to low impedance coupling; an easier manufacturing assembly and final test alinement since a test jig can have low impedance terminations for a signal generator and output response instrument.

---

[25] Shorter whip is base-loaded with the inductor. Longer whip matching circuit is a series fixed inductor with a variable capacitor from whip to ground. Transmit oscillator tank circuit directly matches a 50 Ohm external antenna load. Exact values vary between the different frequency ranges of the PRC-8, -9, or -10.

[26] At a source impedance of 72 Ohms resistive, Q of L1 about 40, Q of C1 about 500, 10 mA RF at 4.3 MHz will be about 0.15 V at module input. RF voltage across C1 will be about 5.6 V at 4.3 MHz with a phase lag of 90 degrees. The configuration of L1, C1 produces a voltage step-up between module input and pentode grid. This is quite normal for a simple L-section L-C matching network. An analysis model should have both T1 and L1, C1 simulated along with their Qs as an integral unit.

Note: L1 and C1 are connected as a series-tuned circuit to the secondary of preceding module T1. If they were lossless that would represent a zero-Ohm impedance. In reality both L1 and C1 have losses, L1 much lower than C1. The impedance looking into a module's IF input pins will appear to drop noticeably at series resonance but that impedance is still a finite value. T1 must be designed so that its secondary winding matches that impedance at T1's parallel resonance with its fixed capacitor.

A single IF module probably has a voltage gain of about 27 db. With five modules in sequence, that would yield about 135 db total gain or 5.6 million total voltage gain. The last module should go into limiting with the specified 0.5 µV antenna input signal. With stronger signals the 4th IF, then the 3rd IF modules would go into limiting, desired for an FM receiver. R1 has little effect on non-limiting signal levels but it serves to rebias the control grid if it goes into positive swings and starts drawing current. Grid current through R1 to ground will put a negative potential on the 5678 pentode control grid relative to its directly-heated cathode. R1 is a form of self-bias that is probably more of a protection for the pentode's control grid considering it is a very low-power tube.[27]

The end result of the IF module design enabled it to be a plug-in replacement for any IF stage

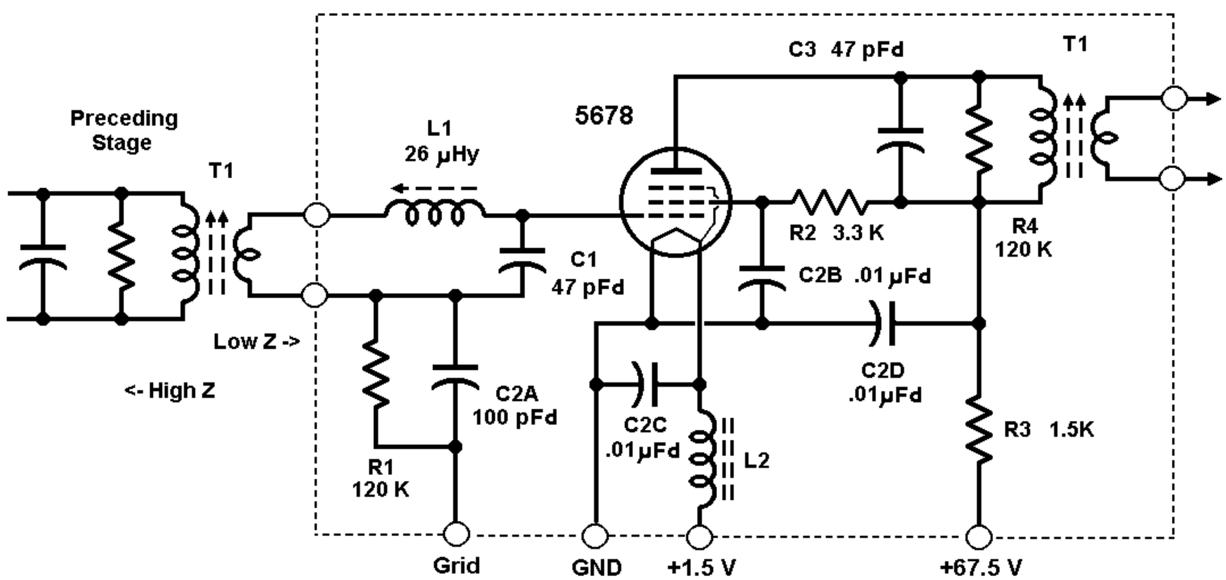


Figure 38-15 IF amplifier module, one of 5 identical modules. C2 is a four-section monolithic capacitor. L1 and T1 estimated by 4.3 MHz resonance.

whether it was used in low-level linear amplification or in high-level limiting stages. The receiver discriminator was also fixed-tuned and sealed at the factory, another plug-in unit. Two germanium

[27] That is conjecture based on not having scarce, half-century old specific vacuum tube data. The 5678 needed only 63 mW of filament power and the total plate and screen dissipation in the circuit would be about 144 mW at no signal input and 67.5 V B+. The 5678 was a military specified tube similar to the commercial 1AD4 but with half the filament consumption. By comparison, a 6AU6 limiter stage with a 100 V B+ would consume a total of 2.60 W, filament, screen, and plate. It is unlikely that a subminiature, low-power consumption tube would tolerate high control grid positive voltages.

diodes were a part of the receiver discriminator assembly, its input tuned with a similar series L-C network as in the input of each IF module.

All other stages besides the IF had tubes in sockets, physically restrained in-place. Adjustment trimmers for those stages were accessible and a small tuning tool was clipped onto the chassis for use when needed.[28] The transmitter oscillator used a 6 VDC filament miniature tube (all others were 1.3 VDC filaments) to insure a 1.0 Watt RF nominal output. The battery pack supplied this at a negative potential, that -6 VDC used as DC bias for the receiver squelch and audio amplifier stages.

The PRC-8 series was made field operational about 1952 and continued to the end of the 1960s, eventually being replaced by the AN/PRC-25, a mostly solid-state transceiver.

Communications Receivers of the early 1950s

Edwin Armstrong's 1918 patent on the superheterodyne had the patent claim that receivers could gain the same selectivity at any part of a tuning band. While that was true, practical radio receiver design up to and including the end of WW2 had multi-band receivers with non-linear frequency tuning and higher frequency bands having a greater tuning span than lower frequency bands.[29] Nearly all multi-band receivers of that era were single-conversion superheterodynes and nearly all used variable capacitors for frequency tuning.

Prior to WW2, variable capacitors had 10:1 to 8:1 capacitance ratios. Since resonance frequency is the square-root of capacitance change, a tuning frequency range of 3:1 was common on each band of multi-band (MF through HF) band-switchable *communications* receivers.[30] Unfortunately, the same physical tuning movement at an MF band of 0.5 to 1.5 MHz became 10 to 30 MHz on a high-HF band. Band-spans were 1 MHz at the low band to 20 MHz wide at the high band. While some receiver designs added shunt-series capacitor networks to attempt keeping the physical frequency tuning to the same band-spans, that added cost and required a very careful alignment of receiver front-ends. Extra components associated with the variable capacitors had to be selected to minimize frequency drift they introduced with changing operating temperature.

The *Set-And-Forget* Receiver Tuning

All of that was about to change with two developments from Collins Radio Company: The

---

[28] This might be considered wasteful by some but they should remember that field conditions do not allow transport of a full electronics shop carrying all possible tools and equipment. All designers had to keep that in mind. The clip-on special tool feature was carried along on nearly all military equipment for a half century.

[29] Not absolutely true. The National Radio Company's HRO series had front-panel plug-ins for fixed L-C values that made band spans approximately equal. While those were purchased by the US military, the manual plug-in bandswitching was a decided nuisance not considered viable for vehicle-mounted receivers. Post-WW2 National HRO designs retained their plug-in *RF drawer* concept into the early 1980s.

[30] The term *communications receiver* came about during WW2, referring mainly to covering HF, then considered the only useable long-distance frequencies for communications. The term stayed in commercial receiver descriptions for decades. By the start of the new millennium microwaves were the predominant long-distance communications frequencies.

PTO or Permeability Tuned Oscillator allowed a *linear* tuning position versus frequency.[31] In combination with multiple mixing, the first LO having fixed quartz-crystal-per-band selection, the rate of tuning was the same anywhere on the MF-to-HF receiver tuning range. This was ideal for any HF radio circuit that required several carrier frequency changes per 24 hour period. Operators could simply dial-in the carrier frequency and trust the displayed reading to be true and stable.[32]

The first of these in the U.S. military was the R-388, a militarized Collins 51J series receiver. Becoming operational about 1951, it could tune any 1 MHz wide band from 0.5 to 30.0 MHz, each band tunable at the same frequency rate to within 1 KHz of the desired frequency. The 1st IF was variable with coupled tuning over 2.5 to 3.5 MHz, that mixing with the actual variable tuning device, the PTO at 2.0 to 3.0 MHz. The last IF was fixed at 500 KHz. While appearing straightforward in a few sentences, what was absolutely required in the design was to mechanically couple the front end tuning with the PTO and variable IF.

Collins Radio had perfected the slug-tuned coil winding scheme for linear frequency versus slug position for any given frequency range. However, the inductance change results in octave band tuning. There would be 30 separate, equal-rate bands over the whole tuning range. The solution was to incorporate a large set of 26 gears, 4 cams, and a geneva wheel with a rotary bandswitch as the tuning subsystem.

Front end inductors were switched in for 5 octave-band tuning ranges. Those variable inductors were moved by gear drive and cams such that their positioning corresponded to both the selected band and the variable tuning. The ten 1st LO quartz crystals were switch-selected and the crystal oscillator output circuit designed to tune to fundamental, second, and third harmonics. In reducing the number of quartz crystals, some of the tuning bands are in *reverse*. This required a drum display on the tuning dial to indicate the coarse frequency setting.[33] While not optimum, the R-388 was a definite plateau-jump to establish the *set-and-forget* frequency tuning system that would become standard in the U.S. military for the next half century.

An Almost No-Tune Transmitter

Transmitter T-195 is the output half of the AN/GRC-19. It had a more conventional transmitter architecture to cover 1.5 to 20.0 MHz with 80 to 100 Watts of carrier power: A master oscillator, up to three frequency multiplier stages, a driver, and a power amplifier. Amplitude modulation was possible using the conventional brute-force method of an audio power amplifier

---

[31] A *PTO* is simply an inductor tuning with a powdered-iron core moving into and out of the coil center by an axial screw drive. Linearity of frequency setting is assured by careful inductor wire winding pitch and an accurate fixed resonating capacitance. That linearity is true only for a single frequency range of tuning.

[32] For 12 KHz multi-channel SSB there would be more operator settings of the *SSB Converter* units that worked in conjunction with the receiver unit. By having true and stable receiver tuning, frequency change times, or *QSY*s in radio communications jargon, could be minimized. The point was to maximize actual message throughput rather than spending a lot of time fussing with signal tuning during a QSY.

[33] The 1.5 to 2.5 and 2.5 to 3.5 MHz bands are in reverse positioning for coarse tuning via the *Megacycle* selector control (corresponding to a bandswitch of older receivers). For that reason the coarse dial scale is marked both 5 to 95 and 95 to 5. The *Kilocycle* knob was directly coupled to the PTO and at the top of the gear-cam drive chain. In 1951 the terms KiloHertz and MegaHertz had not yet become standard.

output in series with the RF power amplifier plate supply. While the R-392 operated with only a 3 Ampere drain on the 28 VDC supply, the T-195 needed 9 Amps in Standby, up to 44 Amps in maximum power demand. That is an extraordinary power range of about 200 to 1000 Watts in a vehicle!

The especial feature of the T-195 was the capability of storing up to eight preset transmit frequencies and the ability to *automatically tune* both the power amplifier output and matching of the integral 15-foot whip antenna. That gave it both excellent operator convenience and an added ability for remote operation, the operator some distance from the transmitter itself. It would have ancestors in the commercial market in the new millennium as the ubiquitous ***automatic antenna tuner*** appliance.

Vietnam War Era and Channelization

One of the near-final versions of small-unit transceivers was the AN/PRC-25, a very simple to operate VHF transceiver that had only a channel selector switch, volume, and squelch controls. Except for the transmitter final amplifier tube, it was all solid-state. A slightly later version, the AN/PRC-77 was entirely solid-state. For a while, that PRC-25 and PRC-77 held the record for the most-produced military small-unit radio, made by a few licensed foreign producers. It only lacked security in voice. Voice security needed another carry-along adapter, something not desired by ground troops in Southeast Asia.

Portable Radios Enter a New Era Past Vietnam

The AN/PRC-104 *IHFR* or Improved HF Radio (manpack version) would cover the HF frequency range with 280,000 selectable channels, have a 20 W PEP transmitter with a 3-Second automatic antenna tuner for its whip antenna. This 1974 design was entirely solid-state with a battery pack roughly one-third of the total size. At 14 pounds it did not weigh down the average soldier. Controls were simplified, frequency selected by push-button up/down switches with decimal digits of frequency easily seen. The basic Receiver-Transmitter (RT-1209) was adaptable to a variety of different systems for vehicular and airborne use.

The significant item was incorporation of an automatic antenna tuner that would match its decade-wide frequency range with a single whip antenna. There was no need to detach sections or plug in another whip for large frequency deviations. The tuner simply switched in the nearest L and C to match the output. All within three seconds.

In a Decade Complete Field Security Was Available

The 1983 design of the ***SINCGARS*** or SINgle Channel Ground-Air Radio System was a breakthrough in field-environment quartz crystal oscillator stability. It meant that a complete encryption was now possible for small-unit operations, along with air or artillery support, all without the other side being able to eavesdrop. With a single Receiver-Transmitter (RT-1523) this would become a combination of a manpack to a 400 W vehicular radio system, with variations in-between.

All mounting combinations would share the same internal logic so an AN/PRC-119 manpack could talk to an air-support craft that was SINCGARS-equipped on the same *network.* SINCGARS would be less the radio and more of a computer terminal. The network keying allowed in-clear or

frequency-hopping over its 30 to 88 MHz operating range. To further confuse the other side, hopping would happen 10 times per second. It was nearly impossible to do a fast field decryption of the system, certainly so until long after an engagement was over.[34]

About the same manpack size as the PRC-104, the PRC-119 was the subject of an *improvement program* that ***halved*** the weight and bulk to half of its original size. Final weight of about 8 pounds, the improved version would now allow anyone on the ground to carry one with relative ease. Nothing was lost of the original SINCGARS flexibility or modes, it simply tracked the lowering of power drain of modern CMOS integrated circuits.

Front-panel design has ***no*** frequency tuning. Instead there is a keypad and an LCD unit plus an assortment of function controls. Operators ***don't need to know frequencies.*** Those are set in by a master keying unit at the start of a patrol or other engagement. Supposedly an internal frequency synthesizer will set the frequency so it is up to the internal clock to keep everyone on a specific network synchronized. It does have some re-synchronization controls to put one back onto a network but its main timing comes from the always-on clock that holds through the rough temperature environment of the field.

Early SINCGARS R/Ts had connections for an AN/PSN-11 handheld GPS unit to supply accurate timing. There is no mention of later models or the *improved* version needing timing updates. Once on a network, the digital modulation scheme allows all net-participants to stay in-synchronism. That includes brief blockages from other networks who might land on a hop frequency occasionally. Such instances might result in a slight pop-sound for demodulated audio. It is possible to have many different networks all operating in the same 30 to 88 MHz frequency range, all without any interference with one another.

Future of Small Radios is Digital

The SINCGARS operating parameters were successful enough to encourage other radio systems to use it, both military and civilian. Harris Corporation's *Falcon II* or AN/PRC-117 is one such that operates over a 30 to 512 MHz frequency range in a manpack unit similar to the PRC-119. Several handheld two-way radios have been born which are compatible with SINCGARS.

By the end of the first decade of this new millennium, about 400,000 SINCGARS R/Ts had been produced, including those for airborne radios. That's a record for one particular two-way radio system of any kind.

Other Radio Systems

While this has been a history of sorts on transceivers of the small kind, the bigger sets should not be forgotten. The progression in hardware dictates that solid-state will predominate and vacuum tube architectures no longer used. Wideband, untuned transmitter modules will be there, coupled together with passive combiners and having *hot-swap* features to allow maintenance and exchange

---

[34] Digital modulation, by itself, can be decrypted in the field, provided the carrier frequency is constant. However, jumping the frequency in a pseudo-random manner ten times per second makes it most difficult to even find where the carrier ***was,*** almost impossible to find where the carrier will be in the next frequency-hop. The combination of digital modulation, encryption of the digital data stream, and frequency-hopping makes the communications security system extremely robust.

of modules while still operating.

Receivers will be essentially untuned since variable capacitors were already on the way out by the 1980 period. Tuning will be solely by Local Oscillators, themselves set digitally.

A single quartz crystal unit will be the controlling element for ***all*** frequencies. It may come to pass that non-quartz-crystal-controlled oscillators (such as today's cesium beam type) will be the future. Cost will be the factor.

SAW filters will determine bandwidth. Commonplace in TV sets and cable boxes and mobile telephones today, their cost is already down and reproducibility is simple and cost-effective.

There will probably be another round of specialized ICs, similar to the peak of digital devices, to make up the bulk of whole sub-assemblies. Cost of such specialized ICs is also down and reproducibility is excellent. If nothing else, this means headaches for electronics distributors who must determine ahead-of-time what to stock.

There is already a change in packaging, in inter-wiring, in mechanical systems, in general physical structures and this is bound to increase.

The so-called *old days* of radio with relatively huge and heavy cabinets, chassis, tubes, and hole cutters for tube sockets are already gone. May they remain so, to be curiosities of a bygone era, objects for museums, whatever their nostalgic allure to new retirees.