THE EVOLUTION OF THE MODERN LOCK-IN AMPLIFIER

DL INSTRUMENTS

IAN 35

by Edwin H Fisher

233 Cecil A. Malone Drive, Ithaca, New York 14850; Phone 607-277-8498; Fax 607-277-8499

The need to measure small signals in the presence of high levels of interference exists in virtually every field of engineering. In these cases, it is reasonable to assume that whatever the signal source is, it is producing the maximum signal that can be generated, and when this is still less than the noise, we have to use some sort of specialized attack to measure the signal. Obviously if we cannot increase the signal power then we have to reduce the noise power, so a brief consideration of noise sources and their spectra is in order to see how we might proceed.

First, we define noise. Noise is any variation in time of some physical parameter that is caused by any combination of all the possible processes than can influence the value of that parameter, producing no useful information to the observer about the event he wishes to observe.

Likewise, we define signal. Signal is any variation in time of some physical parameter that is caused by only one of the processes than can influence the value of that parameter, conveying useful information to the observer about the event he wishes to observe.

Comparing the definitions shows that we have defined noise as any unwanted signal. We have also deliberately required that both signal and noise be time varying. We must also assume that we can convert the time variation of the parameter of interest into an equivalent electrical variation in time by means of a suitable transducer, for we wish to work in the realm of electrical quantities.

Transducers are real, physical devices. They have resistance, inductance, and capacitance. They are often placed in environments with time varying conditions of light, temperature, humidity, pressure, electromagnetic fields, various categories of radiation, and all sorts of people. All transducers exhibit some form of internal impedance, and all contribute noise to a measurement.

The most common types of noise encountered are thermal or Johnson noise resulting from the random motions of the electrons due to their thermal energy; Shot noise caused by the random variation in the velocity of the discrete charge carriers in a conducting path, Flicker, or 1/f noise, of uncertain cause, but largely associated with systems that are charge carrier limited; discrete frequency noise, like 60 Hz; and amplifier noise which contains all of the previous noises and extra noise because of the real components used in their construction. With a little thought, the list of noise sources can be expanded considerably, but we can learn all we need to know from considering a few already mentioned.

If we have two matched resistors, one at room temperature and the other held near aboslute zero, and if they were connected in parallel with wires having zero resistance, energy would be transferred from the warm resistor to the cold one. A zero resistance ammeter connected across the terminals of the warm resistor would read a current of:

$$(1) \qquad I_{Nsc} = \sqrt{\frac{4KT\Delta f}{R}}$$

K = Boltzman's Constant
T = Absolute Temperature
Δf = Bandwidth of the Ammeter in Hz
R = Resistance in Ohms of the Warm Resistor

This is the short circuit current due to the motion of the electrons in the warm resistor caused by their thermal energy. By opening one of the ammeter leads we would expect to see an open circuit emf having an rms value of:

$$(2) \qquad E_{Noc} = I_{Nsc}R = \sqrt{4KTR\Delta f}$$

Connecting the cold resistor now to the warm one, the noise current that will flow between the two is:

$$(3) \qquad I_N = \frac{E_{Noc}}{2R} = \sqrt{\frac{KT\Delta f}{R}}$$

and the average noise power transferred to the cold resistor becomes:

$$(4) \qquad W_N = I_N^2 R = KT\Delta f$$

This is the maximum noise power than can be transferred from the warm resistor to the cold one. The available power unit bandwidth can be written as:

$$(5) \qquad \frac{\Delta W_N}{\Delta f} = KT$$

and when Δf approaches zero:

$$(6) \qquad \frac{dW_N}{df} = KT = w_N(f)$$

The function $W_N(f)$ is the power spectral density. In this case, it is clearly independent of frequency, consequently, we say that thermal noise is "white" by analogy to white light.

A similar analysis would show that Shot noise is white with a power spectral density derived as shown:

$$(7) \qquad I_N = \sqrt{2qI_o\Delta f}$$

q = Charge on the Electron
I_o = Average dc Current

$$(8) \qquad W_N = I_N^2R = 2qI_oR\Delta f$$

$$(9) \qquad \frac{\Delta W_N}{\Delta f} = 2qI_oR$$

$$(10) \qquad \frac{dW_N}{df} = w_N(f) = 2qI_oR$$

Flicker noise, however, has the power spectral density which is a function of frequency:

$$(11) \qquad w_N(f) = \frac{K}{f^n} \qquad 0 \leqq n \leqq 2$$

Several typical noise spectra are shown in Figure 1. Every combination of noises can be characterized by a specific noise power spectral density function from which the total noise power that can be delivered to a matched load within some frequency interval can be calculated by integrating the function over the frequency interval of concern.

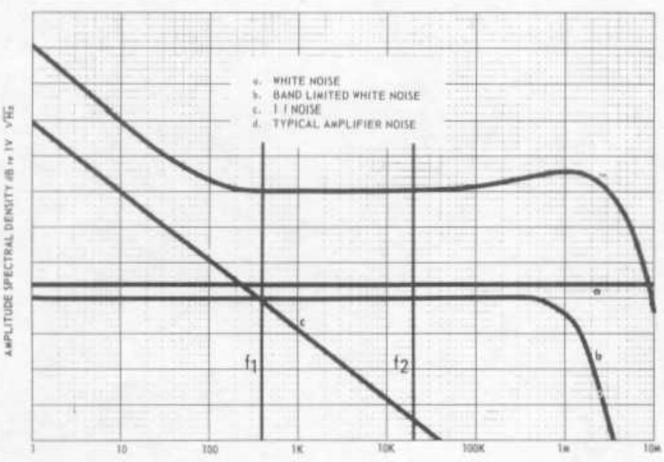


Figure 1

$$(12) \qquad W_N = \int_{f1}^{f2} w_N(f)\, df$$

Graphically this represents the area under the curve in the frequency interval indicated.

Returning to our measurement problem, if we can locate a frequency interval where the noise power spectral density function approaches a minimum and if we can confine our measurement within this frequency interval in the noise "valley" our chances for getting a useable result are substantially improved, so lets examine some of the devices and techniques that help us to accomplish just that.

The response of any physically realizable electrical network to an input stimulus is generally dependent upon the frequency of the stimulus. Take any "black box" with two input and two output terminals and apply an alternating potential difference to its input of:

$$(13) \qquad v_i(t) = V_i \sin(2\pi ft)$$

The resulting output potential will have the general form:

$$(14) \qquad v_o(t) = V_o \sin(2\pi ft + \theta)$$

The function relating these two voltages is the transfer function of the network and will generally be a complex, dimensionless quantity. The effect of the black box is to change the amplitude and shift the phase of the input voltage in a manner that is dependent only on the frequency. The complex transfer function is therefore a function of frequency, and can be written as:

$$(15) \qquad \overline{g}(f) = \frac{V_o(t)}{V_i(t)}$$

If the black box is subject to an excitation consisting of a number of individual frequencies, its output can be determined by summing the response to each of the individual frequencies in the excitation. If the input spectrum is a continuous one, the rms output voltage density function is related to the rms input voltage density function just like a single frequency:

$$(16) \qquad v_o(f) = \left|\overline{g}(f)\right| v_i(f)$$

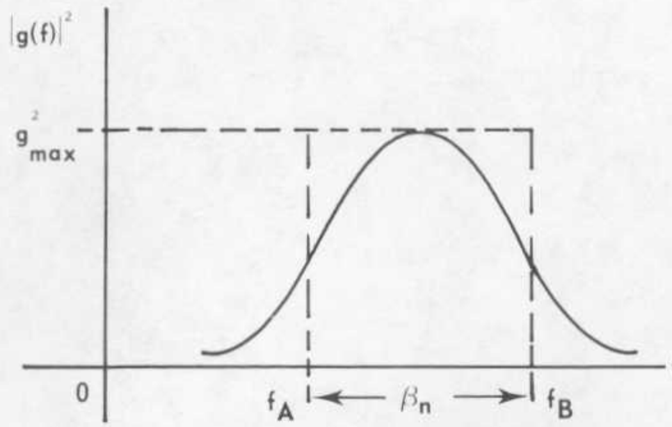
and the rms total voltage over any given frequency interval is found by integrating over that frequency interval.

$$(17) \qquad V_o = \int_{f1}^{f2} v_o(f)df = \int_{f1}^{f2} \left|\overline{g}(f)\right| v_i(f)df$$

If this black box is interposed between a noise source having a noise power per unit bandwidth of: $W_N(f)$, and a load which would otherwise be matched to the source, the noise power delivered to the load via the black box can be expressed as:

$$(18) \qquad w_{No}(f) = w_{Ni}(f) \left|\overline{g}(f)\right|^2$$

since the noise power is proportional to the square of the voltage. Clearly we can reduce the total transferred noise power if we can contrive a black box which gives nearly unity transmission within the range of frequencies where we want the signal to appear but which gives nearly zero transmission at all other frequencies. This is the time honored technique of narrow–banding the information system. By reducing the bandwidth to whatever extent is required to make the available signal power detectable, we have reduced the noise power enough to be able to make a useful measurement. Combining bandwidth reduction with a sensible location of the signal frequency in the noise power spectrum, we come close to doing the best we can to recover the signal information.



EQUIVALENT NOISE BANDWIDTH: β_n

Figure 2

For comparison purposes among a variety of narrow banding black boxes, the concept of an equivalent noise bandwidth is useful. It is a theoretical bandwidth that is rectangular in shape and that has the same amplitude as the peak value of the real network's squared amplitude curve, and the same total area as exists under the real network's curve over all frequencies from zero to infinity. An example is shown in Figure 2. The value of the equivalent noise bandwidth for any network can be calculated from the network's transfer function as:

$$(19)\quad \beta_N = \frac{1}{g^2_{max}} \int_0^{\infty} \left| \bar{g}\,(f) \right|^2 df$$

This is the same as saying that a white noise source would deliver the same noise power to a matched load through either the real network or through its equivalent noise bandwidth. That total noise can be calculated simply by:

$$(20)\quad W_N = w_N(f)\; g^2_{max}\, \beta_N$$

Remember that for white noise, $w_N(f)$ is a constant.

Having defined the yardstick, we are now able to compare a variety of real networks to see if there is a best choice. Perhaps the first that comes to mind is the simple bandpass filter. Intuitively, the closer together the upper and lower cutoff frequencies are set, the smaller the equivalent noise bandwidth. However, once the two frequencies are set closer together than about a factor of four, the filter begins to produce an insertion loss. For a four pole Butterworth filter set so there is no insertion loss:

$$(21)\quad \beta_N = 1.55\, f_o$$

$$f_o = \text{center frequency} = \sqrt{f_{hp}\, f_{lp}}$$

If the cutoff frequencies are made equal, so that $f_{hp} = f_{1p} = f_o$:

$$(22)\quad \beta_N = .513\, f_o$$

The insertion loss here is now 6 dB. We can also relate this to the quality factor Q which is defined as:

$$(23)\quad Q = \frac{f_o}{f_u - f_1} = \frac{f_o}{\Delta f}$$

For the first case, Q = .666; for the second Q = 2.26. We can see from these values that the bandpass filter is not a very selective network.

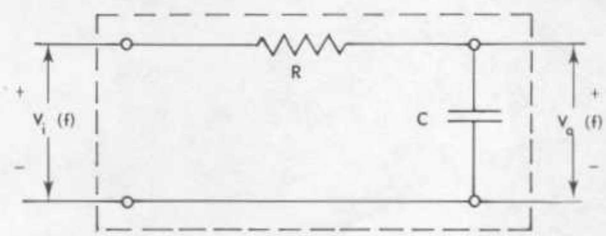
The resonant RLC network can have values of Q that are much larger. Values as high as 100 are realizable in circuits where temperatures are stable and currents are low so that component self–heating and consequent aging are minimal. For the tuned RLC circuit:

$$(24)\quad \beta_N = \frac{\pi f_o}{2Q}$$

With Q = 100:

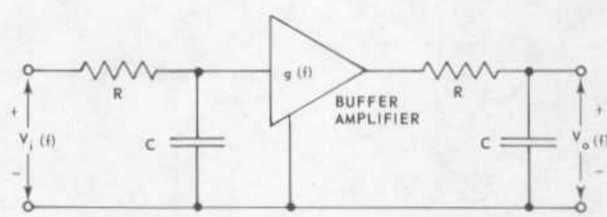
$$(25)\quad \beta_N = 0.016\, f_o$$

The RLC circuit is a much better choice than the bandpass filter and for this circuit, there is no insertion loss at f_o.



SINGLE SECTION LOW–PASS FILTER: $\beta_n = \frac{1}{4RC}$

Figure 3



DUAL SECTION LOW–PASS FILTER: $\beta_n = \frac{1}{8RC}$

Figure 4

Finally, consider a simple RC low pass filter as shown in Figure 3 and also a dual section with an impedance isolating amplifier between sections as shown in Figure 4. For the single section:

$$(26)\quad \beta_{N_1} = \frac{1}{4RC}$$

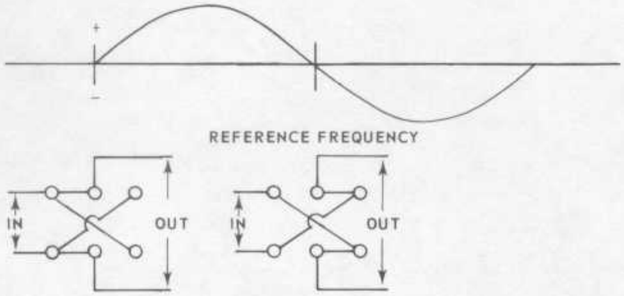
For the dual section filter:

$$(27)\quad \beta_{N_2} = \frac{1}{8RC} = 0.5\, \beta_{N_1}$$

With real values of R and C it is easy to achieve equivalent noise bandwidth values of as little as: $\beta_N = 0.00125$ Hz. Note also that this is an absolute bandwidth, not a percentage bandwidth of an center frequency f_o. To get the same equivalent noise bandwidth with a tuned RLC circuit tuned to 1000 Hz would require a Q of 1.26×10^6.

The low pass filter looks like an ideal narrow banding device so long as all the signal appears at zero frequency. That is the worst place in which to work from the point of Flicker noise, plus the fact that dc amplifiers present all sorts of problems of drift, etc. Fortunately there is a technique which avoids all of the dc problems and still allows us to use the low pass filter as the narrow banding device. It is phase sensitive detection. Assume that we are given a signal of the form:

$$(28) \quad v_s(t) = V_s \sin(2\pi f_s t + \theta)$$



DPDT SWITCH ACTING AS A PHASE SENSITIVE DETECTOR

Figure 5

If we apply this signal to the terminals marked "in" of Figure 5, and if the reversing switch is driven by a reference frequency that is identical to the signal frequency in such a way that for every zero crossing of the reference frequency the switch is transferred instantly to its other position as is illustrated in the figure, the output from the switch has the waveform illustrated in Figure 6. The switch is acting as a synchronous rectifier. It produces an output that has an average or dc component and a series of ripple frequencies. The dc voltage out is:

$$(29) \quad V_{dc} = \frac{2}{\pi} V_s \cos\theta$$

The ripple components are given by the sum of terms:

$$(30) \quad \frac{2}{\pi} V_s \sum_{k=1}^{m} \frac{2}{(2k)^2 - 1} \cos(2k(2\pi f_s t))$$

The process is called phase sensitive detection because the dc output of the device is dependent upon the cosine of the phase angle θ between the signal and reference waves, and it is the dc component we are interested in. This dependence is illustrated in Figures 6 – 10. The phase differences illustrated go from 0° through 180°.

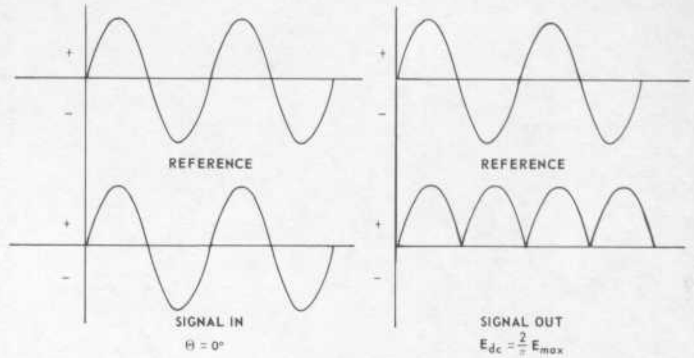


Figure 6

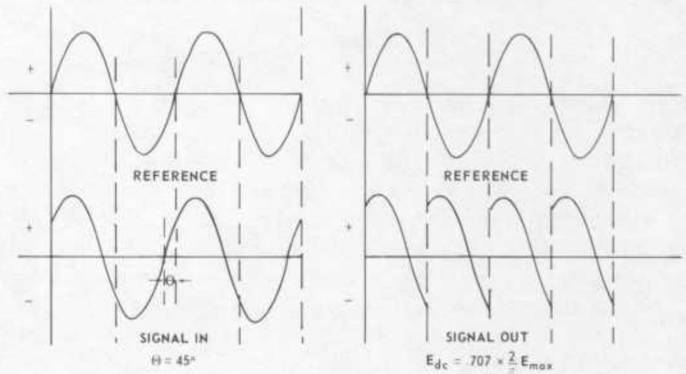


Figure 7

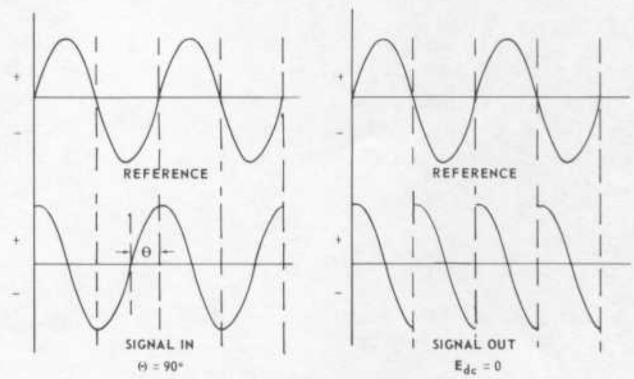


Figure 8

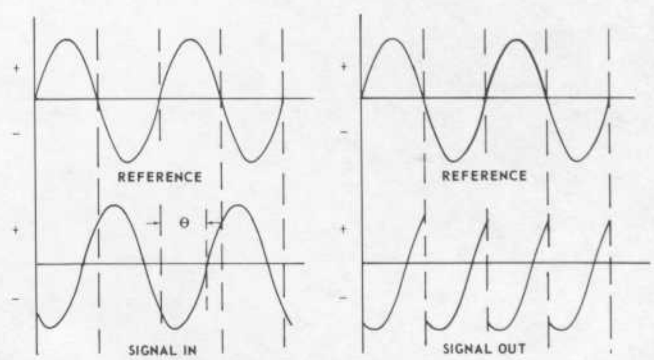


Figure 9

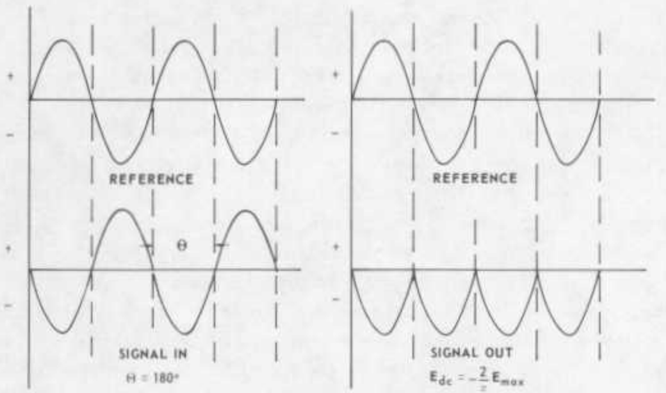


Figure 10

The dc component of the output is related to the peak value of the signal input at frequency f_s so we can now take advantage of the low pass filter as the bandwidth limiting device by simply feeding the output of the phase sensitive detector into the input of the low pass filter. The low pass filter will also prevent any of the ripple components from the phase sensitive detector from getting through. So, in effect, we have created the equivalent noise bandwidth of the low pass filter around the signal at frequency f_s. Moreover, we have no limitations on the signal frequency, so we can still take advantage of the possibility of operating at a low noise portion of the noise power spectrum.

Keeping in mind that it is the dc output of the phase sensitive detector that is of value to us, we now must consider what effects noise has on the device. There are two classes of noise to consider. The first is non–coherent noise, where the frequencies and phases of the individual noise components drift randomly with respect to the signal frequency f_s. The second is coherent noise, a simple harmonic voltage variation whose frequency f_n is some rational multiple of the signal frequency f_s. This can be written:

(31) $$f_n = \frac{p}{q} f_s$$

Here p and q are integers. For the special case where the quotient p/q is itself an integer, the noise components are the Fourier harmonics of the signal frequency.

One way of visualizing the operation of a phase sensitive detector is to consider that it is multiplying the input signal by a square wave at reference frequency that has an amplitude that is plus one for one half cycle and minus one for the other half. The process of multiplying one wavetrain by another creates a new wavetrain whose frequency components are determined by forming all of the possible sums and differences of the frequencies of the one wavetrain with all of the frequencies of the other. Clearly multiplying all of the noise components in a non-coherent noise input with a unit square wave generates more noise frequency components but they all still vary randomly in frequency and phase with respect to the signal frequency and so cannot yield any net dc output from the phase sensitive detector, and the total noise output is then limited by the equivalent noise bandwidth of the output filter. The situation is somewhat different for coherent noise components. A symmetrical square-wave contains only odd harmonics of its fundamental frequency. Therefore the output wavetrain from the phase sensitive detector with a coherent noise input is expressable as:

(32) $$f_{out} = \left(\frac{p}{q} f_s \pm (2k-1) f_r\right) \qquad k = 1, 2, 3, \ldots n$$

There is no physical significance to negative frequencies in this analysis, so that all terms yielding frequencies less than zero can be ignored. What we are concerned with is whether this new wavetrain has any dc terms. The sum terms cannot equal zero so we need consider only the difference terms. Recalling that the signal frequency and the reference frequency are the same, when we set (31) equal to zero, we get:

(33) $$\frac{p}{q} = (2k-1)$$

Which says that there is a dc output only when the quotient p/q is an odd integer – an odd Fourier harmonic of the signal frequency. For any other coherent noise component, the output of the phase sensitive detector will be a ac frequency component varying from a positive maximum to a negative maximum at a rate which is the difference between the signal frequency and the noise frequency. This is shown in Figure 11. Unless the difference frequency is very small the ac component will be eliminated in the low pass filter just like the ripple components. Figure 12 shows what happens when the quotient p/q is an even integer, and as predicted in the analysis, there is no dc output. Figure 13 shows the net dc output resulting from a value of p/q of 3.

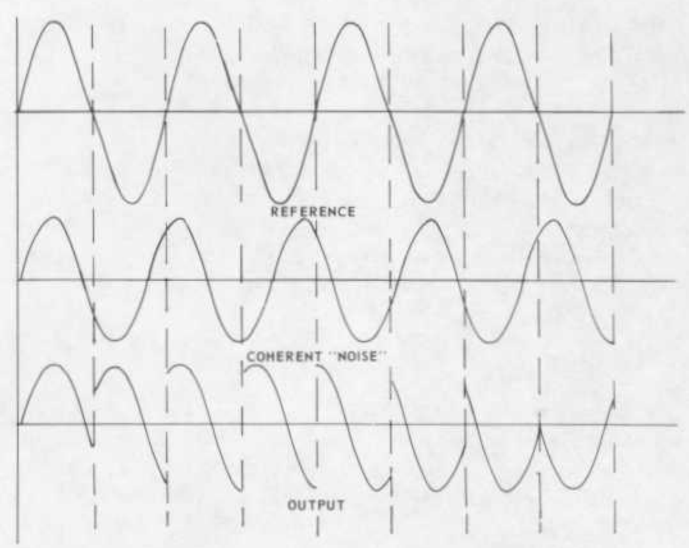


Figure 11

The entire discussion so far has been based on how a phase sensitive detector and low pass filter operate when fed an input signal of fixed frequency and constant

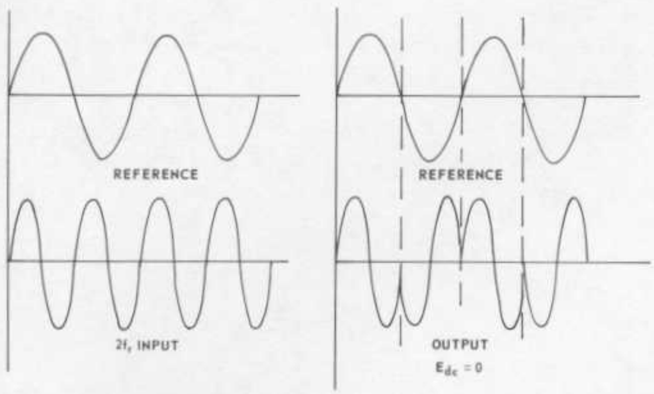


Figure 12

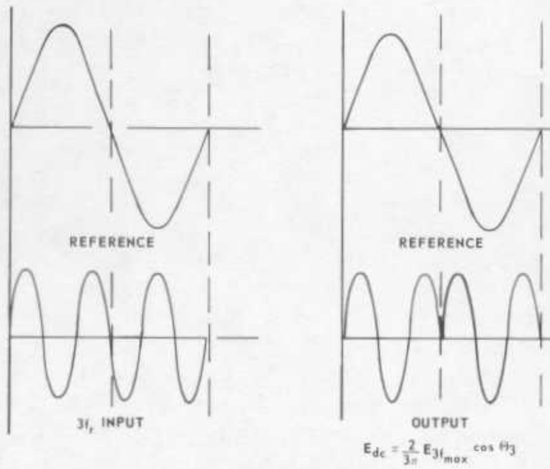


Figure 13

amplitude. In the real world such signals rarely occur, so how does one make use of the detector filter combination in real situations and of equal importance, how does one have any control of the frequency of operation in order to take advantage of the quiet spots in the noise spectrum. The solution is to make use of a carrier technique, and amplitude modulate the carrier frequency with the signal of interest. The carrier frequency then becomes the reference frequency at which the phase sensitive detector operates. Assume we have some complex signal. At any instant of time it can be represented by a Fourier series having the form:

$$(34) \quad v_s(t) = V_s \left(1 + \sum_{k=1}^{m} V_k \cos(2\pi k f_s t + \theta k)\right)$$

When we use this signal to amplitude modulate a carrier of frequency f_R there are upper and lower sidebands created. Analytically the wave can be expressed as:

$$(35) \quad v_s(t) = V_s \left(\cos 2\pi f_r t + \sum_{k=1}^{m} \frac{m_k}{2} \cos(2\pi (f_r \pm k f_s)t + \theta_k)\right)$$

The first term in this familiar expression is the carrier. The terms in $(f_r + kf_s)$ are all in the upper sideband and the difference terms are all in the lower sideband. Only those terms in the upper sideband are significant, since the process of phase sensitive detection translates the carrier frequency down to zero frequency. Note that by discarding the lower sideband, the detection process causes a loss of one half of the available signal strength. In most cases, this is a small penalty to pay when compared to the noise reduction possible with the narrowed bandwidths attainable. An example will illustrate the capability. Synchronous detection of a carrier frequency at 1 kHz that is in a white noise background will provide a noise power reduction of 10,000 if the output bandwidth of the filter is 0.1 Hz, or the noise voltage reduction of 100, since voltage is proportional to the square root of the power.

Finally, in using a phase sensitive detector and low pass filter combination for real measurements, some attention must be paid to the configuration of the filter. As we have seen, the filter's bandwidth must be narrow enough to detect the signal of interest, but the bandwidth set also determines the highest frequency in the actual signal that can be passed without unacceptable distortion. If an attenuation of 3 dB is the acceptable limit as is usually specified in amplifier systems, then the highest frequency in the upper sideband of the modulated carrier must be no greater than the −3 dB frequency of the filter. For a single section filter:

$$(36) \quad \beta_{N_1} = \frac{1}{4RC} \quad \text{and} \quad f_{C_1} = \frac{1}{2\pi RC}$$

For a dual section filter:

$$(37) \quad \beta_{N_2} = \frac{1}{8RC} = 0.5\beta_{N_1} \quad f_{C_2} = 0.643\, f_{C_1}$$

For a three section filter:

$$(38) \quad \beta_{N_3} = \frac{3}{32RC} = 0.75\beta_{N_2} \quad f_{C_3} = 0.793\, f_{C_2}$$

This shows that there is a real advantage in using a dual section filter, since the equivalent noise bandwidth is reduced to half that of a single section unit, but the undistorted signal bandwidth is only reduced to approximately two thirds of a single section. Going to an additional section, however, reduces both the equivalent noise bandwidth and the undistorted signal bandwidth by the same amount, so the addition of the third section is not justified. Another example will serve to emphasize how severe the signal bandwidth limitation is, take the 1 kHz reference frequency assumed previously and the 0.1 Hz output bandwidth that gave a noise voltage reduction of 100:1. This says that RC is 1.25 and therefore the highest frequency in the signal can be only 0.082 Hz, and the filter has a risetime to 95% of full output of more than six seconds. This illustrates the penalty paid in time required to make the measurement that must be accepted if the noise is so great in comparison to the signal.

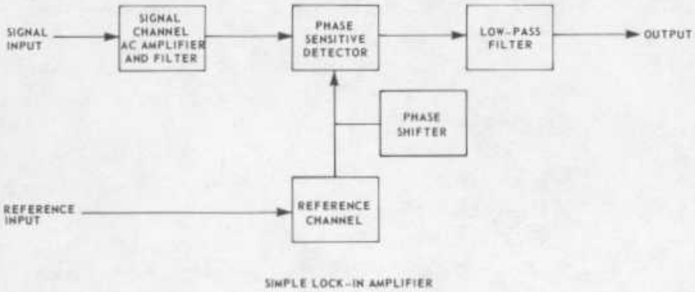


Figure 14

A practical instrument incorporating phase sensitive detection and low pass filtering is called a Lock–In Amplifier. A simple instrument has the block diagram of Figure 14. It has both a signal and a reference input, with the reference frequency driving the phase sensitive detector as we saw earlier. The phase shifter in the reference channel is required to be able to adjust for maximum dc output from the detector. The signal channel has both ac amplification and bandpass filtering that prevents any odd harmonics of the reference frequency from getting through the signal channel to the phase sensitive detector where they would produce a dc output of error in the measurement. Early lock–in amplifiers incorporated manually tuned, high Q amplifiers for this purpose, and so limited the lock–in to operation at a single frequency. Where measurements were desired over a range of frequencies, it was necessary to re–tune and re–phase the instrument on a point by point basis. The tuned amplifiers also have a severe phase shift around their center frequency which can cause substantial phase errors if drift or mis–tuning occur. A practical instrument provides a choice of single section or dual section output filtering; with the dual section filtering used in all cases except when the instrument is in a closed loop system. The choice of output bandwidths is usually from 100 Hz to .001 Hz, and the full scale dc output is usually ten volts.

Because there are countless situations where it is desirable to measure over a range of frequencies, the first evolution in practical lock–in amplifiers took place. Reference channels that could follow a changing reference frequency were developed. However, keeping odd harmonics of the reference frequency out of the signal channel was a problem. Many instruments were built with broadband input amplifiers, some still are. This left the matter of dealing with the errors caused by harmonic responses entirely up to the user; yet this matter is of paramount importance in making measurements where the interferences are larger than the signal to be measured. Figure 15 clearly illustrates why. Here a measurement is being attempted over the frequency range from 89 Hz to 112 Hz. A single coherent noise component exists at a frequency of 10 kHz with an amplitude 100 times greater than the signal to be measured. Dividing 10,000 by all the odd integers from 1 to 9999 will define all the frequencies for which 10,000 is an odd harmonic. There are eleven of these frequencies between 89 Hz and 112 Hz, so making any kind of a meaningful measurement over this frequency range is impossible so long as that 10 kHz noise component can get into the single channel.

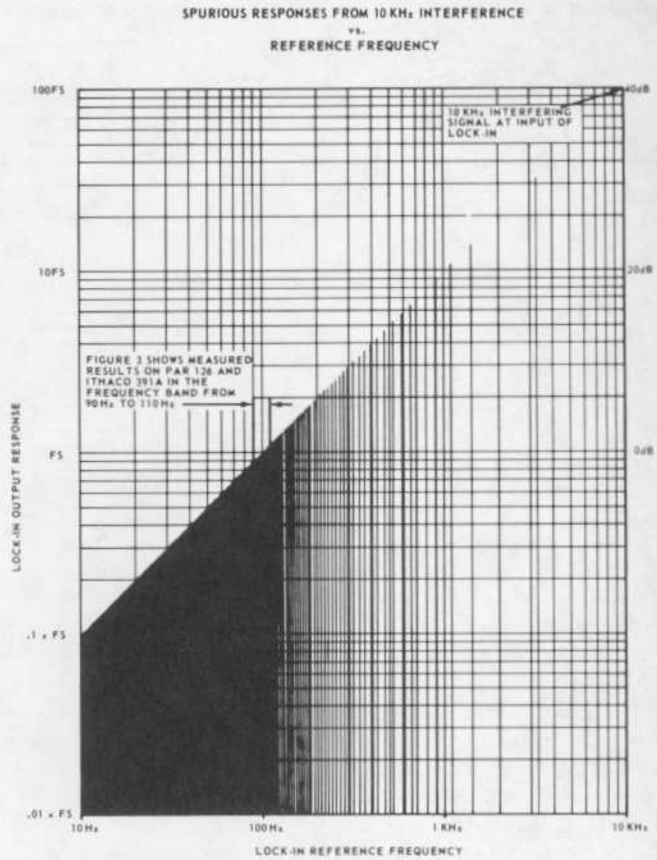


Figure 15

The next evolution, which solved the problem, was ***heterodyning***. Here the incoming signal is translated up in frequency to a fixed intermediate frequency that is above the reference frequency range being use. Fixed frequency filtering can be used at this intermediate frequency preceding the phase sensitive detector which also operates at the intermediate frequency. The block diagram of a heterodyning instrument is shown in Figure 16. The intermediate frequency is f_o; it is generated in a sinewave oscillator in the instrument's reference channel. It drives the phase sensitive detector at f_o via the phase shifter as shown. The frequency f_o is also fed into a reference channel mixer where it is mixed with a frequency f_v generated by a voltage controlled oscillator. As we discussed earlier, the mixing process produces both sum and difference frequencies. The difference frequency $f_v - f_o$ is compared to the actual reference frequency f_r in a phase comparator. Any phase difference between f_r and the difference frequency $f_v - f_o$ causes an error voltage that is used to change the frequency f_v from the voltage controlled oscillator. The only time there is no phase error is when $f_r = f_v - f_o$ with the result that $f_v = f_r + f_o$. The incoming signal to the lock–in is first amplified in a low noise pre–amplifier and then filtered in a sharp cutoff low pass filter whose cutoff frequency is set to approximately 0.5 f_o. The signal information at the output of this filter is then mixed with the frequency f_v from the voltage controlled oscillator. Any desired signal at the input will be coherent with f_r, and when mixed with f_v in the signal channel mixer, the sum and difference frequencies

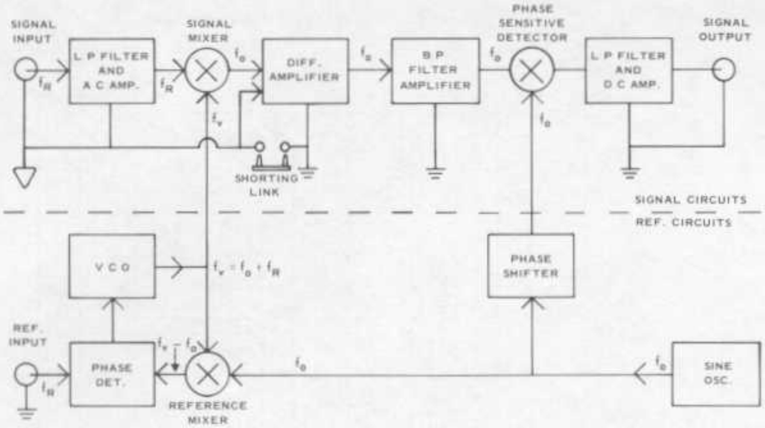


Figure 16

$f_v + f_s$ and $f_v - f_s$ are created, but since $f_v = f_o + f_r$, the sum frequency is $f_o + 2f_r$ and the difference frequency is f_o. The difference frequency is detected by the phase sensitive detector and produces the desired dc output proportional to the signal size. The sum frequency may or may not get through the bandpass filter ahead of the detector, depending on the value of f_r, but under no circumstances can the sum frequency component overload the detector, so it is just like any othercoherent noise component and produces no dc output from the detector as we saw earlier.

The heterodyning process produces a first order immunity to odd harmonics of the reference frequency because there is no direct relationship between f_r and the higher intermediate frequency f_o. Frequencies which when mixed with f_v that would produce a component at $3f_o$ are $4f_o + f_r$ and $2f_o - f_r$. Both of these frequencies are greater than f_o and consequently cannot get to the signal mixer because of the low pass filter in the input.

The heterodyning process utilizing a higher intermediate frequency does not absolutely destroy all response to odd harmonics of the reference frequency f_r. Because it is necessary to maintain very close phase coherence between the reference frequency f_r and the difference frequency $f_v - f_o$, it is better to generate a square wave in the voltage controlled oscillator and consequently there are mixing components at the odd harmonics of f_v. The largest of these is $3f_v$ and when mixed with $3f_r$ in the signal channel mixer results in frequencies of $3f_o + 6f_r$ and $3f_o$. The difference frequency $3f_o$ would produce a spurious dc output from the detector. The bandpass filter ahead of the detector attenuates this component nearly completely, resulting in at least a 55 dB rejection of odd harmonics in the signal input.

The heterodyning technique has accomplished two things: first, it provides a means for operating a lock–in over a range of frequencies without the need for manually tuning either the signal or the reference channel. Secondly, by operating with a higher intermediate frequency than the reference, it avoids all the problems of harmonic responses since fixed frequency filtering can be used as required in the signalchannel. The net effect is the same as having a tracking filter in the input. Figure 17 illustrates how well the technique works, showing the difference in responses of both a broadband lock–in and an ITHACO DYNATRAC® heterodyning instrument to the situation discussed earlier where a measurement is required over the frequency interval from 89 Hz to 112 Hz with a coherent noise component at 10 kHz. It is interesting to note that there are even harmonic responses as well as odd harmonics from the broadband unit. While even harmonic responses do not exist in a perfectly balanced phase sensitive detector, they are clearly present in this instrument's output. They are usually caused by imperfect half cycle to half cycle symmetry in the detector.

The heterodyning lock–in has two operating limits imposed on it. First, the phase detector–voltage controlled oscillator loop in the reference channel sets a limit on the rate at which the reference frequency can be slewed. The phase difference between f_r and $f_v - f_o$ must be compared at the zero crossings of the respective waves. When frequencies are low, there are few zero crossings per unit time so slewing rates must be slow. Secondly, the reference frequency cannot be allowed to become greater than the cutoff frequency of the low pass filter in the signal channel. Consequently, the instrument's overall operating frequency range must be broken up into a series of frequency intervals. As the upper end of a particular interval is reached, a new set of operating parameters must be established. Figure 18 tabulates the various operating conditions for each frequency interval. Notice that each range overlaps the next by one decade of frequency. When using an external reference source, each interval covers a frequency range of 400:1, which is adequate for most experimental work.

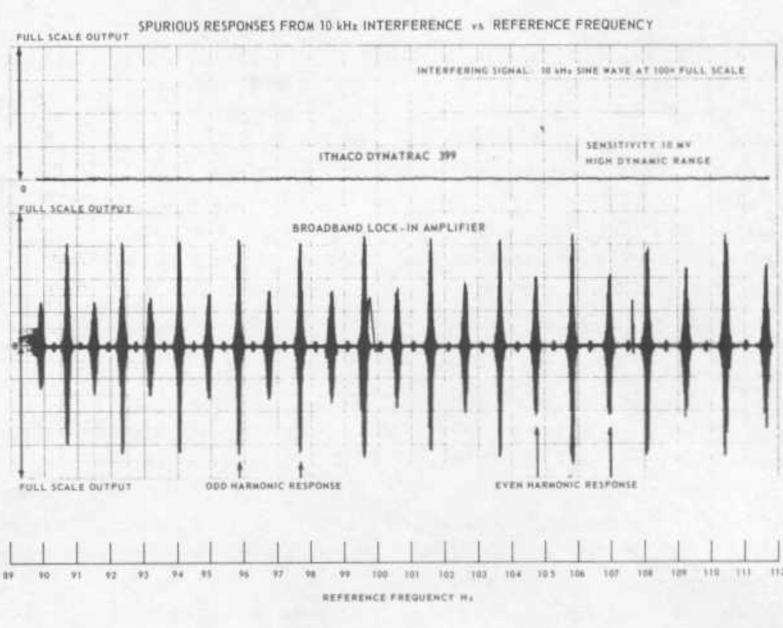


Figure 17

COLOR CODE	NOMINAL FREQUENCY RANGE (Hz)	OPERATING FREQ RANGE EXT MODE (Hz)	OPERATING FREQ RANGE INT MODE (Hz)	24 dB/OCTAVE L P CUTOFF FREQUENCY (Hz)	I F FREQ (Hz)	APPROX I F BW (Hz)	MAX SETTLING TIME	MAX SWEEP RATE Hz/Sec	MIN SWEEP TIME f TO 10f (Sec)
Brown	.1 – 1 1 – 10	.1 – 2 .5 – 20	.1 – 1.1 1 – 11	25	55	2	15 min 100 sec	7×10^{-5} .007	13000 1300
Red	1 – 10 10 – 100	.5 – 20 5 – 200	1 – 11 10 – 110	250	550	20	100 sec 10 sec	.007 .7	1300 130
Orange	10 – 100 100 – 1K	5 – 200 50 – 2K	10 – 110 100 – 1.1K	2.5K	5.5K	200	10 sec 1 sec	.7 70	130 13
Yellow	100 – 1K 1K – 10K	50 – 2K 500 – 20K	100 – 1.1K 1K – 11K	25K	55K	2K	1 sec .1 sec	70 7K	13 1.3
Blue	3K – 30K	1.5K – 60K	3K – 33K	75K	165K	6K	.05 sec	49K	.5
Green	10K – 100K	5K – 200K	10K – 110K	250K	480K	18K	.02 sec	300K	.3

Figure 18 (Model 391A LIA)

There are a lot of low level measurements where the largest available signal is still very small relative to the total noise, so narrow output filter bandwidths are required under best case conditions. Making the required phase adjustment for maximum dc output from the detector is tedious. Still another broad class of measurements occurs where there is a phase shift between the signal and reference during the course of the measurement. The next evolution was the development of the phase insensitive detector or lock–in analyzer. This is accomplished by having two phase sensitive detectors that are phased 90° apart, and a vector summing circuit to take the square root of the sum of the squares of the output from each, as shown in Figure 19. This eliminates the need for making any initial phase adjustments and follows a changing phase angle. As we discussed earlier, the dc output of a phase sensitive detector is:

$$(39) \qquad 2/\pi A\cos\theta$$

The output from a second phase sensitive detector operating 90° out of phase with the first would be:

$$(40) \qquad 2/\pi A\cos\left(\theta + \frac{\pi}{2}\right) = 2/\pi A\sin\theta$$

whereas before θ is the phase angle between signal and reference. If we now take the square root of the sum of the squares of the two outputs we get a value that is:

$$(41) \qquad \sqrt{\frac{4}{\pi^2} A^2\cos^2\theta + \frac{4}{\pi^2} A^2\sin^2\theta}$$

or

$$(42) \qquad \frac{2}{\pi} A \sqrt{\cos^2\theta + \sin^2\theta} = \frac{2}{\pi} A$$

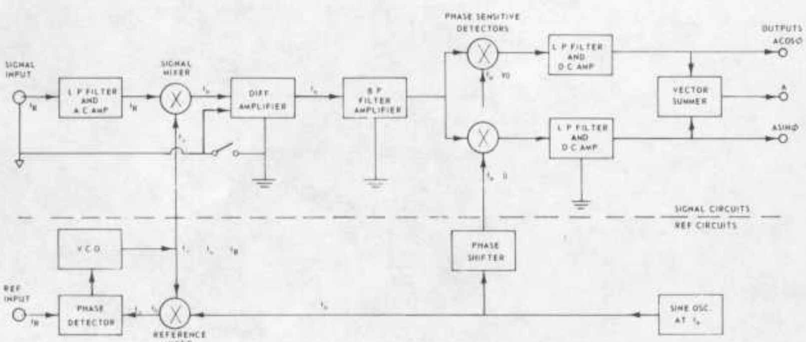


Figure 19

This is the maximum dc output available regardless of the phase angle θ. The combination of the two phase sensitive detectors and the vector sum circuit thus creates a phase insensitive detector. However, the vector sum is correct only so long as the individual phase sensitive detectors are operating on signals without odd harmonics. If there are odd harmonic inputs reaching either of the two detectors, the resulting vector sum is again a function of the phase angle between the signal and the reference. This is clearly illustrated in Figure 20 which shows the amplitude error when measuring a square wave input with a lock–in analyzer with a wideband signal channel. The amplitude error can be as large as 30% as a function of the phase angle θ.

Phase insensitive detection is a genuine convenience in most conventional applications. In addition, it opens up a whole new range of measurements including spectral analysis and operation without a reference input. Moreover when using the DYNATRAC method of heterodyning as developed at ITHACO, the output is accurate for any input signal waveform and there is no additional noise penalty.

Lock–ins with phase insensitive detection are usually called Lock–In Analyzers. They can be equipped with accessories like impedance matching preamplifiers, noise, ratio or phase angle measuring options and dynamic reserve mode switching to increase immunity to large overloads. These are the end result of several generations of product evolution, and find application in virtually every facet of low level or noise limited measurements.

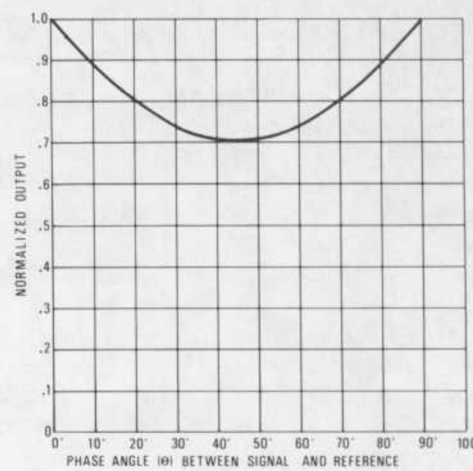


Figure 20