NATIONAL BUREAU OF STANDARDS REPORT

NBS PROJECT | **NBS REPORT**

251.00-40-2510101   October 29, 1964   8400

A NEW MODE OF OPERATION

OF A PHASE SENSITIVE DETECTOR

by

Yardley Beers

NBS

U.S. DEPARTMENT OF COMMERCE

NATIONAL BUREAU OF STANDARDS

A New Mode of Operation of a Phase Sensitive Detector*

Yardley Beers

The P.S. detector plays an important role in the instrumentation for spectroscopy and radio astronomy and in atomic frequency standards. Usually the phase sensitive detector is preceded by a narrow band-pass amplifier centered on the reference frequency or one of its harmonics. Such a procedure discards useful information contained in the other harmonics. It is suggested that the narrow-band amplifier be replaced by a wide-band one and that the reference waveform be shaped to be identical with that of the desired signal. First order calculations show that this method gives slightly better sensitivity than the conventional one. Also, discrimination against coherent signals of other waveforms can be obtained. If a differentiating circuit is inserted into either the input or the reference circuit, the detector output is an odd function of the phase angle. Therefore, this arrangement can be used as a sensor for a servo system used to lock the phase of the reference oscillator to the signal.

* A portion of this work was performed while the author was at New York University and working under contract with the U.S. Office of Naval Research.

1. INTRODUCTION

The phase sensitive detector [1, 2] has become a widely used device for improving the sensitivity in such diverse fields as microwave spectroscopy and radio astronomy. It may also be used to generate an error signal as part of a servo system for stabilizing the frequency of an oscillator to that of some reference signal. The purpose of this paper is to suggest a new mode of operation which can give better discrimination against spurious coherent signals while giving at least as good, if not slightly better, discrimination against noise as the conventional method.

2. THE DETECTION PROBLEM USING THE CONVENTIONAL MODE

The conventional mode is illustrated by Fig. 1. The word "phenomenon" refers to a spectral line, a radio star, or something else which is to be observed. If the phenomenon were sufficiently intense, it could be observed by some detecting device, which is here denoted as the first detector. However, because of the presence of noise and coherent spurious signals, the effect often is too small to be observed. In such cases, a phase sensitive detector is added. With it, the phenomenon is modulated at an angular frequency ω . In microwave spectroscopy the phenomenon itself is modulated by applying a Stark or Zeeman field to the sample giving rise to the line. In astronomy, however, the primary effect is not directly accessible for modulation, but modulation is accomplished in the radio receiver which serves as the first detector by switching the frequency from that of the source to some other one or by switching the input from the antenna to a dummy resistor. In either case, the output of the first detector generally is represented by a Fourier series involving ω and its harmonics. In the conventional mode of operation, this output is amplified and fed through a narrow-band filter to the signal input of the phase sensitive detector. This filter selects the k'th harmonic (most often $k=1$) so that the input to the detector is equal to $A_k \cos (k\omega t - \varphi_k)$. The detector also receives a reference input from the modulator through a phase shifting network. Generally this is also represented by a Fourier series, but if the signal input is confined to the k'th harmonic, only the k'th harmonic of the reference signal contributes to the dc output of the detector. The average output current then is, according to the first order theory,

$$i = gA_k B_k \cos (\varphi_k - \varphi_k'), \tag{1}$$

where B_k and φ'_k are respectively the amplitude and phase angle of the k'th harmonic of the reference voltage and g is a constant. Usually the phase shifter is adjusted to make the cosine factor unity when the phenomenon is at its peak.

In principle, the output consists not only of a steady current as given by Eq. (1) but also of fluctuations arising from noise produced in the first detector or in its input. The proper fluctuation output can be considered as being due to noise voltage components at the input of the phase sensitive detector beating against the reference voltage. However, the phase sensitive detector employs a non-linear element, which also can act as a rectifier. Therefore, there is also a fluctuating output resulting from noise components beating against one another. According to the first order theory, increasing the reference voltage increases the signal output and the proper noise output in the same ratio while leaving the noise output due to rectification unchanged. Therefore, proper operation requires that the reference voltage be large enough that the proper noise output be large compared to the rectified noise output, but this requirement is usually satisfied automatically whenever B_k is large compared to A_k. Usually this condition can be satisfied easily. There is, of course, an upper limit to the magnitude of the reference signal, which results in saturation effects and the failure of first order theory. However, under normal conditions, the band width of the low pass filter at the output is considerably smaller than that of the narrow band filter at the input, the output signal-to-noise ratio is independent of the width of the narrow band filter.

Another consequence of the non-linearity is the generation of harmonics of both signal and reference voltages. For simplicity, in this paper the effects of these harmonics are neglected. Many detectors employ balanced detectors. With these the effects of rectification noise and harmonics are eliminated or reduced.

One function of the narrow band filter, then, is to choose the harmonic of the input voltage which is to give rise to the output. As far as this primary purpose is concerned, the narrow band filter could just as well have been placed in the reference circuit as in the signal circuit, and if either the signal or reference signal or both should consist essentially only of the k'th harmonic, this filter could be omitted. However, there are some secondary advantages to having this filter. In the first place, it is useful in discriminating against spurious signals, especially if reference voltage harmonics other than the k'th have appreciable amplitudes.

Secondly, by limiting the pre-detection bandwidth, the filter reduces the number of noise components which beat together to give noise output by rectifier action. Often when the device is used with marginal signals there is danger of overloading the amplifier by noise. In Fig. 1, for simplicity, the filter is shown at the output of the amplifier. In practical design it usually is incorporated in one of the interstage coupling networks. Then it performs the very important function of reducing the danger of overload in the stages which follow.

To satisfy these purposes, it is merely necessary that this filter have sufficient selectivity to reject adjacent harmonics. In particular, it is undesirable to attempt to achieve unnecessarily small pre-detection bandwidth by cascading a large number of filter sections since a small fluctuation in the frequency ω causes a large shift in phase and results in difficulty in keeping the phase shifter set to its optimum position.

3. THE DETECTION PROBLEM USING THE NEW MODE

The previous discussion, points out the fact that the narrow-band filter is not fundamental to the operation of the detector but greatly assists it. It also implies that useful information contained in the other harmonics of the signal is being discarded. (It should be remarked in passing, however, that Shimoda[3], in a circuit for stabilizing a laser, has used several detectors, each with a filter tuned to different harmonic, to utilize the available information more fully.) Accordingly, in the new mode, the narrow-band filter is omitted; and the reference voltage wave form is controlled to be the same as that of the signal, for then the contributions of all the harmonics add constructively at the detector output. To show this, it is assumed for the moment that they are of different wave forms and that the voltages are given respectively by

$$V_S = \Sigma A_n \cos (n\omega t - \varphi_n) \tag{2}$$

and

$$V_R = \Sigma B_n \cos (n\omega t - \varphi'_n) \quad , \tag{3}$$

where n ranges from one to infinity. (It is assumed that the dc components are blocked out.)

Each harmonic contributes to the average output current in accordance with Eq. (1), and the total average out is given by the sum of such terms:

$$i = g \Sigma A_n B_n \cos (\varphi_n - \varphi_n'). \tag{4}$$

If the signal and reference voltages have arbitrary wave forms, generally some of the cosine factors are positive and some are negative, and there is a tendency for the terms to annul one another. However, if the wave forms are identical,

$$B_n = bA_n \quad , \tag{5}$$

and

$$\varphi_n - \varphi'_n = n\Phi \quad , \tag{6}$$

where b and Φ are independent of n. Then, if the phase shifter is adjusted to make Φ equal to zero, the cosine factors all become plus one and the terms add constructively to give

$$i = bg \Sigma A_n^{\ 2} . \tag{7}$$

Suppose that there is present a coherent spurious signal with the same fundamental frequency but different wave form. Such a situation exists in a frequency-modulated microwave spectrometer. Ideally, in the absence of a sample, the output is frequency independent, and no modulated output should be present. In practice, however, this condition is never fulfilled, and there is an output which can mask the effect of the sample, if the effect of the sample is small. This voltage may be written in the form of Eq. (2) as

$$V_D = \Sigma D_n \cos\,(n\omega t - \varphi_n''), \tag{8}$$

and it gives an output of

$$i_D = bg\Sigma\, A_n D_n \cos(\varphi_n - \varphi_n'') . \tag{9}$$

The supression ratio S may be defined as i/i_D. Then

$$S = \frac{\Sigma A_n^{\;2}}{\Sigma A_n D_n \cos\,(\varphi_n - \varphi_n'')} . \tag{10}$$

One possible advantage of the new mode can be seen by examination of Eq. (10). By proper selection of the type of modulation it may be possible to cause some of the cosine factors to be positive and some negative so that the terms of the denominator may combine destructively while, as has been pointed out, those in the numerator add constructively. Thus the suppression ratio may exceed the value it would have with the conventional method, where there would be only one term in numerator and denominator.

This advantage is of little interest to potential users unless it can be shown that the signal-to-noise ratio is at least as good as with the conventional mode. Therefore, this ratio will be computed now.

For the moment, it will be supposed that there has been inserted into the reference circuit a filter which passes only the k'th harmonic, and it is assumed that the system is operated in such a way that noise contributions to the output due to rectification can be neglected. Then the noise bandwidth f of the system is determined by the low-pass filter at the output. (Actually f is twice the noise bandwidth of the low pass filter since a detector with such a filter of width $f/2$ accepts noise from a band of width f centered at $(k\omega)/(2\Pi)$ after beating with the reference voltage.) The rms equivalent noise voltage at the input of the first detector is given by

$$v_N = (4FkTRf)^{1/2} \quad , \tag{11}$$

where F is the noise figure of the system, k is Boltzmann's constant (not to be confused with the index k) and T is the reference temperature. R is the resistance associated with the first detector input. The noise voltage at the input of the phase sensitive detector is found by multiplying this by the gain G between the two detectors, which is assumed to be frequency independent. The rms output noise current i_{Nk} is found by replacing A_k in Eq. (1) by this voltage and by replacing the cosine factor by its rms value $(1/2)^{1/2}$. With the use of Eq. (5) this becomes

$$i_{Nk} = bgGA_k(2FkTRf)^{1/2} . \tag{12}$$

If the filter in the reference circuit is removed, each harmonic of the reference voltage makes a contribution given by Eq. (12). These contributions are independent. Therefore the total noise current i_N is found by taking the square root of the sum of these squares of the individual contributions. The output current signal-to-noise ratio N is found by dividing Eq. (7) by i_N. In making the computation it is convenient to define a quantity, which has the dimensions of power

$$P = (\Sigma A_n^{\,2})/(RG^2), \tag{13}$$

in which the index k has been replaced by n. If this calculation is made, after some cancellation, it is found that

$$N = \left[P/(2FkTf) \right]^{1/2} . \qquad (14)$$

This result, of course, is based upon the assumption, usually valid, that noise generated in the reference circuit can be neglected. The derivation of the signal-to-noise ratio for the conventional mode is identical except the summations reduce to a single term. Therefore Eq. (14) applies to both modes.

The quantity P can be recognized as the total power dissipated in the input circuit of the first detector associated with the modulation of the primary phenomenon. Therefore, the signal-to-noise ratio depends upon how completely the phenomenon can be modulated. With the conventional mode, to obtain the maximum sensitivity, it is necessary to completely modulate the phenomenon and yet throw all of the power into a single harmonic. With the new mode it is merely necessary to completely modulate the phenomenon, but it is unimportant how the power is distributed among the harmonics. Generally it is impossible to obtain a high modulation index and confine the power entirely to a single harmonic. Therefore, it can be expected that the new mode is capable of giving at least as good a signal-to-noise ratio and probably a somewhat better one than the conventional mode, although with most waveforms, any such improvement would be small.

In practice, it may be difficult or impossible to achieve this improvement in signal-to-noise ratio. It should be remembered that this is a first order theory which neglects overload effects. If the amplifier has sufficient bandwidth to pass several harmonics, these effects are likely to be severe, while, as has been mentioned previously, with the conventional mode the inclusion of the filter greatly reduces the danger of overload.

Whether the new mode is to be preferred to the conventional one can be determined only by consideration of details in individual cases. The new mode is unlikely to be preferred in situations where there are no coherent spurious signals. The new mode is more likely to be useful in the phase servo problem, which will be discussed herewith, than with the detection problem.

4. THE PHASE SERVO PROBLEM

In the previous discussion it has been assumed that the objective has been to detect the presence or absence of some weak phenomenon, and the use of a phase sensitive detector greatly increases the sensitivity of the process. There is another type of application in which the phase sensitive detector is very useful. This is when it is desired to synchronize the frequency of an oscillator to a spectral line, as in the case of a cesium beam frequency standard, or to synchronize the phase of an oscillator with some received signal of high stability, as might be the case with the reception of LF or VLF standard frequency and time radio signals.

In the following discussion it will be supposed that the oscillator is connected to a wave shaping circuit which generates the same wave form as that of the signal from the source of stabilization. Equations (4) and (6) suggest that, if the oscillator is connected to the reference input and the stabilization signal is connected to the signal input of the detector, a maximum output is obtained when the two are in phase. However, in a servo system for correcting the phase of the oscillator, it is necessary to generate an error voltage which is zero when the phase angle Φ is zero and which is an odd function of Φ. It will now be shown that such a voltage can be obtained if there is inserted in either the signal or reference line a circuit whose output is proportional to the time derivative of its input.

For the sake of definiteness it is assumed that this is placed in the stabilization input. Then, by the use of Eq. (2), the signal input becomes

$$V_s' = -\Sigma n\omega A_n \sin(n\omega t - \Phi_n)$$

$$= \Sigma n\omega A_n \cos(n\omega t - \Phi_n + \pi/2). \tag{15}$$

Then the output current i' is given by an expression which is analogous to Eq. (7),

$$i' = -bg\,\Sigma\,(n\omega A_n^{\,2} \sin n\Phi). \tag{16}$$

Equation (16) shows that i' has the desired property of being an odd function of Φ. In contrast to Eq. (7), it is to be noted that the terms that are being summed in Eq. (16) contain a factor of n. Thus the higher harmonics play a much more important role in the phase servo problem than in the detection problem.

It can usually be assumed that noise in the oscillator circuit can be neglected. A differentiating circuit is one whose amplitude response increases with frequency. Therefore, the gain G is no longer frequency independent. As an approximation, it may be assumed that the gain at the n'th harmonic is given by

$$G_n = G_o n. \tag{17}$$

Then by analogy to Eq. (12) and by reference to the paragraph which follows it, the noise output current is

$$i_N' = bg\, G_o (2Fk\, TRf)^{1/2} (\Sigma n^2 A_n^{\ 2})^{1/2}. \qquad (18)$$

Again the importance of the higher harmonics is to be noted.

The factors b and G_o are arbitrary as far as the present discussion is concerned, but it is to be recognized that there exists practical limits to their values. A discussion of these limits is beyond the scope of this paper. If the differentiating circuit is placed in the oscillator line, it can be shown easily that forms identical with Eqs. (16) and (18) are obtained except that the gain in the signal channel G appears as an arbitrary factor. Thus, as far as these first order considerations are concerned, it is immaterial where the differentiating circuit is located. However, when practical limits upon the values of these arbitrary factors are considered, a preference may be indicated. Also, if a coherent spurious signal is present, a greater suppression ratio may be obtainable with one choice than with the other.

The signal-to-noise ratio may be computed by dividing Eq. (16) by Eq. (18). In this case, algebraic cancellation does not simplify the expression to allow the substitution of the power P. Therefore, the explicit relation will not be given here.

Noise can be considered to be equivalent to a fluctuation in phase angle. The rms value may be determined by substituting numerical values of the coefficients in Eqs. (16) and (18), equating i' and i_n', and solving the resulting transcendental equation. By doing this for several alternative types or degrees of modulation, the optimum choice can be determined.

5. ACKNOWLEDGMENT

The writer wishes to thank Mr. Neil T. Larsen for helpful discussions.

6. REFERENCES

1. C.H. Walter, Zeits f. Tech. Physik, 13, 363-367 (1932).
2. C. R. Cosens, Proc. Phys. Soc. (London) 46, 818-823 (1934).
3. K. Shimoda, Buteri, 19, 120 (1964). Also 14th General Assembly of U.R.S.I., Tokyo, September 1964.

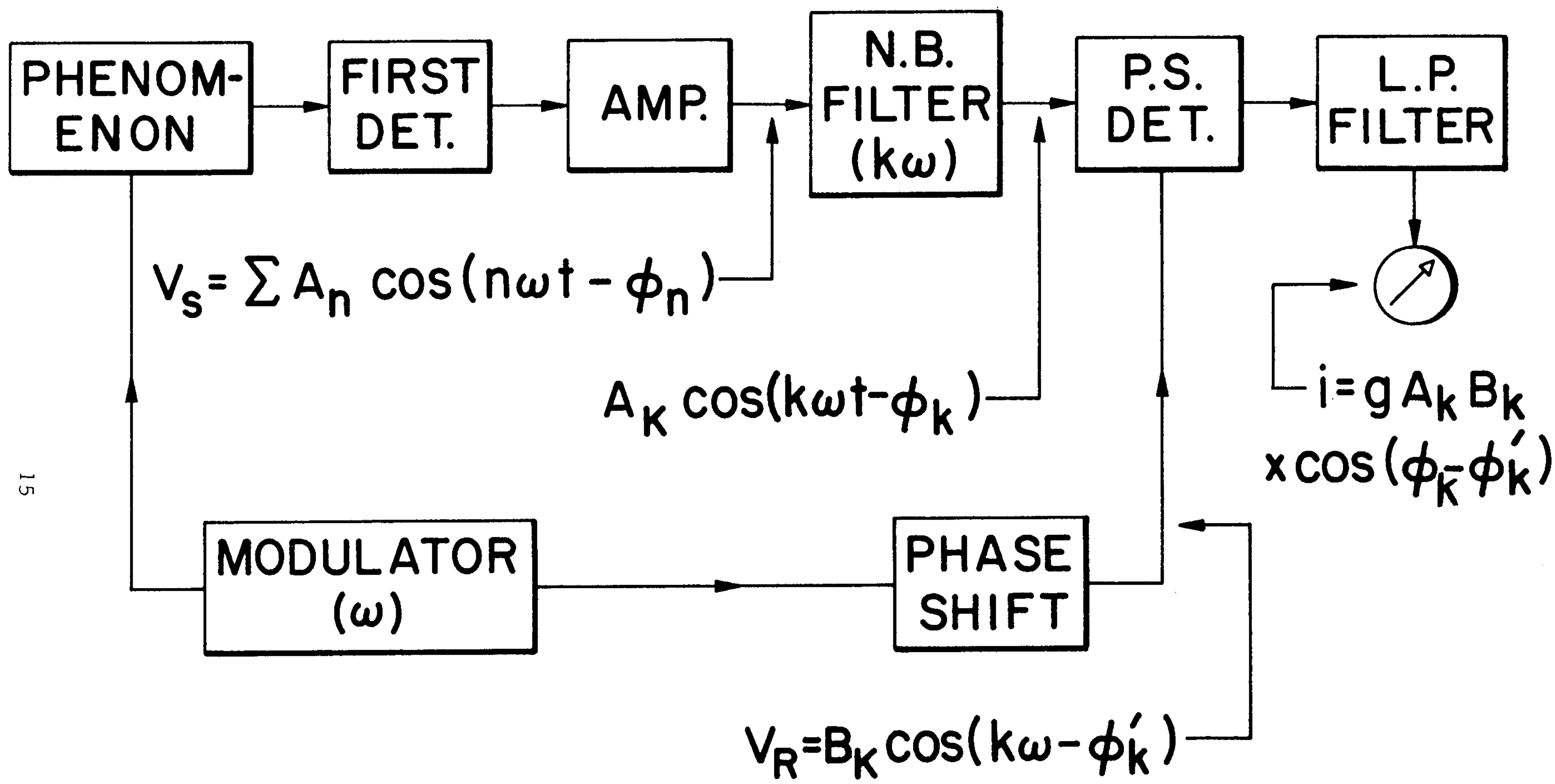

Figure 1. Block Diagram of Conventional Mode

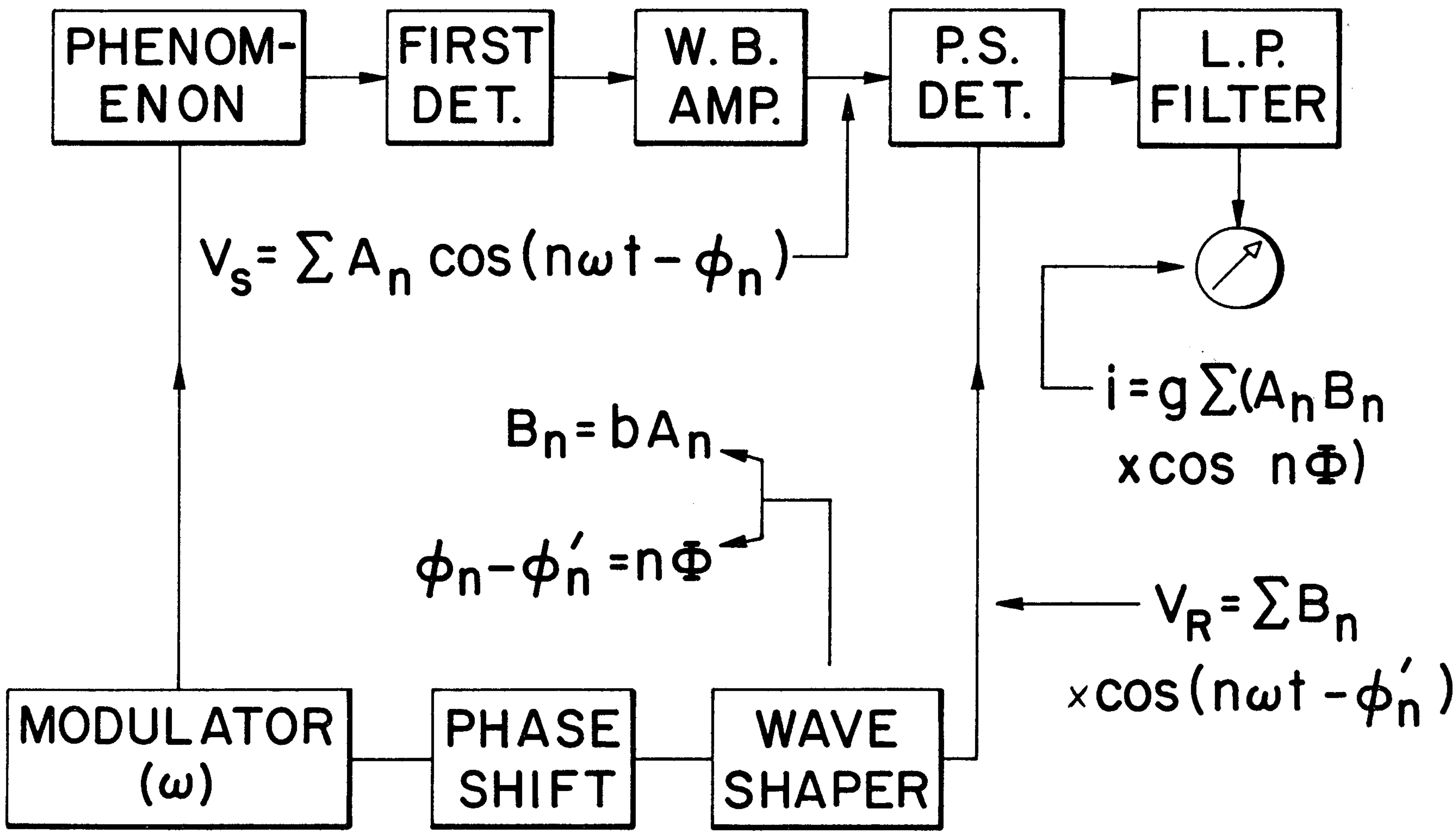

Figure 2. Block Diagram of New Mode. The narrow-band filter has been removed, and a wave-shaping circuit has been introduced into the reference line to make its wave form identical with that of the signal.