THE REDUCTION OF ATMOSPHERIC DISTURBANCES*

By

JOHN R. CARSON

(American Telephone and Telegraph Company, New York City)

Summary—*In the decade or so during which the problem of eliminating or at least reducing atmospheric disturbances has been given serious and systematic study we have learned, more or less definitely, what we can and cannöt do in this direction. For example, we know that there are definite and cannot do in this direction. For example, we know that there are definite limits to what can be accomplished by frequency selection. We know that directional selectivity is of substantial value, particularly when the predominant interference comes from a direction other than that of the desired signal, and we can calculate pretty well the gain to be expected from a given design.*

The object of this note is to analyze another arrangement which provides for high-frequency selection plus low-frequency balancing after detection. The broad idea of balancing out the interference is old, but I know of no general analysis of the arrangement. Furthermore the principle of balance has recently acquired fresh interest due to the system disclosed by Armstrong[1] *in which high-frequency selectivity and low-frequency balancing are essential features. Armstrong's scheme is treated in more detail in the latter part of this paper.*

The conclusions of this study are entirely negative, that is, no appreciable gain is to be expected from balancing arrangements. This is quite in agreement with the conclusion drawn over ten years ago by John Mills as a result of a rather extended experimental study made for the Bell System. In fact, as more and more schemes are analyzed and tested, and as the essential nature of the problem is more clearly perceived, we are unavoidably forced to the conclusion that static, like the poor, will always be with us.

I.

IN any theoretical analysis of the static problem we have to face, at the outset, the difficulty inherent in our ignorance of the origin, wave form, and frequency distribution of static. If the problem can be treated as a statistical one this difficulty, as regards practical deductions, can be successfully avoided.[2] When we wish to analyze schemes involving low-frequency balancing after detection, there are serious difficulties in the way of this mode of treatment. Furthermore, it is desirable to have an independent mode of analysis. This is furnished by the following line of reasoning.

Any disturbance, whether signal or atmospheric, can, over

* Original Manuscript Received by the Institute, April 12, 1928.

[1] PROC. I.R.E., **16**, 1, p. 15; Jan., 1928

[2] "Selective Circuits as Static Interference," *Bell System Tech. Jour.*, July, 1925.

any epoch or finite range of time t, be represented by the very general expression[3]

$$f(t) \cdot \sin(\omega t + \alpha(t)) \tag{1}$$

with the further restriction that $f(t) \geqq 0$ everywhere and ω is a constant at our disposal. Now let us suppose that (1) represents the disturbance after passing through an efficient selective circuit which confines the transmitted frequencies to those essential to the signal. Then $\omega/2\pi$ can be taken as any frequency inside the transmission band (preferably the carrier frequency of the signal) and f and α will be relatively slowly varying functions. Let us suppose, therefore, that we have a radio receiving system which employs efficient frequency selection before detection. The wave presented to the detector may then be represented by the general expression

$$W = S(t) \cdot \sin(\omega t + \theta(t)) + J(t) \cdot \sin(\omega t + \phi(t)) \tag{2}$$

where the first component represents the desired signal and the second interference. $\omega/2\pi$ is a constant, taken as the signal carrier frequency, and due to the action of frequency selection or filters,

$$\frac{1}{\omega}\frac{d}{dt}\theta(t) \text{ and } \frac{1}{\omega}\frac{d}{dt}\phi(t)$$

will be small compared with unity. In fact with small error we can write

$$\frac{d}{dt}W = \omega\left\{S(t) \cdot \cos(\omega t + \theta(t)) + J(t) \cdot \cos(\omega t + \phi(t))\right\}.$$

No other restrictions are imposed on the amplitudes or phase angles.

In the following we shall limit explicit consideration to radiotelegraph systems, whereby we are permitted to simplify the analysis somewhat by setting $\theta(t) = 0$ in (2). The extension of the analysis to radio telephony, however, presents no essential difficulties and the conclusions of the present study apply without modification to this case also.

II.

In analyzing schemes directed to solving the static problem, false conclusions, unduly favorable to the schemes under consideration, have been drawn, time after time, by reason of the

[3] See Appendix to this paper.

simple failure to compare the specific arrangement under analysis with a standard of reference. In the present discussion our standard of reference will be defined as follows. After the received wave is passed through a band filter, or a selective circuit such that the frequencies present in W, as given by (2), are confined to those essential for signaling purposes, the wave is demodulated or detected by a "homodyne" of carrier frequency $\omega/2\pi$. The detected output is then represented by

$$W' = \tfrac{1}{2}S(t) + \tfrac{1}{2}J(t) \cdot \cos(\phi). \tag{3}$$

Possibly this "reference system" requires a word more of explanation. At the transmitting station let the low-frequency signal $S(t)$ modulate a carrier wave represented by $\sin \omega t$, where $\omega/2\pi$ is the carrier radio frequency. The modulated output is then given by the expression

$$S(t) \cdot \sin \omega t.$$

Now let us suppose that $S(t)$ is representable by a series of sinusoidal terms (in the limit a Fourier's integral); that is, suppose

$$S(t) = \sum a_m \sin(\omega_m t + \theta_m).$$

Then by a well-known trigonometric formula, the modulated output is analyzable into two side bands

$$\tfrac{1}{2}\sum a_m \cos[(\omega-\omega_m)t - \theta_m] - \tfrac{1}{2}\sum a_m \cos[(\omega+\omega_m)t + \theta_m].$$

In one side band the frequency is less than that of the carrier wave; in the other greater.

We now suppose that by means of filters or otherwise one side band is suppressed and only one transmitted; say the lower side band. At the receiving station the unmodulated carrier wave $\sin \omega t$ is restored and it, together with the transmitted side band, impressed on the input circuit of a square-law demodulator. The demodulated output is then

$$\tfrac{1}{2}\sin \omega t \sum a_m \cos[(\omega-\omega_m)t - \theta_m]$$
$$= \tfrac{1}{4}\sum a_m \sin[(2\omega-\omega_m)t - \theta_m] + \tfrac{1}{4}\sum a_m \sin(\omega_m t + \theta_m).$$

The first summation represents double radio-frequency waves, which can easily be suppressed in the low-frequency circuit by selective means. The second summation is simply the original low-frequency signal. This system of transmission, it may be remarked, is employed in transatlantic radio-telephony.

Returning to (3) it will be observed that in our standard of reference, the output is linear in signal and interference amplitudes, S and J, and that the ratio of interference to signal is

$$\frac{J \cos \phi}{S}. \tag{5}$$

Since the phase angle ϕ is entirely arbitrary, the mean value of the ratio is

$$\frac{2}{\pi} \frac{J}{S}. \tag{6}$$

Any scheme, proposed for the solution of the static problem, must, in order to prove in, show a smaller ratio of interference to signal.

In the case of radio-telephony, demodulation by the homodyne principle or its equivalent is essential for high quality. In telegraphy, however, the same requirement is not present. Suppose, therefore, we examine the case of a *square-law* detector. With such a device the low-frequency output is given by

$$\tfrac{1}{2}S^2+\tfrac{1}{2}J^2+SJ \cos (\phi). \tag{7}$$

Since the phase angle ϕ is uncontrollable and, in general, variable, the last term of (7) is equally likely to be positive or negative. For the case of very strong interference, (7) becomes

$$\tfrac{1}{2}S^2+\tfrac{1}{2}J^2 \tag{8}$$

while for weak interference it is

$$\tfrac{1}{2}S^2+SJ \cos (\phi). \tag{9}$$

Comparing with (6), the corresponding expression for our standard of reference, it is seen that, in both cases, the interference ratio is increased by square-law detection.

In the case of radio-telegraphy there is another mode of demodulation, namely *straight-line* or *linear rectification.* There seems to be a more or less prevalent belief that this method possesses advantages over square-law detection, particularly as regards intermodulation between signal and interference. This belief is not justified; in fact it is an erroneous inference from the relative difficulty of analyzing the rectified output of a complex wave, and the very complex character of that output itself. In order to analyze exactly the rectified output, the wave form of the input wave must be exactly specified, this requiring information which is never available in practice. General and qualitative information, sufficient for our purposes, can, however, be deduced as follows.

Returning to the wave, W, as given by (2), as the wave impressed on the ideal straight-line rectifier, this can be written as

$$W = F \cdot \sin(\omega t + \psi(t)) \tag{10}$$

where

$$F = \sqrt{\{S^2 + J^2 + 2SJ \cdot \cos(\phi)\}} \tag{11}$$

and

$$F \cdot \sin \psi = J \sin \phi.$$

The output wave is then given by

$$W' = M \cdot W = M\{S \sin(\omega t) + J \sin(\omega t + \phi)\}, \tag{12}$$

where M is a *modulating wave* defined as *unity* when $\sin(\omega t + \psi) > 0$ and *zero* when $\sin(\omega t + \psi) < 0$.

In the idealized case when the interference is absent, the modulating function of equation (12) is given by

$$M = \tfrac{1}{2} + \sin(\omega t) + \tfrac{1}{3} \sin(3\omega t) + \tfrac{1}{5} \sin(5\omega t) + \cdots \tag{13}$$

Comparing (13) with our standard of reference system it is seen that M differs from the demodulating homodyne $\sin(\omega t)$ only by the presence of the zero frequency and harmonics of the carrier frequency. Theoretically perfect demodulation, however, results.

When, however, the signal wave is complex and in addition interference is also present, the modulating function M and the demodulated output of the rectifier is vastly more complicated. Making, however, certain ideal assumptions—the most favorable possible to the method of demodulation under immediate consideration—it may be shown that the predominant term of the demodulated output is represented approximately by

$$\frac{m}{2}\{S(t) \cos(\psi) + J(t) \cos(\phi - \psi)\}. \tag{14}$$

Here m is itself a variable depending on the relative amplitudes and phases of the signal and interference. In addition the phase angle ψ is an uncontrollable variable. In practice, in addition to the output, as given by (14) there are terms of distorted frequencies which cannot be filtered out in the low-frequency circuit.

However, the analysis of the linear rectifier is much more simply effected by employing the following approximate expression for the demodulated low-frequency output:

$$W' = \tfrac{1}{2}\sqrt{S^2 + J^2 + 2SJ\cos(\phi)} \tag{15}$$

This is the *envelope* of the high-frequency disturbance and, it will be observed, is proportional to the square root of the output of the square law detector. This formula, while in general approximate, is quite sufficiently accurate for our purposes, and is indeed favorable to the linear rectifier.

III.

The foregoing completes our discussion of the elementary theory of demodulation with particular reference to the simultaneous presence of signal and interference, which is the critical case to be considered. It remains to apply this theory to low-frequency balancing arrangements.

Common to all balancing schemes, the receiving system must include, in addition to the radio-frequency signaling channel, an auxiliary channel of substantially the same frequency-band width and preferably located very close to the former in the frequency scale. This auxiliary channel must be quite sharply selective *against* the signal itself. On the other hand it is supposed to absorb, from atmospherics or other random disturbances, substantially the same amount of energy as the signal channel; indeed this requirement is essential to the theory of the operation of the arrangement.

The received waves, after passing through frequency selective circuits in the two channels are demodulated by separate devices, and the demodulated outputs are differentially combined in a common low-frequency receiver, or receiving circuit.

The idea of the operation of this device is very simple. If the desired signal is alone present the auxiliary channel does not affect reception since, by means of frequency discrimination, it is unresponsive to the signal frequencies. When signal and atmospheric interference, or static, are simultaneously present, the latter is supposed to be balanced out in the low-frequency circuit, since the two high-frequency channels will absorb substantially the same amounts of energy from the interference.

Let us, however, examine the operation of the system in more detail, in the light of the elementary analysis developed above.

Just as in (2), the high-frequency wave in the signal channel, after frequency selection, may be written as

$$W = S(t)\cdot\ \sin\ (\omega t) + J(t)\cdot \sin\ (\omega t + \phi(t)) \tag{15}$$

In the auxiliary channel the corresponding wave is

$$W_1 = J_1(t)\cdot \sin\ (\omega_1 t + \phi_1(t)) \tag{16}$$

ω_1 may be taken as a constant nearly equal to ω, their difference depending on the frequency separation between channels. The relation between J_1 and J and ϕ_1 and ϕ will depend on the wave form of the interference.

Now let us suppose that the waves, as given by (15) and (16) are demodulated by homodyne generators of frequency $\omega/2\pi$ and $\omega_1/2\pi$. Corresponding then to (3) the demodulated output from the signal channel is

$$\tfrac{1}{2}S(t) + \tfrac{1}{2}J(t)\ \cos\ (\phi) \tag{17}$$

while that from the auxiliary channel is

$$\tfrac{1}{2}J_1(t)\ \cos\ (\phi_1). \tag{18}$$

Subtracting (18) from (17) the resultant low-frequency output is

$$\tfrac{1}{2}S(t) + \tfrac{1}{2}J(t)\ \cos\ (\phi) - \tfrac{1}{2}J_1(t)\ \cos\ (\phi_1). \tag{19}$$

Now since the phase angles ϕ and ϕ_1 are both variable and uncontrollable, it follows that the two interference components are equally likely to add or subtract so that no gain by balancing results on the average.

Suppose, however, demodulation in both channels is effected by a square law detector. The differential low-frequency output is then

$$\tfrac{1}{2}S^2 + \tfrac{1}{2}(J^2 - J_1^2) + SJ\cdot \cos\ (\phi). \tag{20}$$

Let us assume further that $J^2 - J_1^2$ can be made negligibly small, a condition at least partially realizable. The output is then

$$\tfrac{1}{2}S^2 + SJ\cdot \cos\ (\phi). \tag{21}$$

It follows from (21) that the interference is effectively suppressed *in the absence of the signal.*[4] Comparison with (3), however, shows that, when the signal is present, the interference to signal ratio

[4] See, however, the concluding paragraph of this paper.

is just twice that obtainable with our reference standard circuit. Whether or not the suppression of interference in the absence of the signal compensates for the increased interference-to-signal ratio in the presence of the signal is an open question which can only be decided by practical experience. Furthermore the balance obtainable in practice will certainly be far from perfect. It is further to be observed that since homodyne demodulation, or its equivalent, is essential for telephonic signals, no gain by balancing is possible in radio-telephone transmission. An added disadvantage, it may be noted, that attaches to balancing schemes is that the receiving system must be responsive to a frequency range double that required for the usual method of reception.

If the analysis is extended to the case of straight-line rectifier demodulation by means of (15) the general conclusions are of the same character. In fact we have for the resultant demodulated output

$$\tfrac{1}{2}\sqrt{S^2+J^2+2SJ\cos(\phi)}-\tfrac{1}{2}J_1 \qquad (22)$$

which, in the absence of the signal (spacing interval) becomes

$$\tfrac{1}{2}(J-J_1) \qquad (23)$$

As in the case of the square law detector, the gain results only during the spacing interval and is obtained at the expense of an increased interference ratio during the marking interval.

IV.

The foregoing reasoning will now be applied to the receiving system proposed by Armstrong. His arrangement specifies demodulation by rectification. Its only essential difference from the corresponding balancing scheme discussed above is that in the normal absence of the signal (spacing interval) a wave of slightly different frequency is transmitted to which the auxiliary channel is responsive. Applying (15), we have in the presence of the signal (marking interval) the following expression for the demodulated output

$$\tfrac{1}{2}\sqrt{S^2+J^2 2SJ\cos(\phi)}-\tfrac{1}{2}J_1 \qquad (24)$$

while during the spacing interval, we have

$$\tfrac{1}{2}J-\tfrac{1}{2}\sqrt{S_1{}^2+J_1{}^2+2S_1J_1\cos(\phi_1)}. \qquad (25)$$

Comparing (22) with (24) it is clear that no gain results, in the marking interval, over that procured with an ordinary balance: In fact, during this interval the outputs are identical. In the spacing interval, however, a comparison is quite unfavorable to the Armstrong scheme. For, in the usual balance system, it is theoretically possible to balance out substantially interference in the absence of the signal. In the Armstrong system, however, interference occurring during a spacing interval may result in a false signal, depending on the intensity of the interference, and on uncontrollable, variable phase angles.

There is another feature of the Armstrong arrangement which must be taken into account in comparing it with standard systems. This is that, as compared with non-balancing arrangements, the Armstrong system requires a doubling of the power radiated and a doubling of the receiving frequency band. *By discarding the balancing feature and the spacing wave, it should be possible to transmit by usual circuits, the signal message and a repetition thereof with the same power and the same frequency requirements of the receiving system.* It would be extremely interesting to have a comparison of such a system with that proposed by Armstrong.

In the foregoing discussion the possibility of balance in the absence of signal (spacing interval) was treated optimistically in order that the conclusions of the analysis should be conservative, and for the sake of simplicity. That the balance ordinarily obtainable in the spacing interval, even with the ideal rectifier, is likely to be quite imperfect may be seen by the following discussion of a case of probably frequent occurrence touched on by Englund[5] in his discussion; namely, where the interference consists of two or more overlapping disturbances. To consider this case, the interference in the receiving channel may be written as

$$J \sin(\omega t+\phi)+J' \sin(\omega t+\phi')$$

while in the auxiliary balancing channel it is

$$J_1 \sin(\omega t+\phi_1)+J_1' \sin(\omega t+\phi_1').$$

The differential rectified output is then approximately

$$\sqrt{J_2+J'^2+2JJ'\cos(\phi-\phi')}-\sqrt{J_1^2+J_1'^2+2J_1J_1'\cos(\phi_1-\phi_1')}.$$

Now even if we grant the possibility of balancing completely the terms $(J-J_1)$ and $(J'-J_1')$ the presence of the uncontrollable

[5] PROC. I. R. E., **16**, 1, p. 27; Jan. 1928.

variable phase term in the preceding makes it clear that the balance will, in general, be imperfect for the case of overlapping disturbances. As stated above, it is reasonable to suppose that this case is of frequent occurrence.

Appendix

Formula (1) can be established as follows: Any disturbance, supposed to exist only over a finite interval of time, t, can be formulated as the Fourier integral

$$I=\int_0^\infty F(\lambda)\sin[\lambda t+\theta(\lambda)]d\lambda.$$

Now write $\lambda=\omega+\mu$ where ω is a constant at our disposal; we get

$$\begin{aligned}I&=\sin\omega t\int_{-\omega}^{\infty}F(\mu+\omega)\cdot\cos\,[\mu t+\theta(\mu+\omega)]d\mu\\&+\cos\omega t\int_{-\omega}^{\infty}F(\mu+\omega)\cdot\sin[\mu t+\theta(\mu+\omega)]d\mu\\&=I_c\sin\omega t+I_s\cos\omega t\\&=f(t)\sin\,[\omega t+\alpha(t)]\end{aligned}$$

where

$$f^2(t)=I_c^2+I_s^2$$
$$\cos\alpha(t)=I_c/f(t).$$

The preceding analysis is entirely formal; its practical significance enters in only when we suppose that the disturbance has passed through an efficient selective network which confines, more or less efficiently, the transmitted frequencies to a finite band, say $\omega_1\leqq\omega\leqq\omega_2$. In this case

$$I=\int_{\omega_1}^{\omega_2}F(\lambda)\sin\,[\lambda t+\theta(\lambda)]d\lambda.$$

If we now write $\lambda=\omega+\mu$ when ω lies within the transmitted band of frequencies, the same analysis shows that if ω is large compared with $\omega_2-\omega_1$, $f(t)$ and $\alpha(t)$ are relatively slowly varying functions.

As stated, the preceding analysis applies rigorously to all types of disturbances. In the case of a radiotelegraph signal, however, it is convenient and permissible to start with the approximate expression

$$S(t)\cdot\sin\omega t$$

where $\omega/2\pi$ is the carrier frequency and $S(t)$ a low-frequency function $\geqq 0$.