2012

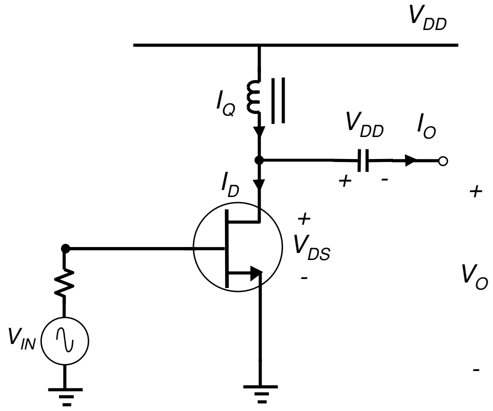
ANNUAL INTERNATIONAL COURSES IN
TELECOMMUNICATIONS
SEMICONDUCTOR TECHNOLOGY
NANOTECHNOLOGY

RF and Microwave Circuit Design: Applications and Theory

Dr. Rowan Gilmore
University of Queensland, and CEO and EM Solutions Pty Ltd., Australia

CEI-Europe Advanced Technology Education
CONTINUING EDUCATION INSTITUTE - EUROPE
www.cei.se
CEI-Europe • Linköping • Sweden • Ph +46 13 100 730 • Fax +46 13 100 731 • cei@cei.se • www.cei.se

Typical transmitter PA output stage



Besser Associates

III - 23

Consider the typical amplifier circuit above, in which a MESFET is biased through an RF choke with a quiescent drain bias voltage of V_{DD} and quiescent drain bias current I_Q . The output blocking capacitor will charge to a steady-state value V_{DD} whenever the output voltage V_O swings low, and assuming it is a large enough DC blocking capacitor, will remain charged at that value throughout the entire RF cycle. We may then write

$$V_{DS} = V_{DD} + V_O$$

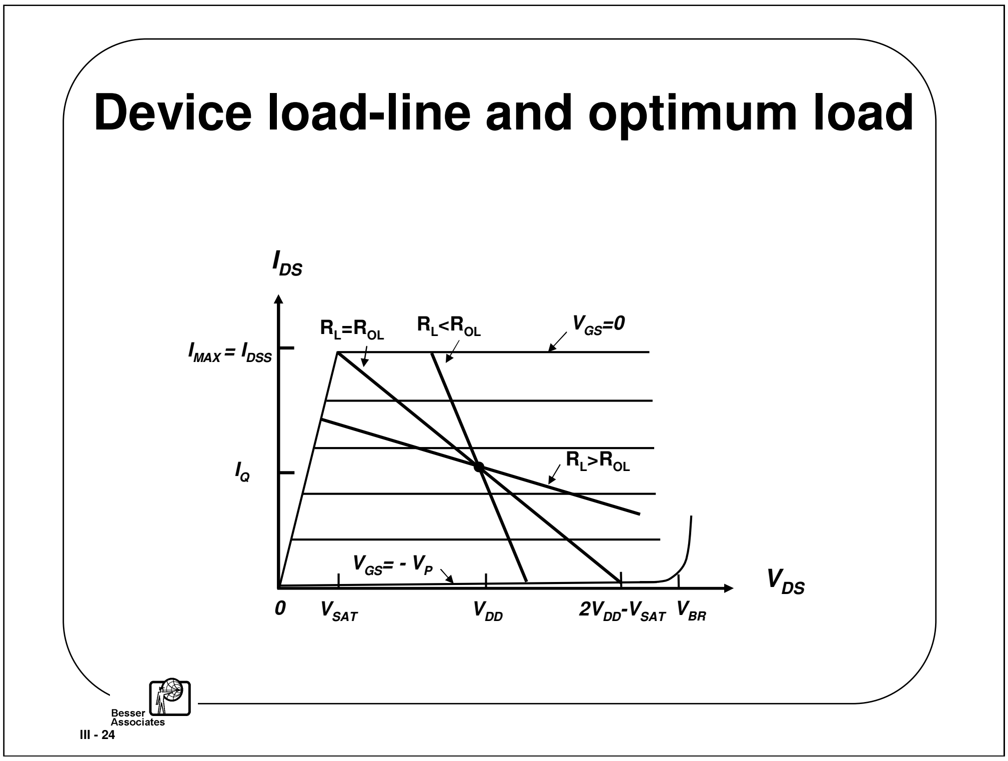
using capital letters to indicate total voltage and current (DC included). If we take incremental quantities instead, we see that $v_{DS} = v_O$ i.e. maximizing the drain voltage swing also maximizes the load voltage swing. The current through the RF choke will also be constant and we can write

$$I_Q = I_D + I_O$$

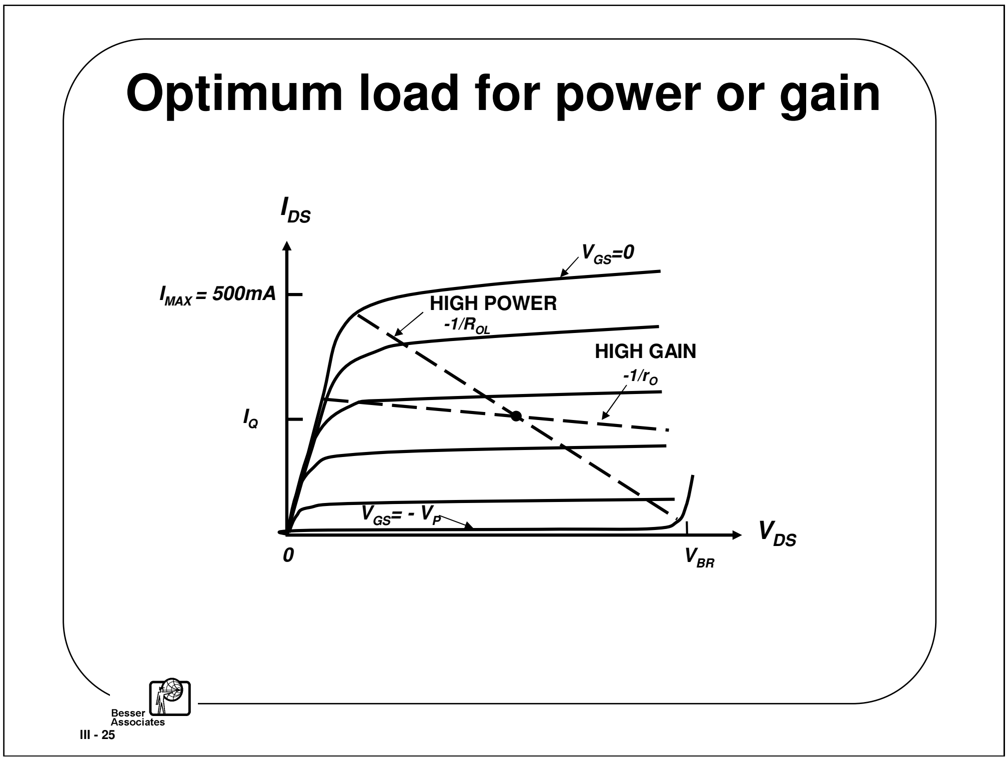
Again if we take incremental quantities $i_D = -i_O$, so maximizing the drain current also maximizes the current in the load.

We now impose a constraint on the device output voltage and current by introducing a load impedance $Z_L = V_O / I_O$, so that we now require

$$\begin{aligned} I_D &= I_Q - I_O \\ &= I_Q - \frac{V_{DS} - V_{DD}}{Z_L} \end{aligned}$$



This last equation is a fundamental equation describing how the transistor output current I_D changes with voltage V_{DS}. The circuit now imposes a boundary condition on the MESFET drain current and voltage, forcing the current and voltage to lie along the trajectory defined by the equation. This trajectory is centered around the bias point of the device, for when $V_{DS} = V_{DD}$, then $I_D = I_Q$. The trajectory is known as the load line, and if Z_L is a real resistance of value R_L, the slope of the load line will equal $-1/R_L$ and its length will be determined by the amplitudes of the current and voltage swings at the drain, since in the above circuit $Z_L = V_O / I_O = v_O / i_O = -v_{DS} / i_{DS}$.
It is normal to superimpose the load line on the I-V curves of the device itself, as shown for an FET above. In the first instance, because the reactive output parasitics of the device are invisible at DC, the load line is essentially a DC relationship, in which the voltage at the intrinsic device terminals i.e. across the output current source, is now constrained to obey the equation.

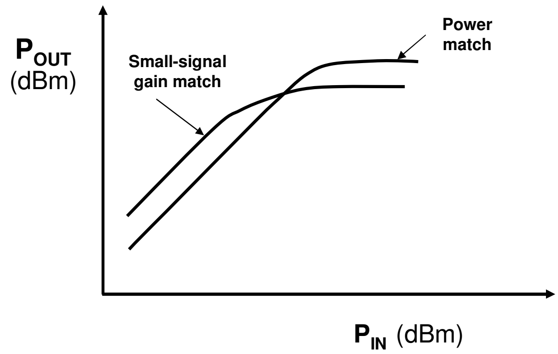


Consider now the case when $R_L = R_{OL}$. If we assume a sinusoidal gate voltage, then as the gate voltage swings positive from $-|V_P/2|$ to zero, the drain current can rise from $I_{MAX}/2$ towards I_{MAX}. The gate voltage can in fact instantaneously swing slightly positive to the point of the gate-source diode entering forward conduction, and although this increases the maximum current swing I_{MAX} beyond I_{DSS} it will increase the distortion and reduce the lifetime of the device. The output voltage is constrained to lie along the load line given by the equation we calculated, and as shown in the figure, will fall from V_{DD} to the knee of the curve V_{SAT}. Assuming sinusoidal output voltage and current across the load can be maintained by a circuit with reasonably high Q at the collector, and neglecting harmonics, the resulting output current and voltage are sinusoidal with zero-to-peak (or 'peak', for short) amplitudes $I_{MAX}/2$ and $V_{DD} - V_{SAT}$ respectively. When the gate voltage swings in the opposite direction down to pinchoff, the drain current falls from $I_{MAX}/2$ to zero, and the output voltage rises from V_{DD} to $2V_{DD} - V_{SAT}$. The RF choke permits the maximum collector voltage to swing to almost twice the rail voltage, unlike the low-frequency DC-coupled audio amplifier where the choke is replaced by a collector resistor. Here we have constructed a typical Class-A amplifier, in which the device is always conducting and the load-line remains within the active region of the device.

The output resistance R_{OL} can be calculated from the slope of the load line derived from its endpoints, giving

$$R_{OL} = \frac{2(V_{DD} - V_{SAT})}{I_{MAX}}$$

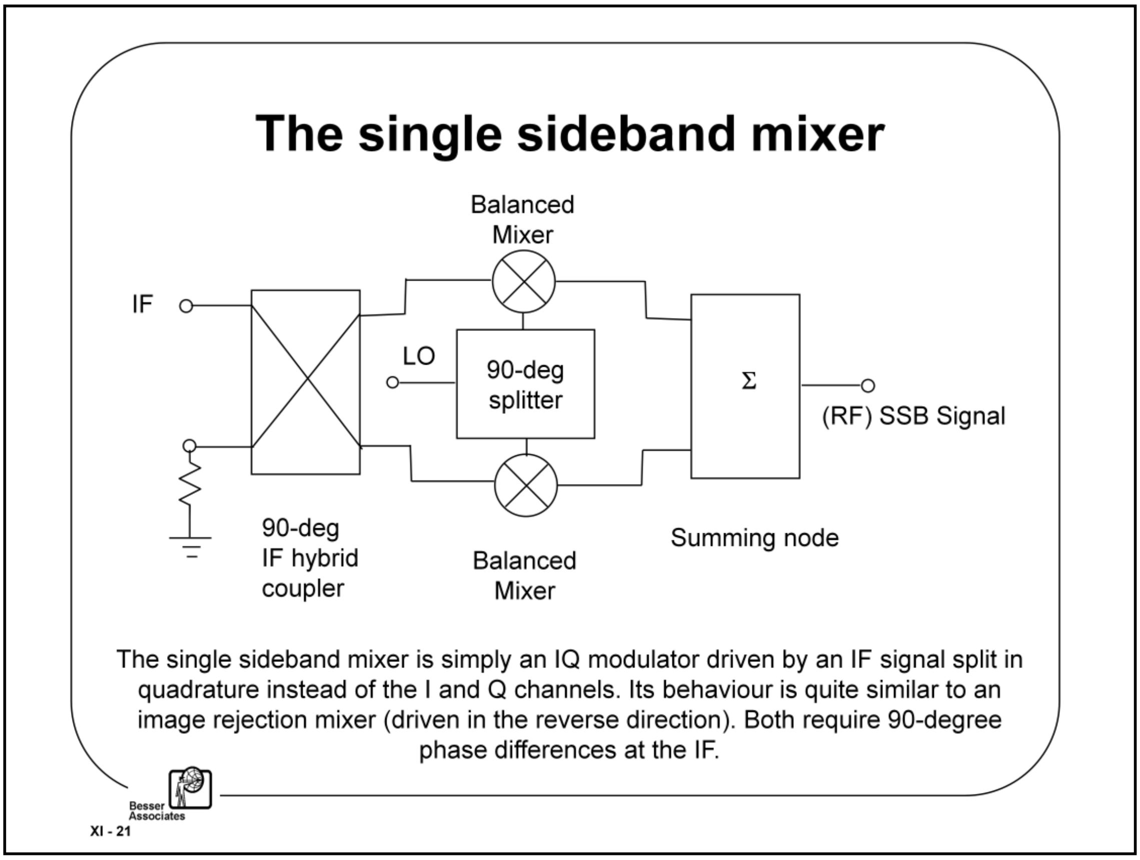
The importance of the correct match



Besser Associates
III - 26

The load line for maximum output power shown on the previous figure had slope equal to $-1/R_{OL}$, and is set by maximizing the drain current and voltage swings for the given bias voltage. In this case, the maximum current is 500mA corresponding to I_{DSS} and the maximum voltage swing is between the knee and the onset of breakdown. It is clear in this example that the slope of the load line for maximum power is considerably steeper than that for maximum small-signal gain, and that here, $R_{OL} < r_O$. What is the difference then between the two lines? The figure above helps clarify.

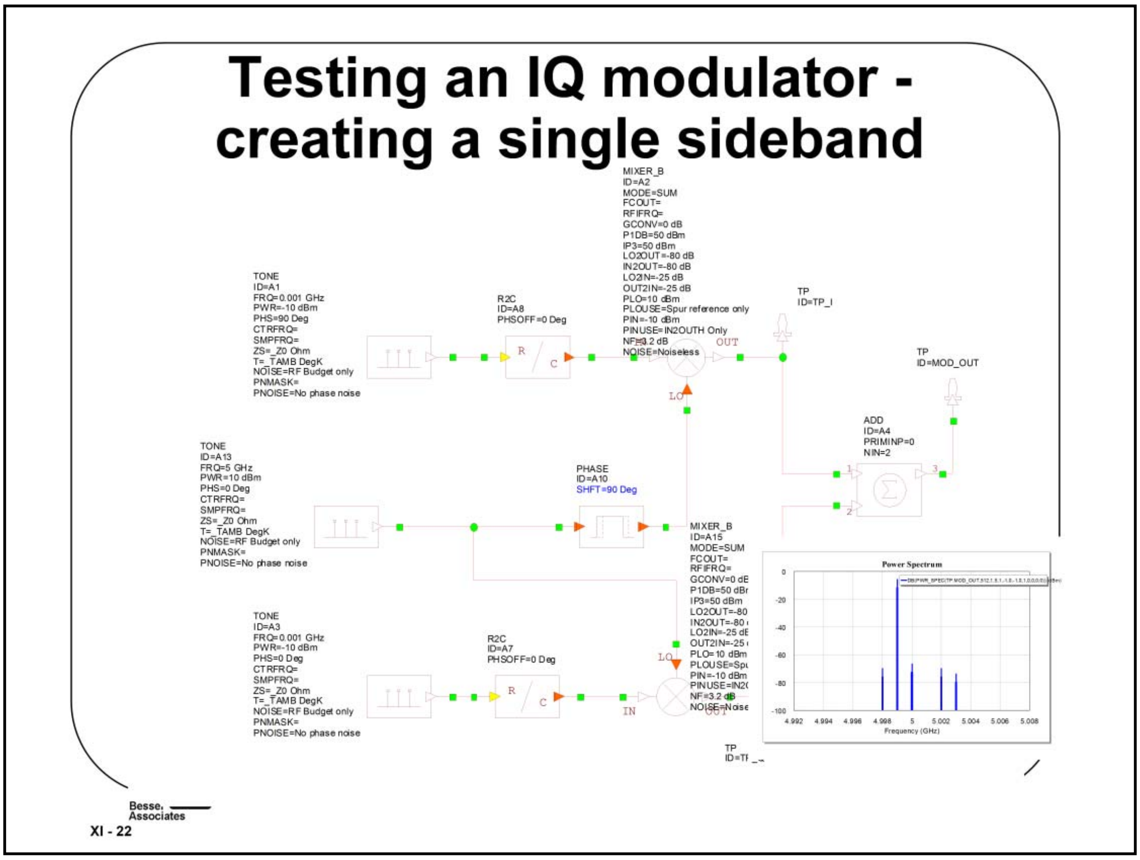
At small signal levels, the load resistor r_0 must give more output power than R_{OL}, because the former is optimized for maximum small signal gain. The output current source appears to have a source impedance of r_0 at small-signal levels, so for maximum power transfer at small signals, the load resistor will also be r_0. However, the picture is quite different at large signal swings, where the ultimate limits on current and voltage across the output current source limit the output power. R_{OL}, by constraining the voltage and current in such a way as to define a trajectory that captures the maximum possible simultaneous voltage swing and current swing, forces the current source to assume an internal impedance of $V_O / I_O = -R_{OL}$, if current is defined as coming out of the device. Thus at high power levels, the output power (and consequently the large-signal gain) with R_{OL} is higher than with r_0. The saturated output power with r_0 is not as high, because (in this example) the output current swing is not as large as it could be. It is important to recognize that for the small-signal design, further increase in drive will not increase the saturated output power because the device is voltage limited. This is fairly typical when a small signal amplifier is driven into saturation, because the small-signal output resistance is typically higher than the optimum load resistance. As a consequence, the saturated output power and even the 1-dB compressed output power of a small-signal amplifier driven with large-signals will be less than the maximum achievable for the device employed.



At first glance, the image reject mixer and the I-Q modulator appear to be similar devices, in that both consist of a pair of mixers driven by a quadrature LO signal and with an RF either applied to, or derived from, both mixers in phase. The key difference is the presence of the *quadrature hybrid coupler at the output* of the image reject mixer to yield the IF, while the IQ modulator has a *sum port at the output* where the two RF outputs from each mixer are added together in phase. As well, an IQ modulator applies two (essentially “quadrature”) channels, an I and a Q signal, as *inputs*. Essentially, one operates in the reverse direction to the other, but both require similar quadrature capabilities.

One way of testing an I-Q modulator is to drive its input with two sinusoidal test tones, in quadrature with each other, in place of the I and the Q channels, as shown in the figure above. You will note from the figure that this is not dissimilar from an image reject mixer driven in reverse, where the RF output is now the output signal of interest. Perhaps then, it is not unexpected that the RF output in such a case will consist of only a single sideband about the LO, offset from it by the frequency of the test signal (i.e. the output will lie at one of either the sum or difference frequency of the test tone and the LO). In fact, a similar analysis to that we used for the image reject mixer shows how the cancellation arises.

Perfect cancellation requires the I-Q modulator to have perfect quadrature and no DC offset. Any amplitude or phase imbalance in the I-Q modulator manifests itself as non-ideal suppression of the unwanted sideband (measured in dBc, with respect to the other sideband). Any DC offset in the IQ modulator shows up as LO feedthrough, since in an ideal mixer there is no output signal at the LO frequency, only at the sum and difference frequencies. A DC component, with zero offset frequency from the LO, therefore shows at the output as the LO frequency itself.



In this example, the I and Q channels of the I-Q modulator have been replaced with two sinusoidal tones (at 1 MHz each), the second with a 90-degree phase offset compared with the first. The LO signal, at 5 GHz, and the sideband at 5.001 GHz are suppressed at the output (they lie in the numerical noise), while the output sideband at 4.999GHz is the dominant sideband.

Try adjusting the quadrature on the LO signal, and the gain on each mixer, to simulate amplitude and gain imbalance. Note the effect on the suppressed sideband. Try reversing the angle of the input phase shift.

Normally an I-Q modulator simply translates the DC-centred I and Q channel spectra up to the LO frequency, and “orthogonally” sums them. Because the inputs are centred at DC, there are no output sidebands around the LO, only the baseband spectrum. However, the “upper” and “lower” portions of the channel spectrum around the LO are not symmetrical. This is because although the individual I and Q spectra are each symmetrical around DC, they are different from each other and when orthogonally summed, possess asymmetry in the two halves of the up-shifted composite channel spectrum.

Therefore, when the I and Q baseband signals are replaced by orthogonal tones at a single frequency (offset from DC), distinct individual sidebands are created either side of the LO at the output of each of the mixers. However, the resultant “orthogonal” summing at the output will cancel one of them.

Bipolar mixer collector waveforms

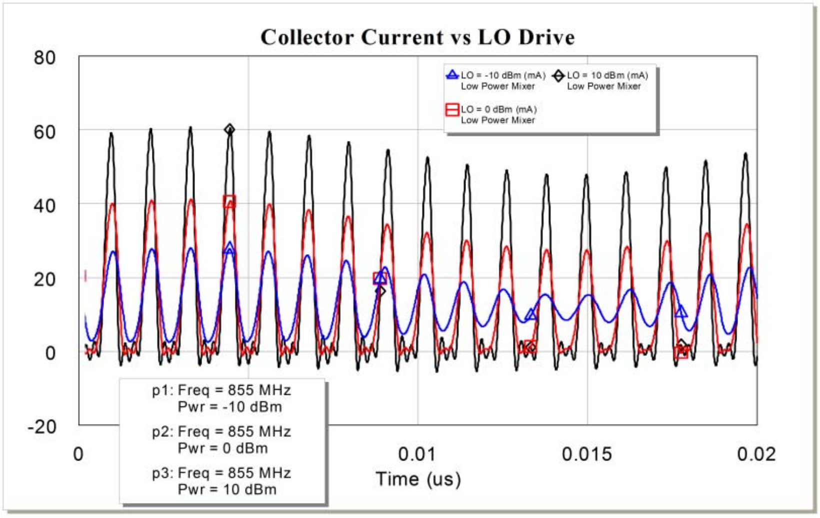
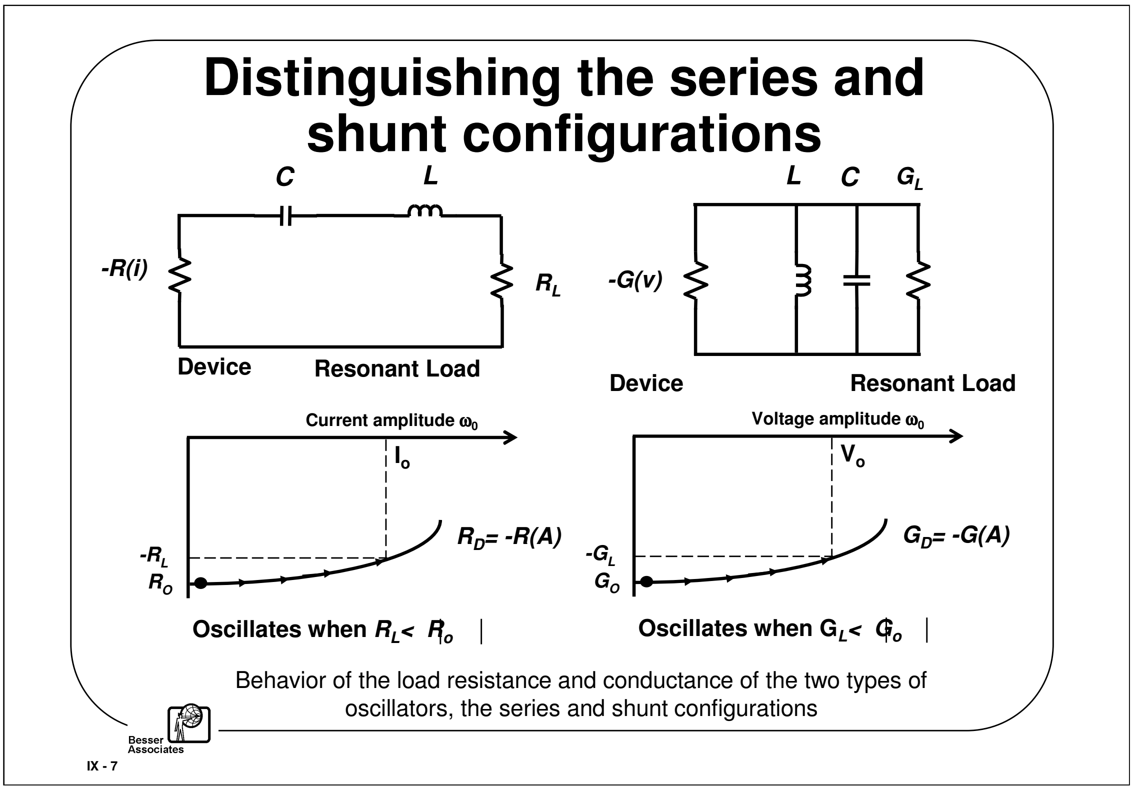


Figure VIII-20. The collector current of the BFP640 mixer at three different LO drive levels. The RF input power remained constant at -30dBm

Besser Associates
VIII - 28

The operation of the mixer is best examined at first in the time domain. The waveforms of the collector current at LO drive levels of -10dBm, 0dBm, and +10dBm are shown in the figure above. The RF input power is held constant at -30dBm. It is apparent from these that as the LO drives the device more large-signal, the average bias current increases, as we noted in the previous figure. At -10dBm, the envelope of the LO output collector current is modulated fairly uniformly by a slowly varying beat frequency, with period of .022 μs, corresponding to the 45MHz IF. The amplitude of the LO is too small to drive the zero-to-peak current swing larger than the bias current of 13mA, so the device is not driven into cutoff on negative half-cycles. However, as the LO input drive increases, the device is driven increasingly in Class-B operation, switching off on the negative peaks of collector current. The amplitude of the envelope also increases together with the IF output power.



We need to be able to discern whether the device behaves as a negative resistance device in which the negative resistance becomes less negative with increasing signal (current); or as a negative conductance device in which the negative conductance becomes less negative with increasing signal (voltage). This figure summarizes these two dual one-port oscillators we have discussed so far: the first a series feedback oscillator, that will not oscillate with large load resistances if the loop resistance $(R_L + R_D)$ is always so positive that $\zeta > 0$ and oscillations can never build up; and the second a shunt feedback oscillator, that will not oscillate with large load conductance if the loop conductance $(G_L + G_D)$ is always so positive that $\zeta > 0$ and oscillations can never build up.

Since the device impedance or admittance is a function of drive level. If the amplitude of the signal swing is characterized by its amplitude *A*, the impedance looking into the output port of an oscillator is a strong function of that amplitude. In the case of the series circuit, the amplitude is taken as the current that flows around the loop. At the frequency of oscillation ω_0, we define

$$Z(A)\big|_{\omega_0} \quad -Z_D \qquad (0.1)$$

where $Z_D = R_D + jX_D$ is the measured or simulated device impedance seen looking into the output of the active device. The terminals of interest at which we choose to split the so-called 'device' from the rest of the circuit can be chosen fairly arbitrarily. However, it is best to associate the bias network and any terminating impedances with the 'device', and to associate the resonant circuit, or main frequency determining components of the oscillator with the 'load' shown as R_L in the figure. It is clear that the figure is overly simplistic in that there the resonant circuit is represented by a simple series LC circuit, and the resonator losses and actual load impedance are lumped into a single 'load' resistor R_L. Usually the device itself will also include a reactive component jX_D that will probably vary with drive level as well. This concept of splitting an oscillator into a 'device' or active part, and a load or resonant part, or into an 'osci' and a 'llator' as one author has suggested [5], proves useful.

Thus (0.1) becomes

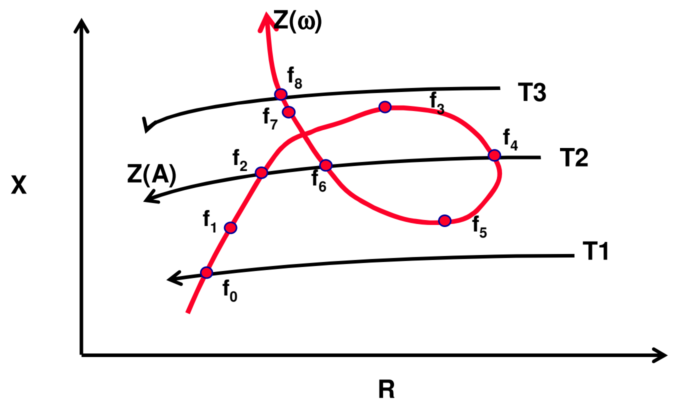
$$Z(A)\big|_{\omega_0} = R(A) + jX(A) = -R_D - jX_D \qquad (0.2)$$

measured at the oscillation frequency. If the total 'load' on the device is then characterized as a function of frequency (since it is a linear element, and presumably frequency sensitive since it contains the resonator), we may rewrite the conditions for steady state oscillation as

$$\begin{aligned} R(A) &= R_L(\omega) \\ X(A) &= X_L(\omega) \end{aligned} \qquad (0.3)$$

The dual equations apply in the case of the shunt oscillator.

The effect of load variation on oscillation frequency



A VCO, in which the device line is tuned with bias or temperature, shifting the oscillation points

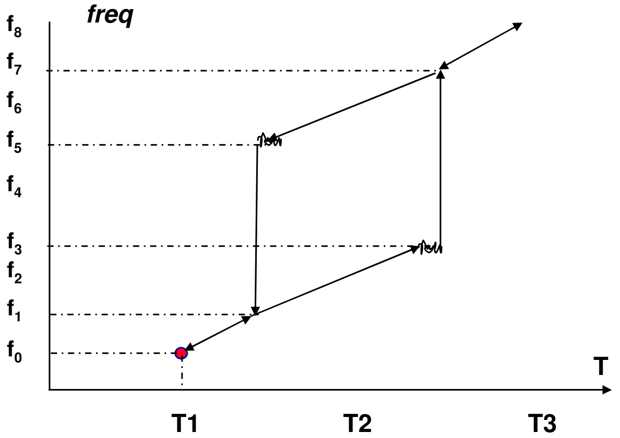
Besser Associates

IX - 9

This figure illustrates this. Suppose the device line $Z(A)$ shifts upward due to some environmental effect. This effect could be temperature, or perhaps changing the bias on the device in order to tune the frequency. In any event, as a result of changing the tuning parameter from T1 to T3, the device reactance changes and the points of intersection with the line $Z_L(\omega)$ shift. Thus at T1, there is a single steady state oscillation frequency f_0, at a relatively high signal level (because *A* is high). This is a stable operating point and because the crossing angle is close to 90 degrees, there is relatively low phase noise at this point. However, as the tuning curve is increased, to a position between T1 and T2, multiple intersections being to occur. However, the frequency will remain along the same section of load line. For instance, at T2, the oscillations will remain at f_2 because this is a stable operating point and there is no perturbation that can push the frequency to f_6 or f_4. Although f_6 is a valid (stable) operating point, the oscillator will remain oscillating at f_2 because it has approached this frequency by tuning from f_1 and is still stable. As the tuning continues to increase, the frequency eventually increases to f_3. At this tuning level, the point of intersection becomes a line of intersection, and any noise in the tuning level violently shifts the frequency about f_3. This point, with a crossing angle of 180 degrees, can be seen intuitively to be very noisy. At higher tuning levels, the frequency makes a hops to a frequency f_7, which is once again a stable oscillating point . As the tuning level continues to increase, the frequency once again rises smoothly through f_8. The oscillator has made a mode hop in jumping from f_3 to f_7, and the frequencies between them cannot be reached through this tuning route.

Unfortunately the tuning history of this oscillator is not as simple as a single mode-hop. As the tuning level is retraced, from T3 to a level back towards T2, the intersection point retraces a different part of the $Z_L(\omega)$ curve. Because f_6 is a stable operating point, the frequency retraces from f_8 through f_6 until f_5 is reached, where once again the 'point' of intersection becomes a line of intersection and the frequency of oscillation is very noisy. As the tuning level is further reduced, a mode hop occurs down to a frequency f_1, so that frequencies between f_5 and f_1 cannot now be reached by this tuning route. As the tuning is further reduced, we retrace the same tuning curve as before down to f_0.

Hystersis and mode hopping



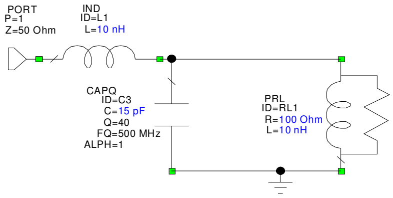
The tuning curve of the oscillator represented by the device and load line locus of the previous slide

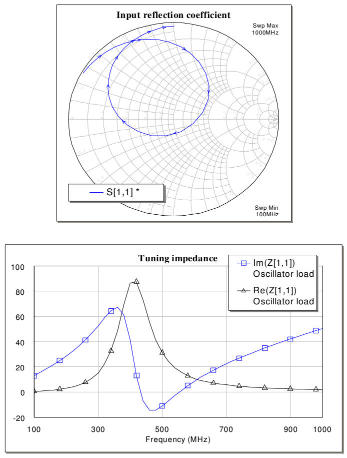
Besser Associates

IX - 10

The tuning curve for this oscillator is shown above. It is apparent that not only does this oscillator suffer from two mode hops, but that it also has hysteresis. This unfortunate situation, which can be observed in many real oscillators, arises because the load on the oscillator offers a load impedance to the device that sustains oscillation at more than one frequency. If in fact the load line $Z_L(\omega)$ were a straight vertical line of constant real part *R* and variable reactance *X* with frequency, there would only be a single possible oscillation point at each tuning level. Such a load could be synthesized with a series RLC circuit: the real part is just the series resistance which is constant, and the imaginary part will be given by $\omega L - 1/\omega C$, which increases monotonically with frequency $f = \omega / 2\pi$.

The effect of a wrong load termination





(a) A simple parallel RLC circuit and a series inductance, modeled in Figure (b)

(b). The equivalent series resistance and reactance of the circuit of Figure (a), showing a 'loop' characteristic through resonance

Besser Associates

IX - 11

In fact, the load impedance just shown is not at all a complicated circuit. Consider the circuit of Figure above, which is a parallel RLC circuit and a small inductance in series, as might occur with a bond wire. On the Smith chart, a parallel RLC circuit on its own traces a line of constant conductance. Adding series inductance simply adds more positive reactance at higher frequencies and skews the constant conductance circle towards the inductive side of the Smith chart, as shown. In terms of R and X, also plotted, the reasons for the ''loop' become obvious, when it is noticed that the equivalent series resistance first increases and decreases, while the equivalent series reactance increases, decreases, and increases again as we pass through resonance.

This exercise illustrates one of the most important, yet least well-known, tricks in oscillator design. The device we are dealing with here behaves like a series representation, because its resistance becomes less negative with increasing drive level A. (We are plotting the negative of Z_D i.e. $Z(A)$ to deal with positive entities). The rule is this: series-type devices, in which the resistance becomes less negative with drive, should be terminated with series resonant circuits. The converse is also true, as was illustrated above, series-type devices should not be terminated with parallel-resonant circuits, since as we have just seen, these can offer multiple frequencies at which the equations for steady-state oscillation are satisfied. Whenever this occurs, the potential for mode hopping and hysteresis exist.

It goes without saying that the dual also applies. When the impedance variation of shunt-type devices are plotted either on the Smith Chart or on Cartesian coordinates with G-B axes, the device conductance becomes less negative with drive. For a reasonably invariant reactive part, the device line $Y(A)$ will plot from right to left on the G-B plot. In this case, which is the dual circuit, it is important to terminate such a device with a parallel resonant circuit, so that the equivalent shunt conductance of the load is constant with frequency and there is only one potential oscillation point as the device is tuned.