BSM # 778 B

SAGE DATA TRANSMISSION SYSTEMS

GENERAL CONSIDERATIONS AND DESCRIPTION

CONTENTS PAGE

1. GENERAL

1.01 The SAGE system is a system of air surveillance and control designed to aid in the air defense of the continental United States. In this system of air surveillance, flight information is transmitted from radars and other sources to a central point where it is evaluated to form a comprehensive picture of the air situation. This picture covers the search area of several large radars, hence centralized surveillance over a sizeable geographical area, called a "subsector," is provided.

1.02 The continuous tracking of aircraft over an area the size of a subsector requires the gathering and evaluation of a great deal of information. Most of the basic flight information is gathered, as noted above, by radars of various types. The evaluation of this information is performed at a central location within the subsector called the "direction center." This evaluation of flight information must necessarily include correlation between the returns of the radars that have overlapping coverage. Further correlation of the received information with the flight plans of friendly aircraft furnished by the CAA and the military must also be performed.

1.03 The end result of this is a continuous picture, at the direction center, of the air situation throughout the subsector. To be of any use this picture must not lag the actual air situation by any very appreciable interval of time. In order to accomplish this the large amount of information on which the situation picture is based must be gathered and evaluated very rapidly. To achieve the required speed both processes have been made largely automatic and the function of human beings in this part of the system correspondingly reduced. This is one big advantage that SAGE has over earlier surveillance systems which largely depended on human beings to handle and evaluate the flight information. This change in technique has resulted in a very fast-acting surveillance system.

1.04 The word "SAGE" is broadly descriptive of the surveillance system. It stands for "Semi-Automatic Ground Environment." "Ground Environment" is a term coined by the Air Force which, in the case of SAGE, stands for a complex of electronic equipments installed at ground locations to handle large volumes of information or data relating to the air situation.

1.05 The various automatic data gathering and processing units in the SAGE system are necessarily geographically scattered. Information flows between the units via special communication systems called data transmission systems. The main interest herein will lie in the general characteristics of these systems and the information or data that is transmitted over them.

2. CHARACTERISTICS OF DATA

2.01 The information that is transmitted between the automatic units of the SAGE system consists mainly of either descriptive or directive information relating to aircraft in the air space over the subsector. The descriptive information conveys the successive geographical position and height of the aircraft being tracked, while the directive information transmits course information, etc., to the aircraft or other weapons that are being controlled. All this information is coded in the form of numbers since these can be more readily handled by the automatic units of the system.

2.02 In this information coding process the end result is not a series of numbers in the conventional decimal system but a series of numbers in the binary system. The binary number system represents number quantities by means of only two digits - 0 and 1. The following table illustrates this method of numbering by tabulating some binary numbers opposite their decimal equivalents:

Table I

Decimal Number	Binary Number	No. of Digits	
		Decimal	Binary
0	0	1	1
1	1	1	1
2	10	1	2
3	11	1	2
4	100	1	3
5	101	1	3
6	110	1	3
7	111	1	3
8	1000	1	4
9	1001	1	4
10	1010	2	4
100	1100100	3	7
1000	1111101000	4	10
10000	10011100010000	5	14
100000	11000010010100000	6	17
1000000	11110100001001000000	7	20

2.03 A comparison of the few numbers presented in the table illustrates the salient differences between these two numbering systems. It is apparent that the representation of numerical quantities above "1" in the bianry system involves the use of greater numbers of digits than are required to represent the same quantities in the decimal system. The representation of the quantity "1000," for example, requires the use of ten digits in the binary system and only four in the decimal system. This is a result, of course, of the fact that the binary system base is two, while that of the decimal system is ten. On the other hand the use of the decimal system requires the handling and recognition of ten different symbols - 0, 1, 2, 3, 4, 5, 6, 7, 8 and 9 - while the use of the binary system requires but two - 0 and 1. Thus any user of the binary system would need to be able to handle a large number of digits for even fairly simple arithmetic operations, but would only need to know and recognize either a 0 or 1. A user of the decimal system would have to manipulate a much smaller total number of digits for the same operation but would have to be able to recognize eight additional symbols. A corollary logically follows - if the two users are to work the same problem in a comparable time, the binary user must handle digits much faster than the decimal system user.

2.04 The fact that the binary system uses only two digits gives it an important advantage over the decimal system for certain users. Human beings are not among these - so far as is known the binary system has never been used by them directly for numbering and computation. The smallest number base known to have had wide use was six and the largest was twenty, in fact practically all the numbering systems known have either six, ten or twenty as a base. Since human beings can readily recognize and remember numerous symbols, numbering systems using six, ten or twenty different symbols are easily used. Furthermore, as the size of the number base is increased, the actual number of digits required to represent a given number decreases. Since the speed with which human beings can handle digits (writing or recognizing them), is limited, their speed of repetitive computing depends largely on the number of digits that have to be handled in the operation - hence the general use of number bases larger than 2. The situation reverses, however, when the user of the numbers is a modern computer. These machines are limited in their capacity to recognize different symbols or digits but they can very rapidly handle large numbers of digits that they can recognize.

2.05 This limited digit recognition ability of modern computers stems largely from the limitations of their components. These computer components are made up of numbers of electrical circuits and associated devices. It is fairly easy to design and construct these so that they are capable of rapidly assuming or recognizing either of two stable conditions or states. Components capable of rapidly assuming or recognizing each of three stable states are more difficult and expensive to construct and these factors increase more than linearly for, say, a six or ten state device. (This assumes comparable speeds of operation for all four devices.) The two-state device is fundamentally adapted to work with the binary system since it can rapidly recognize or represent either of the two digits that this system uses. Increasing the number base to six or ten would involve the more complex circuits mentioned and, while the computing operation can be simplified in that a smaller number of digits have to be handled, the time-saving in the operation still may not make up for the increased cost of the components. Computers have in fact been built for other number bases than 2, but for the mass of information that the SAGE system must process, and the speed with which this must be accomplished, a computer working mainly in the binary system has been found to be most practicable.

2.06 Since the binary number system has only two digits, binary numbers can be represented very readily by electrical means. As an example of this suppose that it is desired to represent the decimal number 45. In the binary number system this is 101101 and it can be represented electrically by a series of "on-off" type of pulses as shown in Fig. 1. Here the number is represented by an array of rectangular voltage pulses plotted against a time scale. The number starts at time t_1 with a marking pulse of amplitude "e." This is the first binary digit, 1. The second digit, 0,is spacing or zero voltage, the third and fourth digits are marking pulses of "e" amplitude, the fifth and sixth digits are represented by a space and marking pulse, respectively. The number is completed at the end of the last pulse, time t_2. It will be noted, that to represent the binary number, the time interval between t_1 and t_2 was divided into six equal time slots and that the number was represented, not by the total number of marks and spaces, but by their sequence of occurrence. Digital data systems are designed in this manner, that is, the information is sent at a fixed rate so every binary digit, either the 1 or 0, lasts the same length of time. The number of these binary digits that can be handled is a good measure of the information capacity of any data system. In this connection the term binary digit has been abbreviated to "bit" and is quite widely used in the data transmission field. It does not always indicate that binary numbers are being transmitted but it does refer to the smallest time interval in any data system that may be occupied by a mark or space.

2.07 The use of binary numbers in connection with a typical air surveillance problem is illustrated in Fig. 2. This is a much simplified diagram of the way in which the air situation in the range of one radar is displayed

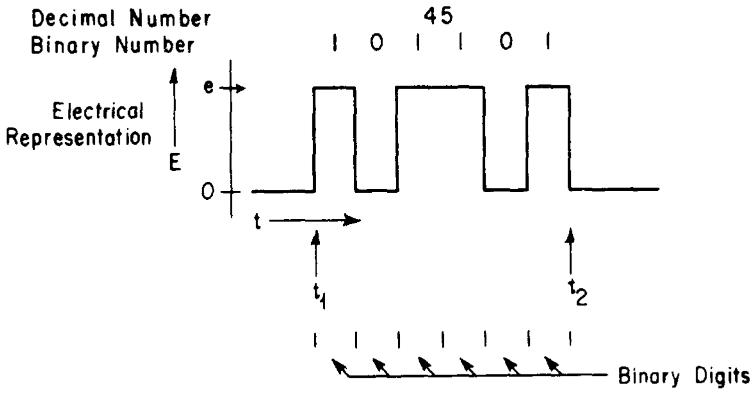


Fig. 1 - Decimal and Binary Numbers

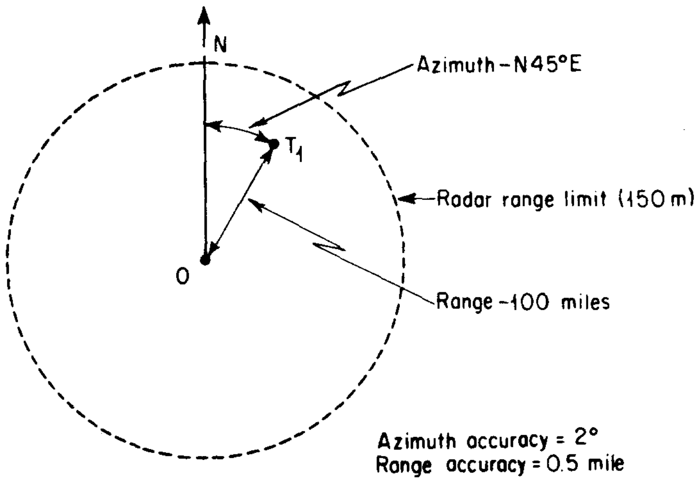


Fig. 2 - Radar Plot

on its monitor. The whole area is shown as a circular plot with the radar at the center point "O." The radar continuously scans this complete area. Its antenna generally rotates about 4 revolutions per minute, and as targets are illuminated by the radar beam they appear as dots of light on the monitor. The position of only one target is shown in Fig. 2, this is at T_1. As long as this target stays within range of the radar, its position will be shown on the monitor whenever the radar beam passes over it - in this case an indication of its position will be shown every fifteen seconds. Assuming that this is a valid target, if effective countermeasures are to be taken, the air surveillance system must provide information from which its course, speed and height can be determined. A radar of the type shown provides only the first two items of information; the target height is usually determined from separate height finding radars. The course and speed of the target, of course, may be readily deduced by observation of the plot in Fig. 2, as the successive target positions are displayed.

2.08 If the course and speed of the target were needed only at the radar location, the air surveillance problem would be largely solved by the radar plot shown. This, however, is only one of many scattered locations that must have this information so it must be relayed to these other points in some manner. There are various method of accomplishing this; in part of the SAGE system the distance of the target from some reference location, usually the radar, and the target bearing from some reference direction are automatically reported at regular intervals. These two quantities are known as target range and azimuth, respectively; they are usually transmitted each time the radar beam illuminates the target. At the instant shown in Fig. 2, the target azimuth is N45°E and its range is 100 miles.

2.09 Of these two items of information that must be relayed only one, the azimuth information, contains both letters and numbers. For background purposes, Fig. 3 illustrates some ways in which this azimuth information may be coded for transmission. The resultant idealized voltage pulses are also shown for each coding. At A the information is coded in International Morse Code with the dots discriminated from dashes in the data transmission system on a time basis - that is the dash pulse is twice as long as the dot pulse. Letters or numbers are separated on a time basis by allocating a space equal to three dots. Assuming that the dot interval is the shortest interval that the system can transmit, this method of coding results in a message length of 33 bits. A second possible method of dot-dash discrimination is shown at B. Here the information is still coded in Morse but the dots and dashes are separated on an amplitude basis - a dot pulse has only half the amplitude of a dash pulse. Letters or figures are separated by spaces equal to two pulses. This method of coding results in a reduced message length, 28 bits. A further reduction is possible, however, when it is considered that, for this information, only two of the four letters, N, S, E, or W will ever have to be transmitted. Furthermore, the largest number that will have to be transmitted is 89, assuming the radar only gives the azimuth to the nearest degree. If the letters are arbitrarily coded as the first four numbers of the binary system, as shown at C, the message structure can then be as shown. The first two bits are reserved for the first letter of the bearing, the next seven bits carry the numbers, (in binary form) and the last two bits carry the concluding letter of the bearing. With this coding and message structure the message length is 11 bits. A fourth method of coding is shown at D. Here the convention used is that the target bearing will be given as the numerical value of the total angle measured clockwise from the reference, or north, azimuth. In this case the transmission system need only carry one out of a range of numbers from 0 to 359. This reduces the total message length to 9 bits. Fig. 4 illustrates the application of this technique to other target positions.

2.10 The transmission of the range information may be accomplished in the same manner except that no letter convention is needed to simplify the process since the range is always a number and no letters are involved. In the plot shown in Fig. 2, the maximum range of the radar is assumed to be 150 miles. For purposes of illustration it is further assumed that the range indications are given in mile increments. Thus the range information to be transmitted will consist only of a single number ranging in value from 0 to 150. For the target shown the range is 100 miles and the complete azimuth and range message for this target position could be coded as shown in Fig. 5. Part of the radar information handled in the SAGE system is coded along these general lines, that is only the pertinent information relating to each target is sent. The actual position message used contains radar location indications and some additional information which results in a considerably longer message. The whole message, however, is still made up of binary digits.

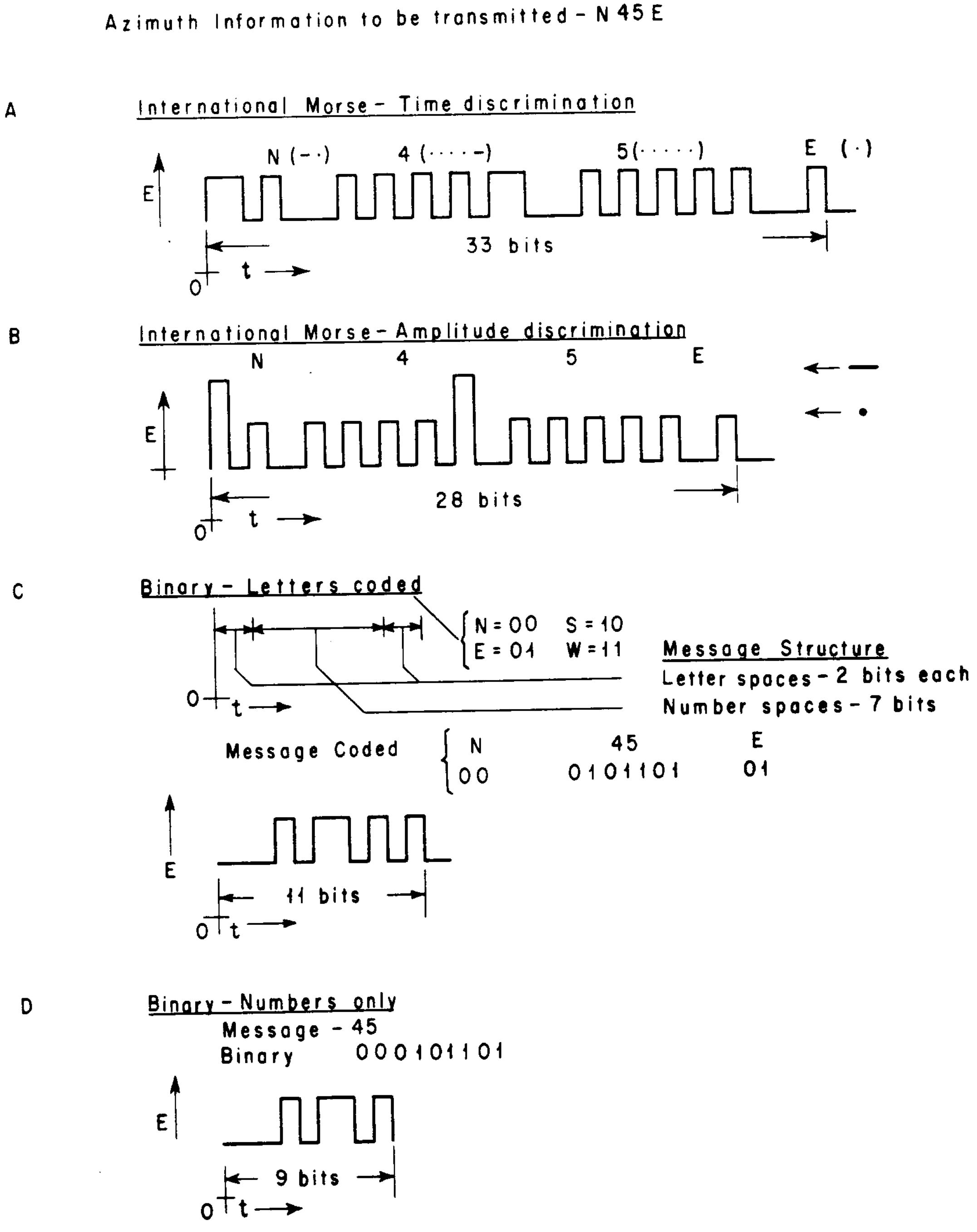
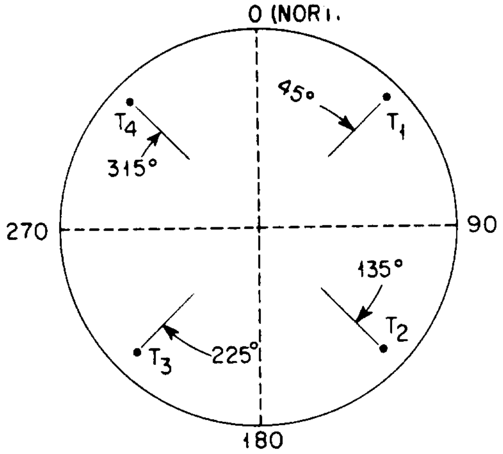


Fig. 3 - Azimuth Information



Target	True Bearing	Angle from Reference (N)	Binary Number
T_1	N45°E	45°	000 101101
T_2	S45°E	135°	010 000111
T_3	S45°W	225°	011 100001
T4	N45°W	315°	100 111011

Fig. 4 - Additional Target Positions

2.11 There is also another way of relaying target positions that is used in the SAGE system. This method consists of sending enough information so that the whole plot of Fig. 2 can be reproduced at the distant points. This reproduction does not give the fine detail available on the original plot but it is good enough so that target range and azimuth can be obtained from it with some accuracy. The accuracy of this method, however, is not equal to that of the first method described.

2.12 With the preceding as a background it is now appropriate to summarize the broad characteristics of SAGE data. These are:

(a) The data is digital in form but is not purely binary.

The term "digital" is a broad term which generally signifies that the information is transmitted over the data system via fixed sequences of bits. In these sequences, the array of marks and spaces carries the information and if only one marking amplitude is used, the system is also purely binary. SAGE data systems, however, use two significant marking levels.

(b) All the digits used in a particular data system have the same time duration.

On the basis of bit rates, it presently appears that the SAGE system will directly use two types of data systems. These systems use bit rates of 1300 and 1600 bits per second. These rates are constant, hence all digits in a system are the same length. This does not hold, however, for data systems owned by other branches of the military with which SAGE will connect.

(c) The data messages are long.

Data "messages" are a collection of significant data sequences that are closely relevant. For example, there are two significant data sequences of 9 and 8 bits, respectively, in the azimuth and range message of Fig. 5. These data sequences are also called "words"; the whole data message is made up of several words. To complete the message of Fig. 5, data sequences containing the radar location, target identification and additional information are added and the whole array of bits thus constructed makes up the complete data message. In the SAGE system these messages will range from 50 to over 300 bits in length. In the data field these are classed as long messages and their use requires that certain special techniques be employed in the transmission system.

(d) There is no information correlation between successive radar data messages transmitted over the data systems.

The radar indications of target positions, for example, are sent at a fixed rate and this rate is high compared to the rate of rotation of the radar antenna. The position of any particular target, however, is only reported once during each antenna revolution. Thus the interval between messages relating to the same target can not be smaller than this and may even be somewhat larger. Successive data messages contain information on different targets and bear no informational relation to one another.

2.13 From the preceding discussion of SAGE data it should be apparent that the satisfactory over-all operation of the system will largely depend on the accurate handling of information by all the system components. In this connection the introduction of errors in the data during its transmission must be held to an absolute minimum. This amounts to close control of circuit noise and is covered in other sections of the Practices. It is realized, however, that despite all precautions, errors will still be introduced. Certain error checking schemes have accordingly been built into the system so that errors will be detected and erroneous data discarded. The main technique that will be applicable and useful in the data transmission portion of the over-all system is called a "parity check." Fig. 6 illustrates this technique applied to the message of Fig. 5. The original message is shown at the top of the figure, it is 17 bits long and 7 of the bits are marking. The parity is applied to the message by adding an extra bit ("P") that is either marking or spacing depending on the number of marking bits preceding it in the message proper. If an even parity is being applied the parity bit is marking if an odd number of message marks precede it; if an even number of message marks precede it the parity bit is spacing. The net result of an even parity is that the total number of marks, message plus parity is always even. An odd parity is the reverse, in this case the total number of marks, message plus parity, is always odd. Fig. 6 shows an even parity; the parity bit "P" is marking. Thus the total number of marks before transmission is 8. At the receiving end the message is held in storage while the marks are counted. If the number of received marks is even the parity is verified and the message is read out of storage with the parity bit removed. It then goes to the data processing unit. If, on the other hand, the mark count results in an odd number, the parity check is not verified and the message discarded. (The erroneous bit in Fig. 6 is designated "E".)

2.14 The parity check is quite widely, but not universally, used in the SAGE system. It is effective when the errors are single errors and fairly widely distributed in time. It is less effective when multiple errors occur. For example, 2 or even 4 additional erroneous marks in the message of Fig. 6 would result in verification of the parity, even though the message itself would be completely wrong.

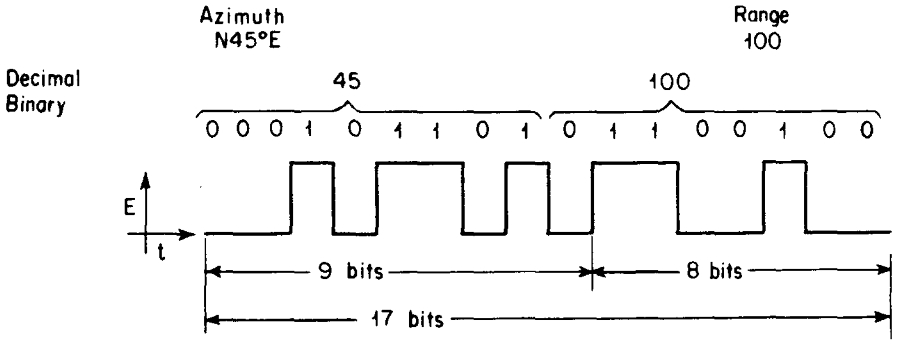


Fig. 5 - Azimuth and Range Message

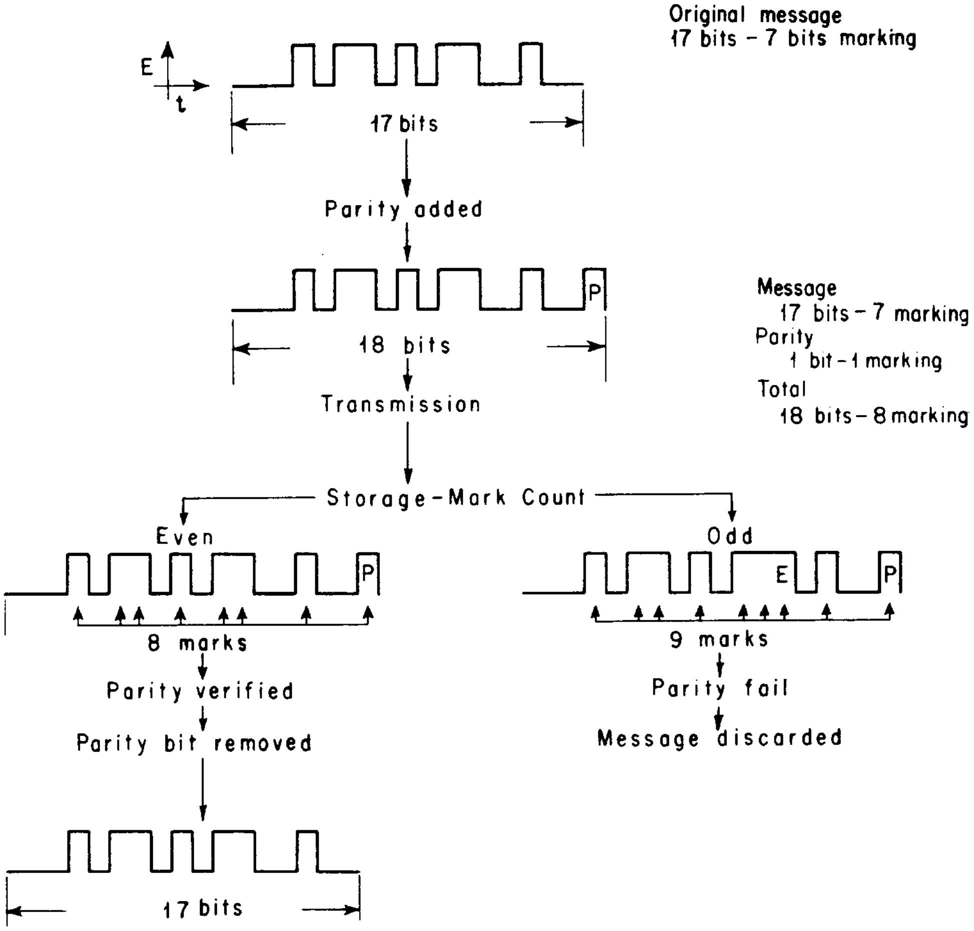


Fig. 6 - Parity Check

3. DATA TRANSMISSION SYSTEMS - INFORMATIONAL ASPECTS

3.01 The preceding material has indicated to some extent the large amount of information that must be handled by an air surveillance system. It has also been shown that significant savings in transmission time and information quantity can be achieved by adopting certain conventions or coding techniques. The main interest herein will be in the general features of a transmission system capable of transmitting information like Fig. 5 rapidly and accurately, commensurate, at the same time, with the fundamental economic elements of communication such as bandwidth and noise.

3.02 The first step toward this objective is to insure that the data transmission system is used only to transmit information that is actually required. Fig. 7 illustrates some of the informational elements of a digital message such as Fig. 5 and how these can be handled. The incident message is shown at the top of the figure; it is 17 bits long and is made up of two words. The first word is 9 bits long and the second 8 - these are azimuth and range, respectively. There are certain elements of this information that are fixed - these can be designed into the terminal gear and need not be transmitted over the system. For example, the form of the input signal is known - it will always be composed of rectangular binary type pulses. The desired output signal is the same so it is obviously wasteful and uneconomic to provide a transmission system that will actually transmit incident waveshapes. The receiving terminal can quite easily be made to regenerate these rectangular pulses at the proper time thus making possible significant reductions of bandwidth provided in the data system. The rate at which these pulses will occur $(1/t_b)$ is also known, as is the length of the message $(17t_b)$. Furthermore, the message make-up is known - the azimuth information will always be in the first 9 bits; the 8 bits following will be the range. These informational elements do not need to go over the transmission system since they will always be the same and can be "built in" both data terminals. This is indicated on Fig. 7 where these elements are shown in the "Not Transmitted" or "Structural" path of the information diagram.

3.03 The element of information that will change with each message and hence can not be predicted and built in are the two numbers indicating range and azimuth. They will change in a definite way, however, and the way they will change is that different bits will be marking. Since this signal is purely binary, bits can only be marking or spacing, therefore, from an information standpoint, only one of these two possible conditions need be transmitted in the "Transmitted" or "Real" path of Fig. 7. Definite indications are transmitted to indicate marks - spaces are automatically assumed by the receiving terminal in the absence of these indications. The mark indications that are transmitted over the system will never bear any obvious resemblance, so far as waveshape

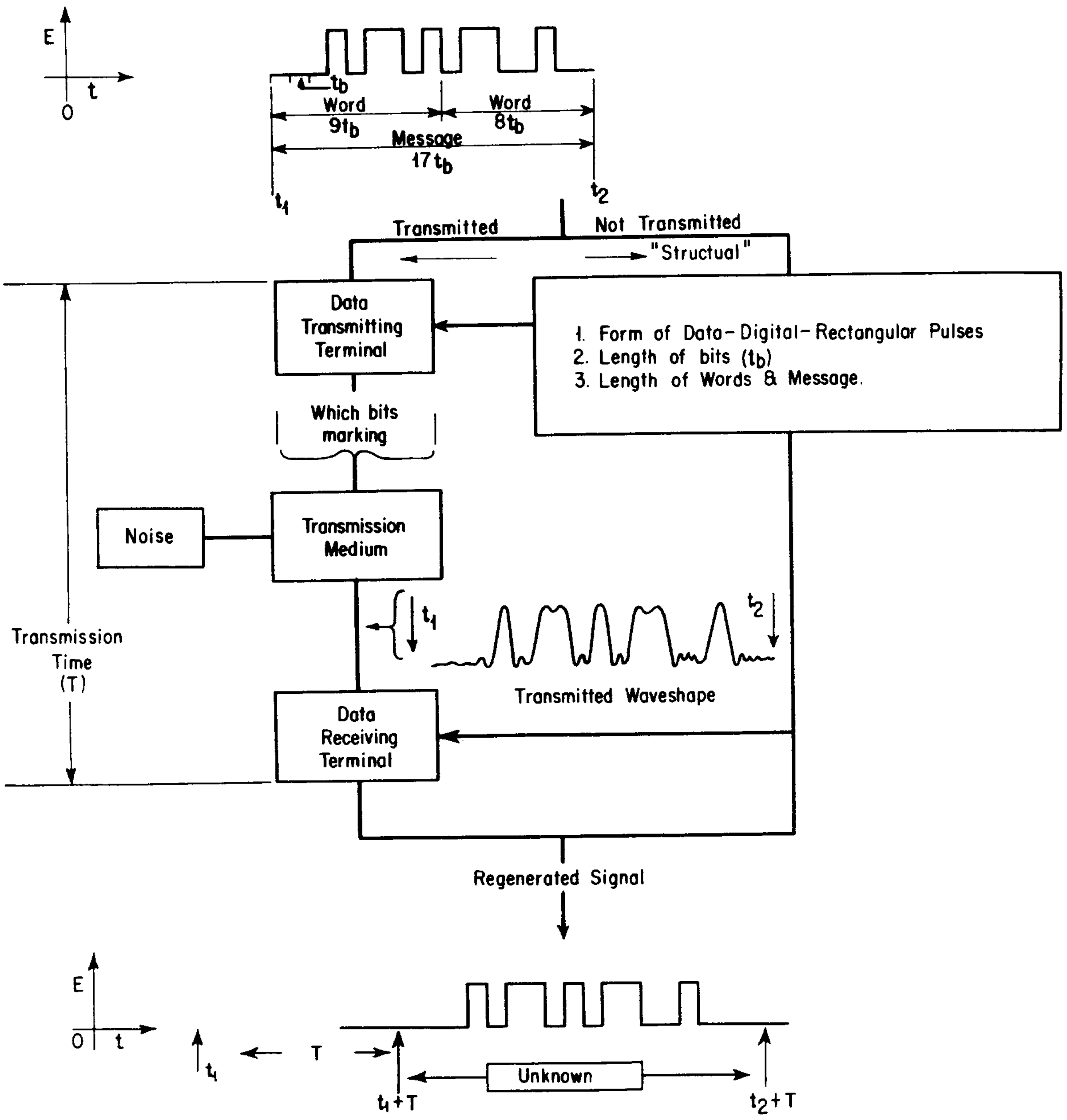


Fig. 7 - Information Elements

is concerned, to the actual mark indications that are generated by the radars and used by the computers of the SAGE system. This will be found to be true in practically all data applications where the speed of transmission approaches SAGE speeds. The form of information which computers can use most advantageously is not well suited to transmission. In the system shown, for example, the mark indications would take the general form shown as the "Transmitted Waveshape" in Fig. 7. This bears little resemblance to the incident rectangular pulses; it does not have to as long as it carries enough information so the receiving terminal can regenerate the rectangular pulses at the right times.

3.04 While the system shown in Fig. 7 illustrates how certain information elements are handled, the system, as shown, is unworkable since one very essential piece of information is not available at the receiving end. The missing information is the exact starting point of the received message. In the illustration, the incident message starts at time t_1 and ends at t_2, 17 bits, ($17t_b$), later. The marking information is sent over the transmission system and is delayed by the time of transmission, "T," so that the start of the received message is at $t_1 + T$. While this is all very clear in the illustration, as a practical matter, with the sending and receiving terminals separated by some distance, the receiving terminal can only register the marks and can not know how many spaces preceded the first one. This system could work, however, if it were possible to establish the instant of time, t_1, at both ends of the system and if the transmission time, T, was known. While the transmission time may be known with some accuracy, t_1 can not be established. To accomplish this would, in the last analysis, require that the exact time of a target's first appearance on the radar scope be known, or, effectively, prior knowledge of the target's existence.

3.05 This difficulty may be eliminated by making each data message "self-contained" to the extent that it contains enough information so the receiving terminal will be able to distinguish its starting point. One way to do this is to precede each message by a combination of marks and spaces that will never occur in the message proper. On receipt of this combination (start signal) the receiving terminal conditions itself to receive the data message which follows it directly. This process is illustrated by the top illustration of Fig. 8 which shows a starting signal made up of a space and 18 marks. This combination can never occur in the information content of a 17-bit message and circuits can, therefore, be built into the receiving terminal that will respond only to this. These circuits can then condition the terminal to receive and use the next 17 bits which is the desired message. Another method of doing the same thing is to use a pulse of different amplitude to precede each message. This is shown in the lower part of Fig. 8, where the message proper is preceded by a single bit which is occupied by a mark of double amplitude (E_s). All other marks are E_m in amplitude. A simple way of using this system would be to precede the receiving terminal with an electronic "gate" that opened on a pulse of amplitude E_s and thus

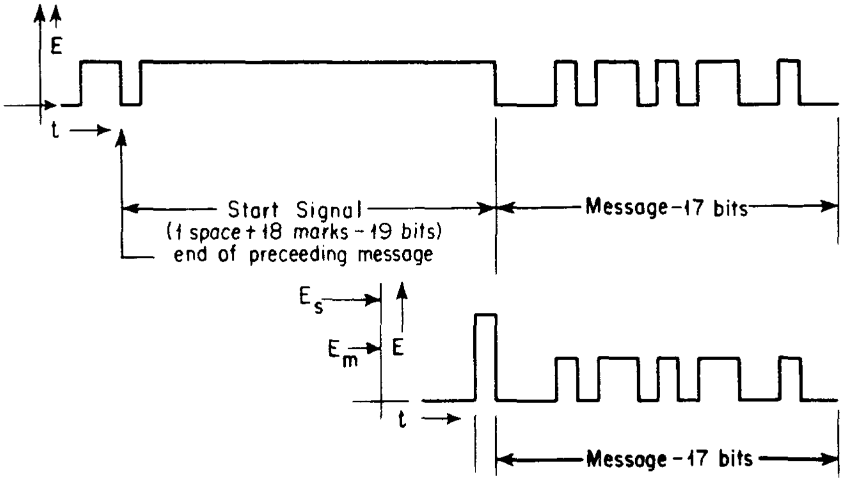


Fig. 8 - Start Signals

allowed the next 17 bits to enter the terminal. This latter technique, amplitude separation to indicate the start of a message, is used in the SAGE system.

3.06 With regard to start signals in general, it may be pointed out here that the first technique of Fig. 8 results in a data system that will be purely binary - only one marking amplitude is ever used. This has a significant advantage in that it will always result in the most lenient signal-noise ratio in the electrical portion of the system. (This is covered in some detail in the next section.) From an information standpoint, however, it has a significant disadvantage in that generally its use will consume a fair number of bits. Amplitude separation, on the other hand, uses up a minimum number of bits, one, but results in a definite impairment of the signal-noise ratio. SAGE data systems use this latter method, however, because they must handle so much data that only a minimum of time may be allocated to start signals.

3.07 With the addition of starting information, the information flow in Fig. 7 is modified to that shown in Fig. 9 - which is somewhat closer to a practical system. The incident information, shown at the left, consists of the same 17 information bits, now however preceded by a single double amplitude start bit, (S).

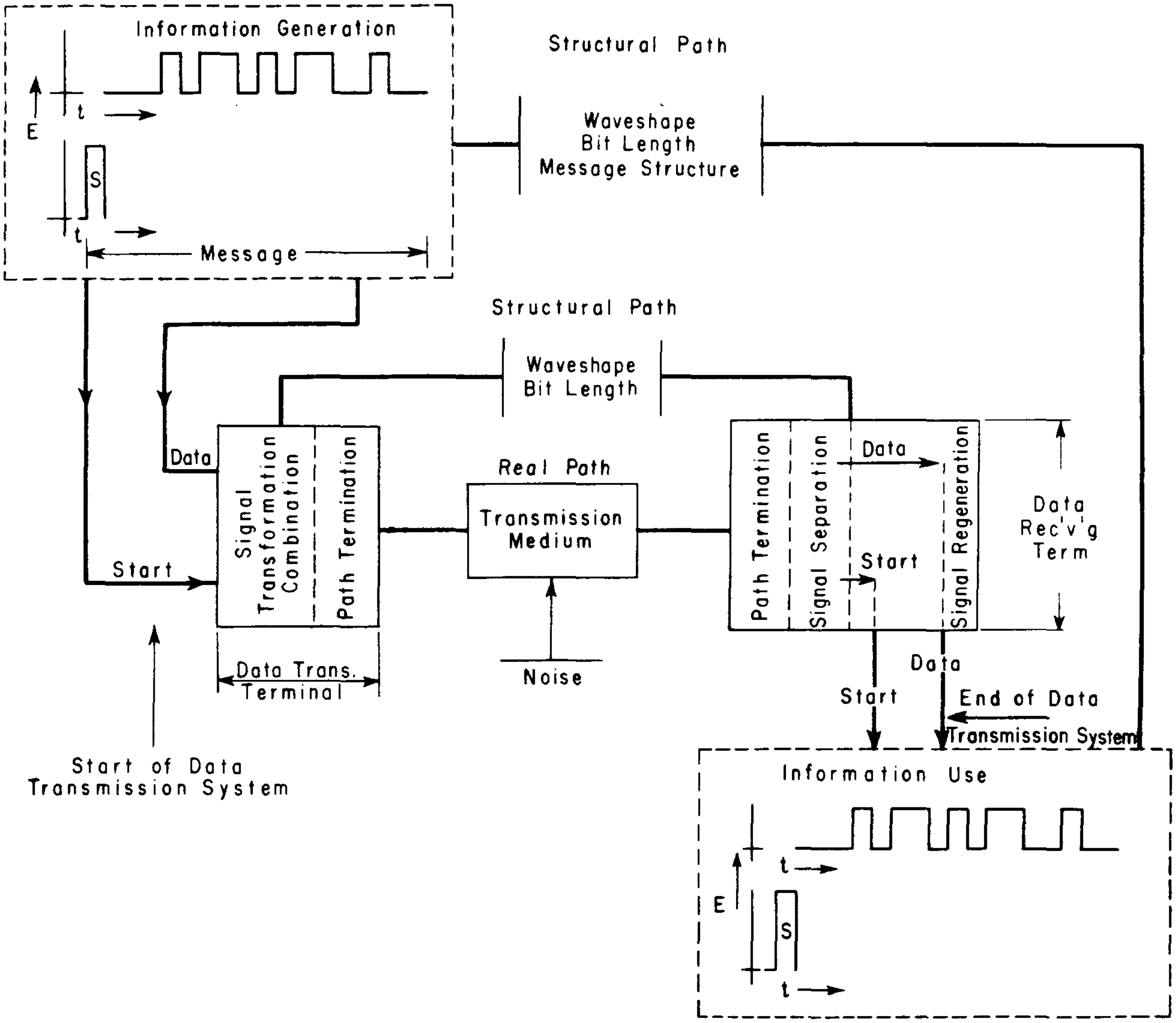


Fig. 9 - Data System

The net result of this is that two separate, though related, information components, a starting component (Start) and an information component (Data) are fed to the data transmitting terminal. The output of the data receiving terminal will present these same two components, delayed by the system transmission time, to the apparatus that uses the information at that end of the system.

3.08 The information diagram of Fig. 9 permits a line of demarcation between the data transmission system, which is the primary interest herein, and the data generating and data using apparatus. As shown, the data transmission system is considered to extend from the two component input of the data transmitting terminal to the two component output of the data receiving terminal. Between these two points the real, or existing, information path carries the start and data information, generally in a combined form. A structural path is also shown which includes the waveshape that will be used at both terminals as well as the bit length that each will handle. A second structural path of information may also be considered to exist between the data generating and using apparatus. (Dashed boxes in Fig. 9.) This is the path where the message structure information usually exists - this is very rarely any concern of the data transmission system. The waveshape and bit information indicated here may also be very different than the corresponding information in the data transmission system. For example, if the data using apparatus is being fed by several data transmission systems, it will be handling information internally much faster than any one transmission system and consequently will probably be working with bits of much shorter duration and different waveshapes.

3.09 With this background information it is appropriate to conclude this portion with a summary of the functions of the component parts of the data transmission system. These are:

(a) A <u>data transmitting terminal</u> which receives the information components from the source, combines them, transforms the result into a form suitable for transmission and finally passes the signal to the transmission medium through suitable path terminal arrangements.

(b) A <u>transmission medium</u> capable of carrying the signal to the receiving terminal

and

(c) A <u>data receiving terminal</u> which receives the combined signal through suitable terminal arrangements, separates it into its original components, and finally regenerates the incident waveform for the data apparatus at that end of the system.

4. DATA TRANSMISSION SYSTEMS - ELECTRICAL ASPECTS

4.01 When it is necessary that modern high speed information processing units exchange information, the data transmission systems by which this is accomplished are almost always entirely electrical in nature. There are exceptions to this, of course, some information processing units exchange information by the use of punched cards that are transported between them. When the information to be exchanged is subject to rapid changes, however, the transport time must be held to a minimum, otherwise the received information will not be valid by the time it reaches the unit that is to use it. This is the situation in the SAGE system, consequently the data transmission systems that interconnect the processing units are entirely electrical since this method of information transmission approaches the minimum transport time that can be achieved.

4.02 The first point of approach to a consideration of the electrical aspects of data transmission systems is to restate in electrical terms the fundamental information requirements outlined in the preceding paragraphs. Neglecting word and message structure, the first requirement is that adequate mark indications be presented to the receiving data terminal so that the incident signal can be correctly regenerated. Fig. 10 illustrates this. Here the incident signal has been simplified to a single binary pulse of "E" amplitude that is received by the transmitting terminal, transformed to some extent, and transmitted over the transmission medium. During this transmission the mark signal will encounter noise so that both signal and noise will be present at the input of the receiving terminal. (It should be noted that noise is present throughout the system - the main noise source of interest, however, is the transmission medium, not the terminal apparatus.) These conditions are shown in the right-hand portion of Fig. 10 where the mark indication is a voltage pulse of "e_r" amplitude with a noise level of "e_n" superimposed. The receiving terminal is arranged so that it reads as a mark any voltage that reaches or exceeds some predetermined level, in this case "e_m." Anything other than this is read as a space. The

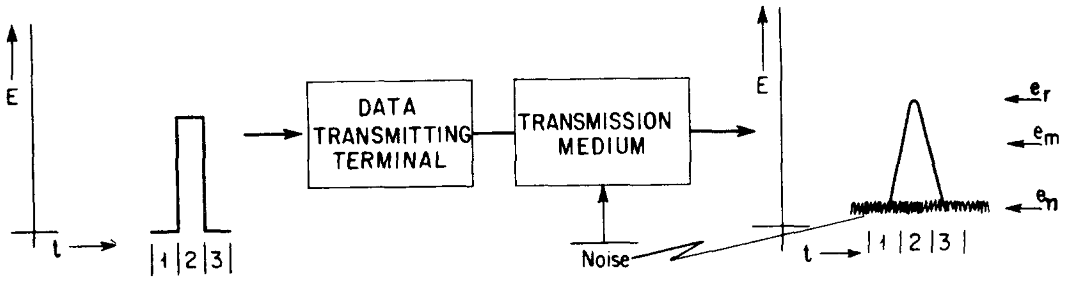


Fig. 10 - Mark Requirement

fundamental system requirement that opened this paragraph can thus be further elaborated to the extent that an "adequate mark indication" is one that can be distinguished in the presence of the noise at the input of the receiving terminal.

4.03 The preceding paragraph related to a broad system requirement that involved signal amplitude and noise level. There is another fundamental system requirement that, while related to the first, mainly concerns waveshape. The incident signal of Fig. 10 was a single pulse in the center of a 3-bit sequence. If this were a data word any combination of the 3 bits could be marking. One of these possible combinations is shown in Fig. 11 where the sequence is mark-space-mark. This is also called a "1-0-1" sequence. The mark indications at the input to the receiving terminal must be such that the terminal can follow this sequence and clearly identify the marks. As discussed below, this amounts to a waveshape criterion that must be met and is primarily a matter of the bandwidth of the transmission medium.

4.04 The factors involved in the reproduction of binary pulses are shown in Fig. 12. Here the 3-bit sequence, or data word, of Figs. 10 and 11 has been shown as a mark-space-space sequence on the left. This is followed at a fixed interval, 't_r', by the next data word which, for simplicity, has also been shown as a mark-space-space sequence. The waveshape presented to the data system is, therefore, a rectangular pulse of "e" amplitude and time duration 't_b' repeated at intervals of time equal to 't_r'. (This is quite representative of actual data systems where significant series of pulses make up data words that are sent at a fixed rate.) The waveshape of Fig. 12 can be represented by the Fourier series shown directly beneath it. This expression indicates that this waveshape has two components. The first term, 'a_0', is a dc component that has a magnitude equal to the average voltage existing over the time interval 't_r'. The second component, the trigonometric series, is the summation of an infinite number of harmonics of $1/t_r$, the pulse repetition rate. The significance of this expression is that the received waveshape is a function of the number of harmonics of the repetition rate that can be passed by the transmission medium. Further, this waveshape will resemble the incident waveshape more closely as the number of transmitted harmonics is increased.

4.05 This improvement in waveshape is illustrated in Fig. 13. The spectral distribution curve is shown at the top as the envelope of the harmonic amplitudes computed from the series of Fig. 12. The pulse width used was 1/1600 sec. and the repetition rate was 25 per second. (This corresponds to the pulse width and word rate of one data transmission system used in the SAGE System.) The form of the distribution curve is largely governed by the expression ($\sin nr\pi / nr\pi$). (Fig. 12.) As shown the harmonic amplitude steadily decreases as the number of harmonics, n, increases. It falls to zero whenever nr equals an integer, in this case at intervals of 64 harmonics, and changes sign after each zero point. Thus, referring to Fig. 13, the harmonics are positive up to the first 0 point, 1600 cycles (64th harmonic of 25 cycles), negative (opposite phase) between 1600 cycles and 3200 cycles (128th harmonic of 25 cycles) and so on. The computed pulse shapes are shown directly beneath the distribution curve. The curve at C, for example, is computed by adding the contributions of all harmonics up to the 64th and none thereafter. This effectively assumes an infinitely sharp cut-off and no resulting

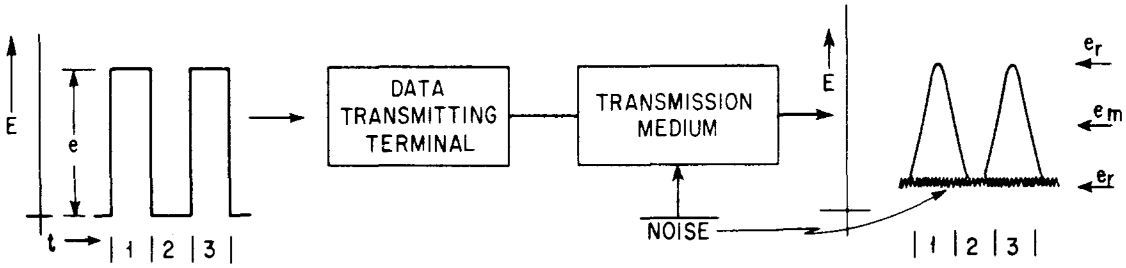
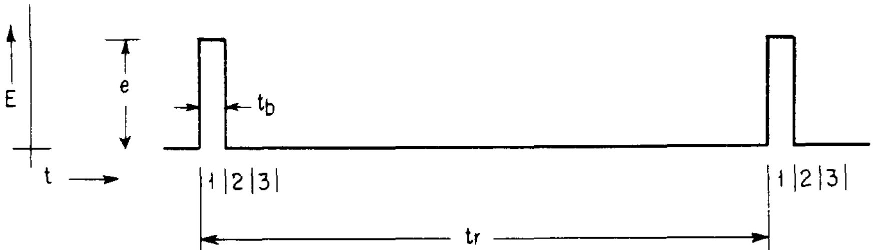


Fig. 11 - 1-0-1 Requirement



$$f(t) = a_0 + \sum_{n=1}^{n=\infty} \left[(2re) \left(\frac{\sin nr\pi}{nr\pi} \right) \right] \left(\cos n \frac{2\pi}{T_o} t \right)$$

$t_b = 1/1600$ Sec., $t_r = 1/25$ Sec., $r = t_b / t_r = 1/64$

n = harmonic number

$\text{Sin} \frac{nr\pi}{nr\pi}$ for n = 64 is $\frac{\text{Sin}\,\pi}{\pi} = 0$

for n = 128 is $\frac{\text{Sin}\,2\pi}{2\pi} = 0$

Fig. 12 - Repetitive Pulses

distortion, hence it is rather theoretical but the curves still illustrate the improvement in waveshape fairly well. It will be noted that even with only 16 harmonics (400-cycle passband) some indication of the presence of the incident pulse is transmitted. (A)

4.06 It is quite apparent that, so far as single pulses are concerned, any of the waveshapes from 'A' through 'G' could be used in a data transmission system although the use of 'A' would require closer noise control than the others. For that matter single pulse indications would be received if the passband were narrowed even further. (This would demand even closer noise control.) The picture changes, however, if pulse combinations are considered, particularly the 1-0-1 sequence. The reproductions of this are shown for the first four passbands at 'H' through 'K' in Fig. 13. It is evident here that the sequence is entirely lost with a 400-cycle passband but adequate indications of it are received with passbands from 800 cycles on up.

4.07 While the 400-cycle passband is not capable of meeting both fundamental requirements of signal transmission, the other passbands shown in Fig. 13 are, and it is now of interest to determine how much bandwidth must actually be provided for the data transmission system. Bearing in mind that the end purpose is to provide an adequate indication to the receiving data terminal so that the signal can be regenerated, the best possible system would provide an indication such as that of Fig. 14A. Here the bandwidth is infinite so the incident pulse is substantially reproduced at the input of the receiving terminal and arrives there with unity amplitude. The receiving terminal will be arranged so that it reads as marks only those indications that exceed a predetermined value. In this case it will read as a mark any indication that exceeds an amplitude of 0.5 (clip level). The clip level is generally set midway between the signal marking and spacing levels to provide maximum noise immunity. The noise at this point can be additive or subtractive so this technique insures equal immunity.

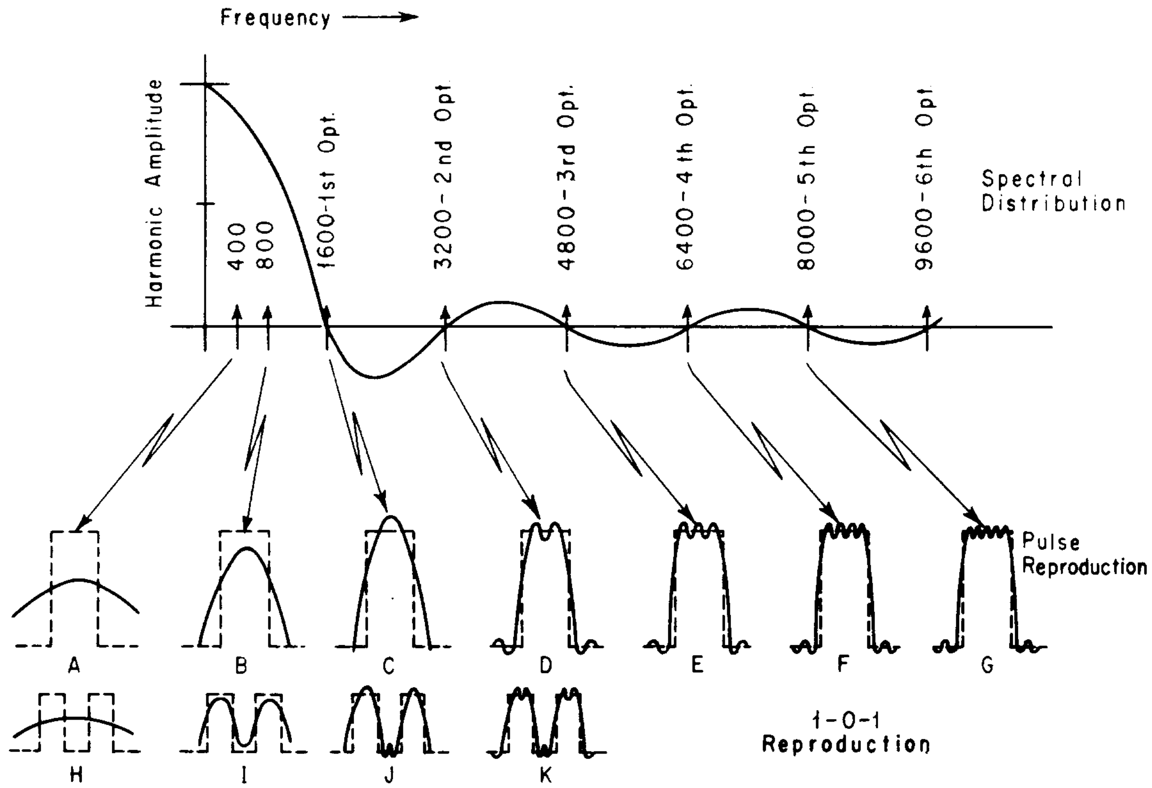


Fig. 13 - Waveshape and Bandwidth

This system will work satisfactorily if the noise does not equal or exceed an amplitude equal to the clip level and the theoretical signal/noise ratio is, therefore, 6.0 db. Furthermore, with this system, the mark indication at unity amplitude is available to the receiving terminal throughout the time interval of the bit and the S/N ratio is, therefore, 6 db whether the receiving terminal uses the whole bit or any part of it as a mark indication.

4.08 The 800-cycle indication is shown at B. Here the limited passband results in a pulse amplitude that is less than unity 0.85 and the clip level (0.42) has been adjusted accordingly. This system will work satisfactorily if the noise level again does not equal or exceed the clip level and the required S/N ratio is 7.5 db. Thus, the restriction of passband results in a noise penalty over the theoretical best of 1.5 db at the receiving terminal input. Furthermore, even this indication is not available over the whole bit interval - it is only available at the exact center of the bit. The effect of doubling the bandwidth to 1600 cycles is shown at (C). Here the peak indication is 1.2, the clip level 0.6, and the S/N ratio 4.4 db at the receiving input. All other things being equal, this is apparently 1.6 db better than an infinite passband would require - but again the indication is only available at the center of the bit. Actually, this gain in S/N ratio would not be realized due to system adjustments - the loss at (B), however, could be made up by bandshaping. The result of this would be about as shown - the top of the pulse would reach 1.0 amplitude at the center of the bit. Thus, no significant improvements in S/N ratio can be obtained by increasing the bandwidth beyond 800 cycles for this data system, provided it is not necessary that anything beyond a small portion of bit width is needed for the mark indication.

4.09 The bandwidth at B, 800 cycles, is, in fact, the minimum bandwidth that can be used to transport 1600 bits without significant degradation of the required S/N ratio. This minimum bandwidth is called the "nominal effective band" and for purely binary systems is always theoretically equal to half the bit rate. Actually it is not possible to provide the

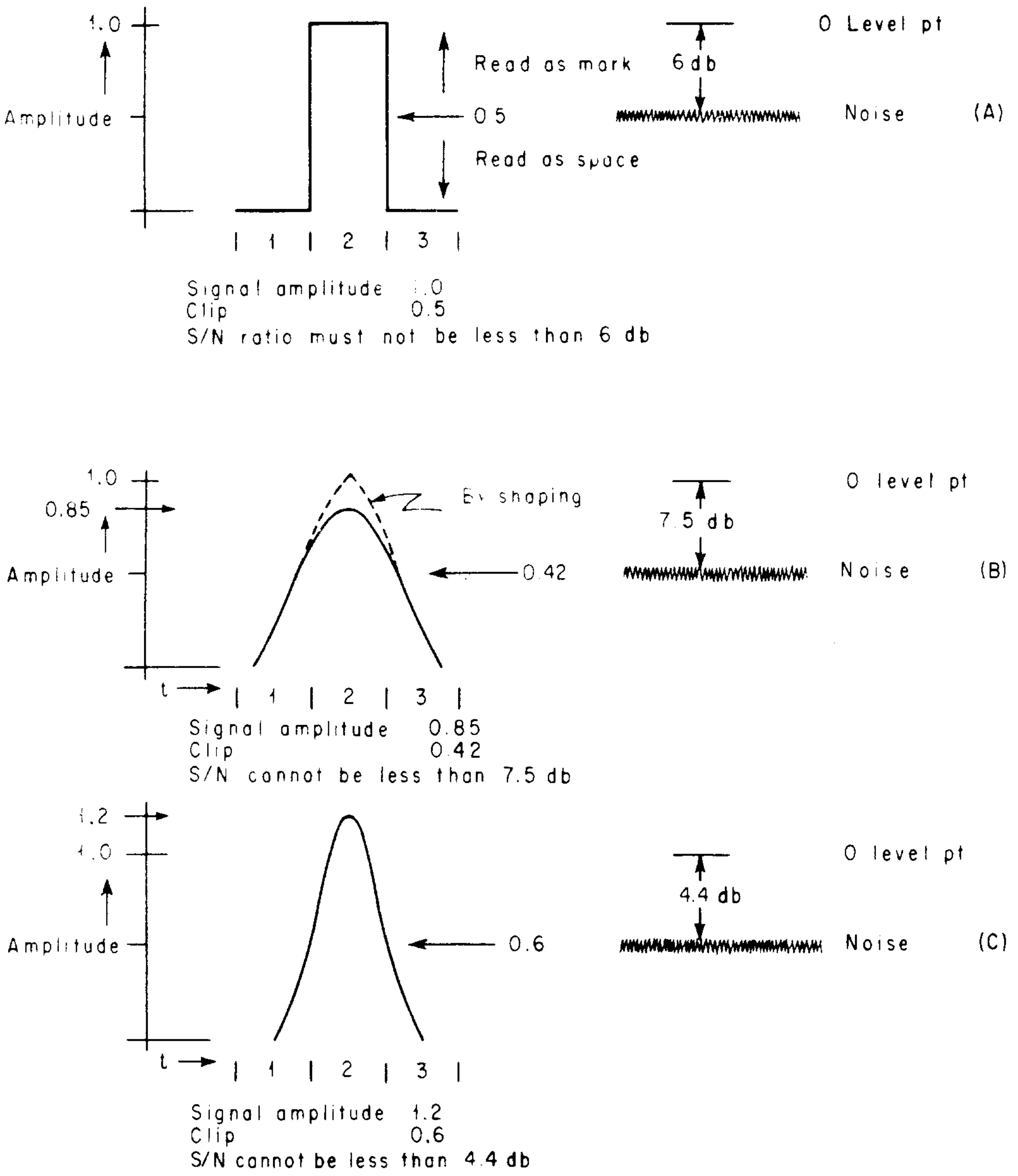


Fig. 14 - Theoretical S/N Ratio

sharp system cutoffs this implies and as a practical matter it is customary to use an equal amount of bandwidth for these regions. The gain and phase characteristics of the transmission medium must be carefully controlled through the nominal effective band but their variations are less important in these roll-off regions.

4.10 It is extremely difficult to control gain and phase deviations in the region below 1000 cycles in the transmission media presently used if any sizeable distance is involved. To avoid this region the SAGE data transmission systems utilize a carrier and single sideband transmission system with the carrier at 2000 cycles. The actual bandwidth used varies with the different bit rates but in no case is anything outside the 1000 to 2500 cycle region used for the transport of the data information. In this region, which includes the lower sideband and both roll-off bands, the gain and phase characteristics are carefully controlled. The phase characteristic is described by the envelope delay distortion and this is held to ±0.4 bit at the highest bit rate used. This results in a total distortion that does not exceed 500 microseconds.

4.11 It will be recalled that the use of an 800-cycle band in Figs. 13 and 14 or more generally, the nominal effective band, resulted in a required 6 db S/N ratio at the receiving input. This ratio only held, however, if the receiving terminal took the indication at the center of the bit interval. The SAGE data systems are arranged so that their receiving terminals do just this, moreover, they only use a very narrow interval at the center of each bit. By this technique the probability of random noise peaks causing errors is reduced since the receiving terminal is only "open" for a small fraction of each bit interval.

4.12 The use of the sampling or interrogation technique at the receiving terminal, while insuring no S/N ratio degradation, carries with it the necessity of extremely accurate location of the sampling interval. If this is not done precisely at the center of the bit the sampling interval will occur when the signal is not at its maximum amplitude. Fig. 15 illustrates the consequences of such a shift in sampling. The correct sampling point is shown by the solid arrow - this is in the center of the bit interval. The dotted arrow indicates faulty sampling - the signal is past full amplitude and is only 0.75. This adds 6.0 db to the required S/N ratio. There is also an additional consequence; the system no longer has equal immunity to additive or subtractive noise.

4.13 To insure correct sampling it is customary to synchronize the sending and receiving data terminals. This may be either a continuous or periodic process. If the latter technique is used the receiving terminal must be capable of accurate "free-running" between the synchronizing intervals. Some of the data systems used in SAGE are continuously synchronized, others are periodically synchronized by means of the start pulse. All the data systems, however, take a timing component from the data generating apparatus at the sending end of the data transmission system and give back a timing component to the apparatus using the data at the receiving end of the system. Fig. 16 shows a complete data system with all three components - start, data and timing. In the receiving terminal the timing has been shown connected to the start component by a dashed line to indicate that it may either be actually sent over the data system or derived from the start component.

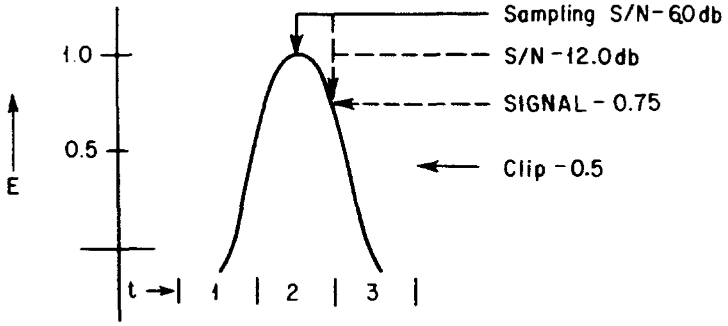


Fig. 15 - Timing Error

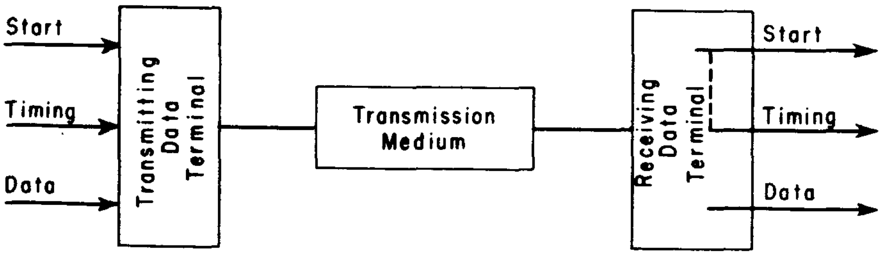


Fig. 16 - Complete Data System