Removing the Distance Limitation in Communication

By WILLIAM M. GOODALL

MANY READERS are familiar with the difficulties of long distance telephony. The signal currents, unless amplified, become attenuated and at a distance of a few hundred miles become lost in a background of noise. In telephone practice, this condition is avoided by amplifying the signal at frequent intervals. The amplifiers have come to be known as repeaters. Even in a system using repeaters, noise ultimately becomes a limiting factor, since the final noise output is the sum of the noise powers accumulated in each link. This is so because the noise from each link becomes part of the signal and is amplified by subsequent repeaters in the same manner as the original signal. In order to prevent the signal from being swamped by the noise, it is necessary to increase the signal power output from each repeater. A constant signal-to-noise ratio can be maintained by increasing the power of each repeater by a factor equal to the number of links in the system. Thus, for a 100-link system, the noise is increased 100 times and the power output of each repeater must be increased 100 times to maintain the same signal-to-noise ratio (S/N).

Since it costs money to provide increased power output from a repeater, it is important instead to decrease the noise arising in each link. There is a fundamental limit, however, to this process, and when the minimum noise has been attained the same limitation on power exists; it is still necessary to increase the power of each repeater in order to transmit over longer distances.

For telephone message circuits, this limitation has long since been overcome and long distance telephony is well established. For wide-band high-fidelity circuits, however, the limitation is more serious. Likewise for microwave radio relay systems, where high power is expensive and noise is more serious, the distance limitation is an important one. With this background, we proceed to a discussion of a method of transmission which avoids the addition of noise for each repeater link.

Before considering this method of transmission, it will be expedient to consider the types of signal waves used in electrical communication. It will be seen that in telegraphy, for example, a system of transmission which avoids the addition of noise at each repeater has been in use for some time. This consideration will indicate how the telephone signal could be modified to obtain the same transmission advantage.

Two broad classes of signals may be used: those with a limited and those with an unlimited number of conditions. In the first category is the On-Off type of signal used in most telegraph circuits. As is well known, perfect reception of telegraph signals can be obtained, provided only that it is possible to recognize whether the signal is on or off. Thus, it is seen that this type of signal is relatively insensitive to noise and distortion arising in the transmission medium. In fact, by regenerating the signal before it has become too badly distorted, it is possible to repeat a telegraph signal indefinitely without impairing the original message content. Regeneration is accomplished by retransmitting a new noise-free signal, controlled by a device which determines the presence or absence of a signal at the input to the regenerating repeater. Thus, in a system using regenerative repeaters, the noise does not accumulate from link-to-link and the output of the final repeater is as good as the output of the first repeater.

The second category includes unaltered signals from a telephone, those produced by a television system, and other similar signals. Here a wide range of finely graduated amplitudes or conditions is necessary for satisfactory reproduction at the receiving end. It has been common practice to transmit signals in this class so that there exists a one-to-one correspondence between the signal wave and either the amplitude or the frequency of a carrier wave. When this is done, the signal wave is transmitted by continuous modulation. Noise and distortion arising in the transmission medium add up for each repeater, as regeneration is not possible for a truly continuous signal.

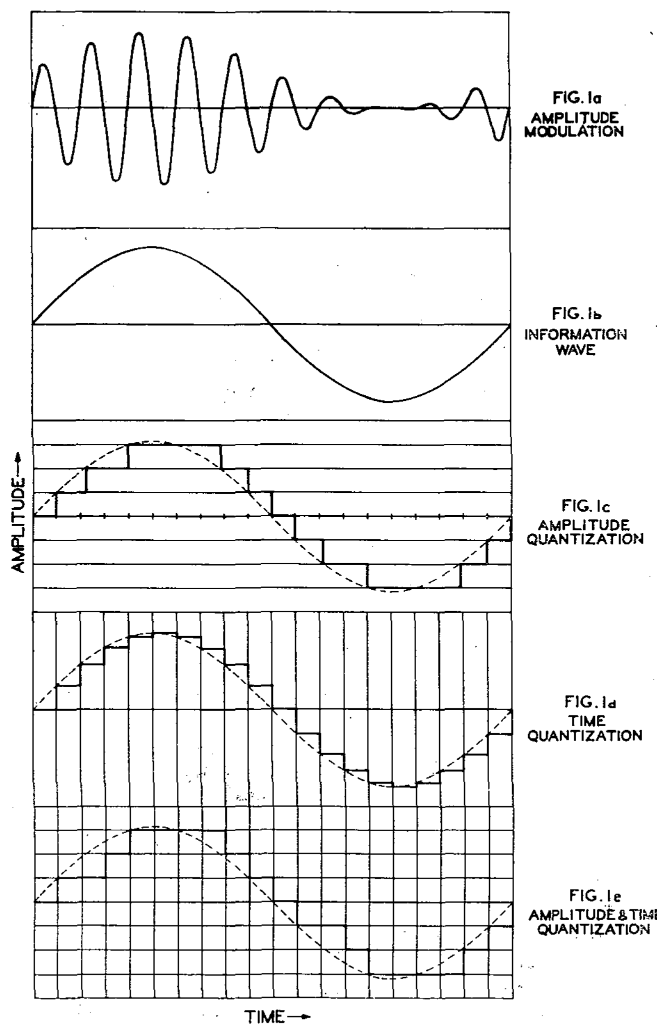


Fig. 1

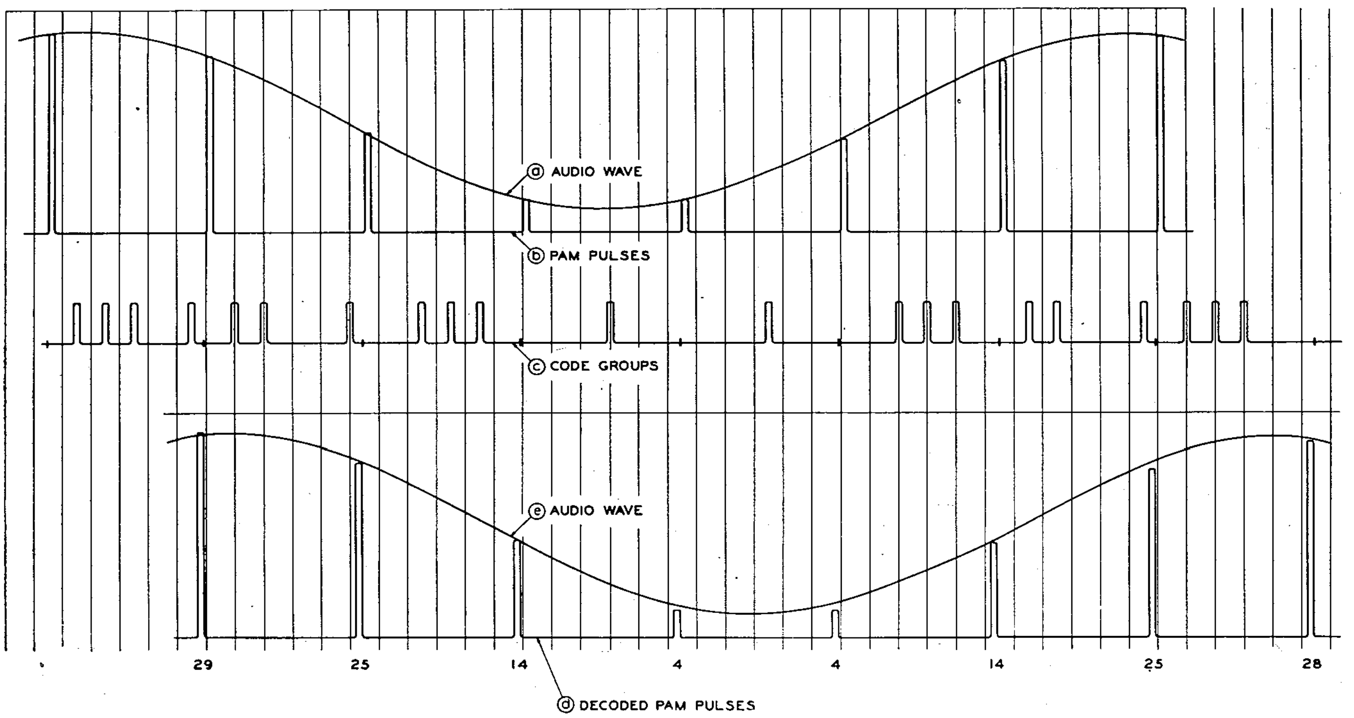


Fig. 2

Since regeneration can be used for signals representing a limited number of conditions, it would appear that it would be profitable to convert the continuous signal into a discontinuous one. The process whereby this is accomplished is termed "quantized modulation." The difference between quantized and continuous modulation is illustrated by waves shown in Fig. 1.

In continuous amplitude modulation, the instantaneous amplitude of the "carrier" varies in a continuous manner with the signal or information wave. This is illustrated by Fig. 1a, which shows the result of amplitude modulation of a carrier by the low frequency wave given by Fig. 1b. This is an example of a method of transmission where the signal-to-noise ratio obtainable in a single link limits the maximum length of the repeater system because of the accumulation of noise.

In quantized modulation, a wave quantized in either amplitude or time or both is generated under control of the information wave. Example of waves of this kind are given in turn by Figs. 1c, 1d, and 1e, where the dashed line represents the same information wave as is shown in Fig. 1b. In order to regenerate the wave, it is necessary to quantize in both amplitude and time. Thus Fig. 1e is an example of a wave that can be regenerated and repeated indefinitely without further distortion.

In Fig. 1c, which is an example of amplitude quantization, it is seen that the quantized wave changes from one level to the next at those times when the information wave has changed a full quantum. In practice, the amplitude quantum would usually be much smaller than the one chosen for this example. Even here, however, it is seen that the quantized wave gives a fair approximation to the original wave. Since it is not possible to represent the original wave exactly, a quantizing distortion results. If the magnitude of this distortion is sufficiently small, it has characteristics similar to random noise. However, the magnitude of this effect depends upon the size of the quantum chosen and not upon noise or distortion arising in the transmission medium.

Fig. 1d is an example of time quantization; here the quantized wave changes its amplitude only at definite time intervals. Pulse Amplitude Modulation, which is mentioned later in connection with Fig. 2, is another example of time quantization.

In Fig. 1e, both amplitude and time are quantized by the same rules as were used for Figs. 1c and 1d. It will be appreciated that had smaller quanta been chosen, the quantized wave would be a closer approximation to the original wave.

The quantized modulation systems just illustrated are not commonly used in practice. At this point, a description of a new system known as Pulse Code Modulation (PCM) will be given, since it is a promising application of quantized modulation. PCM involves the application of three basic concepts. Two of these have been mentioned before; they are the time and amplitude quantization principles. The third concept may be described as the digital coding principle. These concepts will be discussed in the order mentioned.

The particular form of time quantizing used in pulse systems is known as sampling. The essence of the sampling principle is that any input wave can be represented by a series of regularly occurring instantaneous samples, provide that the sampling rate is at least twice the highest frequency in the input wave. This result is well known and is the general basis for time-division or sampling systems.

For present purposes, the amplitude quantization principle states that a complex wave can be approximated by a wave having a finite number of amplitude levels, each differing by one quantum, the size of the quantum jumps being determined by the degree of approximation desired. This process, while comparatively new, appears to be of basic importance in many modulation processes.

Reference should now be made to the curves of Fig. 2. Here the information wave is given as curve (a). The sampling process yields the Pulse Amplitude Modulation (PAM) wave (b). This curve, of course, is an example of a time quantized wave. Each sample is coded to produce the code groups (c). This coding is an application of the digital coding principle discussed below. The decoded pulses yield the amplitude quantized PAM pulses (d). The numbers below the pulses of curve (d) represent the quantized amplitude as an integral decimal number. Curve (d) is an example of an amplitude and time quantized system.

The digital coding principle allows the representation of the quantized amplitude or number by the digits of a number system. Thus, each quantized amplitude is represented by a code group of digits, the number of digits required being determined by the size of the quantum. For example, binary digits can be represented by two position or On-Off pulses. In a binary number system, each digit can have two values, while in a decimal number system, each digit has ten values. For binary numbers, then, 2^n discrete levels can be represented by a binary number of n digits. Thus, binary PCM represents each quantized amplitude of a sampling process by a code group of On-Off pulses, where these pulses represent the quantized amplitude in a binary number system.

The fundamentals of the digital coding principle are illustrated in Fig. 3. Here, beginning at the left, we have the quantized amplitude, the binary representation of that amplitude, the weighted equivalent, and the decoded number. A PCM system codes the quantized amplitude to obtain binary code groups. These code groups are sent over the transmission medium to the receiver, where they are decoded. One process of decoding involves obtaining the weighted equivalents and adding the result to obtain the decoded amplitude shown on the right.*

The binary number system is illustrated by the second column of Fig. 3. With five binary digits, all integral values between 0 and 31 can be represented.

*"An Experimental Multi-Channel Pulse Code Modulation System of Toll Quality" by L. A. Meacham and E. Peterson. Paper presented in BELL SYSTEM TECHNICAL JOURNAL, January, 1948.

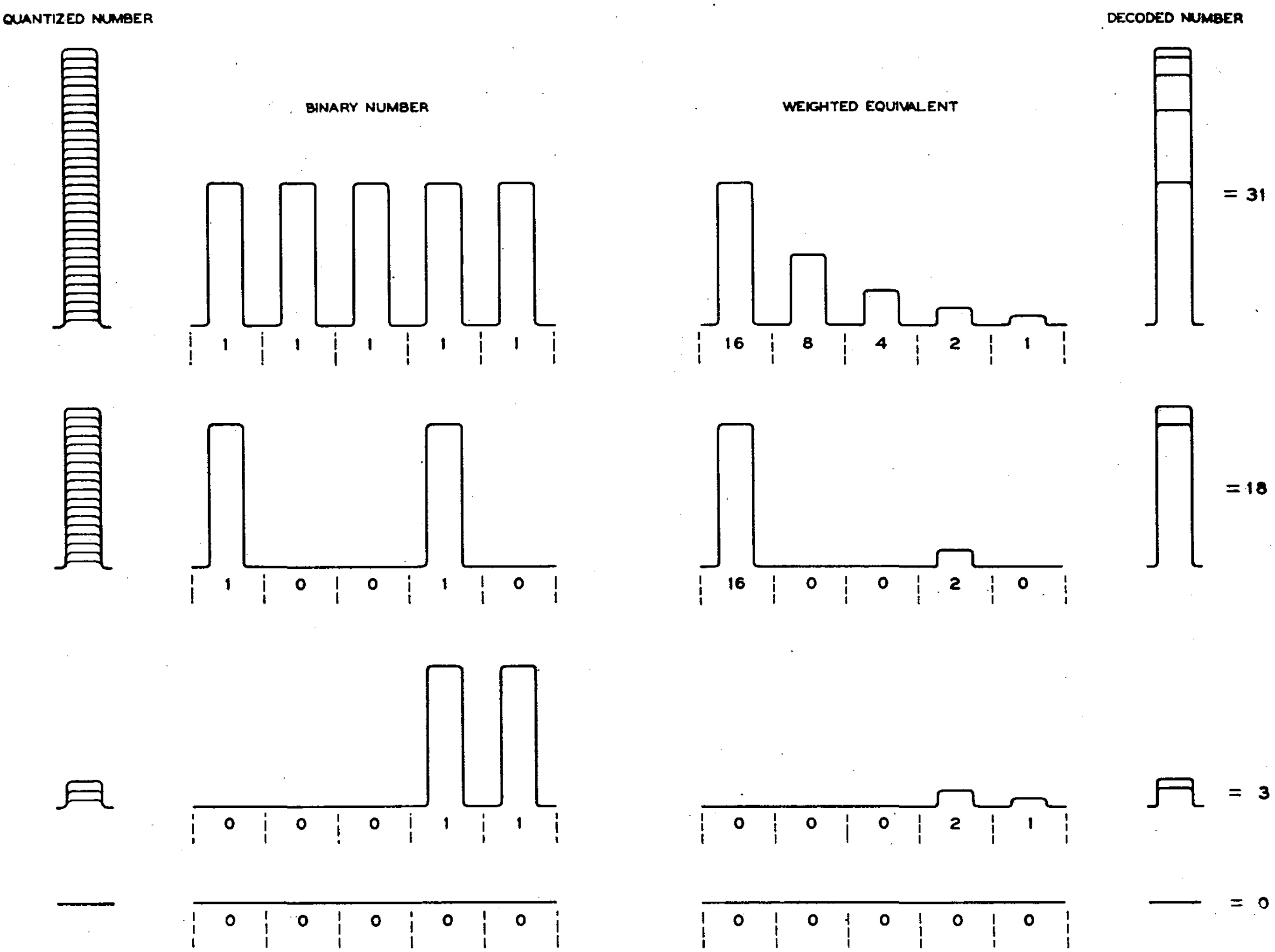


Fig. 3 The four columns in the figure illustrate the steps involved in the representation and subsequent interpretation of numbers in the "binary" system of notation. The first, or "Quantized Number" column, represents the desired number as a sum of "ones" (e.g. 3 = 1+1+1). The second, or "Binary Number" column, shows how the number may be written as a sum of powers of 2 with coefficients zero or one (e.g. $18 = 1\times2^4 + 0x2^3 + 0\times2^2 + 1\times2^1 + 0\times2^0$. This process is completely analogous to the usual or "decimal" representation of a number in which, for example, 5301 represents $5\times10^3 + 3\times10^2 + 0\times10^1 + 1\times10^0$. Thus 31 may be written as 11111 and 18 as 10010 in the binary system. The third or "Weighted Equivalent" column depicts in decimal notation the individual terms involved in the binary representation (e.g. $18 = 1\times2^4 + 0\times2^3 + 0\times2^2 + 1\times2^1 + 0\times2^0 = 16+0+0+2+0$). The last column, of course, represents the actual numbers desired as a total made up of the parts indicated by the previous column.

Thus, the top row of this column represents the number 31 in the binary system. Likewise, the remaining rows represent the numbers 18, 3, and 0 in turn.

Although any signal quantized in time and amplitude can be regenerated, one involving two position or On-Off pulses is almost ideally suited to this process. Thus, it is seen that binary PCM is likely to become a very practical example of a system of transmission that overcomes the distance limitation in the transmission of wide-band high-fidelity material over radio relay systems.

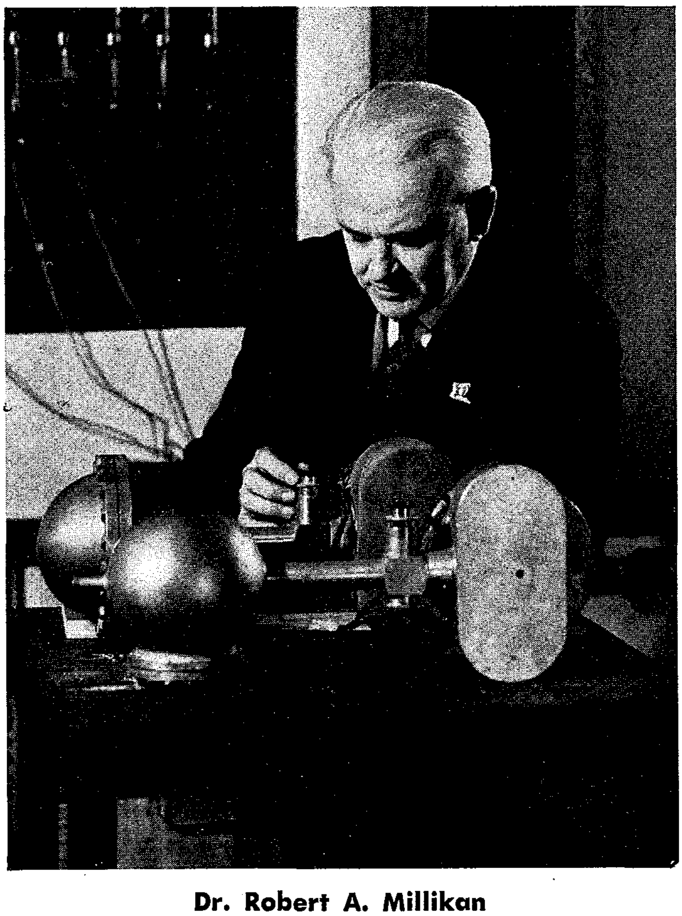
In addition to the regenerative feature of PCM, another important property involved in the digital coding process is brought out by the following discussion. In microwave systems, where relatively wide bands are available, it is desirable at times to trade band width in the medium for signal-to-noise advantage. Wide-band frequency modulation is a well-known example of this method of operation. Here if the band width is doubled, the signal-to-noise power ratio is increased by a factor of 4. For a PCM system, however, a much greater factor is obtained. If we double the band width, we can send twice as many digits. Thus, if we start with a 5-digit system and double the band width, a 10-digit system results. Since a 5-digit system has 32 levels and a 10-digit system has 1024 levels, the signal-to-noise power ratio has increased by $(1024/32)^2 = 1024$ times. From this discussion, it appears that the digital coding principle allows the trading of band width (necessary for the increased number digits) for noise on a much more favorable basis than that realized in ordinary frequency modulation.

Summarizing: We have seen that for continuous modulation systems the accumulation of noise in a repeater system results in a definite limitation to the length of the system. By using quantized modulation where both time and amplitude are quantized, it is possible to use regenerative repeaters and avoid the accumulation of noise in a long system. In addition, PCM, which is an example of a signal that can be regenerated, also has the property of trading band width for noise on a very favorable basis.

The curves shown in Figs. 2 and 3 of this article are taken from a paper by the author, "Telephony by Pulse Code Modulation," BELL SYSTEM TECHNICAL JOURNAL, July, 1947.

ROBERT A. MILLIKAN CELEBRATES 80TH BIRTHDAY

ON MARCH 22, Dr. Robert A. Millikan celebrated his 80th birthday. In honor of this occasion, the Associates gave a dinner on March 15, at the California Club, Los Angeles, and the faculty, under the chairmanship of Professor William Fowler, are planning a dinner early in April at which Dr. Millikan will be presented with a special issue of "Review of Modern Physics", published in his honor. This volume of the "Review", combining both the January and April issues, is comprised of articles on Millikan's scientic interests written by his former students and associates.



Dr. Robert A. Millikan

Dr. Millikan, having retired in 1946 from his administrative positions as director of the Norman Bridge Laboratory of Physics and chairman of the Executive Council at the Institute, is now acting in an advisory capacity as vice-chairman of the Board of Trustees and continuing his research work and writing. Prior to his retirement, Millikan had been with the Institute since 1921 when he left his position as professor of physics at the University of Chicago, where he had been in the Physics Department for 25 years.

A native of Iowa, Dr. Millikan received his A.B. degree from Oberlin College in 1891, his Ph.D. from Columbia University in 1895, and took advanced study in 1895 and 1896 at the Universities of Berlin and Gottingen. Although he took only one semester of work in physics during his undergraduate course, Millikan acquired a lifetime interest in this field when he helped work his way through college by teaching elementary physics during his junior year, and upon graduation, when he accepted a position as a physics tutor at the Oberlin Academy.

A recipient of many medals of honor, among them, the Nobel Prize in Physics from the Royal Swedish Academy in 1923, the Faraday Medal from the Chemical Society of London in 1924, and in 1937 the Franklin Medal from the Franklin Institute of the State of Pennsylvania, Millikan is best known for his work on the isolation and measurement of the electron; the direct photo-electric determination of the fundamental radiation constant known as Planck's **h**; his study of Brownian movement in gases; his more recent study of the nature and properties of cosmic rays, and the 18 books of which he is either the author or joint author.

A former president of the American Physical Society, and the AAAS, Dr. Millikan is an honorary member of many other scientific societies, among them, the National Academy of Science, the American Philosophic Society, and several foreign groups, and has received honorary degrees from 20 colleges and universities in this country and abroad