~~CONFIDENTIAL~~

A METHOD OF CIPHERING

Howard N. Smith

Howard W. Lewis

Palo Alto, California

November 14, 1954

~~CONFIDENTIAL~~

Those authorized to read this paper are requested to reveal its contents to no one without the permission of the authors. Please sign your name and note the date below to indicate that you have read the paper subject to this condition.

Frank E Stone
3/20/55

A METHOD OF CIPHERING

The purpose of this paper is to describe an original method of ciphering in order that it may be properly evaluated by an expert cryptographer.

Briefly, this is a relatively simple method of ciphering in which there is no fixed pattern of any kind and, therefore, no repetitions of a meaningful nature. The cipher is continuously changing in a completely random manner, and never repeats itself in a predictable way, but only as the result of pure chance. The number of possible keys to the cipher is extremely large, the key can be changed as often as desired, and knowledge of both the ciphering principles and procedures and of previously-used keys is of no value whatsoever in breaking an enciphered message. Unless the key that was used is known, a message can only be broken by methods based almost entirely on trial-and-error, and the probability of such methods being successful within a reasonable period of time is infinitesimally small. Furthermore, the solution may not be recognized when found, and can never be known with absolute certainty. It is suggested that for all practical purposes the cipher is completely indecipherable.

In order to describe the cipher, it will first be necessary to describe a simple device which can be used to illustrate the procedures for enciphering and deciphering messages. This device consists of two gears which have twenty-six teeth each or one tooth for each letter of the alphabet. The teeth on each gear are marked from "A" to "Z" in a clockwise direction. One of these gears is called the "message wheel" and the other is called the "cipher wheel." The wheels are free to rotate when engaged, and the cipher wheel can be disengaged from the message wheel, rotated or "indexed" from one to twenty-six teeth in a clockwise direction, and then re-engaged with the message wheel. Thus there are twenty-six possible engagements of the wheels. There are two fixed reference points, one for each wheel, which indicate the message-cipher relationship between the letters on the two wheels. This relationship is fixed for each engagement, and is different for each of the twenty-six possible engagements.

The manner in which the key to the cipher is selected and used will make clear that for the device which has just been described, the cipher has twenty-six times as many keys as there are different ways of arranging the twenty-six letters of the alphabet. Each key consists of one of the twenty-six possible engagements of the wheels plus one of the factorial twenty-six possible arrangements of the alphabet. Thus there are 10^{28} or ten billion billion billion possible keys to the cipher.

One of these keys has been selected for purposes of illustrating the ciphering procedures. The first part of the key, one of the twenty-six possible engagements of the wheels, was selected by placing twenty-six letters in a hat, drawing one of them out at random, and noting it to be "F." Thus the engagement of the wheels which constitutes the first part of the key is such that when the letter "A" is indicated on the message wheel by the fixed reference point for that wheel, the letter "F" is indicated on the cipher wheel by the fixed reference point for that wheel.

The second part of the key, one of the factorial twenty-six possible arrangements of the alphabet, was selected by placing twenty-six letters in a hat, drawing them out *one-by-one at random*, and numbering them in the order in which they were drawn. Thus the arrangement of the alphabet which constitutes the second part of the key is as follows:

M	1	A	10	N	19
O	2	Y	11	J	20
L	3	F	12	W	21
T	4	K	13	I	22
S	5	Z	14	E	23
C	6	X	15	H	24
D	7	G	16	R	25
Q	8	P	17	B	26
V	9	U	18		

Having described the ciphering device and selected a key, it is now possible to illustrate the procedures for enciphering and deciphering messages. The manner in which the key is used will be made clear by means of the procedures themselves.

The essence of the ciphering procedures is that the engagement of the wheels, and thus the relationship between the message and the enciphered text, is changed between every letter of the message in a manner which is completely random. For each engagement of the wheels, two letters are enciphered: the first letter is the next letter of the message, and the second letter is a letter from the key which indicates the amount that the engagement of the wheels is to be changed before proceeding. Thus a message containing x letters will produce an enciphered text containing 2x-1 letters.

Assume that the message to be enciphered is Poe's classic statement, "human ingenuity cannot concoct a cipher which human ingenuity cannot resolve." The initial engagement of the wheels for both enciphering and deciphering is the engagement specified in the first part of the key, so the first step is to engage the wheels in this manner. Then the first letter of the message, "H," is indicated on the message wheel which results in "Y" being indicated on the cipher wheel. "Y" is thus the first letter of the enciphered text.

It is now necessary to change the engagement of the wheels in a completely random manner. This is done by randomly drawing one of twenty-six letters from a hat, noting that it is "K," and then returning it to the hat. Referring to the second part of the key, it is noted that "K" is the thirteenth letter listed. Thus "K" calls for indexing the cipher wheel thirteen teeth. First, however, "K" must be enciphered. When "K" is indicated on the message wheel, "V" is indicated on the cipher wheel, and "V" is thus the second letter of the enciphered text.

The cipher wheel is then indexed thirteen teeth and the next letter of the message enciphered. "U" indicated on the message wheel results in "Y" again being indicated cn the cipher wheel, and "Y" is thus the third letter of the enciphered text. Another letter is then drawn at random from the hat, noted to be "C," and returned to the hat. "C" indicated on the message wheel results in "Q" being indicated on the cipher wheel, and "Q" is thus the fcurth letter of the enciphered text. The cipher wheel is then indexed six teeth as called for by "C," and the next letter of the message enciphered. This procedure is continued until the message has been completely enciphered: the cipher wheel is indexed between every letter of the message a number of teeth determined by randomly drawing a letter from the hat and noting its corresponding number in the key, the letter drawn being enciphered immediately prior to the actual indexing.

The procedure for deciphering the message is essentially a re-enactment of the enciphering procedure except that the cipher wheel now has the initiative and the message wheel is the follower. First, the wheels are engaged in the initial position as specified in the first part of the key. Then the first letter of the enciphered text, "Y," is indicated on the cipher wheel which results in "H" being indicated on the message wheel. Thus "H" is the first letter of the message. The second letter of the enciphered text, "V," is then indicated on the cipher wheel which results in "K" being indicated on the message wheel. Referring to the second part of the key, it is noted that "K" calls for indexing the cipher wheel thirteen teeth, and this is done. The third letter of the enciphered text, "Y," is then indicated on the cipher wheel which results in "U" being indicated on the message wheel. Thus "U" is the second letter of the message. This procedure is continued until the enciphered text has been completely deciphered.

It is now appropriate to consider the problem which this method of ciphering presents to the intercepter of an enciphered message. It is assumed that the intercepter has full knowledge of the ciphering principles and procedures and has a ciphering device such as the one described above, but does not know which one of the ten billion billion billion possible keys was used to encipher the message.

First of all, it must be realized that depending upon the length of the message, thousands, millions, or even billions of the possible keys will produce messages that make sense, and the cipher itself will in no way indicate which of these messages is the message that was enciphered. This can only be determined from the sense of the messages themselves. Thus the only way in which an enciphered message can be broken is to determine keys which produce messages that make sense until a message is produced which, from the particular sense that it makes, is concluded to be the message that was enciphered. Obviously, the message may not be recognized when found, and can never be known with absolute certainty. For purposes of the problem at hand, however, it is further assumed that the intercepter will be able to recognize the message when it is found.

Since the cipher has a finite number of keys, it is theoretically possible to break an enciphered message by testing possible keys on a trial-and-error basis until the key that was used to encipher the message is tested and, presumably, the message recognized. It can be quickly demonstrated, however, that trial-and-error alone is totally impractical. Assuming that the message could be tested at the rate of one billion keys per second, it would take 320 billion years to test all of the possible keys, and the key used to encipher the message would be tested, on the average, at the end of fifty percent of the possible trials or 160 billion years. Through extreme good fortune, the key used to encipher the message might be tested at the end of one millionth of one percent of the possible trials, but even this would require 3,200 years. Thus the probability of breaking an enciphered message within a reasonable period of time by means of trial-and-error alone is so extremely small that for all practical purposes it is zero. In order to significantly increase this probability, the probable number of trials required to find the key must be very greatly reduced by means of supplementary analysis.

There are three basic methods of analysis by which the probable number of trials required to find the key can be reduced. Because of the nature of the problem, however, all three of these methods are necessarily limited to increasing the rate at which messages that make sense will be produced, and thereby increasing the probability of producing the message that was enciphered within a reasonable period of time. It bears repeating that a very large number of the possible keys will produce messages that make sense, and there is no way in which the key that was used to encipher the message can be identified other than by the particular sense of the message that it produces.

The first method of analysis is based on the fact that at any point in an enciphered message, certain engagements of the wheels are more likely than certain other engagements because certain letters of the alphabet appear in the language more frequently than certain other letters. For example, assume that a letter of the message is "B" in the enciphered text. When "B" is indicated on the cipher wheel, the engagement of the wheels which indicates "R" on the message wheel is more likely than the engagement which indicates "K" on the message wheel because "R" appears more frequently than "K" and is more likely to be the letter of the message. Thus at any point in the message, the twenty-six possible engagements of the wheels can be arranged in order of liklihood as determined by the relative frequencies with which the twenty-six letters of the alphabet appear in the language. This means that if a family of keys is defined as one of the factorial twenty-six possible arrangements of the alphabet plus all of the twenty-six possible engagements of the wheels, at any point in the message, and particularly at the beginning, the twenty-six keys in each family can be arranged in order of priority for testing. By testing possible keys in order of overall priority, the rate at which messages that make sense will be produced can be increased.

The second method of analysis is similar to and supplementary to the first method: whereas the first method is concerned with the engagement of the wheels at a particular point, the second method assumes an engagement at that point and is concerned with the change in engagement that immediately follows. It is based on the fact that if at any point in an enciphered message one of the twenty-six possible engagements of the wheels is assumed to be the engagement at that point, certain changes in that engagement are more likely than certain other changes because certain letters of the alphabet appear more frequently in the language than certain other letters. For example, assume that a letter of the message is "B" in the enciphered text, the following letter from the key is "G" in the enciphered text, and the following letter of the message is "S" in the enciphered text. If the engagement of the wheels which indicates "R" opposite "B" is assumed to be the engagement at that point, the letter from the key is "M" since "M" is indicated opposite "G" by that engagement. "M" is more likely to call for indexing twenty-two teeth than ten teeth because when twenty-two teeth are indexed, "E" is indicated opposite "S," when ten teeth are indexed, "Q" is indicated opposite "S," and "E" appears more frequently than "Q" and is more likely to be the next letter of the message. Thus if at any point in the message one of the twenty-six possible engagements of the wheels is assumed to be the engagement at that point, the twenty-six possible changes in that engagement can be arranged in order of liklihood as determined by the relative frequencies with which the twenty-six letters of the alphabet appear in the language. This means that if a family of keys is defined as one of the twenty-six possible engagements of the wheels plus all of the factorial twenty-six possible arrangements of the alphabet, for each letter of the alphabet, each family will consist of twenty-six groups of factorial twenty-five keys each such that the letter will call for indexing from one to twenty-six teeth over the twenty-six groups, and at any point in the message, and particularly at the beginning, the twenty-six groups of keys in each family can be arranged in order of priority for testing. This method of analysis can be used to supplement the first method described above and thereby further increase the rate at which messages that make sense will be produced.

It should be noted that actually the first two methods of analysis would be modified in specific situations to the extent that special frequency distributions of letters or probable patterns of letters would be more significant in that situation than the general frequency distribution of letters in the language. For example, "E" appears in the language more frequently than any other letter, but as the initial letter of a word, the frequency of "E" is to the frequency of "T" as 340 is to 1,194. Thus the first letter of a word is three and one-half times more likely to be "T" than "E." Similarly, "Q" is always followed by "U," and if "Q" is assumed to be the letter at a particular point in a message, the next letter of the message will be "U" regardless of the frequency with which "U" appears in the language.

The third method of analysis is employed after a number of possible keys have been tested, and is undoubtedly the most significant of the three methods. It is based on the fact that in addition to the large number of possible keys which will produce messages that make sense, many times this number will produce messages that contain one or more sequences that make sense, and by comparing the points where such sequences begin and end and the portions of keys operative over such sequences, sequences from different messages can be matched together to produce messages that make sense. For example, assume that one of the possible keys produces a message that contains a sequence from the twenty-fourth letter to the fifty-first letter that makes sense. If another of the possible keys produces a message that contains a sequence that makes sense which begins at the fifty-second letter or ends at the twenty-third letter, and if the sense of this sequence would supplement and extend the sense of the first sequence, the two sequences can be joined together to produce a longer sequence that makes sense provided that the portions of keys operative over the two sequences do not conflict with each other or with the change in engagement required to join the two sequences together. Note also that sequences can be extended a number of letters in either direction to join with other sequences by means of the first and second methods of analysis described above, or, more importantly, by means of trial-and-error guessing based on the sense of the two sequences being joined. In this manner, the rate at which messages that make sense will be produced can be increased, and it is important to note that depending upon the liklihood that the sense that was enciphered will be recognized, the messages produced will be those which, because of the particular sense that they make, have a greater liklihood of being the enciphered message.

It was stated above that in order to significantly increase the probability of breaking an enciphered message within a reasonable period of time, the probable number of trials required to find the key must be substantially reduced by means of supplementary analysis. The three basic methods of analysis which can be employed were then described, but no attempt was made to determine whether or not these methods would substantially reduce the probable number of trials required and thereby significantly increase the probability of successfully breaking an enciphered message. This aspect of the problem has not been fully analyzed, and will not be dealt with in this paper. It seems likely that these methods of analysis would not be of practical significance relative to the total problem presented by the cipher, and that for all practical purposes the cipher, in its present form, is indecipherable. It is possible, however, that this may not be true, and because this point has not been resolved, it is necessary to assume that it is not true, and that the indecipherability of the cipher must be increased.

One method by which this can be accomplished is by increasing the number of teeth on each wheel which, in effect, increases the number of possible keys to the cipher. For example, if ten teeth are added to each wheel to provide for the numbers zero through nine, the cipher will have thirty-six times as many keys as there are different ways of arranging twenty-six letters and ten numbers or 1.3 x 10^{43} possible keys. This is 1.3 quadrillion times as many keys as there were when each wheel had twenty-six teeth, and the time required to test one trillionth of one percent of the possible keys at the rate of one trillion trials per second would be 4.2 billion years. Similarly, if thirty-six more teeth are added to each wheel to provide for twenty-six capital letters and ten punctuation marks, the cipher will have 4.4 x 10^{105} possible keys or fifteen quadrillion quadrillion times as many keys as there are atoms in the known universe. This method will not affect the basic simplicity of the ciphering procedures, and will greatly increase the indecipherability of the cipher in terms of the practical problem which the cipher presents. Note, however, that while the cipher will be changed in degree, it will not be changed in kind, and the three methods of analysis described above can still be employed. Since it was assumed that these methods would be of significant practical value when there were twenty-six teeth on each wheel, it must be further assumed that they would be of significant practical value regardless of the number of teeth on each wheel. This assumption becomes increasingly unlikely as the number of teeth on each wheel increases, but because the actual significance of the methods has not been determined, the assumption must be made. Therefore it is necessary to invalidate one or more of the three methods of analysis and thereby increase the theoretical indecipherability of the cipher.

Because the first and second methods of analysis are based entirely on the fact that the various letters of the alphabet appear in the language in known patterns and with known different frequencies, these methods would be of no value whatsoever when dealing with a language in which the letters of the alphabet appear randomly and with equal frequency. Such a language can be easily constructed, in effect, by constructing a simple code which, in addition to completely invalidating the first two methods of analysis, will make the problem of recognizing the message when it is found considerably more difficult by tremendously increasing the number of possible keys which will produce messages that make sense.

It is important to note, however, that the nature of this code is such that whether or not the code is secret is of little consequence, and it can be assumed than an intercepter has full knowledge of the code in addition to knowledge of both the ciphering principles and procedures and of previously-used keys. Therefore, once the code has been constructed, it will never have to be changed. Actually, use of the code should be thought of as just an additional step in the ciphering procedures because its purpose is not to make the cipher secret or complex in any way, but rather to invalidate methods of analysis which would otherwise be theoretically possible and, presumably, of significant practical value. Thus the code increases the theoretical indecipherability of the cipher but does not affect its fundamental nature: unless the key that was used is known, the cipher defies solution solely on the basis that the probability of finding the key within a reasonable period of time is infinitesimally small, and there is no way in which the key can be positively identified when it is found.

For purposes of illustrating the construction of the code, assume that there are thirty-six teeth on each wheel to provide for twenty-six letters and ten numbers, and, consequently, that the cipher has thirteen sextillion sextillion possible keys. The number of permutations of thirty-six things taken three at a time when each thing may be repeated up to three times is thirty-six raised to the third power or 46,656 which means that it is possible to make a total of 46,656 different three-letter arrangements or code words from the thirty-six symbols included in the cipher. Therefore, for purposes of illustration, assume that the code to be constructed is to include 46,656 different words. Because of the nature of the code, this number of words is not as large as it at first appears since a different code word must be assigned to each grammatical form of each word that is to be included in the code. Thus the code will represent an effective vocabulary in the order of fifteen thousand words which would probably be of adequate size for most purposes. Note, however, that there is no limit to the number of words that can be included in the code: for example, it is possible to make 456,976 different four-letter code words from twenty-six letters alone. A fairly conservative vocabulary was selected because in order to increase the number of possible keys which will produce messages that make sense to the maximum extent possible, it is necessary to assign words to all of the possible code words, and desirable that the words assigned be of fairly common usage.

The first step in the construction of the code is to select from the language the 46,656 words that are to appear in the code and arrange them in alphabetical order. Then 46,656 slips of paper are placed in a hat, each marked with a different one of the possible three-letter code words that can be made from twenty-six letters and ten numbers. The slips of paper are then drawn from the hat one-by-one at random, and the code word on each slip noted opposite the next word in the alphabetical list in the order in which the slips are drawn. In this manner, a language-code dictionary can be constructed. For deciphering messages, a code-language dictionary will be needed, and can be easily constructed br arranging the 46,656 code words in "alphabetical" order ("A" through "9") and noting opposite each its language counterpart.

The manner in which the code is constructed is such that the twenty-six letters and ten numbers will be distributed randomly and with equal frequency throughout the code. Although certain code words will be used more frequently than certain other code words because of the relative frequencies with which their language counterparts appear in the language, if the 46,656 code words are arranged in order of frequency and classified into percentile groups, it is almost a certainty that each of the thirty-six symbols will appear an equal or almost equal number of times in each group, and an equal or almost equal number of times as first, second, and third letters within each group. Thus a symbol-frequency table developed by means of word-frequency analysis will indicate that all of the symbols have the same probability of being first, second, and third letters of a code word.

At this point, again consider the problem which confronts the intercepter of an enciphered message, this time a message which was translated into code before being enciphered. As stated above, it can be assumed that the intercepter has full knowledge of the code in addition to other knowledge previously conceded.

Because the twenty-six letters and ten numbers appear in the code randomly and with equal frequency, it is no longer possible to arrange possible keys in order of priority for testing by means of the first and second methods of analysis. Thus the only procedure available is to test possible keys solely on the basis of trial-and-error and attempt to increase the rate at which messages that make sense will be produced by means of the third method of analysis. Note, however, that because every one of the 46,656 possible code words appears in the code book opposite a real word, every three-letter sequence produced by every possible key will be a word that makes sense. This means that the number of possible keys which will produce messages that make sense will be tremendously increased, and the problem of recognizing the message when it is found (and of recognizing that the message has not been found) will be considerably more difficult if not impossible. The third method of analysis can still be employed, but its ultimate effectiveness will be greatly reduced by the fact that sense will be produced at almost every turn, and the intercepter will essentially be in the position of having to guess the message with a seemingly infinite number of messages to choose from. This effect is maximized by assigning a word from the language to every possible code word, and limiting the words assigned to those of fairly common usage. Use of the code in this manner increases the theoretical indecipherability of the cipher by invalidating the first and second methods of analysis, and increases the practical indecipherability of the cipher by increasing the number of possible keys which will produce messages that make sense.

It should be pointed out that it is possible and perhaps desirable to use the code in a different manner than that described above, and thereby increase the practical indecipherability of the cipher in a different way. For example, assume that the number of teeth on each wheel is increased to five thousand to provide for five thousand different symbols and, consequently, that there are five thousand times factorial five thousand possible keys to the cipher, a number which is almost beyond comprehension. It is possible to make a total of twenty-five million different two-letter code words from five thousand different symbols, but this number is many times larger than the total number of words in the language even when each grammatical form of each word is counted separately. Assume, therefore, that the code consists of 350,000 words and represents an effective vocabulary in the order of 100,000 words. It is almost a certainty that the five thousand symbols will appear randomly and with equal or almost equal frequency in the sample of 350,000 words just as in the total population of twenty-five million words, and thus the first and second methods of analysis will be invalidated as before. However, since most of the possible code words will not be included in the code, most of the two-letter sequences produced by possible keys will not be meaningful words, and the previous effect of producing sense at almost every turn will not be realized even though an extremely large number of possible keys will produce messages that make sense. On the other hand, the practical indecipherability of the cipher will be greatly increased by the fact that the number of possible keys to the cipher will be immeasurably larger. Increasing the number of symbols increases the number of possible keys and the number of possible code words, and while it is desirable to have as many keys as possible, it is also desirable to assign words to all of the possible code words, and there is a limit to the number of words that can be assigned, particularly words of fairly common usage. Therefore it is possible that the indecipherability of the cipher will be maximized by increasing the number of possible keys to a practical or theoretical optimum point beyond the point where words of fairly common usage can be assigned to all of the possible code words. It seems more likely, however, that the number of possible keys quickly becomes sufficiently large for its purpose, and that it is more advantageous to include all of the possible code words in the code than to increase the number of possible keys beyond the point where this is possible. Note that two-letter code words are more advantageous than three- or four-letter code words because a larger number of symbols will be required to provide for a given number of code words, and thus the cipher will have a larger number of possible keys. One-letter code words would maximize the number of symbols for a given number of code words, but the purpose of the code would be defeated since it would then be possible to employ the first and second methods of analysis based on the relative frequencies of words rather than letters. In summary, it seems likely that the theoretical indecipherability of the cipher would be maximized at something in the order of two hundred symbols, 1.6×10^{377} possible keys, and forty thousand two-letter code words fully assigned to an effective vocabulary in the order of twelve thousand words.

The above paragraphs should serve to illustrate the principles and procedures of the method of ciphering which has been developed, and thereby make possible a proper evaluation of this method of ciphering. At every point in its development, the cipher is based on concepts of randomness and probability, and while no attempt has been made to define the resulting indecipherability in mathematical terms, this can be done by the techniques of operations research. It does not seem unreasonable to anticipate the result of such an analysis and suggest that the problem presented by the cipher is so extremely formidable that for all practical purposes the cipher is completely indecipherable.

Before concluding, one important advantage should be mentioned; namely, that the cipher is both simple and practical. Certainly it would not be difficult to develop electronic equipment for use in major communications centers which would rapidly and automatically encipher and decipher messages including the encoding and decoding steps in the ciphering procedures. At the otner extreme, a simple pocket device, essentially a circular slide rule, and a small code book (if necessary) would be perfectly adequate for applications of a tactical nature. Indeed, the cipher is particularly suited for use at the tactical level because it is obviously more difficult to maintain security at this level, and except for the key currently being used, security is not required. This fact, plus the basic simplicity of the ciphering procedures, may make possible an increased number and variety of applications, and may offer potential cost savings of considerable magnitude.