FREQUENCY-SHIFT Radiotelegraph

Frequency-shift transmission of telegraph, teletype, facsimile and radio-photograph signals provides the advantages of f-m over a-m without the usual increase of bandwidth. A special circuit eliminates the effect of mean-frequency drift. A limiter that is free from loading and transients is described

COMMERCIAL radio communication by Morse Code has been standard since the advent of radio. The dot-dash system has never been entirely superseded by voice. Though code has its drawbacks and is subject to error in transmission, and requires highly schooled personnel, it is still the basic means of handling high-speed commercial and press traffic.

Anyone who has operated on noisy long-distance circuits realizes how much more accurate it is to receive by ear than by automatic slip recorders. When the signal is strong and no noise is present, tremendous speeds may be attained by using automatic equipment, but if the noise level rises, automatic equipment fails, speeds decrease and manual operation must take over; traffic piles up.

Developments over the past five years have produced a new system of automatic code transmission and reception which performs electronically what the human ear or brain does with a signal in the presence of noise. This system reaches in and picks out the signal despite the noise.

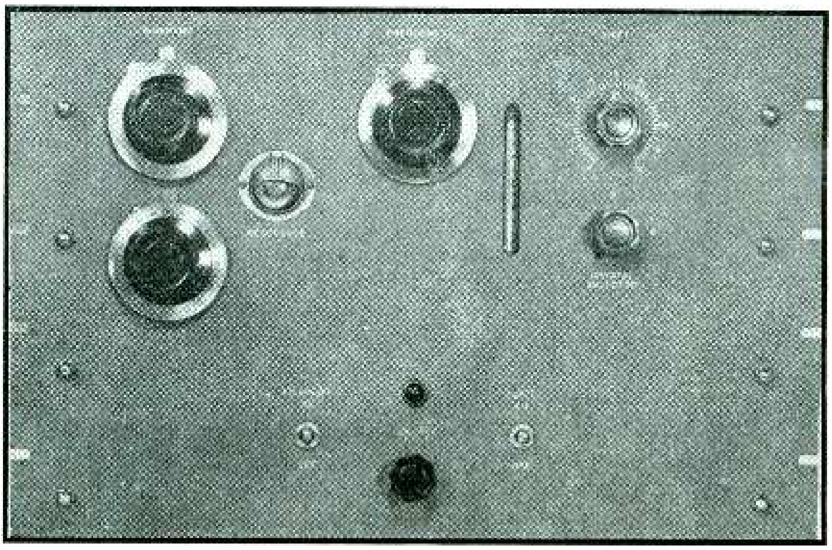
Noise Suppression Without Increased Bandwidth

Everyone is familiar with the improvements in signal-to-noise level in broadcasting that frequency modulation provides over amplitude modulation. Frequency modulation presents only one disadvantage, the required bandwidth is over ten times that required for a-m systems. Developments in the transmission of telegraph signals are now being made which are analogous to f-m broadcasting. Instead of keying a carrier on and off to designate the mark and space of telegraph signals, the transmitter maintains radiation at full power, but the carrier is shifted back and forth between two distinct frequencies to designate mark and space. The same beneficial reduction in signal-to-noise ratio is derived in transmitting telegraphy in this manner as is derived from f-m broadcasting. However, in regard to bandwidth the analogy fails. The bandwidth required by frequency-shift transmission is no greater than that required by the carrier make and break system.

This new type of transmission is not only applicable to the transmission of telegraph signals, but can also be used for transmission of teletype mark and space signals, or for facsimile where one radiated frequency corresponds to black and the other to white. This system can be extended further to the transmission of radio photographs. In this application the half tones are represented by frequencies intermediate between the extreme black and white frequencies.

Early Use of Frequency-Shift

Frequency-shift transmission was used long before it was recognized as such. In the days of arc transmitters it was impossible to interrupt the arc in accordance with telegraph signals, so instead the frequency of the transmitter was shifted during the space. Here



Front panel of frequency-shift transmitter terminal equipment used by Press Wireless, Inc.

By **ROBERT M. SPRAGUE**

Director of Research
Press Wireless, Inc.
Little Neck, New York

and Teletype System

was true frequency-shift transmission, but at the receiver no use was made of the all-important space wave. By the late 1930's the advantages which seemed inherent in this system were being developed in the laboratory.

When the Byrd South-Pole expedition of 1939–1940 was about to leave, equipment from Press Wireless laboratories was rushed into readiness and, with Times Wide World wirephoto equipment, was sent along with the expedition. Pictures published in daily newspapers during the expedition gave ample proof of the effectiveness of this type of transmission. Pictures coming over 8000 miles from a small 500-watt transmitter were received with the clarity of standard wirephoto pictures. Further developments were made over circuits from Berlin, Moscow, London and Chungking. Such transmissions are now commonplace.

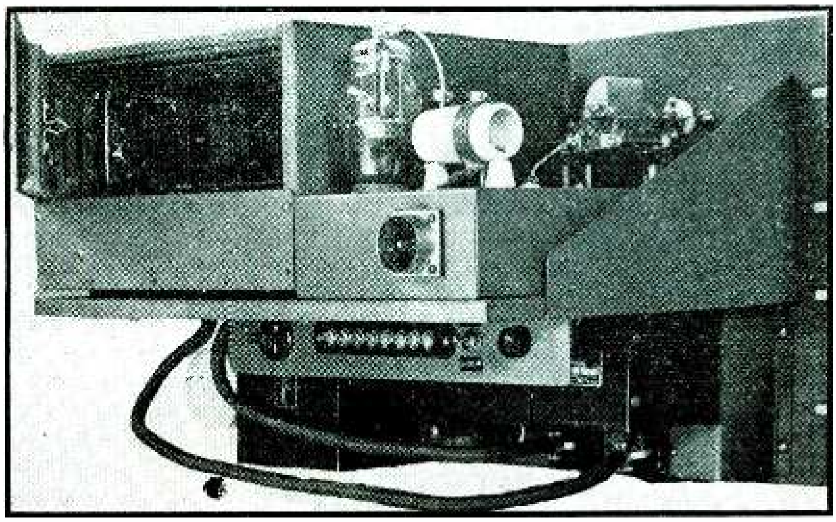
Frequency-shift transmission for radio telegraph was slower in coming, but when it was proved that a 20-db increase in signal-to-noise ratio was available, the system was applied to this service. Now many commercial circuits are operating both telegraph and teletype by means of carrier shift.

Frequency-Shift Transmitter

Several types of carrier-shift equipment are being used. The transmitting equipment developed by Press Wireless, Inc. takes energy from a crystal oscillator and beats it against an extremely stable self-excited 200-kc oscillator. The frequency of the self-excited oscillator is shifted by the signal which is to be transmitted, being increased in frequency on mark and decreased on space.

The incoming signal, usually in the form of a square wave, is filtered to eliminate frequencies higher than the third harmonic of the highest keying frequency required. The filter must be designed so that the fundamental to third harmonic phase relation is not changed. Such filtering introduces only 1.35 percent characteristic distortion, and aids in reducing the bandwidth which must be transmitted.

From the output of the mixer stage of the crystal oscillator and the self-excited oscillator, the upper side band is selected and applied to doublers and power amplifier stages of the transmitter, and radiated on frequencies varying symmetrically about the assigned frequency.



Rear view of frequency-shift transmitter unit showing crystal oven and shifting oscillator

A frequency shift of 850 cycles has been adopted as standard because it gives the best compromise between signal-to-noise level and bandwidth. However, this shift is varied between 400 cycles and 1200 cycles for special services. For example, high-speed facsimile and photograph services use a 1200-cycle shift.

Transmission Bandwidth

Consider the bandwidth for the case of teletype, the signal for which is a 23-cycle fundamental and a third harmonic of one-third the fundamental amplitude. If the carrier were keyed make and break by this signal, the ideal bandwidth would be twice the third harmonic frequency, or 138 cycles. Such a narrow band is never attained because the amplifier stages of the transmitter tend to square the keying signal. The best possible transmitter adjustment will require a bandwidth of approximately 1200 cycles. Only side bands greater than 40 db below the unmodulated carrier level are considered in this value. If, on the other hand, the carrier is shifted 850 cycles by the teletype signal, the emitted bandwidth is 1100 cycles. Were the carrier shift reduced to 250 cycles with the same signal, the emitted bandwidth would be only 480 cycles. Thus frequency shift can result in a much lower bandwidth than carrier make-break.

The calculation of frequency-shift bandwidth is comparatively simple. Assume that the carrier is shifting from mark to space frequencies sinusoidally, for teletype at a frequency of 23 cps. This is equivalent to frequency modulation

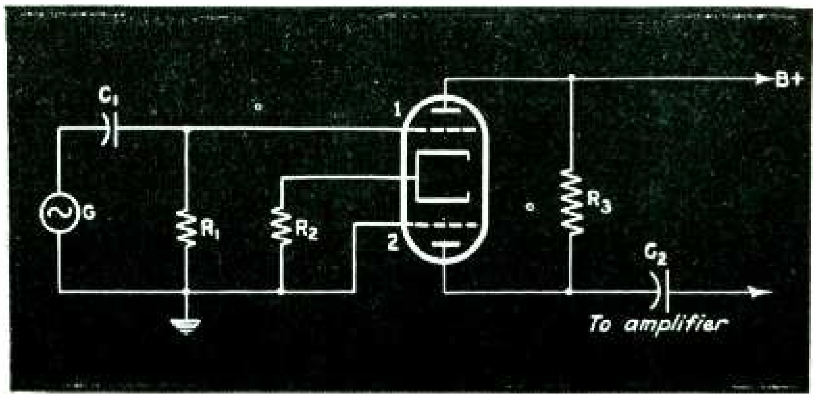


FIG. 1—Transient-free limiter in which the first section of the double triode limits the negative peaks, and the second section, cathode-coupled to the first, limits the positive peaks

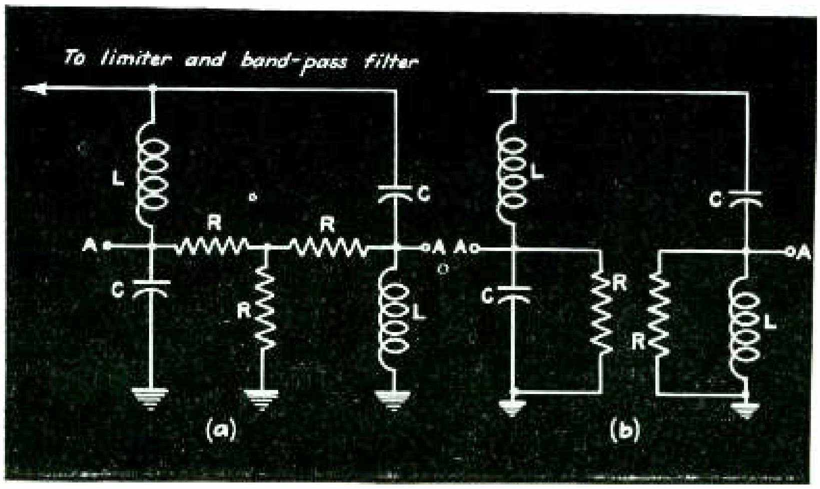


FIG. 2—Two forms of discriminator. The discriminator for frequency-shift reception must be flat over a band far wider than the deviation limits, so as to avoid amplitude modulation in the output produced by the noise in the input

of a carrier whose frequency lies midway between the mark and space frequencies, by a 23 cps signal. The modulation index B is the deviation from the carrier frequency divided by the modulation frequency. For a total shift of 460 cycles, $B = 230/23 = 10$. Under this condition there are 14 pairs of side frequencies greater than 40 db below the unmodulated carrier level. Therefore, the bandwidth is $23 \times 28 = 645$ cycles.

Now consider the third harmonic along with the fundamental. The keying wave is a signal of varying frequency and amplitude. At any instant the transmission band is determined by the instantaneous frequency and amplitude of the keying wave. Therefore, the transmission band varies between maximum and minimum values. The maximum instantaneous keying frequency, with its corresponding amplitude, determines the transmission bandwidth. This maximum instantaneous frequency in this case is 1.5 times the fundamental frequency, and its amplitude is equal to the amplitude of the keying wave. Therefore, the deviation is the same as for the actual keying signal.

The corresponding modulation index is $B' = \Delta F/1.5 f$. This value can be used to calculate all bandwidths where the keying wave is composed of a fundamental and its third harmonic. Returning to the original example, $B' = (B/1.5) = 6.67$, and the maximum instantaneous frequency is $f' = 1.5 f = 34.5$. For a B of 6.67 there are ten significant pairs of side frequencies, and therefore the maximum bandwidth is $20 \times 34.5 = 690$ cycles.

Frequency-Shift Receiver

The receiving systems of various commercial companies using frequency-shift transmission are similar in principle, but use different types of equipment.

Press Wireless, Inc. uses an a-m communications receiver which receives the radio signal and delivers an audio beat note to a band-pass filter. The beat note shifts in frequency about a mean frequency of 2550 cycles in accordance with the transmitter frequency variations. The band-pass of the filter must be wide enough to pass not only the two frequencies between which the audio beat-note shifts, but also all sidebands produced by the frequency modulation which are one-tenth or more of the carrier level. It must also be wide enough to tolerate possible transmitter or receiver drift. Noise signals outside the band are not completely eliminated but, due to filter transients, are reintroduced as signals of much lower amplitude and at frequencies within the filter pass-band.

Transient-Free Limiter

From the filter, the signal goes to a power limiter. The requirements of this limiter are much more stringent than those of the rf limiter used in f-m broadcasting because the carrier and intelligence frequencies are comparable. Thus transients must be extremely short compared to both carrier and intelligence frequencies.

Such a limiter is shown schematically in Fig. 1. The tube is a dual high-mu triode. Consider the effect of a high-amplitude signal applied through capacitor C_1 to the grid of the first section. Small negative voltages cut off the tube and so the voltage across R_2 due to current in the first triode is zero during most of the negative half cycle. The load resistance of the generator is only R_1. As the grid swings positive with respect to ground, the tube space current increases, increasing the voltage across R_2. R_2 is made sufficiently large so that at no time does grid voltage exceed cathode voltage. Thus the grid never goes positive with respect to its cathode, and the generator load is still only R_1. Since no grid current flows, no charge appears on C_1 which must leak off through R_1 in accordance with the time constant $R_1 C_1$. Therefore the circuit is instantaneous in its action and no transients result.

Also, the load presented to the generator remains constant.

As the voltage across R_2 increases because of positive swing of the first grid, the second triode is cut off. This second triode is essentially a cathode drive stage excited by the first triode acting as a cathode follower. The gain of the second triode is low because its plate resistor R_3 is small. It cuts off at about the same positive swing of the first triode grid as does the first triode for negative swings of its own grid. Therefore the action of the limiter is symmetrical about the zero axis, and is both transientless and instantaneous for any abrupt level or frequency change. The limiter gives about 30 db of limiting. Two limiter stages separated by a class A amplifier supply the needed 60 db of limiting.

Discriminator Requirements

Once the signal has been limited, it is fed into a sloping discriminator circuit prior to detection. The requirement of the discriminator is that it be linear over a frequency range far in excess of the deviation band. It need not be the back-to-back variety which is symmetrical about zero, zero voltage corresponding to the mean frequency. Following is the argument for these statements.

In f-m receivers, the tuned circuit following the limiter is usually part of the discriminator circuit. The circuit is linear only over the deviation band. Beyond this band, the response of the tuned circuit falls rapidly to zero. Therefore, any signal outside the deviation band is not received by the detector. With an input level variation of 60 db the limiter delivers a constant output. Whether the input consists of one or a thousand frequencies, the output energy is constant and is made up of one or a thousand frequencies. With one frequency fed into the limiter, the output is of the same frequency, plus its odd harmonics due to the wave squaring of the limiter. But when noise is present, side bands of the original signal become evident, and these side bands extend far beyond the band-pass filter, which accepts noise at the input of the limiter. The greater the intensity of noise in relation to the desired signal, the greater the amplitude of these side bands, and the greater their coverage of the spectrum beyond the deviation band.

In f-m receivers the tuned circuit following the limiter eliminates these frequencies which lie outside the deviation band. Since the energy output of the limiter is constant, the output of the tuned circuit following the limiter must change in amplitude as noise is impressed with the signal on the limiter input. This amplitude change will be as evident after detection as will be the desired change due to frequency modulation. Therefore, a back-to-back discriminator is required.

Such a discriminator operates in push-pull for frequency changes, but in push-push for noise-caused amplitude changes. These latter, therefore, tend to balance and be eliminated. The discriminator need not be linear beyond the deviation band, as no side bands are present, having been filtered out by the tuned circuit following the discriminator. Notice that the back-to-back discriminator is superior to a single-ended non-symmetrical discriminator only in suppressing the amplitude-modulation effects of noise, and not phase or frequency-modulation effects. These latter effects are additive in each half of the back-to-back discriminator and no noise reduction results.

Single-Ended Discriminator Preferred

In many commercial types of converters, the limiter is followed by a single-ended slope circuit. No attempt is made to eliminate noise sidebands outside the deviation spectrum. When such a discriminator is used, it must be of the extended range type, i.e., it must be linear far beyond the deviation band so as not to discriminate against noise components. Otherwise, amplitude modulation resulting in noise will result. Under these conditions, no advantage whatsoever accrues from the use of a back-to-back discriminator. After the discriminator and detector, the signal should be fed through a low-pass filter to eliminate noise caused by phase modulation of the signal at frequencies higher than the desired intelligence frequency. For teletype this filter passes the third harmonic, which may be 70 to 100 cps; for high speed Morse, 250 cps, and for photograph and facsimile, 600 cps.

An extended-range, single-ended discriminator eliminates all amplitude modulation due to noise, but

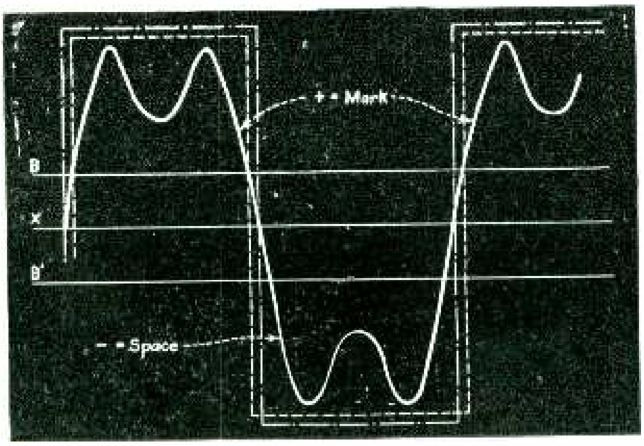


FIG. 3—The pulse consists of the fundamental and third harmonic. Pulses should vary about the central axis

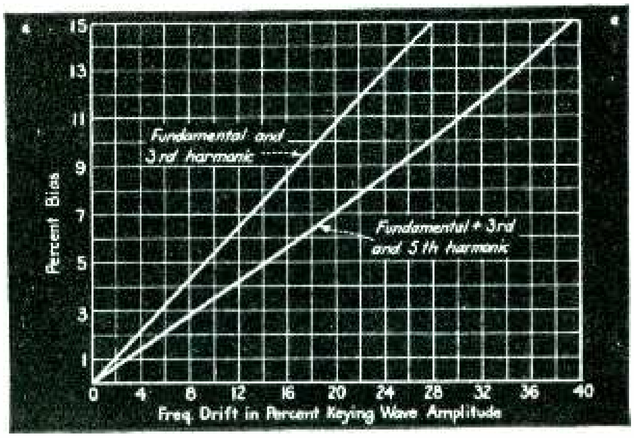


FIG. 4—Bias introduced by transmitter or receiver frequency-drift for two types of keying waves

accepts a wider band of noise to phase-modulate the signal, whereas the pre-filtered back-to-back discriminator only partially eliminates amplitude modulation due to noise, but limits the band of noise to phasemodulate the signal. Experiments have shown a small but definite advantage in the latter type; therefore its use is recommended in all types of terminal equipment for teletype, photo and facsimile.

Such a discriminator is shown in Fig. 2(a). The differential voltages to be detected are taken off at points *A*. In both Fig. 2(a) and 2(b), the input impedance of the discriminator is a constant over the band equal to *R*. *L* and *C* are given by

$$LC = \frac{1}{\omega^2} \qquad \frac{L}{C} = 2R^2$$

where ω is 2π times the cross-over frequency. The configuration of Fig. 2(a), although it does not give linear response over as wide a frequency range as does that of Fig. 2(b), delivers a higher output voltage and is perfectly symmetrical.

Frequency-Drift Compensation

The output of the detector and noise filter is the recreated pulse that originally keyed (frequency-shifted) the transmitter. If the mean frequency of the signal being fed to the converter equals the cross-over frequency of the discriminator, the pulse will alternate symmetrically about a zero axis. The solid curve of Fig. 3 represents such a keying signal, composed of fundamental and third harmonic, oscillating about the *X* axis. A local oscillator may be arranged to key on and off around this zero axis. However, if the frequency of the receiver or transmitter drifts so that the mean frequency no longer corresponds to the discriminator cross-over point, the local oscillator will no longer key on and off symmetrically about the axis, but about some other axis *B*.

When keying takes place about the *X* axis, a square wave shown by the dot-dash line results from the fundamental-third harmonic keying wave. This square wave is unbiased. That is, for equal mark and space of the transmitted wave, the square-wave mark and space are equal. However, when keying about the *B* axis, it can be seen that mark and space are no longer equal and a bias (expressed in percent as 100 times mark-length minus space-length, divided by mark-length plus space-length) results. For an acceptable minimum number of errors, teletype signals can contain no more than 5 percent bias; therefore, any frequency drift will result in more errors.

Figure 4 shows the bias introduced by transmitter or receiver drift expressed as a percent of the frequency shift for the fundamental-third harmonic keying wave and for a fundamental plus third and fifth harmonic keying wave.

To eliminate bias due to frequency shift, the circuit shown in Fig. 5 is used. Here the keying wave is developed across *R*, which is capacitively coupled through C_1 and C_2 to rectifiers T_1 and T_2 connected as shown. The load circuit consists of equal resistors R_1 and R_2, with the keying signal taken off at their junction, point *A*. As the signal across *R* increases positively, capacitor C_2 charges through rectifier T_2 thus holding point *Z* at zero potential. At the same time this positive potential is passed through capacitor C_1 to point *Y*. Rectifier T_1 is non-conducting for this polarity. As the voltage across *R* falls to zero and swings to its negative value, the charge on C_2 is applied to point *Z* and the full peak-to-peak voltage appears at *Z*. The voltage at point *Y* falls to zero and capacitor C_1 is charged through rectifier T_1 to the negative value of the voltage across *R*.

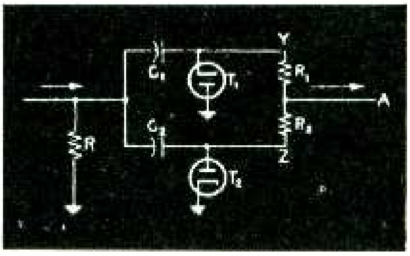


FIG. 5—Balancing circuit used to center the received pulse about the zero potential axis in the event that frequency drift throws it off

On the next positive swing of the cycle, the full peak-to-peak voltage is applied to point *Y*. Thus, point *Y* is varying between a positive value equal to the peak-to-peak voltage across *R* and zero, while the voltage at *Z* is varying between zero and this same peak-to-peak value, but in a negative direction. Therefore, the voltage at point *A* is always varied symmetrically, plus and minus, around zero. Any direct component in the voltage across *R* is not transmitted to point *A*.

The conditions to be satisfied in this circuit are that the time-constant RC_1 or RC_2 be small in comparison with the length of the keying pulses, and that the time constant $C_1(R_1 + R_2)$ is large in comparison with the time interval between pulses. With this circuit, the system is always operating essentially about the *X* axis of Fig. 3.

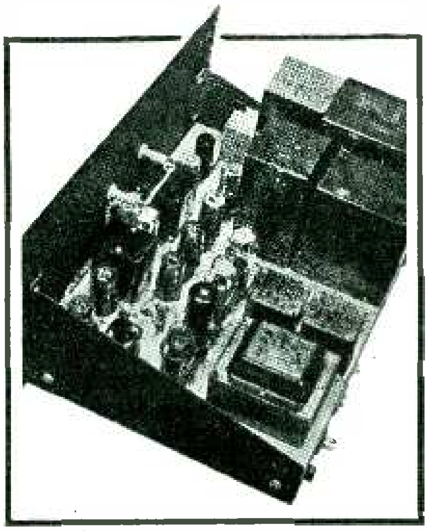
Gas Tubes Block Noise

The keying pulse, as shown in Fig. 3, is now set to key on and off around axis *X*. However, if the crest of the signal has been reached and is holding the local oscillator in its mark position, a noise crash which causes the voltage to drop below the *X* axis will shift the local oscillator, causing it to record a false space. Much greater signal-to-noise ratio can be obtained by using gas control tubes, which strike on and off only after a critical voltage is reached. Under such conditions, these tubes can be adjusted to key the mark signal on at, say, level *B* of Fig. 3, and off at level *B'*. Thus, when the keying voltage is on its positive swing a noise crash must force the keying wave to pass the *X* axis and deviate all the way to the *B'* axis before a false space will be recorded. Similarly, when the keying wave is on its negative half-cycle a noise crash must force the voltage beyond the *B* axis before a false mark will be recorded. The use of gas tubes does not introduce bias, but merely displaces the keying wave slightly, and gives a much greater signal-to-noise ratio.

Combining Circuit for Diversity Reception

Still more faithful reproduction is had by using diversity reception. However, the receiver output in a two-receiver diversity system cannot be combined directly, as fading and phase shift would cause the resultant signal to fluctuate violently in level. Furthermore, a signal fading in relation to noise on one circuit would, because of the limited action, result in a high noise output of the limiter. Instead, the signal from each receiver is fed through two similar channels, each consisting of an input filter, limiter, discriminator, detector and noise filter. Combination of the signals takes place at this point, where the frequency is low and time delay effects are at a minimum.

Referring to Fig. 6, the keying wave appears across R_1 for one channel and across R_1 for the second channel. These keying waves are varying symmetrically around the zero axis. The voltage across R_1 and R_1' are combined through four rectifiers poled as shown in Fig. 6. R_2 and R_2' are the load resistors for these rectifiers. The centers of the rectifier circuit are also connected by resistors R_3 and R_4, and the combined voltage taken out at their junction point *A*. With either a positive or negative voltage simultaneously on R_1 and R_1', half this voltage appears at point *A*. However, if due to diversity action the voltage across R_1' drops to zero, while the voltage across R_1 remains constant, the voltage at point *A* will also remain constant, thereby giving full diversity effect. Even though the voltage across R_1' should go negative when a positive voltage is desired, the voltage at point *A* will not drop to less than zero. Using the gas tube action outlined above, no false signal will be recorded as long as one circuit maintains its full voltage. Use of this type of diversity combination results in 10 to 20-db increase in signal-to-noise ratio. In the circuit shown, R_2 must be large compared to R_1, and R_3 and R_4 large compared to R_2.



Frequency-shift, converter for receivers

Cause of Lower Signal-to-Noise Ratio

The signal-to-noise improvement using frequency-shift transmission occurs for the same reasons that such improvement is obtained in f-m broadcasting. The following remarks may serve to clarify the reasons why there is such improvement. With constant-frequency carrier make-break transmission, noise in the presence of the mark signal tends to amplitude-modulate the signal in direct relation to the signal-to-noise ratio. Noise in the presence of no signal will cause interference depending only on the amplitude of the noise. Where a large mark-to-space signal-ratio is required for automatic recording, a small amount of noise will, therefore, introduce false mark and space signals.

In frequency-shift transmission, however, the signal is first fed through a limiter. This limiter eliminates all amplitude modulation caused by noise. There is always a signal feeding the limiter whether it be of mark or space frequency, and thus noise effects will act the same on mark as on space. If noise is fed into a limiter simultaneously with the signal, only phase modulation of the signal results. If this noise consists, for the purpose of analysis, of only one frequency, the phase modulation of the signal is given by

$$\frac{r^2 + r\cos\phi}{r^2 + 2r\cos\phi + 1}$$

where r is the noise-to-signal voltage ratio and is less than 1, and ϕ equals $2\pi\, ft$, f being the difference between the noise and signal frequencies. In the case of noise consisting of many frequency components, only those components whose vector sums are frequencies lying within the band passed by the noise filter following the detector need be considered. All other components are eliminated by this filter.

From the above expression it can be shown that in the case where the vector sum of the noises is as much as half that of the signal there is only one radian of phase modulation. But frequency-shift transmission employs many radians of shift. The modulation index *B* is often as high as 10 or 20. Therefore, noise, even though it be half that of the desired signal ($B = 1$) will, after detection, be small in comparison to the signal.

As much as 20 db signal-to-noise increase can be expected using frequency-shift transmission. A small, mobile 400-watt frequency-shift transmitter on the beachhead in France is transmitting press traffic to this country at a rate of 500 words a minute, over a million words a month, where in former days a 50-kilowatt transmitter had trouble in maintaining the circuit.

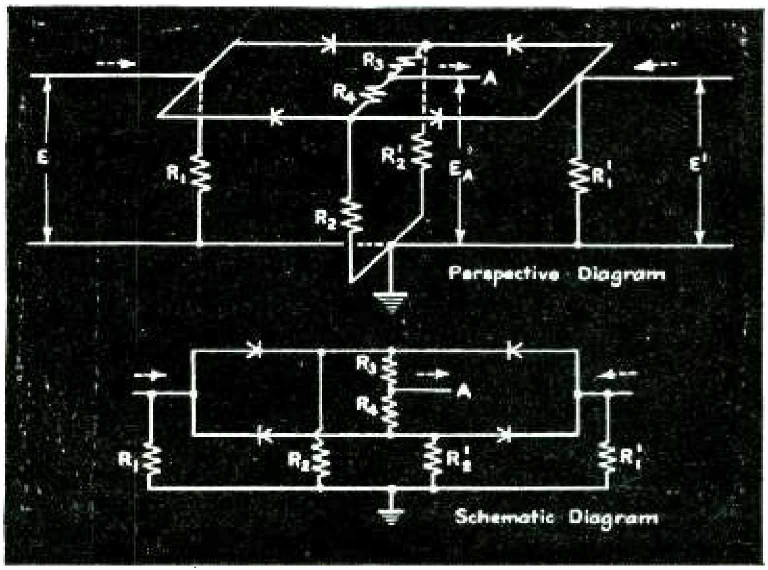


FIG. 6—Coupling circuit used between the outputs of two diversity receivers and the recording apparatus