MILITARY CRYPTOANALYSIS
Part IV
TRANSPOSITION AND FRACTIONATING SYSTEMS

By
WILLIAM F. FRIEDMAN
Principal Cryptanalyst
Signal Intelligence Service

PREPARED UNDER THE DIRECTION OF THE
CHIEF SIGNAL OFFICER

UNITED STATES
GOVERNMENT PRINTING OFFICE
WASHINGTON 1941
The Golden Guess
Is Morning-Star to the full round of Truth
—Poesies

MILITARY CRYPTANALYSIS, PART IV.
TRANSPOSITION AND FRACTIONATING SYSTEMS

CONTENTS

Section | Paragraphs | Page
-------|-----------|-----
I General | 1-3 | 1-2
II Solution of simple transposition ciphers | 4-11 | 3-17
III Incompletely-filled rectangles | 12-16 | 18-36
IV Opportunities afforded by studying errors and blunders made by enemy cryptographers | 17-19 | 37-39
V Special solutions for transposition ciphers | 20-29 | 60-79
VI Principles of matrix reconstruction | 30-31 | 90-94
VII Solution of grilles | 32-34 | 85-93
VIII Combined substitution-transposition systems | 35-36 | 94-96
IX Solution of the ADFGVX system | 37-43 | 97-143
X Solution of bifid fractionating systems | 44-56 | 144-184
XI Analytical key | 57 | 185
Index | 186-188 | 186-188

(III)
SECTION I

GENERAL

Introductory remarks concerning transposition ciphers 1
Basic mechanism of transposition ciphers — 2
Monophase and polyphase transposition... 3

1 Introductory remarks concerning transposition ciphers — a As stated in a previous text, transposition ciphers are roughly analogous to "jigsaw puzzles" in that all the pieces of which the original is composed are present but are merely disarranged. The pieces into which the picture forming the basis of a jigsaw puzzle may be divided are usually quite irregular in size and shape, the greater the amount of irregularity, as a rule, the greater the difficulty in re-assembling the pieces in proper order. In this respect, too, transposition ciphers are analogous to jigsaw puzzles, for the greater the amount of distortion to which the plain text is subjected in the transposition process, the more difficult becomes the solution.

b In jigsaw puzzles there is usually no regularity about the size of the individual pieces into which the original picture has been cut, and this feature, of course, materially contributes to the difficulty in re-assembling the picture. There are, to be sure, limits (dictated by considerations of practicability) which serve to prevent the pieces being made too small, for then they would become unmanageable, on the other hand, there are also limits which must be observed in respect to the upper magnitude of the pieces, for if they are made too large, the puzzle becomes too easy to solve. These features of jigsaw puzzles also have their analogies in transposition methods. In the latter, if the textual units to be subjected to transposition are made quite large, say entire sentences, the difficulties a cryptanalyst will have in reconstructing the text are practically nil, on the other hand, if these textual units are made quite small, even smaller than single letters, then the reconstruction of the transposition text by a cryptanalyst often becomes a very difficult matter. In between these two extremes there may be various degrees of fragmentation, limited only by considerations of practicability.

c It is fortunate, however, that the cryptanalyst does not, as a rule, have to contend with problems in which the size of the textual units varies within the same message, as is the case in jigsaw puzzles. It is perhaps possible to devise a transposition system in which the text is divided up in such a manner that entire sentences, whole words, syllables, individual letters, and fractions of letters form the units for transposition, but it is not difficult to imagine how impractical such a scheme would be for regular communication, and it may be taken for granted that such irregularity in size of textual units will not be encountered in practical communication.

d The days when the simple methods of word transposition were sufficient for military purposes have long since passed by, and it is hardly to be expected that cryptograms of such ineffectual nature will be encountered in the military communications of even the smaller armies of today. However, in time of emergency, when a counter-espionage censorship is exercised over internal communications, it is possible that isolated instances of simple word transposition may be encountered. The solution of such cases should present no difficulties, unless numerous code names and nulls are also used in the cryptograms. Mere experimentation with the cryptograms, trying various types and sizes of rectangles, will usually disclose the secret text. If code names

1 Reference is here made to so-called fractionating systems. See Special Text No. 166, Advanced Military Cryptography, sec. XI
are used and the context gives no clue to the identity of the persons or places mentioned, it may be necessary to wait until additional messages become available, or, lacking such a possibility, there is usually sufficient justification, under the exigencies of war, to compel the correspondents to reveal the meaning of these code names.

Although transposition ciphers, as a general rule, are much less complex in their mechanics than substitution ciphers, the cryptanalyst usually experiences a feeling of distaste and dismay when confronted with unknown ciphers of this category. There are several reasons for his dislike for them. In the first place, although transposition ciphers are admittedly less intricate than substitution ciphers, as a general rule there are not nearly so many cryptanalytic tools and "tricks" to be used in the solution of the former as there are in the latter, and therefore the mental stimulus and satisfaction which the cryptanalyst usually derives and regards as part of the reward for his hard labor in solving a cipher is often missing in the case of transposition ciphers. In the second place, despite their lack of complexity, the solution of transposition ciphers often involves a tremendous amount of time and labor most of which commonly turns out to be fruitless experimentation. Thirdly, in modern military communication, transposition methods are usually not employed alone but in conjunction with substitution methods—and then the problems may become difficult indeed, for usually before the substitution can be attacked it is necessary first to uncover the transposition. Finally, in working with transposition ciphers a much higher degree of accuracy in more mechanical operations is required than in working with substitution ciphers, because the accidental omission or addition of a single letter will usually necessitate rewriting the work sheets applying to entire messages and starting afresh. Thus this sort of work calls for a constant state of concentrated attention, with its resulting state of psychological tension, which takes its toll in mental wear and tear.

3. Basic mechanism of transposition ciphers—

(a) Basically, all transposition ciphers involve at least two processes (1) Writing the plain-text units (usually single letters) within a specific regular or irregular two-dimensional design called a matrix, cage, frame, rectangle, etc., in such a prearranged manner that the said units are distributed regularly or irregularly throughout the various cells or subsections of that design, (2) When the prearranged plain-text units from the design in such a prearranged manner as to change the original sequence in which they followed one another in the plain text, thus producing cipher text. Since the first process consists of arranging the text within the design, it is technically referred to as the process of inscription, and since the second process consists of transcribing the text from the design, it is technically referred to as transcription. Either or both processes may be repetitive, by rearrangement of course, in which case the intermediate steps may be referred to as processes of inscription, or transcription processes.

(b) It is hardly necessary at this point to give the student any indications as to how to differentiate a transposition from a substitution cipher. If a review is necessary, however, he is referred to Section IV of Military Cryptanalysis, Part I.

4. Monophase and polyphase transposition—

(a) As may be inferred from the foregoing definitions, when a transposition system involves but a single process of inscription, followed by a single process of transcription, the system may be referred to as monophasic transposition, commonly called single transposition. When one or more successive processes intervene between the original inscription and the final transcription the system may be referred to as polyphase transposition. As a general rule, the solution of the latter type is much more difficult than the former. The processes of transcription, inscription, and transcription being accomplished with the same or with different keys.

5. Any system which is suited for monophasic transposition is also usually suited for polyphase transposition, the processes of inscription, transcription and transcription being accomplished with the same or with different keys.

Section II

SOLUTION OF SIMPLE TRANPOSITION CIPHERS

(a) Some examples of these simplest types of transposition follow. Let the message be

BRIDGE DESTROYED AT ELEVEN PM

(1) Reversing only the words and retaining original word lengths

Cipher... E G D I R B E Y D E V E L E R T

(2) Reversing only the words and regrouping into false word lengths

Cipher... E G D I R B E Y D E V E L E R T

(3) Reversing the whole text and regrouping into fivees

Cipher... M N E V E L T A D E Y O R T S E G D I R B

(4) Reversing the whole text, regrouping into fivees, and inserting a null in every fifth position

Cipher... T R E M M P N E V E L T A D E Y O R T S L

(b) Writing the text vertically in two columns and taking the resulting digraphs for the cipher text, as shown at the side. The cipher message becomes

B S R T I D E G Y E D D E

6. These simple types can be solved merely by inspection.

(c) The principles of solution of unilateral, route-transposition ciphers—

The so-called unilateral, route-transposition methods are next to be examined. The solution of cryptograms enciphered by these methods is a matter of experimenting with geometric figures, usually rec...
angles, of various dimensions suggested by the total number of letters in the message, then applying these rectangles, searching for whole words or the fragments of words by reading horizontally, diagonally, vertically, spirally, and so on. (See Special Text No 165, Elementary Military Cryptography, 1935, pars 20, 21.)

The amount of experimentation that must be performed in the solution of ciphers of this type may be materially shortened by means of formulæ and tables constructed for the purpose. But because ciphers of this type are of infrequent occurrence today, formulæ and tables are only occasionally useful and hence they have not been included in this text.

6 Keyed columnar transposition with completely-filled rectangles — In practical cryptography, the dimensions of the transposition rectangle, as a general rule, cannot vary between large limits, that is, it can be assumed in practice that rectangles based upon lines of writing containing 5 letters or more will not commonly be less than 10. In other words, the total number of columns, that is, the number of letters, is determined by a key, then the number of rows becomes a function of the length of the message to be enciphered. If the latter is very long, longer than can be conveniently handled by hand, the mystenous practice to break up a message into two or more parts and treat each part as though it were a separate communication. Such parts are commonly termed sections.

When the last row of a transposition rectangle is completely filled, the solution of the resulting cryptogram is considerably more simple than when this is not the case. Consequently, this will constitute the first case to be studied.

1 It is interesting to observe that Daniel, of biblical fame, was apparently the first cryptanalyst in history (as well as one of the earliest interpreters of dreams), for he solved the cryptogram in the "handwriting on the wall," obtaining as his decipherment words which he interpreted as predicting the downfall of Babylon and the dynast (Daniel V 1–8). The following partial account of the episode is not as enlightening as one might wish, but it is the best explanation available. It is taken from Dr. Max Mueller's article on the subject in The Jewish Encyclopedia, vol. 8, pp. 480–481 (1905): "HELE, MEDE, TEREK, ETHABAN, HABER SINEK, PIBARBER, WARM, WINK.

Writings of a mystical nature, found on the walls of Babylon's palace, and in the course of time he solved the cryptogram. Predictions were made by observing the writing, but they were unable to read the mysterious words. Frightened by the apparition, Daniel was then summoned to the Royal Presence. He was told that the King had been changed in the time of versa, so that even the Jews that were found in the royal court could not read an inscription written in cursive characters. But those followed R. Rennson in maintaining that the writing had not changed other solutions for the problem, e.g. it was written in the cryptographer combinations HELE, each letter of which being substituted by its companion. On the other hand, if it is assumed that the words were written in the form of columns, it is obvious that the cryptogram is considerably more simple than when this is not the case. Consequently, this will constitute the first case to be studied.

The first question which presents itself to the cipherist, namely, why could the inscription be deciphered by Daniel? He had the attention of the Talmudists, who advanced various answers. Certain of them concluded that the Hebrew writing had been changed in the time of versa, so that even the Jews that were found in the royal court could not read an inscription written in cursive characters. But those followed R. Rennson in maintaining that the writing had not changed other solutions for the problem, e.g. it was written in the cryptographer combinations HELE, each letter of which being substituted by its companion. On the other hand, if it is assumed that the words were written in the form of columns, it is obvious that the cryptogram is considerably more simple than when this is not the case. Consequently, this will constitute the first case to be studied.

2 In solving a cryptogram of this type the first step taken by the cryptanalyst is to ascertain the dimensions of the rectangle. Close for this are usually afforded by finding the factors of the total number of letters in the cryptogram. Suppose the cryptogram contains 52 letters. The dimensions of the transposition rectangle may be 4 x 38 or 8 x 19, by which is meant that four hypotheses may be made with respect to its dimensions. The rectangle may consist of

1. (1) columns with 38 rows, or
2. (3) columns with 4 rows, or
3. (5) columns with 19 rows, or
4. (19) columns with 8 rows.

In practical work it is rather unlikely to encounter a rectangle that conforms to hypotheses (1) or (2), and for the present these may be discarded. As to choosing between hypotheses (3) and (4), a simple criterion is at hand. If, the rows of letters in the correctly-dimensioned rectangle would read plain text exactly it were not for the transposition which they have undergone within the rows. Therefore, the rows of a correctly-dimensioned rectangle are more likely to manifest the expected vowel-consonant proportions of normal plain text than are the rows of an incorrectly-dimensioned rectangle, because in the latter case there are brought into some of the rows letters which belong to other rows and which are likely to disturb the normal vowel-consonant proportions of plain text. That is, in an incorrectly-dimensioned rectangle some of the rows will have too many consonants and not enough vowels. In other rows this relationship will be reversed, whereas in a correctly-dimensioned rectangle each row will have the proper number of vowels and consonants. Hence in solving an unknown cryptogram of this type, if a count is made of the vowels and consonants in the rows of the rectangle, it is likely that the hypothesis which statement is nearest to the proportions of normal text will be right.

The procedure is to select a column which has a good assortment of high-frequency letters and find another column which may be placed before or after the selected column to build up high-frequency digraphs in the rows, when such a pair of columns has been found, attempts are made at the interchange of the columns with the rows of the rectangle, as if building up more probable values of the digraphs in the rows to assist in discerning combinations which are on the borderline of choice. However, it must be noted that the totals obtained by simple addition of the frequency values of

The Standard Dictionary defines the verb anagram as follows: "noun 1 The letters of a word or phrase so transposed as to make a different word or phrase as, 'none' and 'emn' are anagrams of 'emn' 2 A transposition, interchange" As a verb, it is defined as "to anagrammatise, to make an anagram, of, make anagram" The construction of anagrams was a very widespread pastime in previous centuries. See Whatley's Of Anagrams, London, 1868. In the course of the work, the word would therefore be confused with the other words, the letters of which were therefore rearranged already form bondage words or intelligible phrases. However, this would hardly be broad enough for cryptoanalytic purposes. As used in cryptoanalysis the word is commonly employed as a verb to refer to the process of rearranging the disordered letters of cipher text so as to reconstruct the original plain text.
the digraphs should be considered only as rough approximations or guides in weighing probabilities in favor of one hypothesis as against another, for theoretically the probability of the simultaneous occurrence of two or more independent events is the product, and not the sum, of their respective probabilities. In most cases the calculation of products involves an amount of labor unwarranted by the results to be expected, so that simple addition of probabilities is usually sufficient. However, if tables of the logarithms of the probabilities are readily available, the addition of these logarithms becomes a simple matter and affords a more accurate guide in selection of combinations produced in the anagramming process. Once a set of four or five columns has been correctly assembled it is usually the case that the process may be completed very quickly, for with the placement of each column the number of remaining columns possible for selection diminishes, toward the close of the process, when only two or three columns remain, their placement is almost automatic.

It is desirable, as a final step, to try to reconstruct, if possible, the literal key from which the numerical transposition key was derived

7. Example of solution—Given the following cryptogram, the steps in solution will be set forth in detail

Cryptogram (126 letters)

\[
\begin{array}{cccccccc}
I & L & H & D & T & I & E & O
\end{array}
\begin{array}{cccccccc}
E & U & M & T & S & O & N & S
\end{array}
\begin{array}{cccccccc}
O & S & E & E & E & I & O & F
\end{array}
\begin{array}{cccccccc}
T & R & H & N & E & A & T & N
\end{array}
\begin{array}{cccccccc}
T & N & Y & U & T & L & B & F & A
\end{array}
\begin{array}{cccccccc}
E & D & F & O & Y & C & A & P & D & T
\end{array}
\begin{array}{cccccccc}
R & R & I & I & A
\end{array}
\begin{array}{cccccccc}
R & I & V & N & L & R & R & N & W & E
\end{array}
\begin{array}{cccccccc}
\end{array}
\begin{array}{cccccccc}
\end{array}
\]

The cryptogram contains 126 letters (factors of 126 2, 3, 6, 7, 9, 14, 18, 21), suggesting rectangles of \(7 \times 18\) or \(9 \times 14\). If the former dimensions are taken, the rectangle may have 7 columns and 18 rows or 18 columns and 7 rows, if the latter dimensions are taken, it may have 9 columns and 14 rows or 14 columns and 9 rows. The factors of 126 do not, of course, preclude the possibility that the rectangle may be \(6 \times 21\), that is, with 21 columns and 6 rows or 6 columns and 21 rows. If no good results were obtained by testing rectangles of the dimensions indicated \((7 \times 18\) or \(9 \times 14\)), then one would proceed to test rectangles \(6 \times 21\). In the event that all tests on the basis of a completely-filled rectangle failed, then it would be assumed that the rectangle may be incompletely filled. In making the vowel-consonant test described in paragraph 66, it is advisable to base the count on the columns as well as on the rows of a rectangle, since it is possible that the cryptogram was prepared by marking the plain text in rows and transcribing the text from the columns, or vice versa. After examining a rectangle both horizontally and vertically, it is often possible to discard various arrangements without further tests. For example, at A in figure 1 there is shown a rectangle of 7 columns and 18 rows. Now in a row of 7 letters there should be \((7 \times 40\) percent = 2.8\) either 2 or 3 vowels, but rows 12 and 15 contain no vowels at all and rows 8 and 9 contain 5 vowels, rows 16, 6, 7 vowels. It is concluded at once that this arrangement is highly improbable. If the plain text had been marked vertically in the same rectangle, and then the rows had been transposed in forming the cipher text, then in each column (18 letters) there should be \((18 \times 40\) percent = 7.2\) about 7 vowels, but column 2 contains 11 vowels and column 6 only 4. This likewise indicates that it is highly improbable that the message was marked vertically and the cryptogram formed by transposing the rows. But when the arrangement at

\[\text{Figure 1}\]

B in figure 1 is studied, it is not so easy to say at once that it is improbable. For in 18 letters there should be about 7 vowels and none of the rows of this arrangement shows too great a departure from this expected number. This possibility will have to be explored further and it is for the moment put aside. If it be assumed that the message was marked vertically in the rectangle \(18 \times 7\) and the rows subjected to transposition, there should be \((7 \times 40\) percent = 2.8\) 2 or 3 vowels in each column. But since several of the columns show rather considerable departures from this expected number, it may be concluded that a vertical inscription and horizontal transcription is not probable and this assumption may be eliminated. Then the arrangements at C and D in figure 1 are studied in the same manner, with the result that at the end of the study the situation as regards the various assumptions is summarized as follows.
### Analysis of Vowel Occurrence

#### Rectangular Arrays

**Rectangle 7 × 18**

1. **Improbable**

2. **Very improbable**

3. **Possible**

4. **Impossible**

**Suppose the average amount of deviation from the expected number of vowels in each row is calculated by finding the difference between the actual and expected numbers in each row, adding those differences (neglecting signs), and dividing by the total number of rows.**

For assumptions (3) and (5), the latter are subjected to further scrutiny.

#### Deviation Analysis

**Deviation Analysis**

<table>
<thead>
<tr>
<th>Assumption</th>
<th>Description</th>
<th>Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>(3)</td>
<td>Horizontal inscription, columnar transcription</td>
<td>1.2</td>
</tr>
<tr>
<td>(4)</td>
<td>Vertical inscription, horizontal transcription</td>
<td>1.2</td>
</tr>
<tr>
<td>(5)</td>
<td>Vertical inscription, horizontal transcription</td>
<td>1.2</td>
</tr>
</tbody>
</table>

**Average deviation = 1.2**
The average amount of deviation for assumption (5) is only 0.9 as against 1.2 for assumption (3), therefore the former assumption is considered to be somewhat better than the latter and it will be tried first.

d. The columns of the rectangle shown in figure 16 are now to be cut apart and the procedure of anagramming applied (For this it is best to have the cryptogram written on cross-section paper preferably with 3/8-inch squares for ease in handling) Consider column 7, with the letter J in row 5, this letter, if it is a part of a word, must be followed by a vowel, which eliminates columns 1, 3, 4, and 5 as possibilities for placement on the right of column 7. Here are the digraphs formed by combining column 7 with columns 2, 6, 8, and 9, respectively, and the totals obtained by adding the frequency values of the digraphs formed in the rows.
choice of combination (3) was a bad one, even though it gave the greatest total when frequency values for the digraphs were summed. The second greatest total was for combination (2) in which columns 7 and 6 were put together. The infrequent digraph J U suggests a word such as JUST or JUNCTION. If it were the former there should be a column containing an S in the 5th row, and there is no such column. If the word is JUNCTION, there should be a column containing an N in the 5th row, and there is only one such column, the 3rd. Placing column 3 after columns 7-6 gives the trigrams shown in figure 4-A. All of these trigrams are excellent except the last, and that one may be either an abbreviation of a signature, or possibly nulls added to complete the rectangle. If the word JUNCTION is correct then there should be a column with a C in the 5th row, but none is found. However, column 9 has a C in the 6th row, and if it happened that the last column on the right is No. 3, then column 9 would be the 1st column. Thus, as shown in figure 4-B, the arrangement of columns becomes 9-7-6.

It is believed that the procedure has been set forth with sufficient detail so as to make further demonstration unnecessary. The rectangle can be completed very quickly and is found to be as shown in figure 5.

It will be interesting to see if a calculation based upon the sum of logarithms of the probabilities given in figure 2 would have given the correct combination as the first choice. Note the results shown in figure 6. This calculation gives the correct combination as first choice, viz., 7-6, with a logarithmically-weighted value of 17.35 as against a value of 16.51 for combination 7-8, which was the first one tried on the basis of merely the sums of the frequency values of the digraphs

As a matter of interest, it may be observed that the combination 7-6 is 7 times more probable than combination 7-8, since the difference between 17.35 and 16.51 is 8.4, which is the logarithm of 7. Likewise, combination 7-6 is roughly 15,000 times more probable than combination 7-9, since 17.35-13.77 = 14.8.

8 The probable-word method of solution. — The probable-word method of attack is as important in the solution of transposition ciphers as it is in the solution of substitution ciphers, and if the cryptanalyst is able to assume the presence of such probable words as are usually encountered in military communications, the solution, as a rule, comes very quickly.

As an illustration, looking at the first row of letters in the rectangle shown in figure 1, the letters I S T B R A T F almost at once suggest FIRST BATTALION as the initial words of the message. The rearrangement of the columns of the cryptogram bring the necessary letters into juxtaposition at once discloses the key. Thus

It will be noted that this assumption requires that there be a column headed by P A, another headed by S I, another headed by R I, and so on. Had such columns not been found, then the word BATTALION would not be possible. In that case the word FIRST would still remain as a point of departure for further experimentation.

In the foregoing illustration, the probable word was assumed to appear in the first line of text in the rectangle. If the probable word being sought is in the interior of the message, the steps must be modified somewhat but the basic principle remains unchanged. The modifications are of course obvious.

9 General remarks on solutions. — In solving transposition ciphers advantage should be taken of all the characteristics and idiosyncrasies which are applicable to the language of the

As shown in figure 4-A, the arrangement of columns becomes 9-7-6.
enemy, because they often afford clues of considerable assistance to the cryptanalyst. In all languages there are certain letters, usually of medium or low frequency, which combine with other letters to form digraphs of high frequency. For instance, in English the letter H is of medium frequency, but it combines with T to form the digraph TH, which is of highest frequency in literary text, it also combines with a C, a letter of medium frequency, to form the fairly frequent digraph CH. The letter V is almost in the low-frequency category yet it combines with E to form the digraph VE, which is the 14th in frequency. The low-frequency letter K often combines with C to form the digraph CK. Consequently, in working with transposition ciphers in English, when there is an H, attempts should be made to combine it first with a T or with a C, a V should be combined first with an E, a K should be combined first with a C, and so on.

There is usually in every language at least one letter which can be followed by only one certain other letter, forming what may be termed an obligatory sequence, or an unavertable digraph. In all languages having the letter Q, the combination Q U constitutes such an invariable digraph. In bonafide words of the German language the letter C is never used by itself, when present the letter C is invariably followed by an H, except on rare occasions when the digraph CK is employed. In English, the letter C can be followed only by a vowel, the letter X can only be preceded by a vowel and, except at the end of a word, can only be succeeded by a vowel, or by one of a limited number of consonants (G H P T), and so on. Letters which behave in this manner, that is, letters which have what may be called a limited affinity in combining with other letters to form digraphs, constitute good points of departure for solution and are therefore of sufficient importance to warrant their being designated by the more or less descriptive name of pilot letters.

The presence of pilot letters in a transposition cipher often forms the basis for the assumption of probable words. Obviously, a special lookout should be kept for words of rather high frequency (in military correspondence) which contain letters of low or medium frequency. The frequent word CAVALRY, for example, would suggest itself if the cryptogram has the letters C V L and T, which are all of medium frequency. The important word ATTACK suggests itself if the cryptogram has an H, a letter of low frequency, and a C, one of medium frequency, and so on.

The mechanism of simple columnar transposition makes possible the production of rather long sequences of vowels and long sequences of consonants in the text of the cryptogram. Note, for example, in the cryptogram on p. 6, the sequence of vowels O E E E E E and the sequence of consonants V N L R N R W. If the enciphering or plain-text rectangle is consulted, it will be seen that these two sequences belong together, that is, they are in adjacent columns in that rectangle. It is a characteristic of plain text that consonant-vowel or vowel-consonant digraphs are more frequent than consonant-consonant or vowel-vowel digraphs, and therefore when long sequences of consonants and of vowels are found in transposition ciphers, a good start toward solution may result from assuming that such sequences come from adjacent columns.

It should, however, be noted in connection with tell-tale letters such as Q (entering into the composition of Q U) and C (entering into the composition of CH), that satiric cryptographers who realize the clue which such letters afford often replace invariable digraphs by simple characters, usually those rarely used in the language in question. For example, C H in German may be replaced by Q, Q U in French, by K, and so on. When this is done, solution is made more difficult, but only in those cases where it is dependent upon finding letters forming obligatory sequences in plain text does this sort of subterfuge become of importance.

The letter Q may, of course, be part of an abbreviation, such as SQ, for "square," or it may be used as a null, or as a sign of punctuation. However, unless there are good reasons for believing that this letter is used for such purposes, SQ may be considered to be an invariable digraph.

The CV and VC digraphs constitute about 13 percent of all digraphs.

The presence of many Q's, or K's, or X's in a transposition cipher should not, however, be taken as prima facie evidence of the type of replacement noted in the preceding subparagraph. It is possible that such letters may be used as sentence separators or other punctuation, possibly they may be nulls, although the alert cryptographer would either use nulls not at all or, if he had to, would use letters of medium or high frequency for this purpose.

Because it is important that the cryptanalyst take advantage of every peculiarity specifically applicable to a cryptogram to be solved, especially as regards the presence of low-frequency letters, it is advisable that a uniform frequency distribution be prepared, just as though he were going to deal with a substitution cipher. This is probably the quickest way of bringing to light the peculiarities which may be helpful in solution.

Reconstruction of literal key—The reconstruction of recovery of the literal key from which the numerical transposition key was derived is naturally the last step in the solution of cryptograms of this type. It is often of more than merely academic interest, because it is found that the enemy is employing for this purpose some well-known book, or words or phrases of a simple nature associated with the locales of operations, this fact may be of highest importance in subsequent work.

In this process there are only a few guiding principles to be noted and much must be left to the ingenuity and imaginative powers of the cryptanalyst. Taking as an example the numerical key uncovered in the solution of the cryptogram in paragraph 7, the procedure will be set forth below.

e. The numerical key referred to was found to be 9 1 5 2 8 4 7 6 3. Assuming that this sequence was derived in the usual manner, by assigning numbers to the letters of a key word in accordance with their relative positions in the normal alphabet, the sequence forms the basis for the key-word reconstruction diagram shown in figure 7-A, in which the individual key numbers are written from left to right on different "levels" so that each level contains only numbers normally in succession.

f. It is likely that the digit 1 on the first level in the key-word reconstruction diagram represents a letter at or at least close to the beginning of the alphabet. Since the digits 2 and 3 are on the same level as the digit 1, it is likely (1) that the letter represented by 1 occurs 2 more
times in the key word, or (2) that the digit 2 represents another letter, also near the beginning of the alphabet, and that this letter is repeated, or (3) that the digits 2 and 3 represent two different letters both near the beginning of the alphabet, or (4) that all three digits represent different letters but all near the beginning of the alphabet. The digit 4, on the second level in the reconstruction diagram, must represent a letter beyond the letter represented by the digit 3, the digit 5 must represent one beyond the letter represented by the digit 4, and the digit 6 may represent the same letter as 5, or a letter not much beyond that represented by 5. Assuming that the letters composing the key word are fairly well distributed over the entire alphabet, the digit 7 must represent a letter near or slightly beyond the middle of the alphabet, the digit 8 must represent one further toward the end of the alphabet than does the digit 7, and so on. Assigning several values to the digits, in accordance with the foregoing principle, the results are as shown in figure 7-B.

It is perhaps possible that some students may find the process of reconstructing the literal key somewhat easier if the variant possible letters are merely listed under the respective key numbers as shown in figure 7-C. The candidates for the successive positions in the literal key thus appear in a rather condensed space and the eye is able to pick up "good" combinations very quickly.

\[
\begin{align*}
1 & \quad 2 & \quad 3 & \quad 4 & \quad 5 & \quad 6 & \quad 7 & \quad 8 \\
R & \quad K & \quad A & \quad M & \quad F & \quad L & \quad K & \\
S & \quad B & \quad L & \quad B & \quad N & \quad G & \quad M & \quad L & \quad B & \\
T & \quad C & \quad M & \quad C & \quad U & \quad H & \quad N & \quad M & \quad C & \\
U & \quad D & \quad N & \quad P & \quad I & \quad O & \quad N & \\
V & \quad E & \quad O & \quad E & \quad R & \quad J & \quad P & \quad O & \quad E \\
W & \quad X & \quad Y & \quad Z
\end{align*}
\]

**Figure 7-C**

f. Now comes the trying process of finding a "good" word in this assemblage of letters. The beginning and end of the word are the easiest points of attack, and it is useful to keep in mind the relative frequency order of letters as initial and final letters of the language in question. For English, the data are as follows:

As initial letters: T S A F C O R D N P E M I W B H L U G Y V J Q K X
As final letters: E T D S N Y R O H L A F G P M X C K W U B I Z Q J V

Studying the candidate letters at the end of the key, it is seen that E is one of the possibilities. If this is correct, then a good ending would be one of the type vowel-consonant-vowel, with E as the final letter. There is but 1 vowel in the fourth level in the column under the digit 7, the letter O. This gives O K E, G L E, O M E, or one as possible terminal trigraphs, the best of which from a frequency standpoint is O K E. Seeing the letters P and H in columns 9-4, the ending PHONE and then the word TELEPHONE suggests itself. Checking to see if there are any inconsistencies, none is found and the solution is

Numerical key: 9 1 5 2 8 4 7 6 3
LITERAL key: T E L E P H O N E

*Taken from Tables 9-D (2) and 2-E (2), p. 111, Military Cryptanalysis, Part 1
SECTION III
INCOMPLETELY-FILLED RECTANGLES

<table>
<thead>
<tr>
<th>Paragraph</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>General principles underlying solution</td>
<td>12</td>
</tr>
<tr>
<td>Do selecting the length of the columns of the rectangle, constructing the &quot;hash&quot; diagram</td>
<td>13</td>
</tr>
<tr>
<td>Solution of example</td>
<td>14</td>
</tr>
<tr>
<td>Alternative method of solution</td>
<td>15</td>
</tr>
<tr>
<td>Comparing remarks on simple columnar transposition</td>
<td>16</td>
</tr>
</tbody>
</table>

12 General principles underlying solution—A In the system designated keyed columnar transposition the feature which differentiates an incompletely-filled rectangle from one that is completely filled is a very simple one from the cryptographic point of view. The bottom row of the rectangle in the former case merely lacks one or more letters, a feature which only very slightly complicates the system in practical operation. But the consequences of this simple difference between the two types are, from the cryptographic point of view, quite profound, and the cryptographic effect of this small change in cryptographic procedure is seemingly all out of proportion with the simplicity of the difference.

In cryotirams involving completely-filled rectangles are rather easy to solve because of two circumstances. In the first place, since the rectangle is completely filled, the various possible dimensions of the rectangle can be ascertained by noting the factors of the total number of letters. Usually only a few possibilities are indicated and therefore the probability of the reduction of the amount of experimentation that would be required in the absence of this situation, since it is obvious that when working with incompletely-filled rectangles a good many rectangles of various dimensions become possibilities for trial. In the second place, the columns in a completely-filled rectangle contain the same number of letters, and therefore the anagramming process (matching and assembling of columns) can be performed without any mental reservations such as must be made in working with incompletely-filled rectangles because of uncertainty as to whether the letters which are juxtaposed in form digraphs and trigraphs really come from the same row in the plaintext rectangle. The latter statement cahs for a bit more explanation.

The columns of an incompletely-filled rectangle are of two sorts, which may be conveniently designated as long and short. The long columns are at the left of the rectangle and each one contains just one more letter than the short columns, which are at the right. Thus, for example, from the fact that it is only the last row in such a rectangle which lacks one or more letters to complete the rectangle. The term width, as applied to a transposition rectangle, will be convenient to designate the number of columns, which is, of course, determined by the length of the numerical key or the number of letters in the literal key. Given the width of the rectangle and the total number of letters in the cryptogram, the length and number of the long and short columns may be found by a simple calculation. Multiply the width of the rectangle by the smallest number which will yield a product greater than the total number of letters in the cryptogram. The multiplier gives the length of the long column, this multiplier minus 1 gives the length of the short columns, the excess over the total number of letters gives the number of short columns, the latter deducted from the width gives the number of long columns. Thus, with a cryptogram of 287 letters and a rectangle 15 columns in width [(15X20) - 13 = 287] the long columns will have 20 letters, the short one, 19 letters, there will be 13 short columns and 2 long ones.

d Now if the cryptanalyst were able to cut up the text of a cryptogram produced from an incompletely-filled rectangle into sections corresponding in length with the actual long and short columns, he could handle these columns in exactly the same manner that he handles the equal-length columns in the solution of a cryptogram produced from a completely-filled rectangle. In fact, the solution would be easier because he knows that all the short columns fall at the right, all the long columns at the left of the transposition rectangle, and therefore the amount of experimentation he must undertake in his attempts to juxtapose columns in the anagramming process is considerably reduced. But, unfortunately, there is usually no way in which, at the initial stage of solution, the cryptanalyst can find out, from a single cryptogram, which are the long columns and which the short. This is obviously a matter directly connected with the specific transposition key, and the latter is the sole unknown factor in the whole problem.

If it were practicable to transcribe a cryptogram of this type according to all the possible transposition keys for a given width of rectangle, solution would obviously merely consist in scanning the various rectangles to find the one which is correct—for there will be only one such rectangle. A rectangle 15 columns in width may have been enciphered by any one of factorial 15 transposition keys. While it is conceivable that machinery might be devised for this purpose, so that the production of the millions of possible rectangles could be effected in a relatively short time, in the present state of the art no such machinery has yet been devised. Furthermore, it is problematical whether a solution by such means could be achieved in a reasonable length of time even if the machinery were available, because of the immensity of the task it would have to perform.

f However, this question may be asked: Given a cryptogram of 1 letters and a rectangle of a column in width, is it possible to transcribe the text within a single rectangle so that the latter will show what letters will constitute the respective columns for all possible transposition keys of a element? If so, then such a rectangle would be useful in trying to solve the cryptogram, because the rectangle would then limit the amount of experimentation that would have to be performed by the anagramming process, since it would show whether or not two letters which are brought together in that process to form a digraph could possibly have been in the same row in the plain-text rectangle. If not, of course there would be no use in forming such digraphs, and the number of possible rectangles becomes much reduced. Another way of indicating what is meant is to say that such a rectangle would show the maximum amount that one column may be shifted up or down in trying to match it with another column in the anagramming process. This would be made clearer in a subsequent paragraph. At this point it will merely be stated that it is easy to prepare a rectangle of the nature indicated above for any key, columnar-transposition cryptogram.
12. Delimiting the lengths of the columns of the rectangle, constructing the "hat" diagram — a. Given the following cryptogram of 224 letters and an assumed width of 12 columns in the enciphering rectangle.

**Cryptogram**

<table>
<thead>
<tr>
<th>Column 1</th>
<th>Column 2</th>
<th>Column 3</th>
<th>Column 4</th>
<th>Column 5</th>
<th>Column 6</th>
<th>Column 7</th>
<th>Column 8</th>
<th>Column 9</th>
<th>Column 10</th>
<th>Column 11</th>
<th>Column 12</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDNFP</td>
<td>TIRNT</td>
<td>DURO</td>
<td>EXALN</td>
<td>IETG</td>
<td>NTTME</td>
<td>MZLCN</td>
<td>MEBFC</td>
<td>FTTEE</td>
<td>EMIAL</td>
<td>TEAjj</td>
<td>NOIV</td>
</tr>
<tr>
<td>DSTEQ</td>
<td>ITDMA</td>
<td>NLNOE</td>
<td>BOUHE</td>
<td>NLESE</td>
<td>AACTR</td>
<td>TSDIEF</td>
<td>STFEE</td>
<td>EMIAL</td>
<td>TEAjj</td>
<td>NOIV</td>
<td>EMAI</td>
</tr>
<tr>
<td>FMBIE</td>
<td>HBVTS</td>
<td>ESRSL</td>
<td>LXROR</td>
<td>UMETY</td>
<td>OIKNK</td>
<td>TNDAH</td>
<td>RHRQI</td>
<td>ETETN</td>
<td>OTRA</td>
<td>VRIRS</td>
<td>TSGEF</td>
</tr>
<tr>
<td>TNDAH</td>
<td>RHRQI</td>
<td>ETETN</td>
<td>OTRA</td>
<td>VRIRS</td>
<td>TSGEF</td>
<td>TNDAH</td>
<td>RHRQI</td>
<td>ETETN</td>
<td>OTRA</td>
<td>VRIRS</td>
<td>TSGEF</td>
</tr>
<tr>
<td>EAGOT</td>
<td>HACH</td>
<td>SHEEV</td>
<td>TRESR</td>
<td>AIEIA</td>
<td>TEEAL</td>
<td>EAGOT</td>
<td>HACH</td>
<td>SHEEV</td>
<td>TRESR</td>
<td>AIEIA</td>
<td>TEEAL</td>
</tr>
<tr>
<td>AENE</td>
<td>MYTFI</td>
<td>TANLN</td>
<td>NUACL</td>
<td>RENRT</td>
<td>RATSO</td>
<td>AENE</td>
<td>MYTFI</td>
<td>TANLN</td>
<td>NUACL</td>
<td>RENRT</td>
<td>RATSO</td>
</tr>
<tr>
<td>ALODI</td>
<td>RORYN</td>
<td>NRGY</td>
<td></td>
<td></td>
<td></td>
<td>ALODI</td>
<td>RORYN</td>
<td>NRGY</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**DISTRIBUTION**

A cryptogram of 224 letters and a rectangle of 12 columns (12x19) - 4 = 224 indicates 4 short columns of 18 letters and 8 long columns of 19 letters. The outlines of a rectangle of this specification are drawn on a sheet of cross-section paper and the text is transcribed within it, for the moment assuming only that the transposition key consists merely of the straight sequence of numbers 1 to 12. Thus...

1. Now the capital letters in this matrix or diagram, which is often called a crown or hat diagram, represent the letters which are in the columns in case the first hypothesis (keys numbers 9, 10, 11, and 12 at the extreme right) is true. The capital letters above the heavy black line together with the lower-case letters at the top of the diagram represent the letters which are in the columns in case the second hypothesis (keys numbers 1, 2, 3, 4 at the extreme right) is true. Therefore, since the hat diagram covers the two possible extremes with reference to the positions occupied by the short columns and embraces all possible intermediate conditions by showing what letters may be in the respective columns under any possible arrangement of long...
and short columns, the hat diagram is applicable to any possible numerical key for the cryptogram in question and for the assumed width of rectangle. Therefore, in the analysis process the hat diagram shows the maximum possible amount that any column may be shifted up or down in juxtaposing two columns to form digraphs of letters assumed to come from the same row in the plain-text rectangle. This is because all the letters of the first row of the actual encompassing rectangle will be found in rows 1 to 5, inclusive, of figure 9, all the letters of the second row of the rectangle will be found in rows 2 to 6, inclusive, and so on, as indicated by the braces at the right in figure 9.

Thus there arises the following important principle. Designating the number of short columns in a specific diagram by $n$, only such letters as fall within $(n-1)$ consecutive rows, will be letters that may have appeared in the same row in the actual transposition rectangle. Or, another way of stating the principle is this. Both members of any pair of letters actually in the same row in the transposition rectangle will be found only among the letters appearing in $(n-1)$ consecutive rows in the correct hat diagram. 

In the case under discussion, if the first letter of such a pair is located in row 8, for example, the other letter cannot be in rows 1, 2, 3, or 13 to 20 of figure 9.

The usefulness of this principle in connection with the construction and employment of the hat diagram will soon become apparent. For example, again referring to figure 9, take the letter Q in row 19, column 7, it must be followed by a U in the plain text. There are 4 U's in the message they are in row 13 column 11, row 14 column 3, row 17 column 1, and row 20 column 6.

Note the question is, can any of these 4 U's follow the Q, or may one or more of them be eliminated from consideration at once? Since the U's in rows 13 and 14 fall outside the 4 consecutive rows above that in which the Q is located, it follows that neither one of these U's can be the one that succeeds the Q. Thus two candidates are automatically eliminated from consideration. 

The U in row 17 and the U in row 20 are both possible candidates.

Solution of example—$a$. With the foregoing preliminaries out of the way, the solution of the cryptogram can now be carried forward with rapid progress. It has been indicated that the Q in row 19, column 7 (fig. 9), may be combined with either the U in row 17 column 1, or the U in row 20 column 6. Suppose the columns of figure 9 are now cut apart for ease in assembling. Juxtaposing the indicated columns yields what is shown in figure 10. Since the combination shown at a in figure 10 involves column 1, it obviously begins with the letter 0 and ends with the letter A or L, no other letters can be added to this column. Since column 7 is already the maximum length this column can be under any circumstances, no letters can be added to it at the bottom. Therefore, all the digraphs possible to form by juxtaposing these two columns are indicated in figure 10d. There are only 17 digraphs in all, whereas there should be at least 18.

Hence, combination 7–1 is impossible, and combination 7–6 is the only one that needs to be considered further. There are many excellent digraphs in it, and only one which admittedly looks rather bad, the EX. Seeing the digraphs KB and KS in these columns, a good assumption to make is that the KS are preceded by the letter O. Is there a column with 2 C's in approximately the correct region? Column 4 meets this requirement. Note the excellent digraphs it yields, as shown in figure 10e. It now becomes fairly easy to add columns to this nucleus. For instance, the trigraph R Y B suggests a word ending in R Y, such as INFANTRY, ARTILLERY, CAVALRY, the trigraph M Y V suggests MOVING, the trigraph C K B suggests the word ATTACK, followed by a word beginning with B, and so on. Trial of only a few columns soon yields what is shown in figure 10f, from which it soon becomes probable that the long columns end with column 12, since the letters after L Y yield an impossible sequence (E F F Y). Since it was originally assumed that there are only 4 short columns in the transposition rectangle, and since 4 columns have already been placed at the right (4–7–6–10), the rectangle, with the columns thus far placed, must be as shown in figure 10g. This, then, at once tells what the limits of columns 2, 3, 5, 8, 9, and 11 must be, and the rectangle can now be filled in without further delay. The completed rectangle is shown in figure 11.

![Figure 11](image-url)
The last step, recovering the literal key, is then taken. The key is to be found among the letters of the literal key reconstruction diagram in figure 12.

![Diagram](image-url)

**Figure 12.**

The termination ATIONS seems a likely possibility. If this is correct, assignment of letters becomes modified as shown in figure 13.

![Diagram](image-url)

**Figure 13.**

The word PENETRATIONS will fit and it is taken to be presumably correct. There is no absolute certainty about the matter, for it is conceivable and possible that there are other words which can be made to fit the sequence of key numbers given.

15. Alternative method of solution—a. The foregoing solution will no doubt appeal to the student as being straightforward and simple—if the original assumption as to the width of the transposition rectangle is correct. But, unfortunately, there is no way of knowing whether such an original assumption is correct until solution is well under way. In practice, of course, what
might be done within a well-organized cryptanalytic unit would be to divide up the work among the individuals constituting the unit, each being assigned one or more specific hypotheses to try out with respect to width of rectangle. Then one of these individuals would determine the correct width and he would be joined by the others as soon as an entering wedge had been found in this way. Of, if the cryptanalyst is working alone, he must try out successive hypotheses as to width of rectangle until he hits upon the correct one. In making these hypotheses he must be guided by previous experience with enemy correspondences, which may afford clues as to minimum and maximum widths of rectangles.

However, there is another method of attack which does not necessitate making any definite initial assumptions with respect to the width of the transposition rectangle. This method is a modification of the method set forth in the preceding paragraph. The text of the cryptogram is written out columnwise on cross-section paper, every fifth letter being numbered for purposes of reference. Plenty of space is left between the columns, and about 10 to 15 letters at the bottom of each column are repeated at the top of the next column so that at any point in the transcription there will be a single unbroken string at least one complete column of letters from the transposition rectangle. Then a section of consecutive letters is written on a separate strip of cross-section paper, columnwise of course, and by juxtaposing this strip against the whole text, sliding it to various points of coincidence against the text, an attempt is made to find that position in which the best digraphs are formed of the letters on the movable strip and the fixed sequence. Of course, if there is a Q in the cryptogram, the sliding-strip section is made to contain this letter, and the strip is then placed against the text where a Q is found, so as to form the digraph QU. The digraphs formed above and below the QU are then studied, possibly a written record is made of the digraphs. Then the same thing is done with the Q and all other Qs in the text, to ensure that a correct start is made. It is this initial step which is likely to give the most difficulty (if there is anything difficult at all in the procedure) and it is important that it be correct. If this first step is easy, then solution follows quite rapidly, if the cryptanalyst is lucky, and makes several false starts, the process is likely to be a slow one. In choosing among various possible juxtapositions it may be advisable to calculate the probability value of each possibility by adding the logarithms of the frequency values of the digraphs, as explained in paragraph 3. In the absence of any Qs in the text, recourse may be had to two or more other probable digraphs, based upon the presence of certain other telltale low-frequency letters, such as C, H, J, K, V, and X. The cryptanalyst is fortunate if there are two or three of these low-frequency letters close to one another in a series of letters, for in this case he can search for a place where there are high-frequency letters (in a corresponding sequence) that might be combined with them. For example, there is shown a sequence HEHKK. A sequence such as ARCC must be good to try, for it will yield the digraphs AV, AA, TH, CH, CK. Or if there is a long sequence of consonants, the cryptanalyst should look for a correspondingly long sequence of vowels, since those make the best combinations and are therefore most probable. For these reasons it pays to study the text quite carefully before choosing a starting point, to find all such peculiar sequences as might be useful in affording a good point of departure. It should also be noted that there are at least two correct positions at which the sliding strip can be juxtaposed against the text, since in the enciphering rectangle the letters in one column form digraphs with the letters in the column not only on the right but also on the left. In the absence of any Qs, or other low-frequency letters suitable for a point of departure, the very first 20 or 25 letters of the cryptogram may be used as the starting point, since these letters come from column 1 of the transposition rectangle and therefore there is no uncertainty at least as to the letter which is at the top of that column. Suppose that a good initial juxtaposition has been found for the portion of the text that has been written on the sliding strip, and that a series of excellent digraphs has been brought to light. The next step is, of course, to add to these digraphs on either side by finding sections of text that will yield "good" digraphs and tetragraphs. For example, suppose that the initial juxtaposition has yielded what is shown in figure 14. The digraph PR suggests that it must be followed by a vowel, preferably E, or O, or the digraph AV might be part of the word CALAVER, in which case it will be followed by A, the digraph CR suggests that it might be followed by the vowel A or E, A place is therefore sought in the rest of the text, where there is a sequence of the letter A, or the possibility thus suggested. Suppose such a sequence is found and yields what is shown in figure 15. The skeleton of words are now beginning to appear, suggesting that AV is, in fact, part of the word CAVALRY, there should be an L to follow it, the trigraph T IN suggests the termination Q, the trigraph Z ER suggests the zero. A section of text is therefore sought, which will have the letters L, Q, and O in the order L Q O. Enough has been shown to demonstrate the procedure. In the course of the work it soon becomes evident where these letters of columns are, because the digraphs above and below the nuclear or "good" portion become "bad" quite suddenly, just as soon as letters belonging to nonadjacent columns in the original rectangle are brought together. For example, in figure 15 it is observed that the trigraph at the top, R R S, is highly improbable, as is likely to be the trigraph at the bottom, E O A. This suggests that these letters have been brought together erroneously, that is, that they do not belong in adjacent columns in the enciphering rectangle. If this is true, the "good" portion is composed of the 13 letters between these two trigraphs, and therefore the columns are about 13 letters long. Additional work will soon show exactly how long each column really is, and when this has been ascertained the problem has been practically completed, since at the same time that becomes evident the sequence of columns has also become evident. An example of solution by this alternative method may be helpful. Using the cryptogram of paragraph 13 as an example, figure 16 shows how the text might be transcribed on a sheet of cross-section paper. Noting that the message contains a Q as the 129th letter, a section of text to include the Q is transcribed on a strip of cross-section paper and this strip is juxtaposed against the remaining text to bring the Q in front of a U. How many letters should be included in this strip? The message contains 224 letters, if a width of say 10 to 20 columns is assumed, the columns of the rectangle will be about 12 to 22 letters in length. It will be safer to assume a convenient width and thereby the maximum, consequently a length of 20 letters will be tentatively assumed. Now the Q may be at the top of a column, at the middle, or somewhere in the middle.
letters before a succession of two E's? There is ONE
true, not one of those above the Q is impossible and the same
is true of those below the Q until the digraph VN is reached. Hence,
combination 1 is possible. As for combination 2, this at once appears
to be bad. Digraphs such as I I, and Z H are highly improbable, and
this combination may be discarded with safety. Combination 3 is
possible from the top digraph, O F, to the 12th digraph below the Q,
although the digraph H X looks very bad. However, the X might be a
vowels or a separator, so that this combination cannot be discarded.
Combination 4 looks very improbable, with the one digraph H occurring
but, and other equally bad digraphs showing. Of the four possibilities
then, combinations 2 and 4 are discarded, leaving 1 and 3 for
further study. It is very difficult to choose between these two possi-
blites. All the digraphs in combination 1 down to digraph V N are
possible, many of them are excellent. As for combination 3, all the
digraphs down to V D are also possible and many of them are excellent.
There does not seem to be much use to add the frequency values of
the digraphs (or logarithms thereof) in each combination because it is hard
to know with what digraphs to begin or end, although as a last resort
this could of course be done. However, perhaps it is not essential that
a choice be made at once, possibly further work along the lines now to
be demonstrated will show which combination is correct.
Noting the E K's (in the digraphs K S and K S) among the
combinations before the Q, assume that these E K's are
parts of the digraph C K. Is there a sequence C C in the
text? There is but one such place, at the 86th letter
Suppose the corresponding section as placed in figure 17-B
the combinations 1 and 3 of figure 17, as shown in figure
17-B. It immediately becomes evident that combina-
tion 3 is the correct one, for note the excellent
digraphs it gives, as compared with those in combination
C K. Note that the second trigraph below the E K Q is in
combination 3 consists of 3 E K's, indicating that the
eend of the columns has been reached just before this
digraph. As for the top trigraphs of figure 17-B they are
good all the way up. But now the skeletons of
words are beginning to appear. The T H R immediately
suggests either THREE or THROUGH, the T R O above the T H R suggests FROM or
FRONT. Suppose the word REQUEST is assumed for the E K, and the word
WHERE is assumed for the T H R above it. This requires a section with 2 E K's in
succession.

\[\text{FIGURE 17-B}\]

There are several such places in the text, and further limitation is advisable. The first trigraph from the top is certainly suggestive of the word MOVING, which requires an I to follow the V. There is a place in the text where an I occurs 12 letters before a succession of two E K's. There is one such place, and the corres-
ponding section is juxtaposed at the proper place, yielding what is shown in

\[\text{FIGURE 17-C}\]

The upper and lower limits of the columns are now fairly definite and are
marked by bars, 1 letter is taken at the bottom and A M I E E at the top are
very improbable. The tetragraph C E E S below the top bar is possible, because it may
represent the end of a word like FORCE followed by the beginning of the word ESTIMATED, the
tetragraph above the bottom bar suggests a word ending in E followed by the word IMMEDIATE. It seems hardly necessary to continue with the demonstration, in a few moments the entire
diagram is reconstructed and yields the solution. During this process as soon as a section of
text in figure 16 has been used it is crossed off, so as to prevent its letters from being considered
as further possibilities for addition to the reconstruction diagram. Thus, as the work progresses,
the number of available sections becomes progressively less, and the choice for successive sections
for addition to the diagram becomes a quite easy matter.

\[\text{FIGURE 17-D}\]

If two or three operators are assigned to work upon a cryptogram by this method,
solutions can be reached in a very short space of time, especially if each one of the operators takes a
different point of attack. After a few minutes the fragments of texts obtained may be assimi-
lated into one message which is then completed very speedily.

\[\text{FIGURE 17-E}\]

The additional remarks will be devoted to some remarks of a general nature
concerning columnar transposition of the foregoing type. The degree of cryptography security
afforded by simple columnar transposition methods, especially when incompletely-matched
matrices are employed, is considerably increased if some of the cells of the matrix are occupied by
nulls instead of significant letters. If nulls are employed judiciously their presence serves to
confuse the cryptanalyst by introducing unusual digraphs, trigraphs, and polygraphs which may
lead him to discard correct combinations of columns in the anagramming process and thus
retard solution. Usually, the low frequency of letters such as J, Q, X, or Z as nulls does
not commend itself for this purpose, as such letters would not only distort the normal frequency
distribution and thus give clues to the presence of nulls, but also they would be quickly
"spotted" in the anagramming process.

\[\text{FIGURE 17-F}\]

Another subterfuge, and a good one, to put stumbling blocks in the way of a quick
solution is to leave "blanks" within the transposition matrix, that is, certain cells are left un-
occupied by letters of the text. If only a few cells distributed irregularly within the columns of
the matrix are designated as blanks, the disturbing effect of the anagram-
ning process is quite marked. Thus more or less effectively hinders the cryptanalyst in his
attempts to ascertain the lengths of the columns and considerably increases the difficulty of
the anagramming process.

\[\text{FIGURE 17-G}\]

In order to fix the minimum number of the nulls or of the blanks in the transposition
matrix, definite rearrangements between correspondents are necessary. These may be in
the nature of "forms" outlining the matrix, showing the number of columns and the positions of the
cells to be occupied by nulls, or of the cells to be left vacant in the inscription process, or the
positions of these cells may be derived from the elements of the transposition key itself. If
"forms" are employed, they may be used with varying transposition keys, so that even though
there may be only a few different forms, the use of varying transposition keys serves to
increase cryptographic security to a rather marked degree.

\[\text{FIGURE 17-H}\]

If nulls, or blanks, or both, are distributed irregularly but symmetrically throughout
the transposition matrix (as, for example, blanks are distributed in cross-word puzzles) solution
of single messages produced by simple keyed-columnar transposition from such a matrix
becomes an extremely difficult if not impossible problem. Naturally, if nulls and blanks
are distributed irregularly and asymmetrically the matter becomes hopeless, as far as a single
message is concerned.

\[\text{FIGURE 17-I}\]

Of course, if several messages of identical length and in the same key are available for
superimposition, the presence of the blanks or nulls then makes little difference, because the
general solution to be explained in a subsequent paragraph (par 26) can be applied. Or if messages with similar beginnings or similar endings are available, solution is facilitated here as in the simpler case where nulls or blanks are not employed, as will be explained in subsequent paragraphs (pars 23-24). Considerations of space prevent going into detail in the solution of an example, and the student should undertake a study of these cases for himself.

16 The C-P and the P-C sequences - a Two numerical sequences which constitute the bases for several very important cryptanalytic operations and procedures in the solution of transposition ciphers may be derived from, and are applicable to, most ciphers of this class. They are as follows:

(1) A sequence the successive terms of which indicate the position numbers that the successive letters of the plain text occupy in the cipher text. This sequence will hereafter be designated the plain-cipher sequence, or P-C sequence.

(2) A sequence the successive terms of which indicate the position numbers that the successive letters of the cipher text occupy in the plain text. This sequence will hereafter be designated the cipher-plain sequence, or C-P sequence.

These two sequences are obviously related; one being the inverse or indirect version of the other. Given one of the sequences, the other can be derived from it by the simple operation of reordering, in a normal sequence, the positions occupied by the numbers constituting the sequence on hand. An example will be given presently.

c Note the encipherment shown in Figure 18-A

TRANPOSITION
THE QUICK BROWN
FOX JUMPS OVER
THE LAZY DOG

Table 18-A

Term No | 1 2 3 4 5 6 7 8 9 10 11 12 13
Cipher | X E L B O G O E J A R T C P D W K I M Y H O E U Z K S O T F P H R Y

Now, if, instead of letters, the successive numbers 1, 2, 3, are inscribed in the cells of the matrix, in normal order of writing, the "cipher text" becomes the P-C sequence and is as follows:

12 9 4 14 10 8 6 11 2 13 3 7 5
01 02 03 04 05 06 07 08 09 10 11 12 13
41 14 15 16 17 18 19 20 21 22 23 24 25 26
27 28 29 30 31 32 33 34 35

The student may easily verify that the P-C sequence is what it purports to be by noting that, according to it, the 1st letter of the plain text of the illustrative message, Tm, becomes the 31st letter of the cipher text (since the number 01 occupies position 31 in the P-C sequence shown above), and that in the cryptogram the 31st letter is T, the 2nd letter of the plain text, Hs, becomes the 22nd letter of the cipher text and that in the cryptogram the 22nd letter is H, and so on. In connection with the P-C sequence, it is to be noted that successive terms in the sequence,
The term number is useful merely to facilitate finding and referring to specific terms or numbers in a sequence, whether the latter be a C→P or a P→C sequence. The number simply indicates the locus or position a term occupies in the sequence. In connection with a plaintext message the consecutive term numbers 1, 2, 3, may be used as loci for the successive letters of the message, in connection with a cryptogram the consecutive term numbers 1, 2, 3, may be used as loci for the successive letters of the cipher text.

In single, keyed-columnar transposition an interesting relationship exists between sections of the C→P sequence. Consider the C→P sequence given in subparagraph a above, and note that by adding the integer 1 to the successive numbers thereof, sections of the original sequence show certain identities with sections in C→P sequence +1. Thus

\[ \text{C→P sequence} \ldots \] 
\[ 31 \ 22 \ 01 \ 09 \ 25 \ 19 \ 14 \ 28 \ 04 \ 34 \ 07 \ 17 \ 12 \ 32 \ 23 \ 01 \ 02 \ 26 \ 20 \ 15 \ 29 \ 05 \ 35 \ 08 \ 18 \ 13 \ 33 \ 24 \ 03 \ 11 \ 27 \ 21 \ 16 \ 30 \ 06 \ 36 \ 09 \ 19 \ 14 \ 34 \ 25 \ 04 \ 12 \ 28 \ 22 \ 17 \ 31 \ 07 \]

In fact, if the successive numbers of the C→P sequence are set down in rows to produce sequential numbers in columns, the following interesting diagram is obtained

\[ \text{Diagram 10-2} \]

Reference to figure 18-C will show the identity of this diagram with that figure. Such an arrangement of course indicates at once the number of columns in the transposition rectangle, from which it follows that if the C→P sequence is available it is an easy matter to establish the outlines of the transposition matrix. The phenomena dealt with in this subparagraph are but a reflection of those discussed in subparagraph e above.

The phenomena just indicated may, however, be employed to advantage in another manner in the solution of an unknown example. Referring to the illustrative cryptogram in subparagraph c above, suppose that the cryptanalyst has reason to suspect the presence of the probable word "QUICK." The letters necessary to produce this word (and their numbers in the cryptogram) are as follows

\[ \text{QUICK} \]

The sequence 9, 25, 19, 14, 28 now constitutes a portion of the C→P sequence. Adding the integer 1 successively to these C→P numbers, let the corresponding letters be set down alongside the numbers. Thus

<table>
<thead>
<tr>
<th>Base</th>
<th>Derivative 1</th>
<th>Derivative 2</th>
<th>Derivative 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>25</td>
<td>20</td>
<td>27</td>
<td>28</td>
</tr>
<tr>
<td>19</td>
<td>20</td>
<td>21</td>
<td>22</td>
</tr>
<tr>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
</tr>
<tr>
<td>28</td>
<td>29</td>
<td>30</td>
<td>31</td>
</tr>
<tr>
<td>18</td>
<td>23</td>
<td>33</td>
<td>24</td>
</tr>
<tr>
<td>24</td>
<td>27</td>
<td>34</td>
<td>25</td>
</tr>
<tr>
<td>16</td>
<td>28</td>
<td>35</td>
<td>26</td>
</tr>
<tr>
<td>11</td>
<td>31</td>
<td>41</td>
<td>27</td>
</tr>
<tr>
<td>12</td>
<td>32</td>
<td>42</td>
<td>28</td>
</tr>
<tr>
<td>30</td>
<td>33</td>
<td>43</td>
<td>29</td>
</tr>
<tr>
<td>06</td>
<td>34</td>
<td>44</td>
<td>30</td>
</tr>
<tr>
<td>09</td>
<td>35</td>
<td>45</td>
<td>31</td>
</tr>
<tr>
<td>21</td>
<td>36</td>
<td>46</td>
<td>32</td>
</tr>
<tr>
<td>14</td>
<td>37</td>
<td>47</td>
<td>33</td>
</tr>
<tr>
<td>17</td>
<td>38</td>
<td>48</td>
<td>34</td>
</tr>
<tr>
<td>31</td>
<td>39</td>
<td>49</td>
<td>35</td>
</tr>
<tr>
<td>07</td>
<td>40</td>
<td>50</td>
<td>36</td>
</tr>
</tbody>
</table>

Here it will be seen that portions of "good" plain text become manifest, viz., JUMPS and AZTDO. The 3rd derivative no longer is "good" because the rectangle has but 3 rows and consequently only the 1st and 3rd derivatives from the "base" are valid. It is obvious that the foregoing method of deriving plain-text sections from a correct probable word offers considerable possibilities as a cryptanalytic tool, especially in the case of matrices with more than 2 or 3 rows if sections of text can be reconstructed in this manner and then combined in proper sequence the reconstruction of the complete matrix and the transposition key is a relatively simple matter. The application of the foregoing principle to the solution of unknown examples is, of course, obvious.

There is also an interesting relationship between the sections of the P→C sequence for a cryptogram, though it is somewhat different from that discussed in subparagraphs b and c, the case of the C→P sequence. Consider the P→C sequence set forth in subparagraph d above and note how, by adding the integer 1 to the successive numbers, sections of the P→C sequence become identical with sections of the P→C sequence +1. Thus

\[ \text{P→C sequence} \ldots \] 
\[ 03 \ 16 \ 28 \ 09 \ 25 \ 35 \ 11 \ 24 \ 04 \ 17 \ 20 \ 13 \ 26 \ 07 \ 30 \ 12 \ 26 \]

The equivalence between identities, as indicated above, indicate not only that the encrypting matrix has 15 columns, but also they may be used to establish the actual transposition key or at least a cyclic permutation of the key, by constructing a chain of equivalents. Thus

\[ 1=4, 4=10, 10=8, 8=6, 6=11, 11=2, 2=13, 13=5, 5=7, 7=5, 5=12, 12=9, 9=1 \]

This yields, by eliminating the term common to successive equivalents, the following chain or transposition key

\[ 1 \ 4 \ 10 \ 8 \ 6 \ 11 \ 2 \ 13 \ 3 \ 7 \ 5 \ 12 \ 9 \]

Reference to figure 18-A will show that the foregoing key is a cyclic permutation of the actual key.

There remain only some minor remarks which, being of a general nature, arise from the mechanics of simple keyed-columnar transposition, are worth noting. They are discussed in the subsequent two subparagraphs.

An appreciation of the difficulties introduced by employing only incompletely-filled rectangles indicates that it would be very useful if there were some method whereby in the initial stages of solution the cipher text could be divided up correctly into its component long and short columns, for the subsequent steps of rearranging the columns by the anagramming principle are quite simple. If, for example, there were some feature which provided a means of distinguishing when in encipherment a transit was made from the bottom of one column to the top of the next column, then the location of these transition points or "breaks" would obviously permit of breaking up the cipher text into its correct long and short columns. In later study cases of this kind will be encountered.

It is useful sometimes to be able to ascertain just where breaks cannot occur, that is, where a passage from the bottom of one column to the top of the next one cannot occur in the cipher text, for this will limit the field for experiment. A consideration of the mechanics of the system will afford an excellent clue to the fact that this determination is easy to make. In any transposition rectangle involving simple keyed-columnar transposition the interval, in the cipher text, between two consecutive letters which are in the same row in the matrix is the sum of...
multiple of the length of the short columns and a multiple of the length of the long columns. For example, consider the adjacent letters CK in the plain-text rectangle shown in figure 18-A. In the cipher, C is the 14th letter, K is the 28th and the interval is 28—14=14. The message has a total of 35 letters and the matrix has 13 columns, 9 long ones of 3 letters and 4 short ones of 2 letters. An interval of 14 can therefore be brought about in only one way (4x3)+(1x2)=14, which means that 4 long columns and 1 short one intervene between the C and the K in the plain text, and that the key numbers applicable to the two columns are 5 apart in sequence, that is, if the column in which C is located has key number 1, the column next to it on the right is 6, or if the former is 2 the latter is 7, and so on. Reference to figure 18-A will show that these deductions are correct and that the key numbers involved are 6—11. However, a more general treatment is possible. Given a cryptogram of 26 letters and an assumed width of 6 columns, for example, the matrix can have only 2 columns of 5 letters and 4 columns of 4 letters. Setting down the multiples of the two lengths in tabular form, for convenience, the following is obtained.

| multiples | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 |
|-----------|---|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| 0 | multiple | 0 | 0 | 0 | | | | | | | | | | | | | | | | | | | | | | | | |
| 1 | multiple | 0 | 1 | 2 | 3 | 4 | | | | | | | | | | | | | | | | | | | | | | | |
| 2 | multiple | 0 | 0 | 4 | 8 | 12 | 16 | | | | | | | | | | | | | | | | | | | | | | |
| 3 | multiple | 0 | 5 | 9 | 13 | 17 | 21 | | | | | | | | | | | | | | | | | | | | | | |
| 4 | multiple | 0 | 10 | 14 | 18 | 22 | 26 | | | | | | | | | | | | | | | | | | | | | | |

All the possible positions of breaks in the cipher text, that is, transits from the bottom of one column to the top of the next column, may now readily be ascertained by finding the totals resulting from making all the possible combinations of the indicated multiples taken in pairs. It is convenient to draw up a table to show directly the sum of the combinations. Thus:

<table>
<thead>
<tr>
<th>No of short columns</th>
<th>Length in letters</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0 0</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
</tr>
</tbody>
</table>

Hence, the astute cryptanalyst will at once be able to make a study of the frequency of occurrence of successive letters (such as 8 and 12, 14 and 18, etc.) and determine the length of column, thus narrowing the possibilities of the number of columns. As before, he must be guided by the probabilities of occurrence of letters and words in the language and the characteristics of the cipher text. In this general way, it is possible to reduce the possible number of columns to a more manageable number. But the key to the cipher is not yet known, and the cryptanalyst must now consider another group of deductions.

Section IV

OPPORTUNITIES AFFORDED BY STUDYING ERRORS AND BLUNDERS MADE BY ENEMY CRYPTOGRAPHERS

17 Importance of the study of errors and blunders in early work upon an unknown system—Errors and blunders made by cryptographic clerks in the execution of cryptographic instructions should be rare in a well-trained and well-disciplined cryptographic service. Nevertheless, blunders and mistakes are committed despite all that can be done to prevent their occurrence. Especially in the excitement prior to or during an important action or movement do such mistakes take place and these afford golden opportunities for the enemy cryptanalytic service. This situation exists in respect to all types of cryptographic systems. No cryptanalytic instruction would be complete if cognizance were not taken of the advantages which may be reaped from the blunders, the mistakes, and, occasionally, the downright stupidity of the adversary’s cryptographic personnel.

1. Practically every cryptographic system affords opportunities for the commission of errors in its application, and each system more or less presents a separate case. That is, the errors which may be made in one type of cryptographic system may be peculiar to that type alone and to no other type, hence, the astute cryptanalyst is constantly on the lookout for instances of cryptograms containing the specific type of error by which that system is handicapped. Furthermore, the general types of blunders or errors that may be committed are nearly as numerous as are the general types of cryptographic systems, so that no complete list of such as may be encountered in practice can be drawn up.

2. After the cryptanalyst has by painstaking and more or less arduous labor solved a system and has become thoroughly familiar with its mechanics, he should carefully review the details of the mechanics to learn what things can go wrong, what sort of mistakes the enemy cryptographic personnel are likely to make, and then study the external manifestations of these aberrations so that he may be able to recognize instances of their occurrence in subsequent cryptograms. This sort of study has no value in itself; particularly, its importance lies in the fact that the effects of erroneous treatment may lead to very rapid solution or to quick recovery of keys to subsequent messages.

3. When an unknown system is under investigation and the cryptanalyst is striving to ascertain just how it operates (which is often the most difficult step in solution), a study of the cryptograms representing corrections to previous messages containing errors is a most fruitful source of data. Indeed, at times this sort of intensive study will yield clues for solving a system which might otherwise seem all efforts to break it down for a very long time.

18 Significance of terms "special solution" and "general solution"—Now the importance of the comments made in the foregoing paragraph will be clear if it is noted that a study of the blunders and errors often leads to the elaboration of methods for the rapid breaking down of cryptographic systems. But it must also be realized that in some cases no blunders or errors are essential to a rapid solution of the type alluded to above. Sometimes the very mechanics of
the system itself are such that unavoidable or unpredictable circumstances arise so that special solution become possible. The latter term calls for a bit of explanation.

b When the circumstances surrounding a specific cryptogram or set of cryptograms are such as to present peculiar or unusual conditions that make a solution possible when in the absence of these conditions solution as either impossible or unpredictable, the methods employed in reaching a solution in such cases constitute what is commonly termed a special solution. Some examples will be demonstrated very soon. Systems of this kind may be true in general terms and examples of these may be expected to make evident in practice. Furthermore, in order to obtain solution in the case of unusual conditions necessary for the application of a special solution, in other words, a general solution in cryptanalysis implies a method or procedure which in applied in ordinary cases and under normal conditions will yield the solution. However, the term general solution in cryptanalysis must not be taken too literally. The situation in cryptanalysis is not exactly analogous to that which obtains in the field of pure mathematics, for the circumstances are often quite different in the two sciences. A general solution in mathematics is expected to, and will, solve every case that falls within its province, a general solution in cryptanalysis is likewise intended to solve every case that falls within its province but this usually requires more of the nature of a prayer or hope than an expectation. Much depends upon the amount of traffic available for study, the length of individual cryptograms, and the indefinable element called luck, that is, a set of fortuitous circumstances which happen to make a solution easy or difficult, such as the presence of many or exceptionally long repetitions, etc.

Furthermore, whereas in mathematics a general solution prescribes the exact steps to be followed in arriving at the solution and is applicable in all instances without variation or deviation from a fixed procedure, in cryptanalysis a general solution merely outlines a broad path that may be followed in order to arrive at a solution. Application of a general solution in cryptanalysis in specific instances may involve minor detours to circumvent unexpected obstacles, or may involve quite large changes or modifications in the general procedure.

19 Examples to be studied — As stated above in paragraph 17, a complete list of the specific blunders that cryptographic clerks are prone to perpetrate cannot be drawn up. Certain of them may be described in general terms and examples given of some which have already been encountered in this and in preceding texts. Commonly it is the case that these blunders do not become evident until two or more cryptograms are available for comparison. One of the most frequent sources of circumstance leading to the transmission of cryptograms affording rich material for cryptanalytic comparison is the following: A cryptographic clerk prepares a cryptogram, in the course of which he makes a mistake of such a nature as to render the cryptogram difficult or impossible to decipher by the cryptographic clerk serving the addressee. A request for repetition ensues, whereupon the enunciating clerk reexamines his original work and finds that he has made a mistake. He then omits the grave blunder of reenunciating the identical message (without paraphrasing) and transmitting what to the enemy cryptanalyst is obviously a second version of the original message. The consequences are often fatal to cryptographic security. The least that can happen is that the key for this particular message may be disclosed very quickly, more serious, the basic or primary elements for the entire day's traffic may be wrested from the blunder, but most serious are the consequences if it happens that the blunder has been committed immediately or soon after a new cryptographic system has been instituted and the enemy cryptanalyst is exerting strenuous efforts to learn its mechanics, for then is when the information to be gained is most valuable.

b In his previous studies, the student has observed many opportunities for quick cryptanalytic success afforded by enemy addiction to the use of stereotypic phraseology, especially at the beginnings and endings of messages. Stereotype phraseology affords even more golden opportunities for cryptanalytic success in the case of transposition systems than it does in the case of substitution systems. Only a few special examples of the consequences of cryptographic blunders and mistakes in the case of transposition systems will be studied. These are intended to give the student some ideas of the far-reaching effects such studies may have. It is important that he grasp the fundamental principles they will enable him to develop for himself the methods that he may find necessary in practical work. Incidentally, it may be added that the student should not get the idea that these instances are purely theoretical. It is sometimes almost unbelievable that cryptographic clerks with any common sense would perpetrate the stupid blunders that they do occasionally commit.
The first group of the cryptogram begins with I M. The arrangement shown above gives I U as the top of a column, hence a key length of 10 is not correct. A key length of 11 is then tried.
The same general principles, modified to suit the circumstances, may be followed in the case involving known or suspected endings of messages. The probable words are written out according to various assumed key lengths and the superimposed letters falling at the bottoms of columns are sought in the cryptogram.

21 The case of an omitted column—Sometimes a very careless clerk omits a column in transcribing the text from the enciphering rectangle and fails to check the number of letters in the final cryptogram. Obvioulsy such a cryptogram will be difficult if not impossible to decipher at the other end, and a repetition as requested and sent. If now the identical plain text is enciphered correctly, two cryptograms are at hand for comparison. This will disclose the length of one column, which can be assumed to be either a long one or a short one. The position, in the correct cryptogram, of the column omitted from the incorrect one will often afford direct clues as to the exact dimensions of the enciphering rectangle. For example, suppose the cryptogram in paragraph 208 had first been transmitted as follows:

![Cryptogram](image)

The column which was omitted was N T T S P S D, and falls between columns 1 and 3. Since the omitted column contains 11 letters and column 1 contains 10, the dimensions of the rectangle immediately become known. Thus, uncertainties as to the dimensions of the rectangle are dissolved and a large step forward has been made in the solution. Also, the general whereabouts of columns 1 and 2 are now known, since the former is a short one, the latter a long one.

22 The case of an interchanged pair of columns—a The keying element in the case of columnar transposition is simply a practical means of controlling the order in which the columns of the enciphering rectangle are transcribed in forming the cipher text. Commonly this numeral key is derived from a literal key. Suppose that a cryptographic clerk makes a mistake in the latter step. For example, suppose that the literal key is ADMIRATION and that as a result of a slight relaxation in attention he assigns the number 5 to the letter N and the number 6 to the letter M. A pair of columns will become interchanged as regards their order of selection in the transposition process, and likely as not a repetition will be requested by the addresser. If a second version is sent, enciphered by the correct key, a comparison of the two versions will disclose the width of the enciphering rectangle and possibly the general position (left or right) of the columns that were interchanged.

An example will serve to make the matter clear. Assume the two cryptograms to be as follows:

**First Version**

```
ODN1L NTTHD GS0HA O0QSG TERNF
URNFR RPE NERNI TEPD ETAN 000CO ROGIO S
```

**Second Version**

```
ODN1L NTTHD GS0HA O0QSG TERNF
URNFR RPE NERNI TEPD ETAN 000CO ROGIO S
```

c. The two cryptograms are superimposed as shown in figure 24 and their points of similarity and difference noted.

First version... ODN1LNTTHDGS0HAO0QSGTER[P]SINENN E
Second version... ODN1LNTTHDGS0HAO0QSGTER[NFUEHWR]
[NFUEHWR]IRATPEDETAN000CO ROGIO S
R[P]SINENN EIRATPEDETAN000CO ROGIO S

**Figure 24**

The two versions are alike except for a pair of interchanged sequences, the bracketed sequence PSINE in the first version is matched by the same sequence in the second version, but at a different position in the message, likewise the bracketed sequence NFUEHR in the first version is matched by a similar sequence in the second version, but at a different position in the message. The various deductions which can be made from the situation will now be set forth.

One of these sequences contains 7 letters, the other contains 8. It follows that the columns of the enciphering rectangle are probably 7 and 8 letters in length, hence, with 61 letters, the width of the matrix is 8. Since there are 23 letters from the beginning of the messages to the first point of their difference, it follows that there are 2 columns of 8 letters and 1 column of 7 letters involved in this section $[(2\times 8)+1\times 7]=23$, and that the error made in encipherment does not involve columns 1, 2, or 3, which are therefore properly placed in the first version.

Since...
the sequences which are interchanged are consecutive in the text it means that the numbers 4
and 5 were interchanged in the key for the first version. Since one of these sequences is of 7
letters, the other of 8 letters, one of the numbers, 4 or 5, applies to a long column, the other, to
a short column. Since the second version is presumably the correct version, and since in the
second version the 8-letter sequence comes first, the key number 4 applies to a long column, the
key number 5 to a short column in the correct version. With the foregoing deductions in mind,
the solution and the reconstruction of the numerical key becomes a simple matter.

The text of the correct version is written out as seen in figure 25e. Seeing a Q in column 3
and a U in column 4, these two columns are made adjacent by sliding column 3 one interval
downward, as shown in figure 25f. In the latter, column 7 has also been placed to the second interval
to the right of column 5, because the latter yields good trigrams with columns 3–4. Seeing the
trigram T R O near the bottom of columns 3–4, the letters 0 and P in the same row, suggests
the word TROOP. The columns are to be re arranged to make this word TROOP! There are

\[ \begin{array}{cccc}
1 & 2 & 3 & 4 \\
5 & 6 & 7 & 8 \\
1 & 2 & 3 & 4 \\
5 & 6 & 7 & 8 \\
\end{array} \]

\[ \begin{array}{cccc}
O & N & E & T \\
O & N & E & T \\
O & N & E & T \\
O & N & E & T \\
\end{array} \]

\[ \begin{array}{cccc}
E & T & R & O \\
E & T & R & O \\
E & T & R & O \\
E & T & R & O \\
\end{array} \]

\[ \begin{array}{cccc}
T & R & E & P \\
T & R & E & P \\
T & R & E & P \\
T & R & E & P \\
\end{array} \]

\[ \begin{array}{cccc}
L & O & T & H \\
L & O & T & H \\
L & O & T & H \\
L & O & T & H \\
\end{array} \]

\[ \begin{array}{cccc}
T & W & E & C \\
T & W & E & C \\
T & W & E & C \\
T & W & E & C \\
\end{array} \]

two columns which have an U in the proper row, columns 2 and 8. The trial of combination
3–4–5–6–7, while producing TROOP in the proper row, gives bad pentagrams in the other rows,
but the combination 3–4–5–6–2 shows excellent pentagrams, as will be seen in figure 25f. The
words SQUADRON and HOSTILE are clearly evident, the completion of the rectangle as now a
very simple matter. The result is shown in figure 25f. The recovery of the numerical key now
will enable other cryptograms to be read directly.

25 Messages with similar beginnings—d In military correspondence it is often the case
that somewhat similar instructions or information must be conveyed to a superior commander
to several subordinate commanders simultaneously. Such a situation frequently results in the
situation that two or more cryptograms addressed to different stations will begin with exactly
the same words. When simple columnar transposition is the system used for enciphermen,
then it will result, in such cases as the foregoing, that the first two or more rows of the transposi-
tion rectangle will be identical in the messages which begin alike. Therefore, the cryptograms
will show identical sequences of two or more letters, distributed throughout the texts and by
indicating these identities the cryptanalyst is able at once not only to ascertain the width of the
rectangle but also to divide up the cipher text into sections corresponding with the exact columns
of the rectangle, thus eliminating the only real difficulty in solution, viz., the determination of
which are the long columns, which the short. An example will demonstrate the short cut to
solution which such a situation provides.

\[ \text{Here are two cryptograms which are assumed to have been interspersed within a few minutes}
of each other, the messages being addressed to two battalion commanders by the regimental}
commander.}
and so on. It appears that the identical portion embraces the first four rows of the rectangle and runs over a number of letters on the fifth row. This is because the identical sequences consist of 4 and 5 letters. Figure 27b shows the identities between the first 5 columns of the two transposition rectangles. Only once in the case of this particular example does any uncertainty arise as to exactly where an identical sequence begins or ends, and that is in connection with the seventh pair of identities, involving the series of letters AEPTSNPGNP in No. 1, and AEPTSTONTN in No. 2. These sequences contain 6 identical letters, but even here the uncertainty is of only a moment's duration. The initial letter A does not belong to the identical portion at the top of the transposition rectangle because the A's are needed to complete columns 6 in both rectangles. If the A were placed at the head of column 7 in No. 1, then column 6 would lack a letter at the bottom. Cases of "accidental identities" of course complicate the process of cutting up the text into the respective columns, but they only serve to add a small degree of interest to what would otherwise be a purely cut and dried process.

The final results of the transcription into columns are shown in figure 27c.

It is obvious from a comparison of these two skeletonized matrices, and a consideration of the fact that the long columns must of necessity go to the left side, that the numbers 7 and 10 occupy the first two positions in the key, and that the numbers 2, 4, 11, and 13 occupy the last four positions in the key. By segregating and anagramming columns 7 and 10 as one group, and columns 2, 4, 11, and 13 as another group, the exact positions occupied by these columns are easily ascertained, as shown in figure 27d.

The remaining columns 1, 3, 5, 6, 8, 9, 10, 12, and 14 form a third group of columns to be anagrammed, but this is rather easy now that the columns on either side are fixed. The completed rectangles are shown in figure 27d.

46 Messages with similar endings — a. What has been said at the beginning of the preceding paragraph with respect to the nature of military correspondence and the presence of identical phraseology in the messages sent by a superior commander to his subordinates also operates to produce messages in which the endings are identical. It has been noted that when two messages with similar beginnings are available for comparison, the reconstruction of the transposition rectangles and the recovery of the transposition key is an easy matter. It will now be shown that solution is an even easier matter when two messages having identical endings are available for study.

b. Given the following two cryptograms:

No 1

1

2

EN LION LION
FO COUN COUN
TE ESTO ESTO
FW TTIN TTIN
SX TION SX STES
NE NBDE TP ATET
PR BRIN OV FFLI
GF TREGN NT POSI
NA TROO TI LCOU
PS TD NT

No 2

1

2

ENEMYBATTALION ENEMYBATTALION
FORMINGFORCOUN FORMINGFORCOUN
TERATTACKWESTO TERATTACKWESTO
FWOODSATMOTTIN FWOODSATMOTTIN
SXTAKEPOSITIONSXMOWEHASTKES
NEARLANTZANDBE TPOSSIBLETET
PREDIETORBRIN OVICINITYOFILI
GFLANKINGTREO NTSANDTAKEPOSI
NATTACKINGTROO TIONTEREPLCOU
PS TD NT

Figure 27b

Figure 27c

Figure 27d
The cryptanalyst now carefully compares the two texts, searching for identical sequences of letters, but in this case instead of trying to locate identities in what may be termed a parallel progression (as in the preceding case) he searches for identical sequences of two or more letters appearing in both messages. For example, in the present case, he notes the sequence "TR" forming the initial tri-graph of the 8th group of No 1 and finds a similar sequence forming the initial tri-graph of the 13th group of No 2. Going through both cryptograms in this way, all the identities are marked off in some fashion, by colored crayon or underlining, as shown below. In this search for identities the cryptanalyst bears in mind that when all have been found they should be distributed at quite regular intervals throughout the text. For example, note in the following that the identities in No 1 fall at intervals of 6 letters, with one exception, in No 2 they fall at intervals of 4 letters, with one exception. The intervals between identities serve as a guide in finding them. After they have all been located, the identities in the cryptograms are numbered serially.

No 1

ETRT ERSOA AEUNIYAFLN IAMND RTHV MERNI
EETROUDCCOHTCYMRREAHRTN DREYENRNER
SRBNIGSKAI LRNA NFNAD ALOLT XOMAH HRREI

No 2

TLVYXONPNE MFPDSK YEN UEEBT SREHTIANT
LVYMRVSEIR EENEI NOLMNED TROFPUNARA
CIAAI NSCNNA

e. The identical sequences in the two cryptograms may now be equated, using for this purpose the numbers below the identities. For instance, identity 1 in cryptogram 1 matches identity 7 in cryptogram 2, identity 2 in cryptogram 1 matches identity 6 in cryptogram 2, and so on. Thus

Cryptogram 1: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13

Cryptogram 2: 7, 6, 9, 2, 10, 5, 11, 3, 4, 12, 13, 1, 8

d. Now cryptogram 1 has 105 letters, since the key consists of 13 numbers (indicated by the 13 identities), the rectangle for cryptogram 1 contains 12 columns of 8 letters and 1 column of 9 letters. Cryptogram 2 has 81 letters, and its rectangle contains 10 columns of 6 letters and 3 columns of 7 letters. The rectangle of cryptogram 1 has but 1 long column, whereas that of cryptogram 2 has 3 long columns. Relative to the present the last letter in each rectangle occupies in the last row of the rectangle, it is obvious that the last letter of the rectangle for cryptogram 2 is 2 letters in advance of the last letter of the rectangle for cryptogram 1. Using this difference, ex 2, a cyclic sequence is generated from the series of equivalences given above. Thus, the equivalent of identity 1 of cryptogram 1 is identity 7 of cryptogram 2, and the number 7 is placed two intervals to the right of the number 1, the equivalent of identity 7 of cryptogram 1 is identity 11 of cryptogram 2, and the number 11 is placed two intervals to the right of number 7, and so on until the following sequence is obtained.

1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13

1, 7, 11, 13, 8, 3, 9

e. The equivalent of identity 9 of cryptogram 1 is identity 4 of cryptogram 2, and the number 4 is placed between the numbers 1 and 7 in this sequence, for the sequence may be regarded as partaking of the nature of a cycle or a continuous series. From this point on, the process is the same as before, and finally the following is obtained.

1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13

1, 4, 7, 2, 11, 6, 13, 5, 8, 10, 3, 12, 9

f. After little experiment it becomes obvious that column 8 belongs on the extreme left, because in cryptogram 1 there is only one long column, number 8, ascertainable by counting the number of letters between successive identities in that message. The number 8 being at the extreme left the final actual transposition key is 8 10 12 9 1 4 7 2 11 6 3 5. The completely deciphered messages are shown in Figure 28.

No 1

HEADREDCOLUMN INFANTRYPOINT
NORTHREACHEDSILVERUNCR
PRESUMEDKATSFENTWE
YELLOWRUNCRE NTYAXKREMAINH
EKATSENFORT EREINOBSERVAT
YAXKREMAIHER ION
EINOBSERVATION

No 2

HEAEDREDCOLUMN INFANTRYPOINT
INFANTRYANDAR REDCOLUMN PASS
TILLERYMARCHIE DSILVERUNCR
NGHORTHREACHE EKATSENFTE
DSILVERUNCRE NTYAXKREMAINH
JEKATSENFORT EREINOBSERVAT
YAXKREMAIHER ION
EINOBSERVATION

Figure 28

g. The possibility of the rapid solution of columnar transposition ciphers by means of the method of similar beginnings and endings, constitutes one of the most serious drawbacks to the use of transposition ciphers in military cryptography, because it is almost impossible to avoid such cases where many messages must be sent in the same key each day.

26. Solution of a single message containing a long repetition — Sometimes a lengthy phrase or a series of letters (spelled out in letters) is repeated within a message and if the message is composed by a transposition rectangle of such narrow width (in comparison with the length of the repetition) that the repeated portion forms identical sequences within the text of the cryptogram, a solution somewhat similar in principle to that explained in paragraph 24 may be achieved within a few minutes.

b. Note the following cryptogram, in which identical portions have been underlined.

Cryptogram (109 letters)

OEAE L T RSE DHUFF NRYF NTAE DILSMY
NCETS LETOC AWIAO TISS LEBHN ORIIS
FEBNH UPWE SSMY RCONNO RTTA OGNUG
GTIF RSEOM SWERN RAST OGEAA OSNOO
IBOSDCA YH HLONEM SETFY HLAUX TAOOG
PRSVL YZEBGG TISSO UHPV

Figure 28
There are 18 segments of underlined letters, which means in this case that the rectangle is 9 columns wide, because the repeated portion in the text will give rise to two repeated sequences in each column. This means that the rectangle has 7 columns of 18 letters and 2 columns of 18 letters. The first two segments may therefore be assigned the numbers 19 and 15, since they come from columns 1, the next two segments may be assigned the numbers 26 and 26, since they come from column 2, and so on, as shown above. Identical segments may now be equated thus:

1 2 3 4 5 6 7 8 9 9a 9b 9c 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100

This gives rise to the cycle 1-3-2-4-9-5-7-6-8, which is a cyclical permutation of the actual transposition key.

d) By transcribing the text into a rectangle of proper width, “cutting” the columns so as to bring the identical portions within the same rows, the result shown in figure 29 is obtained.

\[\text{OPTREAO}\]
\[\text{EONICRANR}\]
\[\text{ATCINSOES}\]
\[\text{AASNESMY}\]
\[\text{LEWFOONSIL}\]
\[\text{TDIERMOEY}\]
\[\text{RTABYSONTE}\]
\[\text{SLOONTWIFIE}\]
\[\text{KSTNTEBYG}\]
\[\text{DMSSTRAOHR}\]
\[\text{HLUGONLLT}\]
\[\text{NNSPGRDAI}\]
\[\text{UCSWMACUS}\]
\[\text{FLESAXSUL}\]
\[\text{FTESGTYTO}\]
\[\text{RSDSGSBAU}\]
\[\text{NLHMOLLLOU}\]
\[\text{RSTYSHGOF}\]
\[\text{YOEFAGCV}\]

\[\text{REPORTOFA}\]
\[\text{IRRECONNA}\]
\[\text{ISSANCETO}\]
\[\text{SEVENAMAS}\]
\[\text{FOLLOWSEN}\]
\[\text{EMYTRIEDO}\]
\[\text{BRSTRAVTIO}\]
\[\text{NEWESTOFLY}\]
\[\text{NEGGETYYSB}\]
\[\text{URGDASHMO}\]
\[\text{UNTOLLYS}\]
\[\text{PRINGSAND}\]
\[\text{WASUNSUCC}\]
\[\text{ESPSULXAS}\]
\[\text{STOFGETTTY}\]
\[\text{SBBURGDASH}\]
\[\text{MOUNTHOOLL}\]
\[\text{YSPRINGSH}\]
\[\text{EAVYPFOG}\]

Study of figure 29 shows that columns 2 and 7 are the short columns and belong on the right, either in the sequence 2-7-2 or 7-2. The cyclical permutation of the transposition key obtained in subparagraph 3c in 1-3-2-4-9-5-7-6. In order to bring the 2 and 7 adjacent in a sequence 2-7 or 7-2 one must take intervals of 6 and 4, respectively, and “decrease” the cycle, giving the following: 1-5-3-8-2-7-4-9-6 or 1-9-6-4-7-2-8-3-5. Since columns 2 and 7 belong on the right, the key must be 4-6-9-1-5-3-8-2-7-8 or 8-3-1-9-4-6-7-2. Only a few moments are necessary to establish the correctness of the former alternative and the solution is at hand. It is shown in figure 30.

A good understanding of the principles elucidated in this and the preceding paragraph will enable the student to derive for himself the procedure applicable to cases of somewhat similar nature, such as that wherein a single letter or a whole group has been omitted from the first version of a message and a second (correction message) is sent without paraprasing the original text, or wherein two messages are alike except for a difference in a single word (such as a number) and are cryptographed by identical transposition keys, or that wherein the numerical key has been incorrectly derived from the literal key and two versions of the same plain text are available for comparison, one based on a transposition by means of the incorrect key, the second based on a transposition by means of the correct key, both keys, however, being of the same length.

Solution when several cryptograms of identical length and in the same key are available: c) Although the method to be described in this paragraph is included within the category of special solutions, it is of such general applicability that it might well be treated as a general solution for all transposition systems. It is based upon the very mechanism of transposition as a cryptographic scheme, viz., that the essential feature of the transposition method is the alteration in the positions of the elements (letters, groups of letters, or words) composing the plain text, according to a specific key. It follows, therefore, that the respective elements of two or more messages of identical length, when transposed according to the same key, will undergo identical alterations in position in the course of encryption, and therefore all plain-text elements occupying homologous positions in the original messages will emerge in homologous positions in the cryptograms. The situation is very much like that which may be observed in the movements executed by two symmetrical groups of dancers in a chorus. Suppose each group consists of 8 dancers starting originally in definite positions relative to one another. When a movement is executed each dancer in each group performs certain evolutions, at the conclusion of the movement the 8 dancers in each group may be in quite different positions relative to one another than they were at the beginning of the movement, but the correspondingly numbered dancers in both groups find themselves in identical positions relative to their neighbors. Of course, the fact that in this analogy the groups are based upon 8's as of no significance, if the groups consisted of many more the principle would still apply. Another way of looking at the matter is to call attention to the fact that in any type of transposition the position which a specified letter or element of the plain text will occupy in the final cryptogram is quite definitely a function of the number of letters or elements in the plain text itself. For example, suppose that there is a transposition system which contains exactly 100 letters, and suppose that the transposition system and specific key is such that the 1st plain-text letter appears as the 17th cipher-text letter, the 2nd plain-text letter as the 65th, and so on. Another message of exactly 100 letters, enciphered by the same general system and specific key, is obvious that the 1st plain-text letter must also appear as the 17th cipher-text letter, the 2nd plain-text letter as the 65th, and so on. In short, all correspondingly numbered plain-text letters in both messages will appear in identical positions in the cryptograms.

Granting the obvious truth of the foregoing, to what use can it be put in the solution of transposition ciphers? Simply this: It enables the cryptanalyst to reconstruct the plain texts of cryptograms of identical length without even knowing what the transposition key or system was that produced them.

The process is not at all complicated and if there are several messages the process is very easy. It consists in superposing the several ciphers and anagramming the columns formed by the superposition, for it is obvious that any circumstances which can be used as a guide for rearranging the letters i one of the lines of superimposed text in order to form plain text will require, and can be checked by, the results of an identical rearrangement of the corresponding letters of the other lines of superimposed text.

An example of the method involving the application of this general solution will now be given, using as an example five messages assumed to have been encrypted by an unknown but complex type of transposition. It will now be shown how the security of such a system is demolished when it is used by a large number of intercommunicating commands.
... Let the following be five cryptograms isolated from among many messages intercepted on the same day and therefore suspected of being in the same key. These five cryptograms have been isolated because they contain exactly the same number of letters. They are here shown superimposed (fig. 31) and therefore all the letters in one column have undergone exactly the same evolutions or changes in position in the course of encipherment.

Column No. 1...I A A L N E O S G T O G Y V E N D R M O N L D U O D E E I H I S A T
Column No. 2...T D N M R G K O N A R I E U S E T U N Y I T C O F E A I E I U T T A
Column No. 5...R A E M T M I O N I O D U M A L L Z N X O A T C G N N A I B T N
Column No. 4...I H N L L N R F V W L R E M K A I E A

* * *

### Noting a Q in message 1 column 51, the obligatory sequence Q U is assumed to be present in that message. There is one message 1 but one Q, which is fortunate. Combining columns 51 and 23, the results are found to be fair (fig. 32a). The H T in the third row suggests a word ending in Q H T, such as FIGHT, MIGHT, EIGHT, etc. Searching in message 3 for a Q, two candidates are found columns 10 and 30. The trigrams yielded by each combination are shown in figure 32b. The second of the two possibilities looks much the better. The trigram in the

<table>
<thead>
<tr>
<th>U</th>
<th>H</th>
<th>E</th>
<th>W</th>
<th>D</th>
<th>E</th>
<th>R</th>
<th>H</th>
<th>E</th>
<th>D</th>
<th>B</th>
<th>E</th>
<th>R</th>
<th>H</th>
<th>E</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q</td>
<td>G</td>
<td>Q</td>
<td>S</td>
<td>Q</td>
<td>S</td>
<td>Q</td>
<td>U</td>
<td>A</td>
<td>D</td>
<td>Q</td>
<td>S</td>
<td>Q</td>
<td>U</td>
<td>A</td>
<td>D</td>
</tr>
<tr>
<td>E</td>
<td>N</td>
<td>N</td>
<td>G</td>
<td>T</td>
<td>N</td>
<td>T</td>
<td>O</td>
<td>N</td>
<td>T</td>
<td>O</td>
<td>N</td>
<td>T</td>
<td>O</td>
<td>N</td>
<td>T</td>
</tr>
<tr>
<td>B</td>
<td>G</td>
<td>T</td>
<td>H</td>
<td>G</td>
<td>T</td>
<td>H</td>
<td>G</td>
<td>T</td>
<td>H</td>
<td>G</td>
<td>T</td>
<td>H</td>
<td>G</td>
<td>T</td>
<td>H</td>
</tr>
<tr>
<td>A</td>
<td>P</td>
<td>A</td>
<td>R</td>
<td>A</td>
<td>N</td>
<td>I</td>
<td>A</td>
<td>N</td>
<td>I</td>
<td>A</td>
<td>N</td>
<td>I</td>
<td>A</td>
<td>N</td>
<td>I</td>
</tr>
<tr>
<td>T</td>
<td>O</td>
<td>A</td>
<td>B</td>
<td>A</td>
<td>T</td>
<td>T</td>
<td>A</td>
<td>B</td>
<td>T</td>
<td>T</td>
<td>A</td>
<td>B</td>
<td>T</td>
<td>T</td>
<td>A</td>
</tr>
</tbody>
</table>

### FIGURE 31

This first row suggests the word SQUARE or SQUADRON, that in the last row suggests BATTLE or ATTALION. This means that a column with an A at the top and a Z at the bottom should be sought. There is only one such column, 31. Adding it to the 30–51–23 combination gives what is shown in figure 33a. Looking for a column with a B at the top (for SQUAD) and either an A (for BATTALION) or an L (for BATTLE), there is only one candidate, column 23, yielding the sequences shown in figure 33b. Enough has been shown of the procedure to make further demonstration unnecessary.

Once a good start has been made, progress is quite rapid, unless the cryptanalyst is unfortunate and arrives at a point where all the messages simultaneously terminate in complete words, without a clue as to what follows or precedes in any one of the messages. In such a contingency the only thing he can do is to try all sorts of possible continuations, either "fore" or "aft," that is, in front of the original starting point or after it, until he pecks up another word which will enable him to continue or he may have to search for a new point of entry and build upon that, later joining this structure with the other. In the case under examination no serious difficulties are found and the entire set of five messages is reconstructed.

In the course of this reconstruction the numbers applicable to the columns become assembled in proper sequence, that is, in the correct order to reproduce the plain text. This sequence, constituting the C–P sequence, is shown in figure 34 as the second row of numbers.

<table>
<thead>
<tr>
<th>C–P sequence</th>
<th>C–P sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Message No 1</td>
<td>HAVE ORDERED RATION WAGONS OFF FIRST</td>
</tr>
<tr>
<td>Message No 1</td>
<td>ENEMY D EFEATED DIRECTION OF RETREA</td>
</tr>
<tr>
<td>Message No 1</td>
<td>SECOND 0 DECEP 0 WILL LEAVE HEREATI</td>
</tr>
<tr>
<td>Message No 1</td>
<td>ANIMAL DRAWN VEHICLES OF ENGINEER</td>
</tr>
<tr>
<td>Message No 1</td>
<td>AMMUNITION TRAIN INCLUDING HORSED</td>
</tr>
</tbody>
</table>

### FIGURE 34

The solution by superimposing and anagramming equal-length messages in the case of transposition constitutes a general solution which is applicable in all cases without exception. Indeed, the possibility of solution by this method constitutes the most serious, if not fatal, weakness of transposition as a cryptographic method, for not only is it applicable to the most complex as well as to the most simple types of transposition, but, what is much more serious, the procedure is very simple, requiring very little cryptanalytic ingenuity or expertise. The chief disadvantage of this general solution is, of course, that it is dependent upon the more or less fortuitous availability of messages of identical length, and while this fortunate contingency is quite frequent in a voluminous correspondence, it would naturally be better from the point of view of the cryptanalyst if this requirement were not essential in all cases. Deeper study of the subject will show that the method can still be applied in a modified way to the case of messages of almost the same lengths when the transposition is not too involved. To illustrate, a case of simple keyed-columnar transposition will be used and it will be assumed that several messages of approximately identical lengths are at hand.

A first, take the case of two messages which have been enciphered by completely-filled rectangles, one message having, for example, one more row of letters than the other. In the discussion, the consecutive numbers 1, 2, 3, will be employed as though they constituted the successive letters of a plain-text message that is being enciphered. This method of treatment is very useful in connection with studies of the mechanics of transposition ciphers in general, and especially so in the case of double transposition. Note the P–C sequences that result from the transposition.

<table>
<thead>
<tr>
<th>P–C sequence</th>
<th>P–C sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>A 0 0 0</td>
<td>0 0 0 0 0</td>
</tr>
<tr>
<td>0 1 2 3 4 5</td>
<td>1 2 3 4 5 6 7</td>
</tr>
<tr>
<td>8 9 10 11 12 13 14</td>
<td>0 1 2 3 4 5 6 7</td>
</tr>
<tr>
<td>15 16 17 18 19 20 21</td>
<td>0 1 2 3 4 5 6 7</td>
</tr>
<tr>
<td>22 23 24 25 26 27 28</td>
<td>0 1 2 3 4 5 6 7</td>
</tr>
</tbody>
</table>

### FIGURE 32
It is obvious that the two sequences may be superimposed so as to bring identical sections into superimposition. Thus

A...... 00 15 20 27 02 09 16 23 07 14 21 28 05 12 19 26 03 10 17 24 01 08 15 22 04 11 18 25
B...... 00 15 20 02 09 16 23 07 14 21 05 12 19 03 10 17 24 01 08 15 22 04 11 18 25

The 7 blank spaces in the B line mark the ends of the columns in the transposition rectangle. The regularity in the distribution of the blank spaces follows from the mechanics of encryption. If two messages were superimposed in this manner it is clear that a solution by anagramming becomes perfectly feasible. Moreover, anagramming of columns is perhaps unnecessary, for anagramming merely the letters that would occupy in line A the positions marked by the blanks in line B will yield the transposition key directly. Extension of these principles to the case in which the two rectangles differ by 2, 3, 4, complete rows as obvious.

Taking next a case wherein two rectangles differ by one or two letters in the bottom row, it is clear that by shifting the letters of one message one or two spaces to the right (or left) from a given point will bring most of the text into proper superimposition for a solution by anagramming. Note the P → C sequences applicable to the following transpositions:

<table>
<thead>
<tr>
<th>6</th>
<th>2</th>
<th>5</th>
<th>7</th>
<th>4</th>
<th>1</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>02</td>
<td>03</td>
<td>04</td>
<td>05</td>
<td>06</td>
<td>07</td>
</tr>
<tr>
<td>08</td>
<td>09</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
</tr>
<tr>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
<td>21</td>
</tr>
<tr>
<td>22</td>
<td>23</td>
<td>24</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td>28</td>
</tr>
<tr>
<td>A</td>
<td>B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

P → C sequence for A...... 06 13 20 27 02 09 16 23 07 14 21 05 12 19 26 03 10 17 24 01 08 15 22 04 11 18 25
P → C sequence for B...... 06 13 20 02 09 16 23 07 14 21 05 12 19 03 10 17 24 01 08 15 22 04 11 18 25

It is possible to superimpose these two sequences by shifting the sections in line B after certain numbers. Thus

A...... 06 13 20 27 02 09 16 23 07 14 21 05 12 19 26 03 10 17 24 01 08 15 22 04 11 18 25
B...... 06 13 20 02 09 16 23 07 14 21 05 12 19 03 10 17 24 01 08 15 22 04 11 18 25

In the case of actual messages corresponding to the foregoing P → C sequences, superimposition of the two texts in the manner indicated would at once permit of a solution by anagramming of columns. The unknown factor, of course, is the location of the blank spaces. Where the two messages differ in length by only one or two letters, brief experimental examination will tell the story, where the messages differ in length by a good many letters the process would be much more difficult but not at all hopeless of fruitful results. Only a small section of text reconstructed by anagramming will soon lead to complete solution. Hence, it follows that by regulating the number of blanks to be left here and there and judicious shifting of sections of text, solution by superimposing and anagramming homologous sections of text from several messages in the same transposition key will often be possible.

The foregoing principles will naturally not be applicable to cases where two messages differ in length by but one letter and thus small difference brings about a profound difference in the P → C sequences applicable to the messages. Thus it happens often in the case of true double transposition, but the principle can nevertheless be applied even here. An explanation of the procedure lies beyond the scope of the present text, however, and no more will be indicated hereon concerning the matter in the case of true double transposition. However, in certain cases of combined substitution-transposition to be discussed in a later portion of this text the principles elucidated in these last few subparagraphs may be found to be applicable.

27. Reconstruction of the keys in double transposition. — a Having reconstructed the plain texts of the messages solved by superimposing and anagramming, as explained in paragraph 26 d, e, can the transposition key be found? First, it is necessary to ascertain whether a single columnar transposition had been used and, if not, then the assumption will be that a double transposition had been used.

b) If a single transposition were the case, the relationship pointed out in paragraph 16c, concerning the existence of a constant difference between successive elements of the P → C sequence, should obtain. Having the C → P sequence, the P → C sequence may readily be established by inversion of the former. Hence, the P → C sequence is constructed by inversion, as shown in figure 36a.

---

1 See Special Text No 166, *Advanced Military Cryptography*, sec IV
There is a constant difference between the superposed terms. The reason for its appearance is not hard to understand if reference is made to figure 37a and the manner is studied in the light of the mechanics of the method of encipherment. As for the two terms 28 and 20, while they come from different columns in the T-2 matrix, both come from the same column of the T-1 matrix, as do 14 and 06, 41 and 33, 23 and 15. But the 05 and the 48 not only come from different columns in the T-2 matrix, but also from different columns in the T-1 matrix, this representing a case where there is a transit from the bottom of one column to the top of the next column in the transposition process. Now the constant difference is as in this case 8 because the superposed terms happen to be in segments in the columns in which they fall in the T-1 matrix. If the superposed terms are in the same column in the T-1 matrix but separated by one row, the constant difference will be 16, if separated by two rows, the constant difference will be 24, and so on. Thus, for example

<table>
<thead>
<tr>
<th>Term number...</th>
<th>01</th>
<th>02</th>
<th>03</th>
<th>04</th>
<th>05</th>
<th>06</th>
<th>07</th>
<th>08</th>
<th>09</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>P → C sequence.</td>
<td>20</td>
<td>06</td>
<td>48</td>
<td>33</td>
<td>15</td>
<td>44</td>
<td>30</td>
<td>21</td>
<td>03</td>
<td>39</td>
<td>09</td>
<td>42</td>
<td>32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C → P sequence.</td>
<td>27</td>
<td>47</td>
<td>09</td>
<td>11</td>
<td>31</td>
<td>02</td>
<td>42</td>
<td>44</td>
<td>23</td>
<td>34</td>
<td>50</td>
<td>38</td>
<td>18</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Nothing in the nature of a series of constant differences between successive terms is now discernible in the P → C sequence. But there is, as can readily be seen, a fairly constant relationship between segments or sections of this sequence. For example, take the 1st to 6th terms of the P → C sequence (20 06 48 33 15), set them under the 29th to 34th terms (28 14 05 41 23), and find the differences between superposed numbers (When the modulus is less than the subtrahend the superposed terms are disregarded). Thus

| Term number... | 29 | 30 | 31 | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 40 | 41 | 42 | 43 | 44 | 45 | 46 | 47 | 48 |
|---------------|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| P → C sequence.| 04 | 12 | 20 | 28 | 36 | 44 | 02 | 10 | 18 | 26 | 34 |
| C → P sequence.| 42 | 50 | 06 | 14 | 20 | 38 | 46 | 08 | 16 | 24 |

(4) As in paragraph 26a, the exposition will employ matrices in which the consecutive numbers 1, 2, 3, ... take the place of the successive plain-text letters in the T-1 matrix, because in such handling significant facts arising from the mechanics of encipherment are brought to light.

In order to study the effects of true double transposition on this matter of reconstructing the keys an example will be employed, involving transposition with two different keys. Let the "message" and the keys be as shown in figure 37a.

<table>
<thead>
<tr>
<th>01 02 03 04 05 06 07 08 09 10</th>
<th>01 02 03 04 06 06 07 08 09 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>01 02 03 04 05 06 07 08 09 10</td>
<td>01 02 03 04 06 06 07 08 09 10</td>
</tr>
<tr>
<td>01 02 03 04 06 06 07 08 09 10</td>
<td>01 02 03 04 06 06 07 08 09 10</td>
</tr>
<tr>
<td>01 02 03 04 06 06 07 08 09 10</td>
<td>01 02 03 04 06 06 07 08 09 10</td>
</tr>
<tr>
<td>01 02 03 04 06 06 07 08 09 10</td>
<td>01 02 03 04 06 06 07 08 09 10</td>
</tr>
<tr>
<td>01 02 03 04 06 06 07 08 09 10</td>
<td>01 02 03 04 06 06 07 08 09 10</td>
</tr>
<tr>
<td>01 02 03 04 06 06 07 08 09 10</td>
<td>01 02 03 04 06 06 07 08 09 10</td>
</tr>
<tr>
<td>01 02 03 04 06 06 07 08 09 10</td>
<td>01 02 03 04 06 06 07 08 09 10</td>
</tr>
<tr>
<td>01 02 03 04 06 06 07 08 09 10</td>
<td>01 02 03 04 06 06 07 08 09 10</td>
</tr>
<tr>
<td>01 02 03 04 06 06 07 08 09 10</td>
<td>01 02 03 04 06 06 07 08 09 10</td>
</tr>
<tr>
<td>01 02 03 04 06 06 07 08 09 10</td>
<td>01 02 03 04 06 06 07 08 09 10</td>
</tr>
</tbody>
</table>

As the message is limited to 51 letters, all the correspondents will be 1, 2, 3. It will be seen, in referring to figure 37, that the following repetitions (with the term numbers in each of the sequences indicated) are present

<table>
<thead>
<tr>
<th>Term no in P → C sequence...</th>
<th>32</th>
<th>29</th>
<th>30</th>
<th>31</th>
<th>32</th>
<th>33</th>
<th>34</th>
<th>35</th>
<th>36</th>
<th>37</th>
<th>38</th>
<th>39</th>
<th>40</th>
<th>41</th>
<th>42</th>
<th>43</th>
<th>44</th>
<th>45</th>
<th>46</th>
<th>47</th>
<th>48</th>
</tr>
</thead>
<tbody>
<tr>
<td>P → C sequence...</td>
<td>32</td>
<td>29</td>
<td>30</td>
<td>31</td>
<td>32</td>
<td>33</td>
<td>34</td>
<td>35</td>
<td>36</td>
<td>37</td>
<td>38</td>
<td>39</td>
<td>40</td>
<td>41</td>
<td>42</td>
<td>43</td>
<td>44</td>
<td>45</td>
<td>46</td>
<td>47</td>
<td>48</td>
</tr>
</tbody>
</table>

The width of the T-1 matrix is therefore 12 and its outlines may at once be drawn, since the total number of letters in each message, 51, indicates that there are 3 long columns of 5 letters and 9 short columns of 4 letters.
(1) There is another method of ascertaining the width of the T-1 matrix, which is perhaps a bit shorter and more direct than that described above. Basically both methods are the same, the one now to be presented being but another way of looking at the matter. Suppose that the difference between successive terms in the P-C sequence of figure 37a are calculated and set down as shown below, and then repetitions are sought in the sense of differences, the latter constituting what will hereafter be termed the P-C interval sequence Thus

<table>
<thead>
<tr>
<th>Term number</th>
<th>P-C sequence...</th>
<th>P-C interval sequence...</th>
<th>P-C interval sequence...</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>20 46 33 15</td>
<td>14 + 22 + 18 + 29 - 14 - 9 - 26 + 25 - 10 - 15 = 15</td>
<td>8 + 8 + 8 + 8 + 8</td>
</tr>
<tr>
<td></td>
<td>30 47 34 22</td>
<td>14 + 22 + 18 + 29 - 14 - 9 - 26 + 25 - 10 - 15 = 15</td>
<td>8 + 8 + 8 + 8 + 8</td>
</tr>
<tr>
<td></td>
<td>31 47 34 22</td>
<td>14 + 22 + 18 + 29 - 14 - 9 - 26 + 25 - 10 - 15 = 15</td>
<td>8 + 8 + 8 + 8 + 8</td>
</tr>
<tr>
<td></td>
<td>32 47 34 22</td>
<td>14 + 22 + 18 + 29 - 14 - 9 - 26 + 25 - 10 - 15 = 15</td>
<td>8 + 8 + 8 + 8 + 8</td>
</tr>
</tbody>
</table>

We have ascertained that the constant repetitions of 8 will be found Thus

<table>
<thead>
<tr>
<th>Term numbers</th>
<th>P-C sequence...</th>
<th>P-C interval sequence...</th>
<th>P-C interval sequence...</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>20 46 33 15</td>
<td>14 + 22 + 18 + 29 - 14 - 9 - 26 + 25 - 10 - 15 = 15</td>
<td>8 + 8 + 8 + 8 + 8</td>
</tr>
<tr>
<td></td>
<td>30 47 34 22</td>
<td>14 + 22 + 18 + 29 - 14 - 9 - 26 + 25 - 10 - 15 = 15</td>
<td>8 + 8 + 8 + 8 + 8</td>
</tr>
<tr>
<td></td>
<td>31 47 34 22</td>
<td>14 + 22 + 18 + 29 - 14 - 9 - 26 + 25 - 10 - 15 = 15</td>
<td>8 + 8 + 8 + 8 + 8</td>
</tr>
<tr>
<td></td>
<td>32 47 34 22</td>
<td>14 + 22 + 18 + 29 - 14 - 9 - 26 + 25 - 10 - 15 = 15</td>
<td>8 + 8 + 8 + 8 + 8</td>
</tr>
</tbody>
</table>

The other repetitions will show the same constant difference The terms which produce the repetitions will be found to be located in the same columns of the T-2 matrix in figure 37a, and reference to that figure will show that the constant difference between the sets of terms producing repetitions in the P-C interval sequence is merely the result of the mechanics of encryption

(2) In similar manner, if the interval sequence is constructed for the P-C sequence of figure 36a, the repetitions underscored in figure 36b are noted

<table>
<thead>
<tr>
<th>Term number</th>
<th>P-C sequence...</th>
<th>P-C interval sequence...</th>
<th>P-C interval sequence...</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>20 46 33 15</td>
<td>14 + 22 + 18 + 29 - 14 - 9 - 26 + 25 - 10 - 15 = 15</td>
<td>8 + 8 + 8 + 8 + 8</td>
</tr>
<tr>
<td></td>
<td>30 47 34 22</td>
<td>14 + 22 + 18 + 29 - 14 - 9 - 26 + 25 - 10 - 15 = 15</td>
<td>8 + 8 + 8 + 8 + 8</td>
</tr>
<tr>
<td></td>
<td>31 47 34 22</td>
<td>14 + 22 + 18 + 29 - 14 - 9 - 26 + 25 - 10 - 15 = 15</td>
<td>8 + 8 + 8 + 8 + 8</td>
</tr>
<tr>
<td></td>
<td>32 47 34 22</td>
<td>14 + 22 + 18 + 29 - 14 - 9 - 26 + 25 - 10 - 15 = 15</td>
<td>8 + 8 + 8 + 8 + 8</td>
</tr>
</tbody>
</table>

Here it is seen that the numbers in the columns of the T-2 matrix are consecutive. Obviously, if the columns of the T-2 matrix were retranscribed in a matrix of the same outlines as the T-1 matrix, the numbers would be consecutive in rows and would represent the plain-text sequence 1, 2, 3, , inserted within a T-1 matrix in the normal fashion Thus (fig 37d)
The result of these three steps has been to effect a return to the original $T-1$ matrix containing plan text in normal sequence. The principal point to be noted here is the appearance of the $T-2$ matrix in figure 37e, for it is the basis of the first method for recovering the transposition keys, as well as for other operations in connection with double transposition.

(2) To demonstrate the procedure, the data afforded by figure 37a again will be employed. Let the C→P sequence be inscribed within a matrix of 6 columns (the watch having been established by one of the methods set forth in subpara d-f) Thus

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>27</td>
<td>47</td>
<td>09</td>
<td>11</td>
<td>31</td>
<td>02</td>
</tr>
<tr>
<td>23</td>
<td>34</td>
<td>50</td>
<td>38</td>
<td>18</td>
<td>30</td>
</tr>
<tr>
<td>14</td>
<td>45</td>
<td>37</td>
<td>01</td>
<td>08</td>
<td>17</td>
</tr>
<tr>
<td>41</td>
<td>25</td>
<td>46</td>
<td>29</td>
<td>09</td>
<td>07</td>
</tr>
<tr>
<td>04</td>
<td>21</td>
<td>28</td>
<td>16</td>
<td>36</td>
<td>48</td>
</tr>
<tr>
<td>32</td>
<td>12</td>
<td>24</td>
<td>06</td>
<td>45</td>
<td>35</td>
</tr>
<tr>
<td>19</td>
<td>39</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Find the column in which term 01 appears and set that column down horizontally, placing a vertical bar before and after the series of numbers to set them off as belonging to one column

Step (1)

Thus

| 11 38 01 29 16 06 |

Then find the column in which the term 02 appears and set it down under the row of numbers given in Step (1)

Step (2)

Thus

| 11 38 01 29 16 06 |

Note the ascending superimposed numbers 01, 02, 29, 30, 16, 17, 06, 07. Continue to build up on this structure in the manner depicted in successive steps as follows

Step (3)

Thus

| 11 38 01 29 16 06 |

Step (4)

Thus

| 44 26 22 13 40 03 |

Step (5)

Thus

| 44 26 22 13 40 03 |

The numbers 01 to 05, inclusive, have here formed the basis for building up the structure shown as Step (5). The next term in the sequence is 06 but it is already in the structure, as is also 07.

In the column in which 08 and 09 appear there is just room enough for 08 and 09, since the term 10 is already shown at the bottom of the column. Hence

| 11 38 01 29 16 06 |

The process is continued in this manner until, as shown in figure 37f(9), all the numbers of the C→P sequence have been placed.

The $T-2$ matrix may now be drawn within the confines of the structure shown in this last figure. The positions of vertical lines to be placed at the left and right to mark the outlines of the matrix may now readily be found by referring to the matrix in figure 37c. It is obvious that the column with the terms 11-15 belongs at the extreme left of the $T-2$ matrix, the column with the terms 21-24 belongs at the extreme right. The transposition key for the matrix may then be established directly from the matrix itself, by following the sequence of numbers in the columns.
The process must be as follows

This transposition key and T-1 matrix are identical with the key and T-1 matrix of figure 37e. Note the application of the foregoing method to the C→P sequence shown in figure 36b in connection with the messages solved in paragraph 26d. It has already been found that the width of the T-1 matrix is 12. The C→P sequence of figure 36b is therefore inscribed within a matrix of 12 columns

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>28</td>
<td>04</td>
<td>14</td>
<td>46</td>
<td>19</td>
<td>47</td>
<td>25</td>
<td>47</td>
<td>26</td>
<td>45</td>
<td>41</td>
<td>28</td>
<td>10</td>
<td>06</td>
</tr>
<tr>
<td>02</td>
<td>34</td>
<td>27</td>
<td>12</td>
<td>36</td>
<td>45</td>
<td>17</td>
<td>16</td>
<td>40</td>
<td>18</td>
<td>09</td>
<td>24</td>
<td>41</td>
<td>08</td>
</tr>
<tr>
<td>01</td>
<td>46</td>
<td>01</td>
<td>50</td>
<td>44</td>
<td>11</td>
<td>01</td>
<td>24</td>
<td>31</td>
<td>22</td>
<td>16</td>
<td>07</td>
<td>21</td>
<td>12</td>
</tr>
<tr>
<td>42</td>
<td>10</td>
<td>49</td>
<td>04</td>
<td>43</td>
<td>15</td>
<td>05</td>
<td>39</td>
<td>29</td>
<td>20</td>
<td>38</td>
<td>00</td>
<td>18</td>
<td>28</td>
</tr>
</tbody>
</table>

The process explained in subparagraph h (2) above is now applied. The successive steps have been omitted but the final result is shown herewith

| 14 | 27 | 01 | 32 | 06 | 19 | 36 | 44 | 10 | 44 | 48 | 23 | 40 | 25 |
| 41 | 24 | 36 | 29 | 03 | 34 | 28 | 35 | 28 | 40 | 29 | 41 | 04 | 45 |
| 25 | 17 | 00 | 45 | 09 | 22 | 39 | 47 | 13 | 51 | 43 | 26 | 18 | 41 |

All the numbers from 01 to 51, inclusive, are contained within this structure. Extending it to the left or right to make the T-2 matrix complete, by referring to the T-1 matrix, it is found that the structure must be made as shown herewith

| 14 | 27 | 01 | 32 | 06 | 19 | 36 | 44 | 10 | 44 | 48 | 23 | 40 | 25 |
| 41 | 24 | 36 | 29 | 03 | 34 | 28 | 35 | 28 | 40 | 29 | 41 | 04 | 45 |
| 25 | 17 | 00 | 45 | 09 | 22 | 39 | 47 | 13 | 51 | 43 | 26 | 18 | 41 |

The transpose key for the T-1 matrix is now found to be as indicated at the top of figure 36e. A second method for reconstructing the keys will now be explained. To demonstrate this method, the data afforded by figure 37e will again be employed. Going back to the point where the P→C interval sequence for this example was established in subparagraph f (1) above, the terms, in figure 37e, which gave rise to the three appearing repetitions in the interval sequence (−10 −15 +34) are found to be as follows.

1st appearance (term numbers 12–16).
2nd appearance (term numbers 21–25).
3rd appearance (term numbers 25–29).

These sequences may be rearranged so as to bring the numbers in ascending order within columns. Thus

| 26 | 16 | 01 | 33 |
| 17 | 24 | 09 | 43 |
| 42 | 53 | 17 | 01 |

The constant difference, 8, within the columns of this structure is, of course, the same constant difference as was found before, and corresponds with the width of the T-1 matrix. It derives from the T-1 matrix, as may be seen on referring to figure 37e. The columns of the structure in figure 37e are seen to be portions of the T-1 matrix, lying in the following positions in that matrix

| 01 | 09 | 17 |
| 26 | 34 | 42 |
| 32 | 01 | 09 |
| 17 | 35 | 43 |
| 51 | 26 | 16 |

In the T-2 matrix these numbers fall in the following positions

| 26 | 16 | 01 |
| 17 | 24 | 09 |
| 42 | 34 | 55 |
| 01 | 09 | 17 |
| 35 | 43 | 51 |

Now if the dimensions of the T-2 matrix were unknown, these numbers could nevertheless be placed in a skeletoned T-2 matrix as follows

| 26 | 16 | 42 |
| 16 | 24 | 32 |
| 01 | 09 | 17 |
| 35 | 43 | 51 |
and the block of numbers could be extended on both sides by referring to the T-1 matrix in figure 37a. Thus

| 02 10 18 26 34 42 50 |
| 08 16 24 32 40 48 |
| 01 09 17 25 33 41 49 |

Thus this structure may next be extended by referring to the P→C sequence in figure 37a, by completing the partial columns of the structure:

44 | 02 10 18 26 34 42 50 06 14 22 30 38 46 08 16 24 32 40 48 05 13 21 29 37 45 07 15 23 31 39 47

And this structure may be extended by referring to the T-1 matrix. Thus

04 12 20 28 36 44 02 10 18 26 34 42 50 06 14 22 30 38 46 08 16 24 32 40 48 05 13 21 29 37 45 07 15 23 31 39 47

Noting the appearance of the term 06 in the 1st row and also in the 2d row of the structure, the latter may be transcribed as follows:

04 12 20 28 36 44 02 10 18 26 34 42 50 06 14 22 30 38 46 08 16 24 32 40 48 05 13 21 29 37 45 07 15 23 31 39 47

By referring to the T-1 matrix of figure 37a and the foregoing structure, the key for T-1 can be recovered. It is 0-2-7-1-5-3-8-4 By referring to the P→C sequence in figure 37a the key for the T-2 matrix just constructed may be established. It is 3-9-1-7-4-2-11-8-10-6-5

The width of the T-1 matrix has been established as being 12. The P→C interval sequence in figure 36b shows the repetition +3 +5 -14 -11 -3 appearing two times, the repetition +3 +5 -14 appearing three times, and the repetition -14 -11 appearing three times. The terms giving rise to these repetitions are arranged in a structure with ascending numbers within the columns. Thus

07 10 15 01 31 34 39 25 14 11 43 46 51 37 26 23

The constant difference, 12, indicates a T-1 matrix of 12 columns. The matrix is prepared

| 01 02 03 04 05 06 07 08 09 10 11 12 |
| 13 14 15 16 17 18 19 20 21 22 23 24 |
| 25 26 27 28 29 30 31 32 33 34 35 36 |
| 37 38 39 40 41 42 43 44 45 46 47 48 |

The terms within the columns of the structure in figure 36f are transcribed into rows (of the skeletonized T-2 matrix)

07 10 15 01 31 34 39 25 14 11 43 46
13 14 15 16 17 18 19 20 21 22 23 24
25 26 27 28 29 30 31 32 33 34 35 36
37 38 39 40 41 42 43 44 45 46 47 48

This structure is extended by referring to the T-1 matrix (figure 36g)

Noting that the mutual terms of the P→C sequence in figure 36a (27 13 02) are present in this structure (in the 3d column) thus gives the top of the T-2 matrix as concurring with the 3d row of the structure. The P→C sequence in figure 36a reads 27 13 02 43 46, the 43 and 46 are also in the structure in figure 36a (2) in the 1st and 2d rows, column 5, hence the structure in figure 36a(2) can be rearranged thus

03 15 27 39 51 01 17 37 39 51 03 15 27 39 51 07 19 31 43 55 07 19 31 43 55 10 22 34 46

The structure may now be extended by referring to the P→C sequence in figure 36a

35 03 15 27 49 51 03 17 33 45 01 13 23 35 47 24 36 48 02 14 26 38 50 07 19 31 41 11 23 35 47
10 22 34 46

FIG. 36G

FIG. 36B(2)
Thus, by referring alternately to the P−C sequence and the T-1 matrix, the structure is extended to the following:

```
   35 03 15 27 39 51 | 05 17 29 41 | 09 21 33 45 |
   23 45 | 01 13 25 37 49 | 06 18 30 42 | 12 24 36 48 |
   24 05 48 | 02 14 26 38 50 | 04 16 28 40 | 07 19 31 43 |
   07 19 31 43 | 11 23 35 47 | 08 20 32 44 | 10 22 34 46 |
```

Figure 37a

It will be noted that the first number to the right of each vertical bar is one of the numbers from 1 to 12, indicating that all the columns of the T-1 matrix are now represented in the T-2 structure.

It is now easy to write the transposition key over the T-1 matrix 4-7-1-8-2-5-4-11-3-12-10-6. By following the numbers in the P−C sequence the transposition key for the T-2 matrix is given directly, it is the same as for the T-1 matrix.

(1) A third method for reconstructing the transposition key will now be set forth. It will be explained in connection with the artificial example in figure 37a. It has been noted how the width of the T-1 matrix can be ascertained from a study of the P−C sequence, the work in connection with figure 37a and subparagraph a give an indicated width of 8 for the T-1 matrix in this case.

(2) Let the additive 8 (found in subpass d and f) be applied to the entire P−C sequence of figure 37a, and then let the identifiers between the two sequences be underscored and numbered, as shown in figure 37a.

(A) P−C sequence
```
   20 06 48 33 15 44 30 21 03 29 04 42 32 17 51 |
   36 22 13 49 31 14 24 09 41 26 16 01 35 |
   28 14 05 41 23 10 46 37 19 12 50 40 25 07 |
   18 08 45 27 02 38 29 11 47 |
```

(B) P−C sequence+8...
```
   26 14 06 41 23 52 38 29 11 47 12 50 40 25 39 |
   44 30 21 57 39 42 22 17 51 34 24 09 47 |
   36 22 13 49 31 18 54 45 27 20 58 48 33 15 |
   26 16 03 35 10 46 37 19 55 |
```

Figure 37b

If now the procedure explained in paragraphs 16a, 24c to f, and 25a to e is applied to the repetitions noted in figure 37a, it becomes clear that the T-2 matrix in this case must have 11 columns. The transposition key for that matrix is then established, as follows

```
   B | 1 2 3 4 5 6 7 8 9 10 11 |
   A | 7 11 9 2 4 5 6 10 1 6 8 |
   C | 7 4 2 11 8 10 6 5 3 9 |
```

It is to be noted that the B sequence (that is, the P−C sequence plus the additive) must be used as the base, others use the chain of equivalents will be a reversal of the correct chain.

This is a cycle permutation of the key for the T-2 matrix, to obtain the actual key it is necessary merely to fix the position of one of the key numbers with respect to the matrix. It is easy to find which number belongs at the extreme left or extreme right of the matrix. Only a few minutes experimentation with the key and the T-2 matrix gives the correct starting point for the key, which is found to be 3-9-1-7-4-2-11-8-10-6-5.

(3) The recovery of the transposition key for the T-1 matrix is now a simple matter. Its width having been established as 8 columns, a mere transposition of the P−C sequence numbers from the T-2 matrix into the T-1 matrix gives the key 6-2-7-1-5-3-8-4. The two keys and matrices are found to be different.

(4) The procedure set forth in this subparagraph is applicable without modification to the case where the two transposition matrices are the same and have the same transposition key. Thus will be noted in the following demonstration of the recovery of the matrices and keys for the messages solved in paragraph 26d and e by anagramming. It has already been shown how the width of the T-1 matrix was reconstructed as being 12 columns (paragraph f). The additive 12 is then applied to the entire P−C sequence, identifiers are established between sections of the original sequence and sections of the sequence +12, and these identical sections are equated in the usual manner, leading to the establishment of a cycle permutation of the transposition key for the T-2 matrix. Thus (fig 36b)

```
   A (P−C sequence) .................................................. 27 13 02 43 46 51 27 26 23 41 30 16 20 |
   B (P−C sequence+12) ................................................. 29 25 14 52 59 63 49 28 35 53 42 26 25 |
```

Figure 38

Since sections 1, 4, and 7 of the P−C sequence contain 5 terms (=long columns), the other sections only 4 (=short columns), it follows that the key numbers 4−7 go to the left and the actual key for the T-2 matrix is 4−7−1−8−2−5−9−11−3−12−10−6. Since the number of elements in this key is the same as in the key for the T-1 matrix, it is likely that the same key is employed for both transpositions. Simple experiment will quickly verify this assumption and the transposition matrices for the first of the 5 messages of paragraph 26d may be seen in the following (figure 36b).
A casual examination of these three rows of numbers discloses an interesting invariant relationship between any pair of superimposed numbers in rows 1 and 2 and in rows 2 and 3. For instance, take the very first pair, 01 in rows 1 and 2, in rows 2 and 3 the same pair of superimposed numbers will be found (under term No 14) This same relationship exists between all the superimposed pairs in rows 1-2 and 2-3.

(3) Given only the third row of numbers in figure 36g, that is, the P~C2 sequence (which has herefore been designated merely as the P~C sequence), obtained as a result of a solution by superimposing and anagramming several messages, it is not difficult to reconstruct the second row, the P~C sequence. The width of the T-1 matrix can be ascertained by either of the two methods indicated in subparagraphs a and f. It is now known to be 12. A 12-column matrix is therefore constructed, containing 51 cells numbered in the normal manner. This will, of course, give the T-1 matrix seen in figure 36b, but without the transposition key or the letters in the cells. Thus

\[ \text{Suppose key number 15,27,39,51 of rows 1 and 2, 01 in rows 2 and 3, also, 01 in rows 1 and 2, 02 in rows 2 and 3. The conclusion is obvious that the key number 1 does not occupy the 1st position in the transposition key. Suppose key number 1 belongs in the 2d position in the key. The superimposed sequences are then as follows} \]

**Table 9**

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>02</td>
<td>03</td>
<td>04</td>
</tr>
<tr>
<td>05</td>
<td>06</td>
<td>07</td>
<td>08</td>
</tr>
<tr>
<td>09</td>
<td>10</td>
<td>11</td>
<td>12</td>
</tr>
</tbody>
</table>

**Form 9**

The invariant relationship pointed out in subparagraph (1) above may now be used to establish the T-1 key. Since the key is known to contain 12 elements, a start may be made with any one of 12 possibilities. Suppose that the key begins with 1. The first five terms in the P~C sequence would be as indicated herewith.

1. Term number. 01, 02, 03, 04, 05.
2. P~C sequence. 06, 07, 08, 09, 10.
3. P~C sequence. 11, 12, 01, 02, 03.
4. P~C sequence. 04, 05, 06, 07, 08.
5. P~C sequence. 09, 10, 11, 12, 01.

Form 10

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>02</td>
<td>03</td>
<td>04</td>
<td>05</td>
</tr>
<tr>
<td>06</td>
<td>07</td>
<td>08</td>
<td>09</td>
<td>10</td>
</tr>
<tr>
<td>11</td>
<td>12</td>
<td>01</td>
<td>02</td>
<td>03</td>
</tr>
</tbody>
</table>

**Form 11**

Here again two conflicts are noted. 01 in rows 1 and 2, 02 in rows 2 and 3. Only a single contradiction is sufficient to permit of discarding an hypothesis. The key number 1 does not occupy the 2d position in the key. A trial is made of the 3d position for key number 1. The results are as follows.

1. Term number. 01, 02, 03, 04, 05.
2. P~C sequence. 06, 07, 08, 09, 10.
3. P~C sequence. 11, 12, 01, 02, 03.

**Form 12**

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>02</td>
<td>03</td>
<td>04</td>
<td>05</td>
</tr>
<tr>
<td>06</td>
<td>07</td>
<td>08</td>
<td>09</td>
<td>10</td>
</tr>
<tr>
<td>11</td>
<td>12</td>
<td>01</td>
<td>02</td>
<td>03</td>
</tr>
</tbody>
</table>

**Form 13**

The key number 1 occupies the 3d position in the key. A trial is made of the 4d position for key number 1. The results are as follows:

1. Term number. 01, 02, 03, 04, 05.
2. P~C sequence. 06, 07, 08, 09, 10.
3. P~C sequence. 11, 12, 01, 02, 03.

**Form 14**

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>02</td>
<td>03</td>
<td>04</td>
<td>05</td>
</tr>
<tr>
<td>06</td>
<td>07</td>
<td>08</td>
<td>09</td>
<td>10</td>
</tr>
<tr>
<td>11</td>
<td>12</td>
<td>01</td>
<td>02</td>
<td>03</td>
</tr>
</tbody>
</table>

**Form 15**

The end result is therefore obtained for the T-1 matrix. A single conflict is noted on the 1st row, 2nd column. With the T-1 matrix thus obtained, a message is solved by the T-1 system. Suppose the message is: "I am an old soldier, and will turn out at the sound of the bugle."
Here there are no contradictions and one check at corroboration. If key number 1 really occupies the 3d position in the key, then the superimposition data given in the list of rows of superimposed numbers may be employed, by transforming the data to the proper positions in the skeletonized figure 36m(1)

1. Term number:...
2. P-C1 sequence:...
3. P-C2 sequence:...

If key 1 ln<:t "ll't only numb'<rs, it I'> only ne<'1<:1uv pPrm1t at once
It then becomes at once possible, by referring to the T-1 matrix, to insert more numbers in the P-C1 sequence... Thus

The new placements now permit of placing numbers in the P-C2 sequence... For example, in rows 1 and 2 permit of placing the number 07 above the number 17 in the P-C2 sequence, and so on. In only a few moments the entire P-C1 sequence can be established. Thus

The application of the foregoing phenomena in the case under study is obvious...
specific matrices and keys, but the general principles here presented may nevertheless be applied. In some cases it may be necessary to juxtapose two different chains obtained by equating terms from rows 1 and 3, rather than juxtaposing one chain against itself. Only a few minutes experimentation will be necessary to establish contradictions which will permit of discarding fallacious hypotheses.

(1) In the foregoing explanation, the two transposition keys and matrices were identical. Even when they are different the same principles, with minor modifications, may be applied. The matrices and keys of figure 37a will again be employed to demonstrate the necessary modifications.

(2) First, prepare the two matrices with consecutive numbers in the cells of both matrices, as shown at A and B in figure 37c and then prepare the T-2 matrix shown at C.

(3) Write the P-C sequence for T-1 under it, write the P-C1 for T-2, and under the latter write the P-C1 sequence for the final cryptogram. Thus:

1 Term number
2 P-C sequence for T-1 ....... 01 02 03 04 05 06 07 08
3 P-C sequence for T-2 ....... 04 12 20 28 36 44 02 10 18 26 34 42 50 06
4 P-C sequence ... ... ... ... 01 14 25 36 47 06 17 28 39 50 01 12 23 34
5 18 02 03 04 05 06 07 08
6 09 10 11 12 13 14 15 16
7 17 18 19 20 21 22 23 24
8 25 26 27 28 29 30 31 32
9 33 34 35 36 37 38 39 40
10 41 42 43 44 45 46 47 48
11 49 50 52

Note: now, the invariant relationship between rows 1-2 and 3-4. The same phenomenon is here manifested as was encountered in the preceding case where the T-1 and T-2 matrices and keys are identical. It follows, therefore, that the principles elucidated under subparagraph (1) may be applied, with some modifications, also to the case where different keys and matrices are employed for double transposition. The width of the T-1 matrix may be ascertained in the manner already indicated; the assumption is made as to the position occupied by key number 1 of the T-1 key, this assumption provides data for making an assumption as to the width of the T-2 matrix. When the correct pair of assumptions is made, the data in rows 1 and 2 are corroborated by those in rows 3 and 4. From that point on the rest is easy and follows along the same lines as before.

(1) The procedure will be illustrated by employing the P-C sequence in figure 37b (which is the same as that labelled P-C1 sequence in figure 37p), it being assumed that nothing is known about the matrices, and that the sequence was obtained from a solution by superposing and arranging several messages of identical length.
<table>
<thead>
<tr>
<th>Turn number</th>
<th>01</th>
<th>02</th>
<th>03</th>
<th>04</th>
<th>05</th>
<th>06</th>
<th>07</th>
<th>08</th>
<th>09</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 P→C₃ sequence for T-1</td>
<td>04</td>
<td>12</td>
<td>20</td>
<td>28</td>
<td>36</td>
<td>44</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 P→C₅ sequence for T-2</td>
<td>03</td>
<td>06</td>
<td>07</td>
<td>08</td>
<td>09</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 P→C₅ sequence</td>
<td>20</td>
<td>06</td>
<td>08</td>
<td>33</td>
<td>15</td>
<td>44</td>
<td>30</td>
<td>02</td>
<td>03</td>
<td>39</td>
<td>04</td>
<td>42</td>
<td>32</td>
<td>17</td>
</tr>
</tbody>
</table>

(5) Here there are found no contradictions of the nature of those pointed out above. The T-2 matrix appears to have columns of 4 and 5 letters, since the interval between 04 and 02 in row 3 can accommodate a short column of 4 and a long column of 5 letters, the interval between 05 and 04 can accommodate 2 short columns of 4 letters and 1 long column of 5, the intervals between 04 and 06, 06 and 01, 01, and 05 can accommodate long columns of 5 letters each. Only 2 matrices can be constructed of 51 letters with long columns of 5 and short columns of 4 letters. They are:

Key of 11

- f (long) × 5 = 35
- (short) × 4 = 16

Key of 12

- f (long) × 5 = 25
- (short) × 4 = 20

Each of these T-2 matrices is tested as a possibility.

(6) If matrix A is correct, then the numbers in columns 3, 6, 1, 5, 4, and 2 can be transferred to row 3 in figure 37f (2), these will permit of inserting numbers in row 2. No contradictions and many checks are found. Here is the diagram.

(7) In the first place note, in row 2, the constant difference 8, giving many corroborations that the width of the T-1 matrix as 8, in the second place no conflicts whatever become manifest between the pairs of rows. Thus, the validity of the assumption of a T-2 matrix with 11 columns is well established. The test follows quite readily, with the final result that figure 37f becomes completed, and the recovery of both keys is a simple matter. In fact, both keys may be established from a simple study of rows 2 and 3 of the final figure (which would, of course, be identical with that shown in fig 37p and need not here be repeated).

q. A careful study and good grasp of the principles and methods elucidated in this paragraph will be sufficient to indicate to the student that when, as a result of a close study of several messages in the same keys, partial C→P sequences become available, the entire C→P sequence or sequences can usually be reconstructed from the partial sequence or sequences and the messages solved without too much difficulty. For instance, suppose it has developed that the enemy has become addicted to stereotyped sequences or to several messages can be reconstructed with some assurance of certainty. The construction of partial C→P sequences and their completion by means of the principles set forth, especially those presented in subparagraphs m-p, may result in reconstruction of the complete C→P sequences and ultimate recovery of the transposition key or keys.

28 Special cases of solution of double transposition ciphers — When the double transposition system is employed in the field and is useful for a voluminous traffic it is almost inevitable that certain situations will arise which make possible a rather easy solution. Aside from the case in which several cryptograms of identical length and in the same key are intercepted, other cases of a special nature may arise. Some of these will be discussed in this paragraph.

b. First, there is the case in which an inexperienced cryptographer clerk fails to execute the double transposition properly and causes the transmission of a cryptogram which is only a single transposition. The solution of this message will be a simple matter and will, of course, yield the key. If the key is the same for both transpositions, it is obvious that thus will permit the reading of all other messages even though the latter have been correctly cryptographed. The only difficult part of the matter is to find among a large number of intercepted cryptograms one which involves a blunder of this sort. When the cryptanalyst has, as a result of considerable experience, become adept in the solution of transposition ciphers the work of testing cryptograms to ascertain whether or not they involve single column transposition is not difficult and goes quite
rapidly For only a few minutes are sufficient to give him the "feeling" that the cryptogram is or is not solvable by single transposition. He might not be able to point out any specific indications which give him this feeling if asked to do so, nevertheless it must be recognized that his intuition is alone sufficient to tell him when there is hope of solution along this line and when further work upon the hypothesis of single transposition is useless.

e (1) Next comes the case in which the enciphering rectangles of a double transposition cryptogram happen to be perfect squares (that is, both T-1 and T-2 rectangles are perfect squares) In this case, not only is such a cryptogram detectable at once, since the total number of letters is the square of the number of elements in the key, but also the cryptogram can be solved in a very simple manner. For the cryptogram now represents a case in which a wholly-filled rectangle has been employed, and moreover there is no need even to assume various widths.

(2) Given the following cryptogram of 49 letters (7x7), as an example, the text is transcribed as shown in figure 39a and retranscribed as in figure 39b.

<table>
<thead>
<tr>
<th>Cryptogram:</th>
<th>UCTRNO GESHTOLR GASOE DWDDD NOEDE RONDRF PENCEE R N W E</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>USRRORE UXTRNOE COUNTER HOSTILE</td>
</tr>
<tr>
<td></td>
<td>CHGWEFE SHIETOL HOSTILE FORCEN</td>
</tr>
<tr>
<td></td>
<td>TIAEDREM RGAOED GEROADS COUNTER</td>
</tr>
<tr>
<td></td>
<td>RESDDN NWDDNOE WOUNDED EDONRlD</td>
</tr>
<tr>
<td></td>
<td>NTGNCH OERNDN EDONRlD GEROADS</td>
</tr>
<tr>
<td></td>
<td>GOEDOY RPFNCOE FORCEN EVENMEN</td>
</tr>
<tr>
<td></td>
<td>ELDEE FENMNYE EVENMEN WOUNDED</td>
</tr>
</tbody>
</table>

(3) The columns of figure 39b are now anagrammed, as in figure 39c, and the rows rearranged, as in figure 39d.

d (1) When the enciphering rectangle is not a perfect square but nevertheless a complete rectangle, solution of a single cryptogram becomes somewhat more difficult. Here the columns are all equal in length, since the last row of the rectangle is completely filled. Two cases will be considered, first, when the width of the rectangle is a multiple of the depth, or number of letters in the column, and second, when the depth is a multiple of the width.

e (1) Taking up the first case, note the following encipherment.

<table>
<thead>
<tr>
<th>W</th>
<th>H</th>
<th>N</th>
<th>I</th>
<th>W</th>
<th>L</th>
<th>P</th>
<th>I</th>
</tr>
</thead>
<tbody>
<tr>
<td>S</td>
<td>T</td>
<td>S</td>
<td>Q</td>
<td>A</td>
<td>D</td>
<td>R</td>
<td>O</td>
</tr>
<tr>
<td>N</td>
<td>R</td>
<td>E</td>
<td>A</td>
<td>C</td>
<td>H</td>
<td>G</td>
<td>O</td>
</tr>
<tr>
<td>V</td>
<td>I</td>
<td>L</td>
<td>L</td>
<td>E</td>
<td>T</td>
<td>O</td>
<td>N</td>
</tr>
<tr>
<td>T</td>
<td>I</td>
<td>G</td>
<td>H</td>
<td>A</td>
<td>D</td>
<td>V</td>
<td>D</td>
</tr>
</tbody>
</table>

P-C sequence: UCTRNO GESHTOLR GASOE DWDDD NOEDE RONDRF PENCEE R N W E

Cryptogram: UCTRNO GESHTOLR GASOE DWDDD NOEDE RONDRF PENCEE R N W E

If the P-C sequence is examined it will be found that sections thereof fall into two categories, as follows.

<table>
<thead>
<tr>
<th>Section</th>
<th>Key</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>24</td>
</tr>
<tr>
<td>4</td>
<td>34</td>
</tr>
<tr>
<td>5</td>
<td>44</td>
</tr>
</tbody>
</table>

Category A: 1 5 4 3 2 9 8 7 6
Category B: 2 3 4 5 6 1 7 8 9

(2) There is obviously a definite regularity in the composition of the sections whereby, if the letters corresponding to the numbers in one section can be assembled properly, all the letters corresponding to the numbers in the other sections belonging to the same category (A or B, respectively) will be assembled correctly too. For example, in category B the letters corresponding to the numbers occupying the third, first, and fifth positions in each section are sequent in the plain-text rectangle, in category A the letters corresponding to the numbers occupying the first and fourth positions in each section are sequent. Moreover, all the letters in each section come from the same row in the T-1 rectangle. Consequently, if two sections coming from the same row can be identified, there will be 10 letters which may be rearranged by experiment to form plain text, and the key for this rearrangement will apply to all other pairs of sections. For example, the message in this case has a Q and only one U. The Q (P-C sequence No 15) is in the second section, the U (P-C sequence No 16) is in the seventh section. These two sections come from the same row and the letters may be anagrammed.

<table>
<thead>
<tr>
<th>S</th>
<th>R</th>
<th>O</th>
<th>Q</th>
<th>A</th>
<th>S</th>
<th>U</th>
<th>R</th>
<th>D</th>
<th>T</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>S</td>
<td>T</td>
<td>S</td>
<td>Q</td>
<td>A</td>
<td>D</td>
<td>R</td>
<td>O</td>
<td></td>
</tr>
</tbody>
</table>

Experiment may now be made with two other sections, applying the same transposition. Thus.

<table>
<thead>
<tr>
<th>I</th>
<th>O</th>
<th>N</th>
<th>L</th>
<th>E</th>
<th>N</th>
<th>L</th>
<th>E</th>
<th>T</th>
<th>V</th>
</tr>
</thead>
<tbody>
<tr>
<td>O</td>
<td>N</td>
<td>I</td>
<td>N</td>
<td>E</td>
<td>S</td>
<td>W</td>
<td>V</td>
<td>E</td>
<td>L</td>
</tr>
</tbody>
</table>

Obviously the proper key for rearrangement is 8-9-10-1-4-7-5-9-2-3. By continuing this procedure the following additional rows of the T-1 rectangle are reconstructed.

<table>
<thead>
<tr>
<th>I</th>
<th>L</th>
<th>E</th>
<th>G</th>
<th>H</th>
<th>T</th>
<th>D</th>
<th>Y</th>
<th>S</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>S</td>
<td>E</td>
<td>A</td>
<td>V</td>
<td>G</td>
<td>D</td>
<td>I</td>
<td>H</td>
</tr>
</tbody>
</table>

The fact that the length of the sections corresponds to 6-letter groups has, of course, no bearing on the validity of the method. In this case it just happens that the rectangle contains 6 letters per column.

20\1006-61-8
The various rows are now assembled in sequence, giving the following:

WHENWILLFI
RSTSQADRO
NREACHGOLD
ENVILETTE
IGHTADVISE

The transposition key can now be reconstructed with ease.

(3) The cryptanalyst in this case must, of course, make an assumption as to the width of the enciphering rectangle before he can apply the method. With a number such as 50, the dimensions 10 x 5 or 5 x 10 suggest themselves. The process of finding cipher groups which form pairs on the same row as one of “cut and try.” If there is a single Q and a single U in the message, the initial pair of groups is obvious.

When the depth of the rectangle is a multiple of the width, solution follows along the lines of the preceding case. Taking the same message as before, note what happens in encipherment with a rectangle of 5 columns containing 10 letters each:

<table>
<thead>
<tr>
<th>1 2 3 4 5</th>
<th>1 2 3 4 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 2 3 4 5</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>6 7 8 9 10</td>
<td>6 7 8 9 10</td>
</tr>
<tr>
<td>11 12 13 14 15</td>
<td>11 12 13 14 15</td>
</tr>
<tr>
<td>16 17 18 19 20</td>
<td>16 17 18 19 20</td>
</tr>
<tr>
<td>21 22 23 24 25</td>
<td>21 22 23 24 25</td>
</tr>
<tr>
<td>26 27 28 29 30</td>
<td>26 27 28 29 30</td>
</tr>
<tr>
<td>31 32 33 34 35</td>
<td>31 32 33 34 35</td>
</tr>
<tr>
<td>36 37 38 39 40</td>
<td>36 37 38 39 40</td>
</tr>
<tr>
<td>41 42 43 44 45</td>
<td>41 42 43 44 45</td>
</tr>
<tr>
<td>46 47 48 49 50</td>
<td>46 47 48 49 50</td>
</tr>
</tbody>
</table>

It is obvious that if the 3d, 9th, 1st, 7th, and 5th columns are made sequent, good text will be produced within the 5 rows. Thus

<table>
<thead>
<tr>
<th>1 2 3 4 5 6 7 8 9 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>T T R L Q N S O S E</td>
</tr>
<tr>
<td>R S T S Q</td>
</tr>
<tr>
<td>E N O W H W D N L H C</td>
</tr>
<tr>
<td>W H E N W</td>
</tr>
<tr>
<td>E I N D C E A S R V</td>
</tr>
<tr>
<td>N R E A C</td>
</tr>
<tr>
<td>D H U I O A R T A G</td>
</tr>
<tr>
<td>U A D R O</td>
</tr>
<tr>
<td>L V I E I L F I N</td>
</tr>
<tr>
<td>I L L F I</td>
</tr>
</tbody>
</table>

The subsequent steps are obvious. Here again in solving an unknown example it would be necessary to test various assumptions with respect to the dimensions of the rectangle before attempting to apply the method outlined.

Whenever this simple relationship between the width and depth of the rectangle obtains, that is, when one dimension is a multiple of the other, solution of a single cryptogram is relatively easy. The reason for this is not hard to see. When the enciphering rectangle is a perfect square, every column of the T-2 rectangle is composed of letters which all come from the same row of the T-1 rectangle. Hence solution is as in this case the same as though a false double transposition were in effect, with merely the columns and the rows of a single rectangle shifted about. When the width of the transposition rectangle is twice the depth, a column of the T-2 rectangle contains half the letters appearing on one row of the T-1 rectangle, two columns therefore contain all the letters belonging in the same row of the T-1 rectangle. If the width were three times the depth, then three columns of the T-2 rectangle would contain all the letters belonging in the same row of the T-1 rectangle, and so on.

When the width is half the depth, a column of the T-2 rectangle contains all the letters appearing in two rows of the T-1 rectangle, when the width is one-third the depth, a column of the T-2 rectangle contains all the letters appearing in three rows of the T-1 rectangle, and so on.

But when this multiple relationship no longer obtains, solution becomes more difficult because each column of the T-2 rectangle is composed of letters coming from several columns of the T-1 rectangle, in an irregular distribution. Solution is, of course, most difficult when incompletely filled rectangles are used. However, although solvable, even in the case of a single message, the solution will not be dealt with in this text.

T-1

<table>
<thead>
<tr>
<th>P-C sequence...</th>
<th>T-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>P-C sequence...</td>
<td>T-2</td>
</tr>
</tbody>
</table>

Cryptogram......T T R L Q N S O S E E O W H W D N L H G E I N D C E A S R V

It is obvious that if the 3d, 9th, 1st, 7th, and 5th columns are made sequent, good text will be produced within the 5 rows.

<table>
<thead>
<tr>
<th>1 2 3 4 5 6 7 8 9 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>T T R L Q N S O S E</td>
</tr>
<tr>
<td>R S T S Q</td>
</tr>
<tr>
<td>E N O W H W D N L H C</td>
</tr>
<tr>
<td>W H E N W</td>
</tr>
<tr>
<td>E I N D C E A S R V</td>
</tr>
<tr>
<td>N R E A C</td>
</tr>
<tr>
<td>D H U I O A R T A G</td>
</tr>
<tr>
<td>U A D R O</td>
</tr>
<tr>
<td>L V I E I L F I N</td>
</tr>
<tr>
<td>I L L F I</td>
</tr>
</tbody>
</table>

The subsequent steps are obvious. Here again in solving an unknown example it would be necessary to test various assumptions with respect to the dimensions of the rectangle before attempting to apply the method outlined.

Whenever this simple relationship between the width and depth of the rectangle obtains, that is, when one dimension is a multiple of the other, solution of a single cryptogram is relatively easy. The reason for this is not hard to see. When the enciphering rectangle is a perfect square, every column of the T-2 rectangle is composed of letters which all come from the same row of the T-1 rectangle. Hence solution is as in this case the same as though a false double transposition were in effect, with merely the columns and the rows of a single rectangle shifted about. When the width of the transposition rectangle is twice the depth, a column of the T-2 rectangle contains half the letters appearing on one row of the T-1 rectangle, two columns therefore contain all the letters belonging in the same row of the T-1 rectangle. If the width were three times the depth, then three columns of the T-2 rectangle would contain all the letters belonging in the same row of the T-1 rectangle, and so on.

When the width is half the depth, a column of the T-2 rectangle contains all the letters appearing in two rows of the T-1 rectangle, when the width is one-third the depth, a column of the T-2 rectangle contains all the letters appearing in three rows of the T-1 rectangle, and so on.

But when this multiple relationship no longer obtains, solution becomes more difficult because each column of the T-2 rectangle is composed of letters coming from several columns of the T-1 rectangle, in an irregular distribution. Solution is, of course, most difficult when incompletely filled rectangles are used. However, although solvable, even in the case of a single message, the solution will not be dealt with in this text.

T-1

<table>
<thead>
<tr>
<th>P-C sequence...</th>
<th>T-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>P-C sequence...</td>
<td>T-2</td>
</tr>
</tbody>
</table>

Cryptogram......T T R L Q N S O S E E O W H W D N L H G E I N D C E A S R V

It is obvious that if the 3d, 9th, 1st, 7th, and 5th columns are made sequent, good text will be produced within the 5 rows.

<table>
<thead>
<tr>
<th>1 2 3 4 5 6 7 8 9 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>T T R L Q N S O S E</td>
</tr>
<tr>
<td>R S T S Q</td>
</tr>
<tr>
<td>E N O W H W D N L H C</td>
</tr>
<tr>
<td>W H E N W</td>
</tr>
<tr>
<td>E I N D C E A S R V</td>
</tr>
<tr>
<td>N R E A C</td>
</tr>
<tr>
<td>D H U I O A R T A G</td>
</tr>
<tr>
<td>U A D R O</td>
</tr>
<tr>
<td>L V I E I L F I N</td>
</tr>
<tr>
<td>I L L F I</td>
</tr>
</tbody>
</table>

The subsequent steps are obvious. Here again in solving an unknown example it would be necessary to test various assumptions with respect to the dimensions of the rectangle before attempting to apply the method outlined.

Whenever this simple relationship between the width and depth of the rectangle obtains, that is, when one dimension is a multiple of the other, solution of a single cryptogram is relatively easy. The reason for this is not hard to see. When the enciphering rectangle is a perfect square, every column of the T-2 rectangle is composed of letters which all come from the same row of the T-1 rectangle. Hence solution is as in this case the same as though a false double transposition were in effect, with merely the columns and the rows of a single rectangle shifted about. When the width of the transposition rectangle is twice the depth, a column of the T-2 rectangle contains half the letters appearing on one row of the T-1 rectangle, two columns therefore contain all the letters belonging in the same row of the T-1 rectangle. If the width were three times the depth, then three columns of the T-2 rectangle would contain all the letters belonging in the same row of the T-1 rectangle, and so on.

When the width is half the depth, a column of the T-2 rectangle contains all the letters appearing in two rows of the T-1 rectangle, when the width is one-third the depth, a column of the T-2 rectangle contains all the letters appearing in three rows of the T-1 rectangle, and so on.

But when this multiple relationship no longer obtains, solution becomes more difficult because each column of the T-2 rectangle is composed of letters coming from several columns of the T-1 rectangle, in an irregular distribution. Solution is, of course, most difficult when incompletely filled rectangles are used. However, although solvable, even in the case of a single message, the solution will not be dealt with in this text.

T-1

<table>
<thead>
<tr>
<th>P-C sequence...</th>
<th>T-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>P-C sequence...</td>
<td>T-2</td>
</tr>
</tbody>
</table>

Cryptogram......T T R L Q N S O S E E O W H W D N L H G E I N D C E A S R V

It is obvious that if the 3d, 9th, 1st, 7th, and 5th columns are made sequent, good text will be produced within the 5 rows.

<table>
<thead>
<tr>
<th>1 2 3 4 5 6 7 8 9 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>T T R L Q N S O S E</td>
</tr>
<tr>
<td>R S T S Q</td>
</tr>
<tr>
<td>E N O W H W D N L H C</td>
</tr>
<tr>
<td>W H E N W</td>
</tr>
<tr>
<td>E I N D C E A S R V</td>
</tr>
<tr>
<td>N R E A C</td>
</tr>
<tr>
<td>D H U I O A R T A G</td>
</tr>
<tr>
<td>U A D R O</td>
</tr>
<tr>
<td>L V I E I L F I N</td>
</tr>
<tr>
<td>I L L F I</td>
</tr>
</tbody>
</table>
CHAPTER VI

PRINCIPLES OF MATRIX RECONSTRUCTION

Section 1

29 Special designs or geometric figures

It is impossible here to elucidate and demonstrate by example all the methods which may be used for the solution of cryptograms produced by the various types of transposition designs or geometric figures other than the simple rectangular ones thus far treated. Reference may be made to such matrices as triangles, trapezoids, and polygons of various symmetrical shapes. Most of these matrices, however, are unpractical for military correspondence in any case, so that no attention need be given them in this text.

If such designs were used, although it might be difficult to solve a single or even a few messages in the same key, the general solution described in paragraph 26 is applicable whenever two or more messages of identical length but in the same key are available for study. Since most of these designs are of a fixed or inflexible character with regard to the number of letters that can be accommodated with one application of the design to the plain text to be enciphered, the production of several cryptograms of identical length in the same key is by no means an unusual circumstance. The general solution can usually be depended upon to yield the answer to cryptograms of this category but it then becomes advisable to try to ascertain the exact nature of the specific design or geometric figure employed, that is, to reconstruct the transposition matrix.

For this purpose a general method will be indicated by means of a specific example, leaving other cases to the ingenuity of the student after he has learned the general method.

30 Reconstruction of transposition matrix

A. Assumed that the enemy is employing an unknown geometric figure of rather small dimensions so that it appears from a study of the traffic that it accommodates a maximum of 85 letters. A long cryptogram has been intercepted and it is broken up into sections of 85 letters, which sections are then superimposed, as shown below.

It will be noted that there are 3 complete sections of 85 letters each, plus a final section of but 49 letters.

<table>
<thead>
<tr>
<th>12345678910</th>
<th>1112131415161718192021</th>
<th>222324252627282930</th>
</tr>
</thead>
</table>

The final section will be dealt with later.

| 1 T D N P R A O I S J P E R O C E E A R Y O L E P T L TH A V N |
| 2 W R T D L U O S P C I N N T U I N M O X L H O N P A T S I F |
| 3 M A I S V I T S O T H L T E E R I O V I Y W N G P E O I |
| 4 G R U T S O E B M L R M O G E T C N N D Y E E H T Q C N T |
| 5 A N N C T S Y O A A C E M E N I E I B I H A D E X T C U T R |
| 6 W D H E B R N D T T D I Y A F A D A G R D O E C A A J T R E |
| 7 A T U A C O D F O B I M N R T I N E S H O Y N F L I H N R O |

From this skeleton of what may be termed the matrix-reconstruction diagram it is possible to derive direct clues for the continuance and completion of the C-P sequence and the text of the message. For example, it would appear that the very next column to the left should be 78, the one to the left of 78 should be 68, the one to the left of 68 should be 9. Trial gives the following.


Examining the numbers forming this partial C-P sequence, note the following sections of the sequence:


They show a quite definite relationship, leading to the suspicion that the C-P sequence is systematic in its composition. The numbers are then written down on a cross-section paper so that consecutive numbers appear on the same level, as shown in figure 40-A.

The anagramming process is applied to the superimposed complete sections, using the letter J in the first section as a starting point and building up text on either side, until the following partially reconstructed text is obtained:


Examining the numbers forming this partial C-P sequence, note the following sections of the sequence. They show a quite definite relationship, leading to the suspicion that the C-P sequence is systematic in its composition. The numbers are then written down on a cross-section paper so that consecutive numbers appear on the same level, as shown in figure 40-A.


Text


Matrix reconstruction diagram


Examiner the numbers forming this partial C-P sequence, note the following sections of the sequence. They show a quite definite relationship, leading to the suspicion that the C-P sequence is systematic in its composition. The numbers are then written down on a cross-section paper so that consecutive numbers appear on the same level, as shown in figure 40-A.


Examining the numbers forming this partial C-P sequence, note the following sections of the sequence. They show a quite definite relationship, leading to the suspicion that the C-P sequence is systematic in its composition. The numbers are then written down on a cross-section paper so that consecutive numbers appear on the same level, as shown in figure 40-A.


Examining the numbers forming this partial C-P sequence, note the following sections of the sequence. They show a quite definite relationship, leading to the suspicion that the C-P sequence is systematic in its composition. The numbers are then written down on a cross-section paper so that consecutive numbers appear on the same level, as shown in figure 40-A.


Examining the numbers forming this partial C-P sequence, note the following sections of the sequence. They show a quite definite relationship, leading to the suspicion that the C-P sequence is systematic in its composition. The numbers are then written down on a cross-section paper so that consecutive numbers appear on the same level, as shown in figure 40-A.


Examining the numbers forming this partial C-P sequence, note the following sections of the sequence. They show a quite definite relationship, leading to the suspicion that the C-P sequence is systematic in its composition. The numbers are then written down on a cross-section paper so that consecutive numbers appear on the same level, as shown in figure 40-A.


Examining the numbers forming this partial C-P sequence, note the following sections of the sequence. They show a quite definite relationship, leading to the suspicion that the C-P sequence is systematic in its composition. The numbers are then written down on a cross-section paper so that consecutive numbers appear on the same level, as shown in figure 40-A.


Examining the numbers forming this partial C-P sequence, note the following sections of the sequence. They show a quite definite relationship, leading to the suspicion that the C-P sequence is systematic in its composition. The numbers are then written down on a cross-section paper so that consecutive numbers appear on the same level, as shown in figure 40-A.


Examining the numbers forming this partial C-P sequence, note the following sections of the sequence. They show a quite definite relationship, leading to the suspicion that the C-P sequence is systematic in its composition. The numbers are then written down on a cross-section paper so that consecutive numbers appear on the same level, as shown in figure 40-A.


Examining the numbers forming this partial C-P sequence, note the following sections of the sequence. They show a quite definite relationship, leading to the suspicion that the C-P sequence is systematic in its composition. The numbers are then written down on a cross-section paper so that consecutive numbers appear on the same level, as shown in figure 40-A.
This, of course, speeds up the work involved in the anagramming process and when completed the text, the C→P sequence, the P→C sequence, and the matrix reconstruction diagram are as shown in figure 40-B. In the cells of the diagram there have been inserted in the upper left hand corner small numbers in italics, the latter numbers being merely the term numbers applying to the C→P sequence.

The matrix-reconstruction diagram in figure 40-B shows a total of 7 levels of numbers. Let the term numbers corresponding to the consecutive C→P sequence numbers on the same level in the diagram be set down. Thus, for the C→P sequence numbers 4 to 16, inclusive, on the first level the term numbers are

<table>
<thead>
<tr>
<th>C→P sequence number</th>
<th>4 5 6 7 8 9 10 11 12 13 14 15 16</th>
</tr>
</thead>
<tbody>
<tr>
<td>Term number</td>
<td>1 3 7 13 21 31 43 55 65 73 79 83 85</td>
</tr>
</tbody>
</table>

On the second level there are two sets of consecutive C→P sequence numbers, those from 48 to 58, inclusive forming one set, those from 64 to 74, inclusive forming the other set. Two series of term numbers are therefore derived

<table>
<thead>
<tr>
<th>C→P sequence number</th>
<th>48 49 50 51 52 53 54 55 56 57 58</th>
</tr>
</thead>
<tbody>
<tr>
<td>Term number</td>
<td>2 6 12 20 30 42 54 64 72 78 82</td>
</tr>
</tbody>
</table>

What has been said of the 2d level applies also to the remaining levels, and the term numbers are therefore set down in the following tabular form

<table>
<thead>
<tr>
<th>C→P sequence numbers to which applicable</th>
</tr>
</thead>
<tbody>
<tr>
<td>(4-16)</td>
</tr>
<tr>
<td>(48-58)</td>
</tr>
<tr>
<td>(64-74)</td>
</tr>
<tr>
<td>(24-32)</td>
</tr>
<tr>
<td>(75-83)</td>
</tr>
<tr>
<td>(17-23)</td>
</tr>
<tr>
<td>(38-44)</td>
</tr>
<tr>
<td>(59-63)</td>
</tr>
<tr>
<td>(33-37)</td>
</tr>
<tr>
<td>(1-3)</td>
</tr>
<tr>
<td>(45-47)</td>
</tr>
<tr>
<td>(85)</td>
</tr>
<tr>
<td>(84)</td>
</tr>
</tbody>
</table>

There are in all 13 sets or series of consecutive C→P sequence numbers, indicating that the transposition matrix has 13 columns, the number of letters in each column corresponding with the number of different terms in each series. Thus, there is a column of 13 letters, 2 columns of 11 letters, 2 columns of 9 letters, and so on. This leads directly to the idea of a very symmetrical matrix of the form shown in figure 42-A.
<table>
<thead>
<tr>
<th>Term number</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>C→P sequence</td>
<td>4</td>
<td>48</td>
<td>36</td>
<td>34</td>
<td>42</td>
<td>24</td>
<td>14</td>
<td>97</td>
<td>86</td>
<td>27</td>
</tr>
<tr>
<td>Plan text</td>
<td>FIRST BATTALION FOURTH INFANTRY HAS REACHED ROAD ADJU</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Matrix reconstruction diagram</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Term number</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C→P sequence</td>
<td>38</td>
<td>30</td>
<td>38</td>
<td>30</td>
<td>38</td>
<td>30</td>
<td>38</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Placing of the text is not clear due to the image quality.
The recovery of the transposition key (for the columns of figure 42-A) is now a simple matter. Referring to the P→C sequence shown in figure 40-B, and noting the various columns in figure 42-A in which successive numbers of the P→C sequence fall, the key number 1 of the transposition key obviously applies to the column containing P→C sequence numbers 26-38-50; the key number 2 obviously applies to the column containing P→C sequence numbers 1-3-7-13-21-31-43-55-65-73-79-83-85; and so on. The complete transposition key and the matrix are shown in figure 42-B.
1. The solution of the incomplete section of the message (the last 79 letters) now becomes a simple matter, since the matrix and the transposition key are both known. The matter can be handled as if simple transposition were involved, by outlining the matrix to contain exactly 79 letters and inscribing the cipher text in the columns in key-number order.

c. The foregoing principles and procedure will be found quite valuable not only in facilitating the anagramming of the text in its initial stages (as in subparagraph d) but also in reconstructing various types of matrices based upon symmetrical designs used with single transposition (subparagraphs e-g). It should be noted that the number of leads in the reconstruction diagram corresponds with the number of different-length columns in the matrix; the number of different categories of term numbers (as in figure 41) corresponds with the number of columns in the matrix.

84

13 1 9 3 4 8 2 10 11 6 5 7 12
1 2 3 4
5 6 7 8 9
10 11 12 13 14 15 16
17 18 19 20 21 22 23 24 25
26 27 28 29 30 31 32 33 34 35 36
37 38 39 40 41 42 43 44 45 46 47 48 49
50 51 52 53 54 55 56 57 58 59 60
61 62 63 64 65 66 67 68 69
70 71 72 73 74 75 76
77 78 79 80 81
82 83 84
85

FIGURE 42-B.

SECTION VII

SOLUTION OF GRILLES

Revolution grilles... 31
Solution of example... 32
Concluding remarks on the solution of revolving grilles... 33
Indecipherable or continuous grilles... 34

51. Revolving grilles.—a. In this type of grille the apertures are distributed among the cells of a square sheet of cross-section paper in such a manner that when the grille is placed upon a grid (a sheet of cross-section paper of the same size as the grille) certain cells of the grid are disclosed; then when the grille is turned three times successively through angles of 90° from an initial position upon the grid, all the remaining undislosed grid cells (for all but the central cell) are disclosed in turn. Correspondents must, of course, possess identical grilles and they must have an understanding as to its initial position and direction of rotation, clockwise or counterclockwise. There are two procedures possible in using such a grille. (1) The letters of the plain text may be inscribed successively in the grid cells through the apertures of the grille; when the grid has been completely filled the grille is removed and the letters transcribed from the grid according to a prearranged route. (2) All the letters of the plain text may first be inscribed in the grid cells according to a prearranged route and then the grille applied to the completely-filled grid to give the sequence of letters forming the cryptogram. The two methods of using the grille are reciprocal; if the first-described method is used to encipher a message, the second is used to decipher the cryptogram, and vice versa. The first of the two above-described methods, the one in which the plain text is inscribed through the apertures, will here be referred to as the alpha method; the second method will be referred to as the beta method.

b. The number of letters in a cryptogram enciphered by such a device is either a perfect square, when the grille has an even number of cells per side, or is 1 less than a perfect square, when the grille has an odd number of cells per side, in which case the central cell of the grid is not disclosed and hence remains unfilled.

c. The manner of construction and the method of use of a grille entails certain consequences which can be employed to solve the cryptograms and to reconstruct the grille itself. The student who wishes to get a thorough grasp of the underlying principles to be explained will do well to prepare a grille and study the properties which characterize cryptograms produced by its use. Three principles will be brought to bear in the solution of grille ciphers of this type and they will be demonstrated by reference to the grille and message shown in figure 43.

1 See Special Text No. 166, Advanced Military Cryptography, sec. V.

2 Of course, the cryptograms may consist of the letters produced by several applications of the same grille. For example, if a message of 170 letters is to be enciphered by a grille accommodating only 36 letters a time, the message is divided up into sections of 36 letters each (10 nulls being added to make the total a multiple of 36). The total number of letters (180) here shows no properties of the type noted. Again, if the grille has a capacity greater than the number of letters to be enciphered, certain of the grid cells may be cancelled, so that the number of letters in the final cryptogram will not be a perfect square or 1 less than a perfect square.

3 Detailed instructions for the construction of revolving grilles will be found in Special Text No. 166, Advanced Military Cryptography, sec. V.
YOUR LINES TO THIS COMMAND POST CUT BY SHELL FIRE REQUEST YOU CHANGE THE ROUTE.

Grille: 8x8.

The first principle may be termed that of symmetry. When a revolving grille is in position 1 a certain number of cells of the underlying grid are disclosed (uncovered). For each such disclosed cell of the grid there is a symmetrically-corresponding cell on the same grid which is disclosed when the grille is turned to positions 2, 3, and 4, because the apertures of the grille remain fixed—only their positions change as the grille is turned in the process of encipherment. Now two successive apertures in position 1 will, of course, be occupied by a plain-text digraph (alpha method of encipherment). When the grille reaches position 3, after a turn of 180°, the two
apertures concerned will disclose two cells which will also be occupied by a plain-text digraph, but the letters comprising the digraph will be in reverse order in the plain text. This property is true also of two successive apertures in position 2 when they turn up in position 4. Let the student verify this by means of the grille which he has constructed. Thus, referring to figure 43, at A is shown the grille in position 1. In the first row are shown 2 apertures, at coordinates 1-4 and 1-8. At B are shown the results of the first application of the grille to the grid. Note the letters Y O (first 2 letters of messages) in cells 4 and 6. Now note that the symmetrical-corresponding cells disclosed when the grille is in position 3 are cells 67 and 61 and these correspond to cells 4 and 8 in the reverse order. The letter T in cell 57 therefore symmetrically corresponds with letter O in cell 8; the letter Y in cell 61 corresponds with letter Y in cell 4. The same is true of all other letters in positions 1 and 3. As a consequence of this property of grilles, a single cryptogram can be handled as though it were really two cryptograms of identical length having certain characteristics by means of which an assumption made in one text may be verified by what it yields in the other text. That is, when the cryptogram is transcribed as a series of letters in one line and the same text is written in another line under these letters but in reversed order, then the superimposed letters will bear the symmetrical relationship pointed out in this paragraph. If two letters in the upper line of such a transcription are taken to form a digraph, the two corresponding letters in the lower line must form a digraph but in reverse order in the plain text. For example, if the cryptogram of figure 43 is written out as explained above, the results are as shown at figure 44. Now the presence of the Q in position 39 suggests that it be combined with the U. If the U in position 43 is taken, then the symmetrical digraph corresponding to U would be L I; if the U in position 56 is taken, the symmetrical-corresponding digraph would be W I. Furthermore, two apertures which are in the same column and which do not have an intervening aperture between them, will yield a good digraph in all 4 positions of the grille. For example, note apertures 2-5 and 3-6 at A in figure 43. When the grille is turned to positions 2, 3, and 4 they will disclose two sequential letters in each case. An analysis of the symmetries produced by an 8×8 grille yields the following table, which shows what cells are disclosed in the other 3 positions when an aperture is cut in any one cell in 1 of the 4 positions of the grille. For example, an aperture cut in cell 11 (position 1) will disclose grid-cell 23 when the grille takes position 3, grid-cell 54 when the grille takes position 3, and grid-cell 42 when the grille takes position 4.

### Figure 44

<table>
<thead>
<tr>
<th>Position</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
</tr>
</tbody>
</table>

### Figure 45

<table>
<thead>
<tr>
<th>Position</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
</tbody>
</table>

The second principle may be termed that of exclusion. On account of the system upon which the construction of a revolving grille is based, a knowledge of the location of an aperture in one of the bands brings with it a knowledge of 3 other locations in which there can be no apertures. For example, referring to A in figure 43, the presence of the aperture at coordinates 1-4 precludes the presence of apertures at coordinates 4-8, 5-9, and 5-1. By virtue of this principle of exclusion, the number of possibilities for choice of letters in solving a cryptogram prepared by means of a revolving grille becomes much reduced and the problem is correspondingly simplified, as will be seen presently.

j. The third principle may be termed that of sequence. When trying to build up text, the letters which follow a given sequence of plain-text letters will usually be found to the right and below, that is, if the normal method of writing was used (left to right and from the top downward). For example, referring to figure 44, if the trigraph Q U E is to be built up, neither the U in position 5 nor the U in position 10 is very likely to be the one that follows the Q; the U in position 43 is the most likely candidate because it is the first one beyond the Q. Suppose the U is in position 43 is selected. Then the E for Q U E cannot be the one in position 40, or in any position in front of 40, since the E must be beyond the U in the diagram.

k. In solving a grille, it will be found advisable to prepare a piece of cross-section paper of proper grille size for the grille inscribed at any location in the grille becomes quite definite. In this way not only will the problem be simplified but also when completed the proper grille is at hand.

**Solution of example.** Suppose the cryptogram shown in figure 43 is to be solved. It has 64 letters, suggesting a grille 8×8. The cryptogram is first transcribed into a square 8×8, yielding what has already been obtained at F in figure 43. The Q in position 39 suggests that it is part of a word inscribed when the grille was in position 3, since there will be 16 plain-text letters inscribed at each position of the grille. Then a piece of cross-section paper is prepared for making the grille as shown in figure 45-A, and an aperture is cut in the proper position to disclose, in position 3, cell 39. It will be found that this is the aperture located at coordinates 3-2 of the grille shown in figure 45-A. At the same time the other 3 are numbered 1, 6, and 7.

![Diagram](attachment:image.png)
All of the symmetrical correspondents of these 3 Q U’s are good digraphs and it is impossible to eliminate any of the three possibilities. The U in position 49 would place an aperture at coordinates 6-0; the U in position 46 would place an aperture at coordinates 6-0; and the U in position 66 would place an aperture at coordinates 7-8. All of these are possible, none being excluded by principle 2. Suppose the Q U is followed by E. There are only two possibilities: an E in cell 61 and E in cell 63. The following possibilities are presented:

<table>
<thead>
<tr>
<th>22</th>
<th>26</th>
<th>30</th>
<th>34</th>
<th>38</th>
<th>42</th>
<th>46</th>
<th>50</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q</td>
<td>Q</td>
<td>Q</td>
<td>Q</td>
<td>Q</td>
<td>Q</td>
<td>Q</td>
<td>Q</td>
</tr>
<tr>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
</tr>
</tbody>
</table>

I (or L I) I (or E I) I A (or A I) I M (or M I)

None of the symmetrical correspondents of the Q U E’s are impossible sequences in plain text, although O A I is not as probable as the others. The O could be the end of a word, the A the beginning of the word AID, AIM, AIR, etc.) Each of these possibilities would be tested by principle 2 to see if any conflicts would arise as to the positions of apertures. As in all cases of transposition ciphers, the most difficult part of the solution is that of forcing an entering wedge into the structure and getting a good start; when this has been done the rest is easy. Note what the results are when the proper apertures are assumed for QUIST in this case, as shown in figure 46-C. In position 1 this yields QUI LI ... ; in position 3 it yields two digraphs AN and UT; in position 4 it yields two digraphs HA and RO. The student should note that the indicated digraphs A N and R O in positions 2 and 4, respectively, are certain despite the fact that there is a space between the two apertures disclosing these letters, for the principle of exclusion has permitted the crossing off of this cell as a possibility for an aperture.

---

8. Enough has been shown of the procedure to make further demonstration unnecessary. Given the sequence QUI LI one begins to build on that, assuming a word such as LINE. This yields possibilities for the placement of additional apertures in the grille; these are tested in positions 2, 3, 4, and so on. When any 16 consecutive letters of plain text have been established all apertures have been ascertained and the problem has been completed. Subsequent cryptograms prepared by the same grille can be read at once.

---

91

33. Concluding remarks on the solution of revolving grilles.—a. There is nothing about the mechanics of revolving grilles which prevents their employment in enciphering complete words instead of individual letters. However, the assembling of whole words in intelligible sequences and thus the reconstruction of the original plain text is a much easier matter than assembling single letters to form the words of the original plain text.

b. In case the same grille has been employed several times with separate grids to encipher a message that is considerably longer than a single grid will accommodate (see footnote 2, par. 318), the several sections each representing the set of letters enciphered on one grid may be super-imposed and the general solution described in paragraph 26 may then be applied.

c. In case the capacity of a grille is in excess of the number required by the length of the text to be enciphered, either of two procedures may be agreed upon. The grid cells which would otherwise be unoccupied may be filled by nulls, or the grid may be left incomplete. As in all cases of partial cryptograms, the former procedure, little more need be said than that the presence of a few nulls will only delay solution a bit until the fact that nulls are being employed for this purpose becomes established. But the second type of procedure calls for more comment. If the grid is to be left incomplete it is necessary, before applying the grille, to count the number of plain-text letters and to cancel from the grid a number of cells equal to the number of cells in excess of the total number required. The position of the cells to be cancelled must be agreed upon: commonly, they are those at the end of the grid. Such cells are marked so that when they become exposed during the rotations of the grille they will not be used. Thus, for example, the grille shown in figure 43-A is intended for a grid of 64 letters; if the message to be enciphered contains only 53 letters, 12 cells of the grid must be cancelled, and by agreement they may be cells 53 to 64, inclusive.

The solution of a single cryptogram of this sort, or even of several of them of different lengths, may become a rather difficult matter. First of all, clues as to the dimensions of the grille are no longer afforded by the total number of letters in the cryptogram, so that this information can be obtained only by more or less laborious experimentation. Grilles of various dimensions must be assumed, one after the other, until the correct dimensions have been found. In the second place, the symmetrical relationships pointed out in paragraph 31 no longer obtain, so that a single cryptogram cannot be handled as though it were constituted of two messages of identical length. Of course, in trying out any assumed dimensions, the 64 letters of the cryptogram may be written out in two superimposed lines, blanks being left for those positions which are unfillable. The procedure then follows the normal lines. About the most hopeful clues would be obtained from a knowledge of the circumstances surrounding the transmission and affording a basis for the assumption of probable words. However, were such a system employed for regular communication there would undoubtedly be cases of cryptograms of identical lengths, so that the type of solution given in paragraph 26 will be applicable. Once a solution of this sort has been obtained, the dimensions of the grille may be ascertained. Subsequent cryptograms may then be attacked on the basis of the normal procedure, with such modifications as are indicated by the absence of the number of letters needed to make a completely-dilled grid.

34. Indefinite or continuous grilles.—a. In his Manual of Cryptography, Sacco illustrates a type of grille which he has devised and which has elements of practical importance. An example of such a grille is shown in figure 46. This grille contains 20 columns of cells, and each column contains 5 apertures distributed at random in the column. There are therefore 100 apertures in all, and this is the maximum number of letters which may be enciphered in one position of the grille. The plain text is inscribed serially, from left to right, using only as many columns as may be necessary to inscribe the complete message. A 25-letter message would require...
but 5 columns. To form the cryptogram the letters are transcribed horizontally from the rows, taking the letters from left to right as they appear in the apertures. If the total number of letters is not a multiple of 5, sufficient nulls are added to make it so. In decryptography, the total number of letters is divided by 5, this giving the number of columns employed. The cipher text is inscribed from left to right and top downwards in the apertures in the rows of the indicated number of columns and the plain text then reappears in the apertures in the columns, reading downward and from left to right. (It is, of course, not essential that nulls be added in encryption to make the length of the cryptogram an exact multiple of 5, for the matter can readily be handled even if this is not done. In decipherment the total number of letters divided by 5 will give the number of complete columns; the remainder left over from the division will give the number of cells occupied by letters in the last column on the right.)

b. Such a grille can assume 4 positions, two obverse and two reverse. Arrangements must be made in advance as to the sequence in which the various positions will be employed.

c. The solution of a single cryptogram enciphered by one and only one position of such a grille presents a practically hopeless problem, for the apertures being distributed at random throughout the grille there is nothing which may be seized upon as a guide to the reconstruction of either the grille or the plain text. It is conceivable, of course, that a person with an infinite amount of patience could produce an intelligible text and a grille conformable to that text, the grille having a definite number of columns and a fixed number of apertures distributed at random throughout the columns. But there would be no way of proving that the plain text so obtained is the actual plain text that was enciphered; for it would be possible to produce several "solutions" of the same character, any one of which might be correct.1

d. However, suppose a grille of this sort were employed to encipher a long message, requiring two or more applications of the grille. For example, in the case of the grille shown in figure 46, having a capacity of 100 letters per application, suppose a message of 400 letters were to be enciphered, requiring two obverse and two reverse applications of the grille. It is obvious that symmetrical relationships of the nature of those pointed out in paragraph 21 can be established. Of course, if the grille is used several times in the same position to its full capacity, producing cryptograms of multiples of 100 letters, then the sections of 100 letters may be superimposed and the general solution elucidated in paragraph 26 applied.

1 In this connection, see Military Cryptanalysis, Part III, sec. XI, footnote 8.

e. If the grille shown in figure 46 were used to encipher two messages, one of 80 letters, the other of 85, it would be possible to solve these messages. For by eliminating 5 letters from the longer message, the two cryptograms can be superimposed and handled as in paragraph 26. The difficulty would be in finding the 5 extra letters. Of course, if it should happen that one of the messages required 3 or 4 nulls and letters such as J, X, or Z were employed for this purpose, the nulls would be likely characters for elimination. But regardless of this, even if letters of medium or high frequency were used as nulls, patient experimentation would ultimately lead to solution. The latter, it must be conceded, would be difficult but not impossible.
SECTION VIII

COMBINED SUBSTITUTION-TRANSPOSITION SYSTEMS

35. Reasons for combining transposition with substitution. — a. Transposition methods are, from the cryptographic point of view, rather highly regarded because they are, as "hand methods," rather rapid in operation and usually quite simple. However, from their very nature they entail the disadvantage that a single-letter omission or addition may render their decryptographing difficult if not impossible for the average cryptographic clerk. But from the standpoint of modern cryptography the principal disadvantage of transposition methods is that they can be mechanized only with great difficulty—certainly with greater difficulty than is the case of substitution methods. Only one or two attempts have been made to produce machinery for effecting transposition, and these have not been successful.

b. Pure transposition, that is, transposition by itself, without an accompanying substitution or other means of disguise for the letters of the plain text, hardly affords sufficient guarantees for cryptographic security in the case of a voluminous correspondence which must be kept really secret for any length of time. For no matter how simple the method, or how many transpositions may be applied to the letters of a single message, sight must never be lost of the fact that when there are many messages in the same key there are bound to be two or more of identical length; and when this is the case the type of solution described in paragraph 26 may be applied to the letters of a single message, sight must never be lost of the fact that when there are many messages in the same key there are bound to be two or more of identical length; and when this is the case the type of solution described in paragraph 26 may be applied to the letters of a single message, sight must never be lost of the fact that when there are many messages in the same key there are bound to be two or more of identical length; and when this is the case the type of solution described in paragraph 26 may be applied to these cryptograms, the transposition keys recovered, and then all other messages in the same key translated.

c. A message may undergo monoalphabetic substitution and the resulting text passed through a simple transposition. When this is the case a uniliteral frequency distribution will, of course, exhibit all the characteristics of monoalphabeticity, and the cryptogram will resist all attempts at solution according to straightforward simple substitution principles. It is usually not difficult to detect that a transposition is involved because there will not only be long strings of low-frequency letters or high-frequency letters but what is more important, there will be very few or no repetitions of digraphs, trigraphs, and tetragraphs, since these will be broken up by the transposition. When a uniliteral distribution presents all the external evidences of monoalphabeticity and yet there are no repetitions, it is almost a positive indication of the presence of transposition superimposed upon the substitution, or vice versa. (The former is usually the case.)
d. When confronted with such a situation the cryptanalyst usually proceeds by stages, first eliminating the transposition and then solving the substitution. It is of course obvious that the general solution for transposition ciphers (cryptograms of identical length in the same key) will not be applicable here, for the reason that such a solution is based upon anagramming, which in turn is guided by the disclosure of good digraphs, trigraphs, and polygraphs. Since the letters of a combined substitution-transposition cipher are no longer the same as the original plain-text letters, simple anagramming of columns formed by superimposing identical-length ciphertexts can yield no results, because there is nothing of the nature of plain text to guide the cryptanalyst in his juxtaposition of columns.4

e. Of course, if it should happen that the substitution process involves known alphabets, the cryptanalyst can remove the effects of the substitutive process before proceeding to eliminate the transposition, even if in the encipherment the substitution came first. For example, if a standard cipher alphabet were employed for the substitution the uniliteral frequency distribution would give indications thereof and the cipher letters could immediately be converted to the normal plain-text equivalents. The latter may then be studied as though merely transposition had been applied. But if unknown mixed cipher alphabets were employed, this initial step can not be accomplished and a solution must usually wait upon the removal of the transposition before the substitution can be attacked. The latter may be very difficult or impossible where a good transposition method is used; where simple columnar transposition is used the removal of the transposition can be effected if the message is long enough.

f. Of course if nothing is known about the system of transposition that has been employed, there is hardly anything to do but experiment with various types of transposition in an attempt to bring about such an arrangement of the text as will show repetitions. If this can be done, then the problem can be solved. For example, suppose that a message has been enciphered by a single mixed cipher alphabet and the substitution text has then been inscribed within a rectangle of certain dimensions according to one of the usual routes mentioned in paragraph 5 of this text. Repetitions in the plain text will of course be preserved in the substitution text but will be destroyed after the transposition has been applied. The cryptanalyst, however, in his attempts to eliminate the transposition, may experiment with route transpositions of the various types, employing rectangles of various dimensions as suggested by the total number of letters in the cryptogram. If he perseveres, he will find one route which he will know is correct as soon as he tries it because it will disclose the repetitions in the plain text, although the latter are still covered by a substitution.

g. Practically all the methods of transposition which may be applied to plain text may also be applied to a text resulting from an initial transformation by substitution. As already mentioned, route transposition may be used; reversed and rail-fence writing, columnar transposition with or without keying and with complete or incomplete rectangles are also possible. From a practical standpoint, keyed-columnar transposition applied to a monoalphabetic substitution is not only a popular but also a fairly secure combination because in this case the elimination of the transposition is a rather difficult matter. If the rectangle is completely filled the problem is not insurmountable in the case of a long message transposed by means of transposition with a rectangle of fairly small dimensions. For by assuming rectangles of various dimensions suggested by the total number of letters, cutting the columns apart, and then combining columns on the basis of the number of repetitions produced within juxtaposed columns and between different sets of juxtaposed columns, it is possible to reconstruct the rectangle and thus remove the transposition phase. This, however, is admittedly a slow and difficult process even under the most favorable conditions; and if the rectangle is incompletely filled the process is very difficult. For in the latter case the lack of absolutely clear-cut knowledge as to the lengths of the columns, the juxtaposition of columnar material becomes replete with uncertainties and engenders feelings of confusion, hopelessness, and inadequacy in the mind of the cryptanalyst. However, he need

4 It should, however, not be inferred that anagramming processes are entirely excluded in the cryptanalysis of all combined substitution-transposition systems. In certain cases the anagramming process may be guided by considerations of frequency of letters or fragments of letters. A case of this kind will be encountered in the solution of the ADFGVX cipher, par. 60.
38. Other types of combined substitution-transposition systems.—a. There is no technical
difficulty in the application of a transposition to the text resulting from any type of substitution,
even if the latter is polyalphabetic or polygraphic in nature. The obstacles, or rather objections,
to such combinations are practical in their character—they are too complex for ordinary use and
the prevalence of errors makes them too difficult to handle, as a general rule. However, they
have been and are sometimes used even as field ciphers. For instance, on the southeastern front
during the World War, the Central Powers made use of a somewhat irregular polyalphabetic sub-
istitution involving four standard alphabets and a keyed columnar transposition with incom-
pletely-filled rectangles of a relatively large number of columns. Nevertheless, messages in this
system were solved by taking advantage of the possibility of devising special solutions.

b. A few remarks may be made in regard to the order in which the two processes, sub-
istitution and transposition, are employed in a combined system. It is clear that when the
substitution is monoalphabetic it is immaterial, so far as cryptographic security is concerned,
whether substitution is followed by transposition or vice versa, because the equivalent of each
plain-text letter remains fixed regardless of the order in which the plain-text letters appear in
the plain text. However, if the substitution is polyalphabetic in character it is better that the
transposition process precede the substitution process, and that the number of alphabets em-
ployed be different from the number of elements in the transposition key, if columns trans-
position is the case. The best situation, from a cryptographic security standpoint, is when
the two key lengths (substitution and transposition) have no common factor. If the two keys
are of the same length, the letters in each column are enciphered by the same cipher alphabet
and thus the cryptogram would contain a certain number of sections of approximately equal
length, composed of letters falling in the same cipher alphabet.
c. Digraphic substitution, such as that produced by the Playfair Cipher, may be combined
with transposition to yield cryptograms of fair security. But here again the alternation of the
transposition phase by taking advantage of special circumstances or by rearranging the text so
as to uncover the repetitions which are inevitable in the Playfair Cipher, will result in solution.
d. A particularly fruitful source of combined substitution-transposition is to be found in those
methods generally designated as fractionating systems, in which in the substitution phase each
plain-text letter is replaced by an equivalent composed of two or more components or "fractions"
and these components are subjected to transposition in a second phase. This latter may be
followed by a third phase, recomposition of distributed components, and a fourth phase, the
replacement of the recomposed components by letters. Thus, such a system comprises a first
substitution, a transposition, a recomposition, and a second substitution. In the subsequent
paragraphs we may expect to find in this field the most interesting examples of practical systems of
 cryptography which have been used in the field of military operations in the past and may again be used in the future. The first one to be discussed is particularly interest-
ing for this reason alone; but it is also of interest because it will serve as a model for the student
to follow in his study of methods for the solution of combined substitution-transposition ciphers in
general.

1. See Special Text No. 166, Advanced Military Cryptography, sec. XI.

SECTION IX

SOLUTION OF THE ADFGVX SYSTEM

Introduction.—a. One of the most interesting and practical of the many
methods in which substitution and transposition are combined within a single system is that
known in the literature as the ADFGVX cipher. In this system a 36-character bipartite substitu-
tion checkerboard is employed, in the cells of which the 26 letters of the alphabet and the 10
digits are distributed in mixed order, often according to some key word. The row and column
indicators (coordinates) are the letters A, D, F, G, V, and X. These substitutive pairs are then inscribed within a rect-
tangle and a columnar transposition takes place, according to a numerical key. The cipher text
consists then merely of the 6 letters A, D, F, G, V, and X.

b. The ADFGVX cipher system was inaugurated on the Western Front by the German Army
on March 1, 1918, for communication between higher headquarters, principally between head-
quarters of divisions and corps. When first instituted on March 1, 1918, the checkerboard con-
sisted of 25 cells, for a 22-letter German alphabet (J was omitted), and the 6 letters A, D, F, G,
and X used as coordinates. On June 1 the letter V was added, the checkerboard having been enlarged
to 36 cells, to take care of a 22-letter alphabet. Because of the order in which the plain-text letters appear in
the plain text. These substitutive pairs are then inscribed within a rect-
angle and a columnar transposition takes place, according to a numerical key. The cipher text
consists then merely of the 6 letters A, D, F, G, V, and X.

1. See Special Text No. 166, Advanced Military Cryptography, sec. XI.

The general solution to be described in paragraphs 41–43 was not established until after the Armistice.

Had it been elaborated earlier there would no doubt have been many more solutions than were actually effected by
the methods then available.
Because they are interesting and useful some attention will be devoted to both the general and the special solutions. Since the special solutions are easy to understand and serve as a good introduction to the general solution, they will be taken up first.

23. Special solution by means of identical endings.—a. In paragraph 24 it was demonstrated how the solution of keyed-columnar transposition ciphers can be facilitated and simplified by the comparison of two cryptograms which are in the same key and the plain-text endings of which are identical. It was noted in that case that a study of the irregularly distributed cipher-text identities between the two cryptograms permits of not only cutting up the text into sections that correspond with the long and short columns of the transposition rectangle but also of establishing the transposition key in a direct manner almost entirely mathematical in nature. When this has been accomplished the plain texts of these two messages are at once disclosed, and all other messages in the same key may be read by means of the key so reconstructed.

b. The same method of solution is applicable to the similar situation, if it can be found, in the case of the ADFGVX system, except that one more step intervenes between the reconstruction of the transposition rectangle and the appearance of the plain text in the rectangle: A monoalphabet substitution must be solved, since the text in the rows of the rectangle does not consist of plain-text letters but of pairs of components representing those letters as enciphered by means of a bipartite substitution alphabet. Moreover, this latter step is comparatively simple when there is a sufficient amount of text in the two rectangles; if not, additional material for use in solving the monosubstitution can be obtained from other cryptograms in the same key, if they are available, since the transposition key, having already been reconstructed from the two cryptograms with identical endings, will permit of inscribing all other cryptograms in the same key within their proper rectangles.

c. A demonstration of the application of the principles involved in such a solution will be useful. The following cryptograms have been interpolated on the same date, the 20th:

To CG 22d Brigade:

No. 1

<table>
<thead>
<tr>
<th>X V A A X</th>
<th>V D D A G</th>
<th>D A D V</th>
<th>A D A D</th>
<th>P X G F V</th>
<th>X F A X</th>
</tr>
</thead>
<tbody>
<tr>
<td>X V A A X</td>
<td>A V X D A</td>
<td>G P F X</td>
<td>F G A V</td>
<td>F D D G</td>
<td>D G A F</td>
</tr>
<tr>
<td>A D D D D</td>
<td>X D A V G</td>
<td>G A A D X</td>
<td>A D F V F</td>
<td>P D F X F</td>
<td>G F G A V</td>
</tr>
<tr>
<td>A F A F X</td>
<td>F P X F X</td>
<td>F V D V X</td>
<td>A F F G A</td>
<td>A A A A V</td>
<td>A A A G</td>
</tr>
<tr>
<td>D D P A G</td>
<td>V P A D V</td>
<td>F A V X G</td>
<td>A A A A</td>
<td>P D F A X</td>
<td>X F A A G</td>
</tr>
<tr>
<td>D X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

To CG 23d Brigade:

No. 2

<table>
<thead>
<tr>
<th>F D F F P</th>
<th>F V F A D</th>
<th>D V F V D</th>
<th>G A F D</th>
<th>D A G A D</th>
<th>F D F A F</th>
</tr>
</thead>
<tbody>
<tr>
<td>V A X G V</td>
<td>X V F X V</td>
<td>X V D X A</td>
<td>A A A A G</td>
<td>G X F F</td>
<td>V F A A G</td>
</tr>
<tr>
<td>Y G V F F</td>
<td>F D A F F</td>
<td>F P X D A F</td>
<td>X G A F D</td>
<td>V F V X X</td>
<td>D D F A D</td>
</tr>
<tr>
<td>D A A A X</td>
<td>A A F F A</td>
<td>F P V F X</td>
<td>P A X X A</td>
<td>A D V X A</td>
<td>V D A V F</td>
</tr>
<tr>
<td>D P A V X</td>
<td>V A D F X</td>
<td>A X F F X</td>
<td>X A A V X</td>
<td>X A D X A</td>
<td>A A V V G</td>
</tr>
<tr>
<td>A G D X X</td>
<td>F D F A X</td>
<td>F D G D F</td>
<td>P X D G X</td>
<td>F A G D F</td>
<td>F D D Y D</td>
</tr>
<tr>
<td>X D A F P</td>
<td>A X X A F</td>
<td>G A V</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

d. The delimitation and marking of identities between these two cryptograms is a procedure similar to that explained in paragraph 248, except that a little more study may be necessary in this case because occasionally there may be considerable uncertainty as to exactly where an identity begins or ends. The reason for this is not difficult to understand. Whereas in paragraph 248 the process involves “unfractionated” letters and there are about 18 or 20 different letters to deal with, so that an “accidental identity” is a rather rare occurrence, in the present problem the process involves fractions of letters (the components of the bipartite cipher equivalent), and there are only 6 different characters to deal with, so that such “accidental identities” are quite frequent. Now the cryptanalyst is not able at first to distinguish between these accidental identities and actual identities and this is what makes the process somewhat difficult.

What is meant will become perfectly clear presently.

e. Taking the two illustrative cryptograms, the first step is to ascertain what identities can be found between them, and then mark off these identities. For example, it is obvious that if the messages end alike the last several letters in No. 1 should be found somewhere in No. 2, and likewise the last several letters in No. 2 should be found somewhere in No. 1. The number of letters in identical sequences will depend upon the length of the identical text and the width of the transposition rectangle. Searching through No. 2 for a sequence such as A G D X, or G D X, or at least D X, the tetragraph A G D X is found as letters 151-154. The last column of No. 2 ends with F G A V; searching through No. 1 for a sequence F G A V, or G A V, or at least A V, the tetragraph F G A V is found as letters 87-90. These identities are underlined or marked off in some fashion, and search is made for other identities. It would be a great help if the width of the transposition rectangle were known, for then it would be possible to cut up the text into lengths approximately corresponding to column lengths, and this would then restrict the search for identical sequences to those sections which correspond to the bottoms of the columns. Suppose the key contain 20 numbers. Then the rectangle for No. 1, containing 152 letters, would consist of 12 long columns of 8 letters and 8 short ones of 7 letters; that for No. 2, containing 194 letters, would consist of 14 long columns of 10 letters and 6 short ones of 9 letters. If that were correct then in No. 1 the end of the first column would be either X V D D or X V. Searching through No. 2 for either of these a sequence X V D D is found as letters 84-87. Column 1 is probably a long column in No. 1. The word probably is used because the identity may extend only to the letters X V D D, and the next D may be an accidental similarity, since the chances that D will appear by pure accident are 1 in 6, which is not at all improbable. It must also be pointed out that a certain number of telegraphic errors may be expected, and since there are
only 6 different letters the chances that an F, for example, will be received or recorded as a D are fairly good. Column 1 of No. 2 ends either with VFAD or VFA. Searching through No. 1, a sequence VFAD is found as letters 14-17; a sequence VFAd is found as letters 24-46; a sequence VFAD is also found as letters 126-130; a sequence VFA is found as letters 139-140. Here are several possibilities; which is the one to choose? Two of these possibilities coincide exactly with the full sequence being sought, VFAD. Can one of them be eliminated as a possibility? Perhaps tables to facilitate the location of possible "breaks" will be helpful in making the elimination (see paragraph 16). "Break tables" are therefore constructed for the messages on the basis of rectangles of 20 columns, and are as shown below.

<table>
<thead>
<tr>
<th>Break table for No. 1 (162 letters)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
</tr>
<tr>
<td>7</td>
</tr>
<tr>
<td>14</td>
</tr>
<tr>
<td>21</td>
</tr>
<tr>
<td>28</td>
</tr>
<tr>
<td>35</td>
</tr>
<tr>
<td>42</td>
</tr>
<tr>
<td>49</td>
</tr>
<tr>
<td>56</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Break table for No. 2 (194 letters)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
</tr>
<tr>
<td>9</td>
</tr>
<tr>
<td>18</td>
</tr>
<tr>
<td>27</td>
</tr>
<tr>
<td>36</td>
</tr>
<tr>
<td>45</td>
</tr>
<tr>
<td>54</td>
</tr>
</tbody>
</table>

From these tables it follows that as regards message No. 1 there can be a break after the 7th, 8th, 14th, 15th, 16th . . . letters but not after the 6th letter, nor after the 9th to 14th letters, nor after the 17th to 21st letters, and so on; as regards message No. 2 there can be a break after the 9th, 10th, 11th, 12th, 13th . . . letters but not after the 8th letter nor after the 11th to 18th letters, nor after the 21st to 27th letters, and so on. Referring again to the two VFAD sequences in No. 1 which may correspond with the VFAD sequence in No. 2, it was found that the first candidate would require a break immediately after the 17th letter. But the break table for No. 1 precludes this possibility; hence the first VFAD sequence in No. 1 in position 14-17 may be eliminated as a candidate, leaving the second VFAD, in position 126-130, as a candidate. This would require a break after the 130th letter and reference to the break table for No. 1 shows this to be a possibility. Hence, the VFAD in position 126-130 in No. 1 will tentatively be accepted as matching the VFAD sequence in No. 2. Another section of the text of one or the other cryptogram is next selected, with a view to establishing additional identities. To go through the whole process here would consume too much space and time. Moreover, it is not necessary, for the only purpose in carrying the demonstration this far is to indicate to the student the general procedure and to show him some of the difficulties he will encounter in the identification of the similar portions when the text is composed of only a very limited number of different letters. In this case, after more or less tedious experimentation, the hypothesis of a key of 20 columns is established as correct, whereupon two sets of 20 identities are uncovered and the identities are found to be as shown below.
No. 2

FDPP FVFAD DVFV D GAFD D AGAD FDPAS

VAXGD VXGFX VXDXV VXXDFA XGFDF VFPAG

VGVFF FDAFF FXADF XGAFD VVXV VDFAD

DAAD A AFFF FXFXA AXXA A DVXAV FDPAS

DFAXV DAAXF AXFXA XAAXA XADAX AAVVG

AGDXF FDFA XFDGD FXDAX FAGDF FDDVD

DXDFA F AGXXA F GAV

Form 4—Continued

f. A table of equivalencies is then drawn up:

No. 1: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
No. 2: 9 8 7 10 13 11 17 2 19 15 7 20 14 12 5 18 1 4 3 16

Since the rectangle for No. 2 has two letters in the last row than the rectangle for No. 1, two chains of equivalences at two intervals are constructed. Thus:

\[
\begin{array}{cccccccccccccccc}
9 & 8 & 7 & 10 & 13 & 11 & 17 & 2 & 19 & 15 & 7 & 20 & 14 & 12 & 5 & 18 & 1 & 4 & 3 & 16 \\
\end{array}
\]

These chains must now be united into a single chain by proper interlocking. Since cryptogram No. 1 has 12 long columns, and since the identities of these 12 columns are now known (1, 2, 3, 5, 7, 9, 12, 13, 14, 16, 17, 19, 20), the interlocking of the two chains and hence the transposition key must be this:

\[
\begin{array}{cccccccccccccccc}
\end{array}
\]

The two cryptograms may now be transcribed into their proper transposition matrices, as shown in figure 48.

g. It is necessary to remark that in setting up the table of equivalences, after determining the width of the rectangle, that message which has the lesser number of long columns is used as the basis for the normal sequence 1, 2, 3, ... If the one having the greater number of long columns is employed as the base, the reconstructed key will be reversed.
I becomes evident that the key phrase upon which it is which still remain vacant after the key word mixed sequence is fully developed in the checkerboard, and all the letters which do occur in the two messages are inserted in their correct cells immediately after the letters A, B, . . . .

Speculating upon the disposition of the letters within the enciphering checkerboard, it becomes evident that the key phrase upon which it is based are MILITARY CIPHERS. The fact that the digit 2 follows B and the digit 3 follows C suggests that the digits are inserted immediately after the letters A, B, C, . . . , as they occur in the mixed sequence. Note the cells which still remain vacant after the key word mixed sequence is fully developed in the checkerboard, and all the letters which do occur in the two messages are inserted in their correct cells immediately after the letters A, B, . . . .

The complete checkerboard may therefore be taken almost certainly to be as shown in figure 10c. The date (20th) indicates that the transposition key will have 20 numbers in it. The transposition key was evidently derived from the first 20 letters of the mixed sequence.

GERMANILTYCPHBSDFJKO

20. Special solution by means of identical beginnings.—In paragraph 23 was demonstrated the method of solution based upon finding two cryptograms which are in the same key and the plain texts of which begin with the same words. The application of this method to the corresponding situation in the case of the ADGUVX system should by this time be obvious. The finding of identical sequences is somewhat easier in this case than in the case of identical endings because the identities can be found in parallel progression from the beginning to the end of the two cryptograms being compared. Moreover, the discovery of two cryptograms with similar beginnings is easier than that of two with similar endings because in the former case the very first groups in the two cryptograms contain identities, whereas in the latter case the identities are hidden and scattered throughout the texts of the two cryptograms. On the other hand, the complete solution of a case of identical endings is very much more simple than that involving identical beginnings because in the former case the establishment of the identities carries with it almost automatically the complete reconstruction of the transposition key, whereas in the latter this is far from true and additional cryptograms may be essential in order to accomplish this ease given for the solution.

The following represent 8 cryptograms of the same date, assumed to have been enciphered by the same key.

---

No. 1

VDDFA XFPAX DXXGF FVFXF GXDGG DAAFG AAGDA XDDA DFFDA

A VDXF XAD

No. 2

GXDDA DDGGF VXAX XXGXX AAAAD FADDX AVDXF XAD

No. 3

XDDAA GXXDX FFFDD GDADD XXAAG DFDAG AFDGA VGVDV FDFXA CPFAX AFAAX DDFDS

XAXVA DXFXF DGAGF GGGAD AGDXX AYDDG ADFAX XPAXG FDGDV FXDAD XFDFF

GDXXV DADDV ADDDDG ADAGG AAFDG GDXXA FGVXD GDDDF APAGV AFGXG VDDAX XDYFF

FFFFX GVDGF AVADA XXAFA AFDDG FVXXX

AAGAG AFDGX AFAFX KGAG AAFPA AFDDA GAVXX DGXXF GADDG DADAD XVYAX FVADD

GAFGF GXAXD FDFDX AAAAA

No. 4

APFGX AGXAG XDADF AAXAY GDDDD FAAGY DGDXA FDXXA GFDGD VADAX XGFXA FDADD

GAD

No. 5

XAAAD DAGAG DDDXF FAAVA XDDGD FPAAA DAAXA GXDXD XXXDG VFADA DFFFF VVDGF

XFDGG DAXDG ADP
The cryptograms have been examined for identical beginnings, and Nos. 3 and 6 apparently begin alike, identical portions being underlined as shown in figure 51. Now the number of identical sections in the two cryptograms is 15; this indicates that the width of the transposition rectangle is 15. Therefore, No. 3 sequences, the number of letters between the identical sequences, and the lengths of the columns as calculated from the dimensions of the rectangle. Whenever difficulties in their serial order, as may be many more instances of such accidental identities before and after the case under study, viz, in columns 5 and 12.

The identical sections for column 13 here consist of 5 long columns of letters [(15X21)-13=302]. In this transcription no serious difficulty is usually given by the study of identical sections, writing them merely shown in figure 51, the next step is to transcribe the texts into true identical sequences. Another thing to be noted is that the identical beginnings in this process, they are brought about by accidental identities of letters before the head of column 13, it leaves column 12 one letter short at the bottom in each diagram. This means that the initial A's in these short columns of 20 letters and minimum of 5 in any identical sequence are accidental identities. There are several such accidental identities in the case under study, viz, in columns 5 and 12.

The cryptograms have been marked off as shown in figure 52. In this transcription no serious difficulty is usually encountered in the division into correct column lengths, this process being guided by the identical numbers of letters between the identical sequences, and the maximum and minimum lengths of the columns as calculated from the dimensions of the rectangle. Whenever difficulties are encountered in this process, they are brought about by accidental identities of letters before and after the true or actual identical sequences. In the present case no such difficulties arise except in going from column 12 to column 13. The identical sections for column 13 here consist of the sequence A F F A A F; if these sections are placed at the head of column 13, it leaves column 12 one letter short at the bottom in each diagram. This means that the initial A's in these identical sequences represent an accidental identity; these A's belong at the bottom of column 12 in each diagram, and the true identical sequences are F F A A F, and not A F F A A F. In some cases there may be many more instances of such accidental identities before and after the true identical sequences. Another thing to be noted is that the identical beginnings in this case run along for at least 4 complete rows and part of the fifth row in the transposition rectangle. Therefore, the identical sequences should consist of not less than 4, and not more than 5 letters; any letters in excess of 5 in any identical sequence are accidental identities. There are several such accidental identities in the case under study, in columns 5 and 12.
### FIGURE 62

No. 3

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>D</td>
<td>D</td>
<td>F</td>
<td>D</td>
<td>A</td>
<td>D</td>
<td>D</td>
<td>G</td>
<td>X</td>
</tr>
<tr>
<td>D</td>
<td>X</td>
<td>V</td>
<td>D</td>
<td>D</td>
<td>G</td>
<td>F</td>
<td>A</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td>A</td>
<td>F</td>
<td>X</td>
<td>A</td>
<td>V</td>
<td>F</td>
<td>G</td>
<td>D</td>
<td>X</td>
<td>V</td>
</tr>
<tr>
<td>D</td>
<td>F</td>
<td>G</td>
<td>F</td>
<td>D</td>
<td>A</td>
<td>D</td>
<td>D</td>
<td>G</td>
<td>X</td>
</tr>
<tr>
<td>X</td>
<td>D</td>
<td>G</td>
<td>A</td>
<td>G</td>
<td>A</td>
<td>F</td>
<td>F</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>A</td>
<td>A</td>
<td>D</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>A</td>
<td>F</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>F</td>
<td>D</td>
<td>G</td>
<td>F</td>
<td>D</td>
<td>A</td>
<td>D</td>
<td>D</td>
<td>G</td>
<td>X</td>
</tr>
<tr>
<td>X</td>
<td>D</td>
<td>G</td>
<td>A</td>
<td>G</td>
<td>A</td>
<td>F</td>
<td>F</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>D</td>
<td>F</td>
<td>G</td>
<td>F</td>
<td>D</td>
<td>A</td>
<td>D</td>
<td>D</td>
<td>G</td>
<td>X</td>
</tr>
<tr>
<td>X</td>
<td>D</td>
<td>G</td>
<td>A</td>
<td>G</td>
<td>A</td>
<td>F</td>
<td>F</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>F</td>
<td>D</td>
<td>G</td>
<td>F</td>
<td>D</td>
<td>A</td>
<td>D</td>
<td>D</td>
<td>G</td>
<td>X</td>
</tr>
</tbody>
</table>

### FIGURE 63—Continued

c. Now comes the attempt to place the columns in proper sequence in the respective transposition rectangles. Since No. 6 has only 2 long columns, viz, 5 and 14, it is obvious that these two columns belong at the extreme left of the rectangle. Their order may be 5-14 or 14-5; there is no way of telling which is correct just yet. Since No. 3 has 5 long columns, viz, 3, 4, 5, 7, 14, and since from No. 6 it has been ascertained that 5 and 12 go to the extreme left, it is obvious that columns 3, 4, and 7 occupy the third, fourth, and fifth positions in the rectangles. Their order may be any permutation of the three numbers 3, 4, and 7; their exact order must be ascertained by further study.

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>D</td>
<td>D</td>
<td>F</td>
<td>D</td>
<td>A</td>
<td>D</td>
<td>D</td>
<td>G</td>
<td>X</td>
</tr>
<tr>
<td>D</td>
<td>X</td>
<td>V</td>
<td>D</td>
<td>D</td>
<td>G</td>
<td>F</td>
<td>A</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td>A</td>
<td>A</td>
<td>F</td>
<td>X</td>
<td>A</td>
<td>V</td>
<td>F</td>
<td>G</td>
<td>D</td>
<td>X</td>
</tr>
<tr>
<td>D</td>
<td>F</td>
<td>G</td>
<td>F</td>
<td>D</td>
<td>A</td>
<td>D</td>
<td>D</td>
<td>G</td>
<td>X</td>
</tr>
<tr>
<td>D</td>
<td>F</td>
<td>G</td>
<td>F</td>
<td>D</td>
<td>A</td>
<td>D</td>
<td>D</td>
<td>G</td>
<td>X</td>
</tr>
<tr>
<td>D</td>
<td>F</td>
<td>G</td>
<td>F</td>
<td>D</td>
<td>A</td>
<td>D</td>
<td>D</td>
<td>G</td>
<td>X</td>
</tr>
<tr>
<td>D</td>
<td>F</td>
<td>G</td>
<td>F</td>
<td>D</td>
<td>A</td>
<td>D</td>
<td>D</td>
<td>G</td>
<td>X</td>
</tr>
<tr>
<td>D</td>
<td>F</td>
<td>G</td>
<td>F</td>
<td>D</td>
<td>A</td>
<td>D</td>
<td>D</td>
<td>G</td>
<td>X</td>
</tr>
<tr>
<td>D</td>
<td>F</td>
<td>G</td>
<td>F</td>
<td>D</td>
<td>A</td>
<td>D</td>
<td>D</td>
<td>G</td>
<td>X</td>
</tr>
<tr>
<td>D</td>
<td>F</td>
<td>G</td>
<td>F</td>
<td>D</td>
<td>A</td>
<td>D</td>
<td>D</td>
<td>G</td>
<td>X</td>
</tr>
</tbody>
</table>
d. In this study, to fix the exact order of the columns and thus to reconstruct the transposition key, advantage can be taken of the diverse lengths of other cryptograms that may be available in the same key. In this case there are 6 additional cryptograms, Nos. 1, 2, 4, 5, 7, and 8, suitable for the purpose. The following calculations are made:

<table>
<thead>
<tr>
<th>Cryptogram No.</th>
<th>Total number of letters</th>
<th>Lengths of columns</th>
<th>Number of columns</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>60</td>
<td>4</td>
<td>All same length</td>
</tr>
<tr>
<td>2</td>
<td>38</td>
<td>3 and 2</td>
<td>8 7</td>
</tr>
<tr>
<td>4</td>
<td>62</td>
<td>0 and 4</td>
<td>2 10</td>
</tr>
<tr>
<td>5</td>
<td>74</td>
<td>0 and 4</td>
<td>14 1</td>
</tr>
<tr>
<td>7</td>
<td>124</td>
<td>9 and 8</td>
<td>4 11</td>
</tr>
<tr>
<td>8</td>
<td>34</td>
<td>4 and 3</td>
<td>9 6</td>
</tr>
</tbody>
</table>

Now No. 7 has 4 long columns, and these must consist of 4 columns from among the 5 already ascertained as falling at the extreme left, viz., 3, 4, 5, 7, and 14. Columns 5 and 14 have furthermore been placed in positions 1-2, leaving columns 3, 4, and 7 for positions 3-4-5. Which of these three possibilities is to be omitted as a long column in No. 7? A means of answering this question involves certain considerations of general importance in the cryptanalysis of this type of system.

e. Consider a transposition rectangle in which the number of columns is even, and consider specifically the first pair of columns in such a rectangle. The combinations of bipartite components formed by the juxtaposition of these 2 columns correspond to plain-text letters, and therefore the distribution of the bipartite digraphs in these columns will be monoalphabetic in character. The same is true with respect to the bipartite components in the third and fourth columns, the fifth and sixth columns, and so on. Hence, if a long cryptogram of this nature is at hand, and if the 2 columns which belong at the extreme left can be ascertained, then a distribution of the bipartite digraphs formed by juxtaposing these columns should not only be monoalphabetic, but also this distribution, if it is at all normal, will afford a basis for matching other columns which will produce similar distributions, for the text as a whole is monoalphabetic. In this way, by proper matching of columns, those which really go together to form the pairs containing the bipartite equivalents of the plain-text letters can be ascertained. From that point on, the solution of the problem is practically the same as that of solving a columnar transposition cipher by nonfractionated letters.

f. But now consider a plain-text rectangle in the ADGFV system, in which the number of columns is odd, and consider specifically the first pair of columns in the rectangle. Now only the alternate combinations of bipartite components in these columns form the units of plain-text letters. The same is true of the bipartite components of the third and fourth, the fifth and sixth columns, and so on. In all other respects, however, the remarks contained in subparagraph e apply equally to this case where the width of the rectangle is odd.

g. Returning to the problem under study, it has been ascertained that columns 5 and 14 fall at the extreme left. Whether their correct order is 5-14 or 14-5 cannot at the moment be ascertained, nor is it essential. The thing to do is to make a distribution of the bipartite pairs and see what it is like. Since the width of the rectangle here is odd, only the 1st, 5d, 5th, . . . pairs down the columns can be distributed in a frequency square. The results are shown in Fig. 53.

<table>
<thead>
<tr>
<th>No. 3</th>
<th>No. 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Col. 1</td>
<td>Col. 1</td>
</tr>
<tr>
<td>D A</td>
<td>D A</td>
</tr>
<tr>
<td>D A</td>
<td>D A</td>
</tr>
<tr>
<td>A A</td>
<td>A A</td>
</tr>
<tr>
<td>G F</td>
<td>G F</td>
</tr>
<tr>
<td>5 D</td>
<td>5 D</td>
</tr>
<tr>
<td>G A</td>
<td>G X</td>
</tr>
<tr>
<td>7 X</td>
<td>7 G</td>
</tr>
<tr>
<td>A A</td>
<td>D X</td>
</tr>
<tr>
<td>9 V</td>
<td>9 A A</td>
</tr>
<tr>
<td>11 D V</td>
<td>11 C A</td>
</tr>
<tr>
<td>13 A A</td>
<td>13 F D</td>
</tr>
<tr>
<td>15 A F</td>
<td>15 D G</td>
</tr>
<tr>
<td>17 A</td>
<td>17 V</td>
</tr>
<tr>
<td>19 F D</td>
<td>19 F D</td>
</tr>
<tr>
<td>A C</td>
<td>V V</td>
</tr>
<tr>
<td>21 V F</td>
<td></td>
</tr>
</tbody>
</table>

A. The distribution is fairly good. Five occurrences of AA are noted, 3 of FD. These must represent high-frequency letters. The φ (Pφ) test for monoalphabeticity may be applied.

Expected value of φ for plain text = 0.0667×21×20 = 28.01
Expected value of φ for random text = 0.0385×21×20 = 16.17
Observed value of φ in this case = (5×4)+(2×1)+(2×1)+(3×2)+(2×1)+(2×1)=34
The observed value of φ is considerably greater than the expected value for plain text and more than twice as much as the expected value for random text. Using the distribution in figure 53 as a basis, an attempt is made to add to the 5-14 combination a column selected from among columns 3, 4, and 7, so that the second, fourth, sixth . . . pairs down the second and third columns in the rectangle will give bipartite pairs that will conform to the distribution noted in figure 53. Since the results sought will be very materially affected if the combination 5-14 should really be 14-5, all possible combinations of 5-14 and 14-5 with 3, 4, and 7 must be tried. The various combinations tested are shown in figure 54.

Now consider the second and third columns of plain-text letters. The same is true of the bipartite components of the third and fourth, the fifth and sixth columns, and so on. Hence, if a long cryptogram of this nature is at hand, and if the 2 columns which belong at the extreme left can be ascertained, then a distribution of the bipartite digraphs formed by juxtaposing these columns should not only be monoalphabetic, but also this distribution, if it is at all normal, will afford a basis for matching other columns which will produce similar distributions, for the text as a whole is monoalphabetic. In this way, by proper matching of columns, those which really go together to form the pairs containing the bipartite digraphs formed by juxtaposing these columns should not only be monoalphabetic, but also this distribution, if it is at all normal, will afford a basis for matching other columns which will produce similar distributions, for the text as a whole is monoalphabetic. In this way, by proper matching of columns, those which really go together to form the pairs containing the bipartite equivalents of the plain-text letters can be ascertained. From that point on, the solution of the problem is practically the same as that of solving a columnar transposition cipher by nonfractionated letters.
i. Frequency distributions are now made. If combination 5–14–3 is correct for No. 3, it is also correct for No. 6. Hence, a single distribution is made of the bipartite pairs in rows 1, 3, 5, ..., of columns 5–14, and of the pairs in rows 2, 4, 6, ..., of columns 14–3. Similar distributions are made of the pairs given under each of the other combinations. These distributions are shown in figure 55.

![Figure 55](image)

j. These distributions are now tested for monoalphabeticity, by applying the χ² test. The number of occurrences in each distribution is 41. Then 41×40×0.067=109.4 is the expected value of χ² for plain text; 41×40×0.0385=63.1 is the expected value of χ² for random text. Here are the calculations for the first distribution (combination 5–14–3) yielding the observed value of χ² as 52:

\[
\chi^2 = \frac{(5×4)+(3×2)+(1×9)+(2×1)+(1×0)+(1×0)}{(1×0)+(3×2)+(1×0)+(1×0)+(5×4)+(2×1)+(1×0)+(4×3)+(2×1)+(1×0)+(8×2)+(2×3)+(2×1)} = 52.
\]

The observed values of χ² for all 6 frequency distributions are shown here:

<table>
<thead>
<tr>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>52</td>
<td>63.1</td>
</tr>
<tr>
<td>(2)</td>
<td>120</td>
<td>76</td>
</tr>
<tr>
<td>(3)</td>
<td>110</td>
<td>70</td>
</tr>
</tbody>
</table>
Only two of these distributions give close approximations to 109, the expected value of $\varphi$, and they may be retained for further experiment. They are the ones for combinations (4) and (6), with values of 120 and 110, respectively.

$k$. Selecting combinations (4) and (6) viz., 14-5-3, and 14-5-7, since columns 14, 3, 4, and 5 form the group of 5 columns at the left of the transposition rectangle, the following combinations are possible:

(1) 14-5-3-4-7
(2) 14-5-3-7-4
(3) 14-5-7-3-4
(4) 14-5-7-4-3

l. The following sets of columns correspond to these 4 combinations in the 2 cryptograms (fig. 56):

No. 3

<table>
<thead>
<tr>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A A D D F D</td>
<td>A D D F D</td>
<td>A D D F D</td>
</tr>
<tr>
<td>2</td>
<td>A D D F D</td>
<td>A D D F D</td>
<td>A D D F D</td>
</tr>
<tr>
<td>3</td>
<td>A A F F X F</td>
<td>A A F F X F</td>
<td>A A F F X F</td>
</tr>
<tr>
<td>5</td>
<td>D D F D X</td>
<td>D D F D X</td>
<td>D D F D X</td>
</tr>
<tr>
<td>6</td>
<td>A D F D A</td>
<td>A D F D A</td>
<td>A D F D A</td>
</tr>
<tr>
<td>7</td>
<td>D A A D D</td>
<td>D A A D D</td>
<td>D A A D D</td>
</tr>
<tr>
<td>8</td>
<td>A A A G D V</td>
<td>A A A G D V</td>
<td>A A A G D V</td>
</tr>
<tr>
<td>9</td>
<td>D V F D X</td>
<td>D V F D X</td>
<td>D V F D X</td>
</tr>
<tr>
<td>10</td>
<td>X G X F A</td>
<td>X G X F A</td>
<td>X G X F A</td>
</tr>
<tr>
<td>11</td>
<td>V D A X D</td>
<td>V D A X D</td>
<td>V D A X D</td>
</tr>
<tr>
<td>12</td>
<td>V G F F A</td>
<td>V G F F A</td>
<td>V G F F A</td>
</tr>
<tr>
<td>13</td>
<td>A A A D V</td>
<td>A A A D V</td>
<td>A A A D V</td>
</tr>
<tr>
<td>14</td>
<td>D D D F G</td>
<td>D D D F G</td>
<td>D D D F G</td>
</tr>
<tr>
<td>15</td>
<td>F A A A A</td>
<td>F A A A A</td>
<td>F A A A A</td>
</tr>
<tr>
<td>16</td>
<td>V G X F D G</td>
<td>V G X F D G</td>
<td>V G X F D G</td>
</tr>
<tr>
<td>17</td>
<td>A D F D A</td>
<td>A D F D A</td>
<td>A D F D A</td>
</tr>
<tr>
<td>18</td>
<td>D D D F G</td>
<td>D D D F G</td>
<td>D D D F G</td>
</tr>
<tr>
<td>19</td>
<td>D F D G G</td>
<td>D F D G G</td>
<td>D F D G G</td>
</tr>
<tr>
<td>20</td>
<td>G A A D A</td>
<td>G A A D A</td>
<td>G A A D A</td>
</tr>
</tbody>
</table>

The other squares constructed with adding columns 4-7 to the basic combination 14-5-3 are distributed in the 4th frequency distribution square of figure 55, yielding the distribution shown in square (1) of figure 37. The other squares in figure 37 are constructed in the same way, for the other combinations of figure 56.

$\varphi$. The additional bipartite pairs given by adding columns 4-7 to the basic combination 14-5-3 are distributed in the 4th frequency distribution square of figure 55, yielding the distribution shown in square (1) of figure 37. The other squares in figure 37 are constructed in the same way, for the other combinations of figure 56.
117

10 columns which are to be anagrammed to ascertain their correct sequence. The column to follow column 7 is ascertainment on the basis of the repetitions which are brought about when the selected column is placed on the right. These repetitions should fall into those cells of frequency distribution (1), figure 57, which are of high frequency. In other words, the process is one of selecting from among columns 1, 2, 6, 8, 9, 10, 11, 12, 13, and 15 that column which will yield the most repetitions of bipartite digraphs with the digraphs given by the juxtaposition of columns 14–5–3–4–7, as distributed in frequency square (1) of figure 57. The column thus selected turns out to be No. 10. Then other columns are added by proceeding along the same lines, the work becoming progressively more easy as the number of available candidates decreases. Sometimes the discovery of what appears to be a long repetition within one of the cryptograms or between two cryptograms facilitates the process. In this case the results obtained from the 3 cryptograms under study are shown in figure 59.

<table>
<thead>
<tr>
<th>14</th>
<th>5</th>
<th>3</th>
<th>4</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>AD</td>
<td>DF</td>
<td>VX</td>
<td>AD</td>
<td>DF</td>
</tr>
<tr>
<td>A</td>
<td>D</td>
<td>F</td>
<td>V</td>
<td>X</td>
</tr>
<tr>
<td>D</td>
<td>A</td>
<td>F</td>
<td>V</td>
<td>X</td>
</tr>
<tr>
<td>X</td>
<td>D</td>
<td>V</td>
<td>A</td>
<td>F</td>
</tr>
<tr>
<td>D</td>
<td>A</td>
<td>F</td>
<td>V</td>
<td>X</td>
</tr>
<tr>
<td>A</td>
<td>D</td>
<td>F</td>
<td>V</td>
<td>X</td>
</tr>
<tr>
<td>A</td>
<td>D</td>
<td>F</td>
<td>V</td>
<td>X</td>
</tr>
<tr>
<td>A</td>
<td>D</td>
<td>F</td>
<td>V</td>
<td>X</td>
</tr>
<tr>
<td>A</td>
<td>D</td>
<td>F</td>
<td>V</td>
<td>X</td>
</tr>
</tbody>
</table>

Figure 56.

The combination 14–5–3–4–7, giving the greatest value for φ (a little better than the expected value), is very probably the correct one.

Examination of the other cryptograms that are available, it is seen that No. 7 is the third longest one of the entire set, with 124 letters; moreover, the dimensions of the rectangle (15X9) -- 11 -- 124 are such as to bring about 4 long columns of 9 letters and 11 columns of 8 letters. The first 5 columns are definitely fixed in position, since it is known that the first 5 key numbers are 14–5–3–4–7. The resulting diagram is shown in figure 58. There is now a section consisting of

<table>
<thead>
<tr>
<th>14</th>
<th>5</th>
<th>3</th>
<th>4</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>X</td>
<td>D</td>
<td>V</td>
<td>D</td>
</tr>
<tr>
<td>A</td>
<td>X</td>
<td>D</td>
<td>V</td>
<td>D</td>
</tr>
<tr>
<td>A</td>
<td>X</td>
<td>D</td>
<td>V</td>
<td>D</td>
</tr>
<tr>
<td>A</td>
<td>X</td>
<td>D</td>
<td>V</td>
<td>D</td>
</tr>
<tr>
<td>A</td>
<td>X</td>
<td>D</td>
<td>V</td>
<td>D</td>
</tr>
<tr>
<td>A</td>
<td>X</td>
<td>D</td>
<td>V</td>
<td>D</td>
</tr>
<tr>
<td>A</td>
<td>X</td>
<td>D</td>
<td>V</td>
<td>D</td>
</tr>
<tr>
<td>A</td>
<td>X</td>
<td>D</td>
<td>V</td>
<td>D</td>
</tr>
<tr>
<td>A</td>
<td>X</td>
<td>D</td>
<td>V</td>
<td>D</td>
</tr>
<tr>
<td>A</td>
<td>X</td>
<td>D</td>
<td>V</td>
<td>D</td>
</tr>
</tbody>
</table>

Figure 57.

The combination 14–5–3–4–7, giving the greatest value for φ is 81X80X.0667=432. The observed value for the four combinations of figure 57 are as follows:

(1) For combination 14–5–3–4–7, φ=436
(2) For combination 14–5–3–4–7, φ=276
(3) For combination 14–5–3–4–7, φ=544
(4) For combination 14–5–3–4–7, φ=318

The combination 14–5–3–4–7, giving the greatest value for φ (a little better than the expected value), is very probably the correct one.

n. Again applying the φ-test, the expected value of φ is 81X80X.0667=432. The observed values for the four combinations of figure 57 are as follows:

(1) For combination 14–5–3–4–7, φ=436
(2) For combination 14–5–3–4–7, φ=276
(3) For combination 14–5–3–4–7, φ=544
(4) For combination 14–5–3–4–7, φ=318
What the cryptanalyst now has before him is a monoalphabetic substitution cipher, the solution of which presents no difficulties. The cipher square is reconstructed as completely as possible, blanks being left where there are no occurrences to give clues as to the character involved, usually some of the digits and the very infrequent letters.

In this case the only letters which do not occur in the plain text are Q, X, and Z. The digits 5 and 7 are recovered from the context, in message No. 6, where the caliber of a gun is mentioned and the digits are confirmed at other places in the message. The square that is obtained is seen in figure 60.

Examination of the mixed sequence discloses that it is based upon the phrase THE FLOWERS THAT BLOOM IN THE SPRING. This permits of the establishment of the transposition key and of the position of the digits in the checkerboard (as in par. 381). The results are shown in figure 61. The completely solved messages are shown in figure 62.
40. Special solution by the exact factor method.— a. The student who has comprehended the successive steps in the solution of the example discussed in the preceding paragraph is in a position to grasp at once the mechanics of the special solution by the exact factor method. The latter is based upon the interception of a number of cryptograms, preferably lengthy ones, which have been enciphered by rectangles in which the last row is completely filled with letters. The total number of bipartite components in the case of such a cryptogram will yield clues as to the dimensions of the transposition rectangle. Then the text is transcribed into columns of appropriate length, all being equal in this respect, and the process of combining columns, as explained in paragraph 39e, is applied in order to produce the best monoalphabetic distribution of bipartite
digraphs down the juxtaposed columns. There is nothing to prevent the simultaneous use of the cryptogram that has been enciphered by completely filled rectangles, for it is so clear that if, for example, columns 15 and 4 are to be paired in one cryptogram, the same columns will be paired in all other cryptograms. Hence, even if the rectangles are small in depth they can be used in this process; it is necessary only that all columns of any rectangle be of the same length. Now if only two or three such pairs of columns can be set up correctly, solution follows almost as a matter of course. No additional or new principles need be brought in play, beyond those already possessed by the student.

8. In this special solution, the important step is, of course, the initial one of experimenting with rectangles of various dimensions until the correct size has been hit upon. In some cases, excessive experimentation may not be necessary if the total number of characters is such as to yield only one or two possibilities with regard to the length of the column. For example, suppose that previous work has established the fact that the enemy uses transposition rectangles not less than 15 and not more than 22 columns in width. A message totaling 703 letters would indicate a rectangle of 19 columns of 37 letters, since these two numbers are the only factors of 703. If this then were corroborated by other cryptograms of 159 (19X8), 122 (19X8), 190 (19X10) letters, the probability that 19 is the width of the transposition rectangle becomes quite persuasive. Of course, there will be and there should be other cryptograms of lengths that do not factor exactly; these represent the ones in which the rectangles are not completely filled in their last row. They do not enter into the solution at first, but just as soon as the positions of two or three key numbers become fixed, the data afforded by these messages become available for use in the later stages in the solution.

9. The exact-factor method is a useful one to know. For despite all instructions that may be drawn up insisting upon the advisability of not completing the last row of a transposition rectangle, the tendency to violate such a rule is quite marked, especially where a large cryptographic personnel must be employed. It is not astonishing to find that for lazy or ignorant clerks the temptation to fill the rectangle completely is particularly hard to resist when it happens that a message falls just one, two, or three letters short of forming a completely-filled rectangle: it is so much easier for such clerks to handle a rectangle with equal-length columns than one in which this is not the case. Moreover, the number of errors and therefore the number of times a sender or a receiver of cryptograms must go over his work to correct errors in minutes not upon the actual rectangle, becomes very clear if one studies figure 63. Two transposition rectangles are shown, one with an even number of columns, the other with an odd number. Instead of the actual components (ADFGVX), the symbols 0, 1, 2, and 3 are used to indicate the two classes of columns, initial and final, because in this analysis interest centers not upon the initial and final components, but only upon the component within a given column. Now when the transposition is applied, all the components within a given column are of the same class, either initial or final. No intermixtures or alternations of the two classes is possible. On the other hand, consider a rectangle with an odd number of columns. Each of its rows contains an odd number of bipartite components, the 1st row containing one more initial component than final components, the 2d row containing one more final component than initial components, and so on, this arrangement alternating regularly in the successive rows of the rectangle. When one studies the various columns of the rectangle, it is seen that in each column there is a perfectly regular alternation of initial and final components, the odd columns (1st, 3d, 5th, ... ) beginning with an initial component, the even columns (2d, 4th, 6th, ... ) beginning with a final component. This alternation in components remains true even after the transposition is applied. These remarks become very clear if one studies figure 63. Two transposition rectangles are shown, one with an even number of columns, the other with an odd number. Instead of the actual components (ADFGVX), the symbols 0, 1, 2, and 3 are used to indicate the two classes of columns, initial and final, because in this analysis interest centers not upon the actual identity of a component but upon the class to which it belongs, initial or final. At the top of each column is placed a "plus" to denote a column occupying an odd-numbered position in the rectangle, or a "minus" to denote a column occupying an even-numbered position.

124

125

10. The attention of the student is directed to the comments made in paragraph 18, with regard to the significance of the term general solution in cryptanalysis. He must be cautioned not to expect that in practical work a general solution will, in the cryptanalytic as in the mathematical field, immediately lead to a solution. If there is a sufficient amount of text and if the text contains no abnormalities, the attempt to apply the general solution will usually be successful. But the cryptanalyst must remember that the ADFGVX system is by no means a simple one to solve even under the best of conditions and if there is only a small amount of text, if it happens that the transposition key is unusually long, or if the text is abnormal, he may not succeed in solving the messages by the straightforward method to be set forth below, and he may have to introduce special modifications. For the latter he can only rely upon his own ingenuity and intuition.

45. Basic principles of the general solution.—a. Every transposition rectangle in the ADFGVX system must conform to one or the other of two and only two fundamental types: the number of columns must be either odd or even. A number of important consequences follow from this simple fact, some of which have already been pointed out in paragraph 39c. They will be elaborated upon in the next subparagraphs.

b. Consider a rectangle with an even number of columns. Each of its rows contains an even number of bipartite components, half of which are initial components, half, final components, alternating in a regular order from left to right in the rows. When the transposition is applied, all the components within a given column are of the same class, either initial or final. No intermixtures or alternations of the two classes is possible. On the other hand, consider a rectangle with an odd number of columns. Each of its rows contains an odd number of bipartite components, the 1st row containing one more initial component than final components, the 2d row containing one more final component than initial components, and so on, this arrangement alternating regularly in the successive rows of the rectangle. When one studies the various columns of the rectangle, it is seen that in each column there is a perfectly regular alternation of initial and final components, the odd columns (1st, 3d, 5th, ... ) beginning with an initial component, the even columns (2d, 4th, 6th, ... ) beginning with a final component. This alternation in components remains true even after the transposition is applied. These remarks become very clear if one studies figure 63. Two transposition rectangles are shown, one with an even number of columns, the other with an odd number. Instead of the actual components (ADFGVX), the symbols 0, 1, 2, and 3 are used to indicate the two classes of columns, initial and final, because in this analysis interest centers not upon the actual identity of a component but upon the class to which it belongs, initial or final. At the top of each column is placed a "plus" to denote a column occupying an odd-numbered position in the rectangle, or a "minus" to denote a column occupying an even-numbered position.
sections in this case will be even. Moreover, all the components in a section corresponding to an odd column in rectangle \( b \) will be \( \Omega \), or initial components, all those in a section corresponding to an even column, \( \Omega \), or final components. The sections or columns are completely homogeneous with respect to the class to which their constituent components belong. On the other hand, when the transposition is applied to odd-width rectangle \( b \), the cryptographic text will consist of an odd number of sections, each corresponding to a column of the rectangle. The components in the sections consist of members of both classes of components in a regular alternation; in a section corresponding to an odd column the order is \( \Omega \rightarrow \bar{\Omega} \rightarrow \bar{\Omega} \cdots \); in a section corresponding to an even column the order is \( \bar{\Omega} \rightarrow \Omega \rightarrow \Omega \cdots \). The sections or columns are not homogeneous in this case as they are in the former.

d. Now if there were some way of distinguishing between initial components as a class and final components as a class it is clear that it may be possible first of all to ascertain whether the transposition rectangle contains an even or an odd number of columns. Secondly it may be possible to identify those columns which are even and those which are odd. Finally, it may be possible to ascertain which are the long columns and which are short, thus yielding the exact outlines of the rectangle in case the last row is incompletely filled. From that point on, solution is possible to ascertain which are the long columns and which are short, thus yielding the exact outlines of the rectangle in case they are in the former.

e. The foregoing depends then upon the possibility of being able to distinguish as a class between initial and final components of the bipartite cipher equivalents in this system, or at least between letters belonging to one or the other of these two general classes of components. Now if the substitution checkerboard has not been consciously manipulated with a view to destroying certain properties normally characterizing its rows and columns, the sort of differentiation indicated above is quite possible. For example, if in the checkerboard shown in figure 61 there are several cryptograms available for study, as would be shown very soon, by studying an example. For the purposes of this study the set of 12 cryptograms given below will be used.

has a frequency value of 284 as an initial component of the bipartite cipher equivalents of plaintext letters, and a frequency value of only 169 as a final component.

Similarly, the letters \( V \) and \( X \) have frequency values of 47 and 69, respectively, as initial components and 190 and 198 as final components. It is obvious, then, that in this checkerboard the weighted frequency values of the letters \( A, V, \) and \( X \) as initial components differ considerably from the values of these same letters as final components, the value for \( G \) as an initial is only a little less than its value as a final, the values of \( D \) and \( F \) as initials are only a little more than their values as finals. But it is the wide variations in the weighted frequency values of certain of the letters as initial components and as final components, exemplified in the case of \( A, V, \) and \( X \), which form the basis of the general solution, because these wide variations afford a means for making the various differentiations noted in subparagraph d.

f. Of course, in working with an unknown example, the composition of the checkerboard is unknown and therefore no accurate frequency weights may be assigned to the \( ADGVX \) components in the cryptograms. However, it is still possible to arrive at some approximations for these weights in case there are several cryptograms available for study, as would normally be true in actual practice. How this can be done will be shown very soon, by studying an example.

As given in fig. 5, p. 18, Military Cryptanalysis, Part I.
43. Illustration of solution.—a. Since the initial letters of all 12 cryptograms are in the same class, that is, either initial or final components, they may all be combined into a single distribution. Furthermore, since it is certain that regardless of whether the transposition rectangle has an odd or an even number of columns the 3d, 5th, 7th... letters of the cryptograms are in

This illustration use the same cryptograms and follows quite closely along the lines employed in a technical paper of the Signal Intelligence Service entitled General Solution for the ADPGVX Cipher, prepared by Messrs. Rowlett, Knillbeck, and Stinker, in 1934.
the same class as the first letter, the 3d, 5th, 7th ... letters may be added to the distribution, so long as these odd letters come from the same section (column). It is, however, necessary to limit the number of letters taken from the beginning of any one cryptogram to a reasonable length of column, depending on the size of the cryptogram. It is assumed that the enemy is using transposition keys of not less than 15 nor more than 25 numbers, the letter could be taken as the maximum possible size. But to be on the safe side it will be here assumed that a transposition rectangle of not more than 25 columns is being used. Hence, so far as concerns cryptogram I, which has 212 letters, on the basis of a key of 25 numbers \(25 \times 9 - 13 = 212\) there will be 12 columns of 9 letters and 13 columns of 8 letters. Since there is no way of telling which are long and which are short columns, it will be safer to work on the basis of columns of 8 letters. Therefore, the first 8 letters of cryptogram I are to be taken. In the case of cryptogram II, with 108 letters, its first 4 letters will be taken, and so on, through the 12 cryptograms, the number of letters to be taken in each case being governed by the length of the cryptogram.

### Table of Odd and Even Letters

<table>
<thead>
<tr>
<th>Cryptogram</th>
<th>Length</th>
<th>Letters taken</th>
<th>Cryptogram</th>
<th>Length</th>
<th>Letters taken</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>212</td>
<td>VOGGIOF</td>
<td>VII</td>
<td>224</td>
<td>GAGGFZPWVF</td>
</tr>
<tr>
<td>II</td>
<td>108</td>
<td>VOGA</td>
<td>VIII</td>
<td>144</td>
<td>DVGZ</td>
</tr>
<tr>
<td>III</td>
<td>116</td>
<td>GAGGAOFX</td>
<td>IX</td>
<td>162</td>
<td>XFXFBXV</td>
</tr>
<tr>
<td>IV</td>
<td>110</td>
<td>AGXY</td>
<td>X</td>
<td>130</td>
<td>XDZDF</td>
</tr>
<tr>
<td>V</td>
<td>202</td>
<td>GFDGFV</td>
<td>XI</td>
<td>248</td>
<td>TFGFVAX</td>
</tr>
<tr>
<td>VI</td>
<td>120</td>
<td>GCOF</td>
<td>XII</td>
<td>224</td>
<td>XFDXFDVD</td>
</tr>
</tbody>
</table>

#### Figure 66

### Table of Even and Odd Letters

<table>
<thead>
<tr>
<th>Even (2d, 4th, 6th...)</th>
<th>Odd (1st, 3rd, 5th...)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A D F G V X</td>
<td>A D F G V X</td>
</tr>
<tr>
<td>H H H H H H</td>
<td>H H H H H H</td>
</tr>
</tbody>
</table>

#### Figure 67

### Letters

**V** and **X** are of high frequency in the odd positions (fig. 66) but of low frequency in the even positions (fig. 67), whereas the letter **F** is of low frequency in the odd positions and of high frequency in the even positions. There can be no question that the two distributions are dissimilar, and the indications are clear that the transposition rectangle involves an odd number of columns.

d. Now the letters in figure 66 may be initial components, those in figure 67, final components, or the reverse may be the case. At the present stage of the study it is impossible to ascertain which of these alternative hypotheses is correct. However, this information is really immaterial at this stage. Suppose the letters in figure 66 are arbitrarily designated as class 1 components, those in figure 67 as class 2 components. Class 1 components (fig. 66) are characterized by a predominance of **F's** (over their frequency in fig. 67); class 2 components (fig. 67) are characterized by a predominance of **F's** (over their frequency in fig. 66).

### Figure 68

#### The two distributions in figures 66 and 67 apply to the letters which come from column 1 of the transposition rectangles for the 12 cryptograms understudy. In the transposition rectangle for the 12 cryptograms understudy. In the transposition rectangle for the 12 cryptograms understudy. Therefore, beginning with position 1, the components in this column show a predominance of **F's** (over its frequency in fig. 66).
g. It has already been noted that beginning with the first letter of any one of the cryptograms, the type of alternation for column 1 is +. It is therefore not astonishing to find, within the first 10 letters, an alternation of the + type. Note how the V's and X's fall in the odd positions, the F's in the even. Thus:

<table>
<thead>
<tr>
<th>Letter</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Odd</td>
<td>A</td>
<td>D</td>
<td>F</td>
<td>G</td>
<td>V</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A</td>
<td>D</td>
<td>F</td>
<td>G</td>
</tr>
</tbody>
</table>

It is seen that there are 2 V's which fall in odd positions (1 and 5), but one V falls in an even position (10). There is an X, which falls in an odd position (7); there are 2 F's which fall in even positions (2 and 6). Unquestionably, then, the type of alternation, at least for the first 10 letters in each of these cryptograms, is +.

1. Take the next section of 10 letters in these two cryptograms. The letters are as follows:

III. D A G A F G A G Y
   XI. Y F D D Y A X G D A

Here there are 4 F's; 3 of them fall in odd positions (13, 17, 17), and one falls in an even position (12). There are 2 X's; one falls in an odd position (10), one in an even position (16). There are no V's among these letters.

2. So far as the evidence afforded by the F's is concerned, it would appear that this section of text shows the type 2 or "+ type" of alternation of components, since in type 1 or "- type" the F's occupy even positions and here the majority of them occupy odd positions. But so far as the X's are concerned, the evidence is equally balanced: one X falls in an odd position, one in an even position. There being no V's, no conclusions can be drawn from this letter. To be guided solely by the evidence afforded by the 3 F's may be unwarranted. Is it not possible to weight the frequencies of the letters so that it will be unnecessary to rely merely upon a few of them and the evidence afforded by all the letters can be taken into account? Why not assign frequency weights according to the two distributions in figures 66 and 67? The figures then become as follows:

Odd (1st, 3rd, . . . letters) Even (5th, 6th, . . . letters)

<table>
<thead>
<tr>
<th>Letter</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Type: + 1 + 1 + 1 = 4 + 1 - 1 = 3 - 1 = 2 letters

Since the odd letters have a total frequency of 42, the even, a total frequency of 85, for purposes of equalizing the distributions in applying the weights it seems advisable to deduct one-sixth from the total when applying the weights to odd letters.

c. Now in applying these weights to the letters, it must be borne in mind that since a transposition rectangle with an odd number of columns is involved, half of the letters are class 1 components, the other half are class 2 components. Hence, in finding the frequency value of the letters it is necessary to apply the weighted frequencies to alternate letters in the sections, as shown in figure 70.

<table>
<thead>
<tr>
<th>Letter</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Odd</td>
<td>A</td>
<td>D</td>
<td>F</td>
<td>G</td>
<td>V</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A</td>
<td>D</td>
<td>F</td>
<td>G</td>
</tr>
</tbody>
</table>

These distributions, when evaluated in accordance with figure 69, yield a total frequency value of 126; when evaluated in accordance with figure 69 reversed, yield a total frequency value of 143. The detailed calculations are as follows:

On the basis of figure 69 normal (odd letters as 01's, even letters as 02's):

\[ 0(3) + 3(11) + 3(3) + 3(4) + 0(11) + 1(0) = 72 \]

\[ 72 - 72 = 60 \]

\[ 1(4) + 3(10) + 1(1) + 4(5) + 0(4) + 1(1) = 66 \]

\[ 66 - 66 = 0 \]

Total = 126

On the basis of figure 69 reversed (even letters as 01's, odd letters as 02's):

\[ 1(3) + 3(11) + 1(10) + 4(8) + 0(11) + 1(0) = 77 \]

\[ 77 - 77 = 0 \]

\[ 0(4) + 3(10) + 3(11) + 3(0) + 0(4) + 1(1) = 79 \]

\[ 79 - 79 = 0 \]

Total = 143

f. Now the frequency sums here obtained (126 vs. 143) indicate that an alteration of the type \( \theta_{1} \rightarrow \theta_{0} \rightarrow \theta_{1} \) is in effect, that is, if a beginning is made with position 11, the type of alternation is \( --+ \). Since the type of alternation for the first 10 letters is \( +-- \) and for the second 10 letters \( --+ \), the reversal in alternation would indicate that column 1 of the transposition rectangle ends somewhere near the 10th letter. This same sort of reversal takes place after the 20th letter, as shown by the calculation in figure 71.

<table>
<thead>
<tr>
<th>Letter</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Odd</td>
<td>A</td>
<td>D</td>
<td>F</td>
<td>G</td>
<td>V</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A</td>
<td>D</td>
<td>F</td>
<td>G</td>
</tr>
</tbody>
</table>

Distribution of odd letters

<table>
<thead>
<tr>
<th>Letter</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Distribution of even letters

<table>
<thead>
<tr>
<th>Letter</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Total = 122
Beginning with the 21st position, the alternation is of type $\theta_1 \rightarrow \theta_2 \rightarrow \theta_3$; hence it is of the $"\pm"$ type. Again the reversal in type of alternation occurs in passing from the 24th set of 10 letters to the 53rd set, and this indicates that column 2 of the transposition rectangle ends somewhere near the 29th letter. But, fortunately, this time the exact location of the break is definitely indicated: The simultaneous appearance of V and X in the sequent positions 22 and 23 leads to the idea that the 22d letter marks the end of column 2 and the 23d letter marks the beginning of column 3.

There is nothing of an absolute nature in this point: It is merely an indication based upon probabilities and does not constitute a conclusive proof by any means. Now if there is this definite break at the end of 22 letters it means that columns 1 and 2 must each contain 11 letters. The calculations have heretofore been based upon sections of 10 letters and the results are therefore modified as shown in the following calculation:

### FIRST SECTION (letters 1-11)

<table>
<thead>
<tr>
<th>A</th>
<th>D</th>
<th>F</th>
<th>G</th>
<th>V</th>
<th>X</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

**Distribution of odd letters**

**Distribution of even letters**

**Weighted values of distributions:**

On the basis of figure 69 normal (odd letters as $\theta_1$'s, even letters as $\theta_2$'s):

$1(3)+5(11)+0(3)+3(8)+0(11)+1(6)=110$

$56 \div 6 = 92$

$1(4)+3(10)+1(11)+4(5)+1(4)+2(1)=71$

$110 \div 56 = 20$

**Total = 153**

On the basis of figure 69 reversed (even letters as $\theta_1$'s, odd letters as $\theta_2$'s):

$1(3)+5(11)+0(3)+3(8)+1(11)+2(6)=94$

$94 \div 6 = 16$

$0(4)+1(10)+5(11)+3(5)+0(4)+1(1)=81$

$81 \div 56 = 15$

**Total = 159**

Since the distribution here begins with an even-numbered position (12), and the greatest total is obtained on the basis of figure 69 reversed, the type of alternation for the second section of 11 letters is therefore again $\theta_1 \rightarrow \theta_2 \rightarrow \theta_3$ or $"\pm"$.

### SECOND SECTION (letters 12-22)

<table>
<thead>
<tr>
<th>A</th>
<th>D</th>
<th>F</th>
<th>G</th>
<th>V</th>
<th>X</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

**Distribution of odd letters**

**Distribution of even letters**

**Weighted values of distributions:**

On the basis of figure 69 normal (odd letters as $\theta_1$'s, even letters as $\theta_2$'s):

$2(3)+1(11)+2(3)+1(8)+1(11)+0(6)=51$

$51 \div 6 = 8$

$1(4)+5(10)+0(11)+3(5)+2(4)+1(1)=78$

$78 \div 56 = 14$

**Total = 120**

On the basis of figure 69 reversed (even letters as $\theta_1$'s, odd letters as $\theta_2$'s):

$1(3)+3(11)+3(3)+1(8)+0(11)+2(6)=65$

$65 \div 6 = 11$

$0(4)+3(10)+3(11)+1(5)+0(4)+2(1)=74$

$74 \div 56 = 13$

**Total = 153**

Since the best values are obtained on the basis of figure 69 normal, the type of alternation for the third section of 11 letters is $\theta_1 \rightarrow \theta_2 \rightarrow \theta_3$ or $"\pm"$.
k. Now if columns 1 and 2 contain 11 letters, and the total number of letters is 186, the transposition rectangle obviously has 17 columns, there being 16 long columns of 11 letters and one short column of 10 letters ([17 x 11] - 1 = 186).  

l. There is another cryptogram which also contains but one short column, viz., VII, of 254 letters [17 x 15] - 1 = 254. The columns of this cryptogram contain 4 more letters than the corresponding columns of XIII and XI. Assuming, momentarily, that the last column is the short one, cryptogram VII may be added to the superposition of XIII and XI, provided these sets of 4 additional letters are accounted for. This has been done in figure 73. In that figure the 4 extra letters pertaining to cryptogram VII are shown as falling under the last letters of the columns of cryptograms XIII and XI, but this is only an arbitrary placement. It is sufficient to place these extra letters in such positions as will make the first one of the series begin in an even position.

m. Since the transposition rectangle is now known to be 17 columns wide, the data in figure 69 may be enlarged to correspond to this information. For example, whereas in originally constructing figure 69 the first column of cryptogram I was assumed to have only 8 letters (to correspond to a key of 25 numbers), it may now be extended to a column of 12 letters, and so on. The additional portions used to make the distributions in figure 74 are shown underlined in figure 73.
Weighted values of distributions:

On the basis of figure 1 normal (odd letters as $O_1$, even letters as $O_2$):

$$
\begin{align*}
2(4) + 4(14) + 1(5) + 3(11) + 7(15) + 3(10) &= 237 \\
4(9) + 4(10) + 5(14) + 2(8) + 0(7) + 2(2) &= 183 \\
\text{Total} &= 413
\end{align*}
$$

On the basis of figure 2 normal (odd letters as $O_1$, even letters as $O_2$):

$$
\begin{align*}
4(4) + 4(14) + 5(5) + 2(11) + 0(15) + 2(10) &= 189 \\
2(9) + 4(15) + 1(14) + 3(8) + 7(7) + 3(2) &= 171 \\
\text{Total} &= 360
\end{align*}
$$

Since the greatest total is obtained on the basis of figure 1 normal, the type of alternation for the third section of letters is $O_1 \rightarrow O_2 \rightarrow O_1$, or "+".

Continuing the foregoing process with the letters beyond position 33, the data are as follows:

### Distribution of odd letters

<table>
<thead>
<tr>
<th>A</th>
<th>D</th>
<th>F</th>
<th>G</th>
<th>V</th>
<th>X</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>7</td>
<td>3</td>
<td>7</td>
<td>3</td>
<td>7</td>
</tr>
</tbody>
</table>

### Distribution of even letters

<table>
<thead>
<tr>
<th>A</th>
<th>D</th>
<th>F</th>
<th>G</th>
<th>V</th>
<th>X</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>7</td>
<td>3</td>
<td>7</td>
<td>3</td>
<td>7</td>
</tr>
</tbody>
</table>

Weighted values of distributions:

On the basis of figure 4 normal (odd letters as $O_1$, even letters as $O_2$):

$$
\begin{align*}
4(4) + 5(14) + 1(5) + 3(11) + 3(15) + 4(10) &= 209 \\
3(9) + 4(15) + 6(14) + 4(8) + 0(7) + 0(2) &= 203 \\
\text{Total} &= 412
\end{align*}
$$

On the basis of figure 2 reversed (even letters as $O_1$, odd letters as $O_2$):

$$
\begin{align*}
3(4) + 4(14) + 6(5) + 4(11) + 0(15) + 0(10) &= 142 \\
4(9) + 5(15) + 1(14) + 3(8) + 3(7) + 4(2) &= 173 \\
\text{Total} &= 315
\end{align*}
$$

Since the distribution begins here with an even-numbered position (34), and the greatest total is obtained on the basis of figure 4 normal, hence the alternation for the fourth section or column is of the type $O_1 \rightarrow O_2 \rightarrow O_1$, or "-".

The types of alternation for the first 5 columns, which are all long columns, is therefore $++--+$.

The calculations for the next or sixth section of 11 letters are shown below:

### Distribution of odd letters

<table>
<thead>
<tr>
<th>A</th>
<th>D</th>
<th>F</th>
<th>G</th>
<th>V</th>
<th>X</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>7</td>
<td>3</td>
<td>7</td>
<td>3</td>
<td>7</td>
</tr>
</tbody>
</table>

### Distribution of even letters

<table>
<thead>
<tr>
<th>A</th>
<th>D</th>
<th>F</th>
<th>G</th>
<th>V</th>
<th>X</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7</td>
<td>1</td>
<td>7</td>
<td>1</td>
<td>7</td>
</tr>
</tbody>
</table>

Since cryptograms III and XI contain but one short column, it is advisable to be on the lookout for it as the work progresses. It is possible to continue with the process detailed above. For example, the calculations for the next or sixth section of 11 letters are shown below:
Weighted values of distributions:

On the basis of figure 74 normal (odd letters as $0_1$, even letters as $0_2$):

\[
\begin{align*}
5(4)+4(14)+4(5)+1(11)+1(15)+2(10) &= 142 \\
1(9)+7(15)+1(14)+2(8)+8(7)+1(2) &= 202 \\
\text{Total} &= 544
\end{align*}
\]

On the basis of figure 74 reversed (even letters as $0_1$, odd letters as $0_2$):

\[
\begin{align*}
1(4)+7(14)+1(5)+2(11)+8(15)+1(10) &= 259 \\
8(9)+4(15)+4(14)+1(8)+1(7)+2(2) &= 180 \\
\text{Total} &= 439
\end{align*}
\]

Since the distribution starts with an even position (56) and the greatest total is obtained on the basis of figure 74 reversed, the type of alternation for the sixth section or column is $0_1=0_2=0_3$, or "++-".

But perhaps advantage should be taken of the availability of additional cryptograms. For example, cryptogram $V$, of 202 letters, has 2 short columns ($1/2 \times 12 = 202$), whereas the cryptograms thus far dealt with each have but one. That is, cryptogram $V$ has one short column in common with cryptograms III, XI, and VII, and one additional short column not possessed by the latter. Can this additional short column of cryptogram $V$ be located?

Suppose column 1 of cryptogram $V$ is the additional short column. Then the letters of column 2 would be $FXFXFFFGF$. These letters when evaluated on the basis of figure 74 normal (odd letters as 0's, even letters as 1's) yields a total of 77; when weighted on the basis of figure 74 reversed, a total of 144. The calculation is as follows:

\[
\begin{align*}
\text{Distribution of odd letters} & \\
A & D & F & G & V & X \\
\frac{1}{51} & \frac{3}{51} & \frac{1}{51} & \frac{1}{51} & \frac{1}{51} & \frac{1}{51}
\end{align*}
\]

\[
\begin{align*}
\text{Distribution of even letters} & \\
A & D & F & G & V & X \\
\frac{1}{9} & \frac{3}{9} & \frac{1}{9} & \frac{3}{9} & \frac{1}{9} & \frac{1}{9}
\end{align*}
\]

On the basis of figure 74 normal (odd letters as $0_1$, even letters as $0_2$):

\[
\begin{align*}
1(4)+0(14)+5(5)+0(11)+0(15)+0(10) &= 29 \\
0(9)+1(15)+1(14)+1(8)+1(7)+2(2) &= 48 \\
\text{Total} &= 77
\end{align*}
\]

On the basis of figure 74 reversed (even letters as $0_1$, odd letters as $0_2$):

\[
\begin{align*}
1(4)+0(14)+5(5)+0(11)+0(15)+0(10) &= 29 \\
0(9)+1(15)+1(14)+1(8)+1(7)+2(2) &= 48 \\
\text{Total} &= 77
\end{align*}
\]

According to this calculation, column 2 of cryptogram $V$ seems to correspond to the type of alternation $0_1=0_2=0_3$, that is "++-". But from previous work it is fairly certain that column 2 is of the "+-" type. Hence, column 1 of cryptogram $V$ is probably not the additional short column of that message. Assuming column 2 to be the extra short column, no such contradiction is obtained, for the calculation is as follows:

\[
\begin{align*}
\text{Distribution of odd letters} & \\
A & D & F & G & V & X \\
\frac{1}{51} & \frac{3}{51} & \frac{1}{51} & \frac{1}{51} & \frac{1}{51} & \frac{1}{51}
\end{align*}
\]

\[
\begin{align*}
\text{Distribution of even letters} & \\
A & D & F & G & V & X \\
\frac{1}{9} & \frac{3}{9} & \frac{1}{9} & \frac{3}{9} & \frac{1}{9} & \frac{1}{9}
\end{align*}
\]

On the basis of figure 74 normal (odd letters as $0_1$, even letters as $0_2$):

\[
\begin{align*}
1(4)+0(14)+5(5)+0(11)+0(15)+0(10) &= 29 \\
0(0)+1(15)+1(14)+1(8)+1(7)+2(2) &= 48 \\
\text{Total} &= 77
\end{align*}
\]

Assuming column 2 to be short, the letters of column 3 are $X A V D A G F D V D G F$.

Weighted values of distributions:

On the basis of figure 74 normal (odd letters as $0_1$, even letters as $0_2$):

\[
\begin{align*}
1(4)+0(14)+5(5)+0(11)+0(15)+0(10) &= 60 \\
1(0)+3(16)+11(14)+1(18)+9(7)+0(2) &= 76 \\
\text{Total} &= 136
\end{align*}
\]

On the basis of figures 74 reversed (even letters as $0_1$, odd letters as $0_2$):

\[
\begin{align*}
1(0)+0(13)+1(14)+1(18)+2(7)+1(2) &= 47 \\
1(4)+3(14)+1(5)+1(11)+0(10)+1(10) &= 62 \\
\text{Total} &= 109
\end{align*}
\]

Since the greatest total is obtained on the basis of figure 74 normal, the type of alternation is $0_1=0_2=0_3$, and column 3 is a "+-" column, which is consistent with the formula $++-++$ for columns 1 to 5, as previously ascertained.

If all the foregoing reasoning is correct, and column 2 is the additional short column for cryptogram $V$, it must be the next to the last column of the transposition rectangle. Since it is a "+-" column, the last column must be a "-" column; one; therefore, there are 9 "++-" columns and 8 "-" columns. This definitely determines that the "-" columns are the odd ones, the "+-" columns the even ones, since in an odd-width rectangle there is one more odd column than even columns.

The single short column which is common to cryptograms III, XI, and VII is one of the columns beyond column 5. Assuming each possibility in turn, there is obtained for the type of alternation in each column the distributions of "++" and "-" as shown in figure 75.

![Figure 75](image)

The correct assumption must satisfy the following conditions:

(a) There must be 9 "+-" and 8 "+-" columns.

(b) The short column must be "-".
Only assumptions (3) and (5), in which column 8 and column 10 are short columns, satisfy these conditions. Therefore, column 2 is followed by either column 8 or 10. Testing the combination 2-8 for monoe-\(/\)alphabeticity of bipartite pairs, the distribution shown in figure 76 is obtained. When combination 2-10 is tested, the distribution shown in figure 77 is obtained. Obviously, the 2-8 combination is the better.

\[ E(\phi) = 0.0667 \times 17 \times 16 = 18.14 \]
\[ \phi = 22 \]

\[ E(\phi) = 0.0667 \times 17 \times 16 = 18.14 \]
\[ \phi = 4 \]

Figure 76.

2. It is possible by introducing cryptograms with additional short columns to determine more of the key. Thus, it was found by using cryptograms XII and VI that the first 3 numbers of the transposition key are 16-5-7. But the process of anagramming will yield the solution at least as rapidly. In this process, of course, advantage may be taken of the fact that the columns have been classified into the "+" and "-" types and no combinations of two "++" or two "--" columns need be tested, since only combinations of the type "++" or "--" are permissible.

3. The transposition key and the substitution checkerboard are shown in figure 78.

\[ 7 \quad 6 \quad 5 \quad 4 \quad 3 \quad 2 \quad 1 \quad 11 \quad 12 \quad 13 \quad 14 \quad 15 \quad 16 \]

\[ \text{FIGURE 78.} \]

A. All the foregoing details concern a case in which the transposition rectangle has an odd number of columns. Now if the rectangle contains an even number of columns, this type of solution is, of course, still applicable, and in fact is easier, since the letters of the text of the re-
spective sections do not have to be distributed into odd and even letters. It is only necessary to identify a section as being composed of initial components or of final components. This analysis then produces a series of sections corresponding in number with the number of columns in the transposition rectangle. This number will, of course, be even. By a careful study of where alternations in composition of components (0, or 02) occur, the division of the text into sections corresponding to long and short columns can be accomplished. The remaining steps are obvious and follow the lines elucidated in paragraph 39e-j.

y. The entire structure upon which this general solution rests is destroyed if the substitution checkerboard has been consciously manipulated to equalize or flatten out the sums of the weighted frequencies of the letters in its rows and columns. For example, note the following checkerboard, which is not "perfect" but gives approximately similar frequencies in its rows and columns.

\[ \begin{array}{cccccc}
A & D & F & G & V & X \\
I & 14 & Q & S & U & 30 \\
T & 20 & 10 & 8 & 15 & 20 \\
C & 50 & 35 & 15 & 10 & 10 \\
D & 20 & 10 & 8 & 15 & 20 \\
G & 10 & 50 & 35 & 15 & 10 \\
X & 20 & 10 & 8 & 15 & 20 \\
Z & 30 & 20 & 10 & 8 & 15 \\
S & 10 & 50 & 35 & 15 & 10 \\
E & 10 & 50 & 35 & 15 & 10 \\
R & 10 & 50 & 35 & 15 & 10 \\
H & 10 & 50 & 35 & 15 & 10 \\
\end{array} \]

\[ \text{FOUR SECTIONS} \]

\[ \begin{array}{cccccc}
A & D & F & G & V & X \\
14 & Q & S & U & 30 & 180 \\
20 & 10 & 8 & 15 & 20 & 180 \\
50 & 35 & 15 & 10 & 10 & 180 \\
20 & 10 & 8 & 15 & 20 & 180 \\
10 & 50 & 35 & 15 & 10 & 180 \\
20 & 10 & 8 & 15 & 20 & 180 \\
30 & 20 & 10 & 8 & 15 & 180 \\
10 & 50 & 35 & 15 & 10 & 180 \\
10 & 50 & 35 & 15 & 10 & 180 \\
10 & 50 & 35 & 15 & 10 & 180 \\
\end{array} \]

\[ \text{FOUR SECTIONS} \]

v. If the statistical calculations upon which this general solution is based make use of the logarithms of the frequencies instead of the frequencies themselves, much more accurate and clear-cut data will be obtained.

* The frequencies indicated as those given in fig. 3, p. 13, Military Cryptanalysis, Part I.
SECTION X

SOLUTION OF BIFID FRACTIONATING SYSTEMS

Review of principles underlying the cryptographic method .......................................................... 44
Example of a simple bifid cipher ........................................................................................................ 45
Principles of solution .......................................................................................................................... 46
Example of solution ........................................................................................................................... 47
Special solutions .................................................................................................................................. 48
Periodic bifid ciphers .......................................................................................................................... 49
General principles underlying the solution ......................................................................................... 50
Ascertaining the length of the period ................................................................................................. 51
Illustration of solution ......................................................................................................................... 52
Special solutions for bifid systems ................................................................................................... 53
Solution of trifid systems .................................................................................................................. 54
Concluding remarks on fractionating systems .................................................................................. 55
Concluding remarks on transposition systems .................................................................................. 56

44. Review of principles underlying the cryptographic method.—Several bifid fractionating systems have been explained in previous texts of this series. In certain of these systems four basic steps are involved, two of substitution and two of transposition. These steps may be briefly described as follows: (1) A process of decomposition (substitution), in which each plaintext letter is replaced by two components, or elements, of a bifid or bipartite alphabet; (2) a process of separation (transposition), in which the two components originally paired together are separated; (3) a process of recomposition (transposition), in which the separated components are combined to form new pairs; (4) a process of recomposition (substitution), in which each new pair of components is given a letter value according to the original or a different bifid alphabet.

45. Example of a simple bifid cipher.—a. One of the simplest bifid fractionating systems is that exemplified in the following subparagraphs. It will be employed to set forth certain principles in the general solution of systems of this and similar nature.

b. Given the 25-cell substitution checkerboard shown in figure 80, let the message to be enciphered be ONE PLANE REPORTED LOST AT SEA. The first step is to replace the plaintext letters by the bipartite equivalents, the two elements or components being set down vertically beneath the plain-text letters. This represents the first two of the four processes referred to in paragraph 44, the first being that of decomposition or substitution, the second, that of separation or transposition, represented by the manner in which the two bipartite elements are set down vertically (instead of horizontally), thus separating the two elements from their normal horizontal juxtaposition.

See Special Test No. 168, Advanced Military Cryptography, sec. XI and Military Cryptanalysis, Part I, sec. IX.

44

Plain text ONE PLANE REPORTED LOST AT SEA
Components

\[
\begin{array}{cccc}
1 & 2 & 3 & 4 \\
M & A & N & U & F \\
C & T & R & I & G \\
B & D & E & H & K \\
L & O & P & Q & S \\
V & W & X & Y & Z \\
\end{array}
\]

The third process, that of recomposition or reconstruction, also involving a transposition, is now to be performed and will consist in combining elements standing in diagonal relationship to the right, that is, as shown by the arrows below:

ONE PLANE REPORTED LOST AT SEA

\[
\begin{array}{cccc}
7 & 9 & 11 & 13 \\
1 & 2 & 3 & 4 \\
5 & 6 & 1 & 2 & 3 & 4 & 5 & 6 & 1 & 2 & 3 & 4 & 5 & 6 & 1 & 2 & 3 & 4 & 5 & 6 & 1 & 2 & 3 & 4 & 5 \\
\end{array}
\]

giving the pairs 21, 33, 34, 11, 21, 33, etc. There are left, at the end of the process, one element in the upper line at the extreme left and another element in the lower line at the extreme right, yielding the pair 24, which may be placed at the head or tail of the resultant combinations, as prescribed. The last or fourth process, that of recomposition or substitution, is to replace the new pairs of components by letters from the original or a new checkerboard. If the same checkerboard is used, it yields the text shown herewith:

Plain text ONE PLANE REPORTED LOST AT SEA

Components

\[
\begin{array}{cccc}
1 & 2 & 3 & 4 \\
M & A & N & U & F \\
C & T & R & I & G \\
B & D & E & H & K \\
L & O & P & Q & S \\
V & W & X & Y & Z \\
\end{array}
\]

Cipher C E H M C E D E H T D E I U W C T I X B I

c. Another and perhaps more simple way of accomplishing the same process is to set down the bipartite equivalents horizontally and recombine them as shown below:

ONE PLANE REPORTED LOST AT SEA

Components

\[
\begin{array}{cccc}
1 & 2 & 3 & 4 \\
M & A & N & U & F \\
C & T & R & I & G \\
B & D & E & H & K \\
L & O & P & Q & S \\
V & W & X & Y & Z \\
\end{array}
\]

The results are identical with those obtained from the preceding manner of operation. The text is of course sent in 5-letter groups.

d. Instead of using the digits 1, 2, 3, 4, 5, as the bipartite components one can use the vowels A, E, I, O, U, or any other characters that are deemed suitable. Perhaps digits are best as they are less likely to be confused with letters of the text.

e. As intimated above, the checkerboard used for the recomposition may be different from that employed in the decomposition. But it will be shown that the additional safety afforded by using two different checkerboards is somewhat illusory, and is by no means as great as may appear on first consideration.
60. Principles of solution.—a. Note the following skeleton encipherments, using the checkerboard shown in figure 80:

<table>
<thead>
<tr>
<th>Plain</th>
<th>Cipher</th>
</tr>
</thead>
<tbody>
<tr>
<td>C E N</td>
<td>L P U</td>
</tr>
<tr>
<td>O S E</td>
<td>R O M</td>
</tr>
<tr>
<td>N R H</td>
<td>T I A</td>
</tr>
<tr>
<td>X E D</td>
<td>E H A</td>
</tr>
<tr>
<td>T E Y</td>
<td>T E N</td>
</tr>
</tbody>
</table>

These five encipherments have in common a plain-text digraph EN. The five cipher versions, however, have only a single letter in common, B. This is, of course, a phenomenon already encountered many times by the student and its cause is easily understood by him: The mechanics of the system tend to reduce by one character the lengths of the repetitions in the cipher text, as compared with their lengths in the plain text, a digraphic repetition in the plain text manifesting itself as a digraphic repetition in the cipher text, a tetragraphic one becoming a trigraphic, and so on. More will later be stated on this phase of the matter.

b. But now study the individual cipher letter immediately preceding and succeeding the cipher letter which those five encipherments have in common. They are as follows:

Letters preceding B —— N, X, E, P, R
Letters succeeding B —— H, E, D, B, K

Reference to the checkerboard discloses the very interesting and important fact that the letters preceding the cipher repetition (B) all come from the same column in the checkerboard, the letters succeeding the repetition all come from the same row in the checkerboard. How this phenomenon is brought about is quite simple to see. Take the first of the five examples, that in which C E N O produces N B H. The N is the result of combining the second component of the bipartite equivalent of C with the first component bipartite equivalent of E, yielding the combination 12, which is N. No matter what the other three letters in the plain-text tetragraph may be, if the second letter is E, the second component bifid equivalent of the first letter of the cipher tetragraph must be 3. This means that this first letter of the cipher tetragraph must come from column 3 of the checkerboard. Exactly which row this letter will come from is determined by the identity of the second component of the bifid equivalent of the first letter of the plain-text tetragraph. Hence, since the 5 tetragraphs in the example all have the same plain-text letter in the second position, the initial letters of the cipher tetragraphs all must come from the same column of the checkerboard. It is unnecessary to go through the reasoning, which is parallel, in the case of the third letters of the cipher tetragraphs: these must all come from the same row of the checkerboard.

c. A good understanding of the phenomenon just noted can certainly be employed to advantage in solving this and similar types of systems, for it becomes obvious that a careful study of the letters immediately preceding and following cipher repetitions should facilitate a reconstruction of the checkerboard employed in the substitution. Indeed, if there were no other phenomena to disturb this very simple relationship, solution would be quite easy. All that would be required would be to study the prefixes and suffixes to all the A's, B's, C's, . . . in the cryptogram, find the letters which belong in the same columns and rows of the checkerboard, and the reconstruction of the latter would follow very simply. Unfortunately, however, there is a disturbing phenomenon which must now be considered.

The principle involved in such reconstruction was, to my knowledge, first pointed out and successfully employed early in 1938 by Associate Cryptoanalyst S. Kullback and A. Rinikov.
A 4-letter plain-text repetition appears as a 4-letter cipher-text repetition because the second component of the TR, and TR happen to be identical (if R and T are in the same columns of the checkerboard).

f. From a study of these phenomena the rule may be deduced that a 4-letter repetition in the plain text is really reduced to an (n-1)-letter repetition in the cipher text, but it can happen fortuitously that the real repetition is extended on either or both ends of the repetition by a pseudo-repetitious letter. Hence, a 3-letter plain-text repetition may appear as a 2-, 3-, or 4-letter repetition in the cipher text.

A. It is therefore possible to make wholly erroneous deductions from some repetitions, especially if the latter are short. Note for instance the following example, still using Fig. 90 as a basis:

\[
\begin{align*}
(A) & \quad I \quad N \quad T \quad H \quad E \quad G \quad 15 \quad 35 \\
(L) & \quad D \quad R \quad P \quad D \quad 15 \quad 35 \\
(B) & \quad R \quad E \quad U \quad 15 \quad 35 \\
(T) & \quad D \quad R \quad P \quad B \quad 15 \quad 35
\end{align*}
\]

Here are 2 sequences of 5 cipher letters, identical save in the central letter, and yet the 6-letter plain text sequences have only 2 letters in common. This example is cited to show that the plain text is really reduced to an (n-1)-letter repetition in the cipher text, but it can happen fortuitously that the accidental repetitions are such as to make wholly erroneous deductions from some repetitions.

First, the 7-letter cipher sequence HI WOTQC is common to all four cases; if only the plain text were available, one could conclude that the plain-text repetition consists of 8 letters. Second, the letters H and Y probably come from the same column in the checkerboard, but as for P, and T, they may or may not come from the same column, most probably not. (Actually, 0 and T do, but P does not come from the same column as these 2 letters.) Third, the letters H and E probably come from the same row in the checkerboard, but as for C, and M, they may or may not come from the same row, most probably not. (Actually, all 5 letters come from different rows.)

e. Note the following cases of encipherment: The fact that the 7-letter cipher sequence HI WOTQC is common to all four cases means that the plain-text repetition consists of 8 letters.
The pseudo-repetitious letters, I in the third position in cases (1) and (2), and the letters Y, and U, in corresponding positions in cases (3) and (4) mean that I, U, and Y, come from the same column of the checkerboard. The I, in position 2, in cases (1), (3), and (4) and the D, in the corresponding position in case (2) indicates that I, and D, are probably in different columns in the checkerboard. In position 11, the H, B, and D, give indications of being in the same row of the checkerboard. In position 14, W, and Y, likewise give indications of coming from the same row. But note now that from position 12 it may be deduced that W, A, and T come from the same column of the checkerboard. These are examples of the type of detailed analysis that the student should follow in his attempt to solve a problem of this sort.

l. In general it may be said that when the repetitions are numerous and fairly lengthy, that is when there is a good deal of traffic all in the same checkerboard, and repetitions of tetragraphs and better are plentiful, solution should be relatively easy. In fact, with a fairly large amount of traffic, most of the work involved would consist in listing the 2, 3, 4 . . . letter repetitions. Then a chart would be drawn up to show the associations which the prefixes make among themselves and the associations which the suffixes make among themselves. For example:

\[
\begin{align*}
X & A B Q \\
N & A B R \\
Z & A B T \\
N & A B Q \\
L & A B I \\
Z & A B T \\
\end{align*}
\]

Here it is noted that L, N, X, and Z appear as prefixes to repetitions. The letter X is “found in company” with N twice; the “association value” of X and N is 2 units. The association value of Z and N is, however, 4 units, for the N occurs twice and so does the Z. The association value of LX or LZ is 1 unit; that of LN or LZ, 2 units. Thus, the association value for each combination can be studied in all the repetitions and, of course, when the value is high for a given combination it indicates that the two letters really belong together, or in the same column of the checkerboard.

m. What can be done with but one or two relatively short cryptograms depends largely upon their lengths, the number of repetitions they happen to have, the exact construction of the checkerboard, and the ingenuity and patience of the cryptanalyst. Once the letters that constitute the columns and the rows of the checkerboard employed in the recomposition are known, the proper assembling of the columns and rows is a relatively simple matter. If a key-word has been used as the basis for the distribution or mixing of the letters, naturally the reconstruction of the checkerboard is much facilitated. If not, then either the original or an equivalent checkerboard may be reconstructed. Having the recomposition checkerboard at hand, the determination as to whether it is the same as that used in the decomposition follows directly. If not the same, the reconstruction of the decomposition checkerboard is a relatively simple matter.
47. Example of solution.—a. Suppose the two following cryptograms suspected of being in
the same key are at hand:

<table>
<thead>
<tr>
<th>Line</th>
<th>No. 1</th>
<th>No. 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>ISDCY IWVEW DNBAX WSAVI WILT K</td>
<td>I WILT GSIHW WAWKI NDCUW WAXLX</td>
</tr>
<tr>
<td>B</td>
<td>XLYIW DAINB DRWVR VSDDT NULQD</td>
<td>D I WIR CVNOD NGSLN GIGWL VFDWV</td>
</tr>
<tr>
<td>C</td>
<td>UBINR KVLQD NLQDT VITIN NDCUW</td>
<td>R VSDD TILLQ DUKDW SSHXS ENCQD</td>
</tr>
<tr>
<td>D</td>
<td>WDCVS ODGS ISSGU HKIRX YIRNG</td>
<td>T QGTE UDVC I OIWTX WVRVS DDTNU</td>
</tr>
<tr>
<td>E</td>
<td>TIWAU UIAAQ CWNBIEWAYU LQSAV</td>
<td>L QDUH KIRXY IRNBA YIIG TEWAY</td>
</tr>
<tr>
<td>F</td>
<td>NIUH QDTXW ITAVN NINDC UWWCG</td>
<td>I I LWN KIRNG STLBD DTITI LUDVL</td>
</tr>
<tr>
<td>G</td>
<td>TLGWW ALWDS NSILN NIRNG STLBD</td>
<td>G VTTAQ AVNIU UHQDT XDHYI WXNIN</td>
</tr>
<tr>
<td>H</td>
<td>DTWAU UIAAQ DGTEL QDTXWD WYIWX</td>
<td>B DRWS CLPLW AHNTL QDTXC WSCIV</td>
</tr>
<tr>
<td>I</td>
<td>FNII UH QDTXW ITAVN NINDC UWWCG</td>
<td>D TNIL VNODN ULDU BGTEW VRVSD</td>
</tr>
<tr>
<td>J</td>
<td>TDTW</td>
<td></td>
</tr>
</tbody>
</table>

b. A careful and detailed listing of significant repetitions is made, these to show the single-
letter prefix and suffix in each case. A partial list of the many repetitions present in the two
cryptograms is given in figure 81.

c. Consider the first set of repetitions listed in figure 81, as extracted and shown in figure 82.
According to the principles elucidated in the preceding paragraph, it would seem that the following tentative deductions may be made from the data contained in the columns of figure 82:

1. From column 1: E, D, R, P, X belong in the same column of the checkerboard.
2. From column 10: W, I, N belong in the same row.
3. From column 14: D and G belong in the same column.
4. From column 16: B and H belong in the same row.
5. From column 16: H and L belong in the same column.
7. From column 25: G and B belong in the same row.
8. From column 27: I and T belong in the same column.
9. From column 33: I and W belong in the same row.
10. From column 36: C and D belong in the same row.
11. From column 37: W and Q belong in the same row.
12. From column 40: Q and V belong in the same row.

It would be most fortunate and unusual for all these tentative deductions to be correct, for the disturbing effects of accidental adventitious repetitions have not been taken into account as yet. But let an attempt be made to assemble the data deduced thus far, to see if they can all be reconciled.

d. Tentative deduction (1) indicates that E, D, R, P, and X belong in the same column of the recomposition checkerboard. If correct, the complete set of 5 letters of one column is at hand. But tentative deduction (3) indicates that D and G belong in the same column and this would mean that the column has 6 letters, which is impossible. Further evidence will be required to corroborate the hypothesis that E, D, R, P, and X are all actually in the same column, or that D and G are actually in the same column. For this purpose, further study must be made, and it is convenient to compile an "association table" showing how often certain letters are associated among themselves as prefixes to the repetitions. A similar association table is made for the suffixes. The tables may be combined in a manner similar to that shown in figure 83, where the prefixes to repetitions appear at the left of the central alphabet, the suffixes to the right.

Take column 1 of figure 82, having D, E, P, R, and X as prefixes to a long repetition. A stroke is placed in the E, P, R, and X cells of row D; a stroke is placed in the P, R, and X cells of row E; a stroke is placed in the R and X cells of row P; and finally a stroke is placed in the X cell of row R. Again, take column 16 of figure 82, reading B H H H L. The B need not be considered, since it is not a prefix to the repetition beginning K I R X Y . . . , but the H and L may be considered. In the L cell of row H three strokes are inserted to indicate that H and L are associated that many times. Each time a datum is obtained, it is added to this table. Figure 83 shows the appearance of the table after all the data obtainable from the repetitions listed in subparagraph b have been inserted. From even this small amount of material a few deductions can be made. For example, it is seen that the B line of the table for prefixes shows 5 strokes at G and 3 strokes at W, from which it would appear that B, G, and W may be in the same column. The letters C and L likewise seem to be in the same column, as do H and L, making C, H, and L appear to be in the same column. Studying the table of suffixes, it would appear that B and H are in the same row; I and N are in a row. After the entire text has been examined and the prefixes and suffixes distributed in this way, the whole table is studied carefully with a view to eliminating the effects of the accidental or pseudo-repetitious letters, trying to locate those letters which represent the prefixes and suffixes of true repetitions.
e. Suppose the data have been reduced to the following:

Letters belonging in same columns
(1) D, G, U
(2) H, L, C
(3) Y, K, N
(4) W, Q, S, T, B

Letters belonging in same rows
(1) G, B, H, K
(2) I, W, N, L
(3) D, C, A, S
(4) Q, V, X, Y, Z
(5) T, U

The presumption that Q, V, X, Y, and Z are all in the same row leads to the assumption that the mixing of the checkerboard is based upon a key word or key phrase. Following up this hypothesis, the data are assembled in the following manner:

```
  1  2  3  4  5
1  W I L N
2  T U
3  S D C A
4  B G H K
5  Q V X Y Z
```

Figure 84.

f. Only 6 letters remain to be placed in the checkerboard. But there are enough letters already placed to warrant an immediate attempt at decipherment. For example, take the first few groups of message No. 2 and replace the letters by their bipartite equivalents:

```
I   W   I   T   G   S   I   H   W   W   A   W   K   I   N   D
12  11  12  13  21  31  12  43  11  11  34  11  44  12  14  32
```

Recombining the bifid elements:

```
.1  21  11  21  32  14  23  11  24  31  11  13  41  14  41  21  43  2.
```

Substituting by means of figure 84:

```
.1  21  11  21  32  14  23  11  24  31  11  13  41  14  41  21  43  2.
```

Obviously the decomposition and recomposition checkerboards are different. But the reconstruction of the former is not at all difficult, since the text is now in monoalphabetic form. The message begins with a group showing a repeated letter in the first and third positions: is the 1st word E N E M Y? Probably it is, for message No. 1 also contains the sequence W I L T. At any rate, a transcription of the cryptograms into the bifid equivalents given by the nearly complete recomposition checkerboard (fig. 84) soon yields sufficient monoalphabetic text to permit of the complete reconstruction of both checkerboards:
11,;1
recomposition checkerboard.

on the box

FABRIC

label on the box containing

INFAM

g.

GI

Speculating upon the key words

~ =~

5 K

4

3

3

1

NUTYPE

(Decomposition)

This yields the sequence for the decomposition checkerboard. The legend

reads: WILLABLE AND FLUID

component

A

B

(Decomposition)

(Recomposition)

of this chemical.

used
typewriter cleaning

to produce the mixed sequences in these checkerboards. The

are as follows:

No. 1

ISDCYIWEWDBAXWSAVIWIILX
12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35
ITISREPORTEDTHATTHEENEMY

XLTYWDAINDBRWRVRSDTNULQD
40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63
ARETIREDTOPOSITIONWESTOF

UBINRKYQLQDNQDVTWININDCUW
64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87
FNEWCHESTERSTOPONEREGIMENT

WDCVSDOCGSISSGUHIRXYIRNG
88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110
TISINVICINITYOFCROSSROAD

TIWAUDIJAACQINWSNIESWAULQASV
111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134
NETWEOIGHTANDNORTWESTTHE

NILUHEDTWTIATVNNINDCUWCG
135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159
REOFSTOPANOTHERREGIMENTCO

TLGWWALWDSNILNNRINGSTLBD
160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184
CENTRATINGNEARROADJUNCTI

DTWAUDIJAAQDGETLQDTXDIWYIX
185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209
ONTWEOIGHTFOURSTOPBEPREPA

NINDBRQTEDWCSLCPTKIRX
210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234
REDTOSUPPORTATTACKONCROSS

YIRNTIWAUIDAACQQTZIWIY
235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259
ROADONETWEOIGHTSTOPKEEPON

ININDCUWDSNILVHLQDTXWS
260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284
EREGIMENTINRESERVESTOPATT

CLPWRVSDTNULQDUH KirbyYI
285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309
ACKPOSITIONWESTOFCROSSROAD

NGTIWAUUAACQQDTY
310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334
DONETWEOIGHTSTOP

The two cryptograms may now be deciphered directly from the checkerboards. The

plain-texts

follows:

1 2 3 4 5

1 NUTFYP

2 EOXFL

3 AMBCS

4 RWGDH

5 KQVXZ

1 WILNO

2 TUREM

3 SDCF AF

4 EGHKP

5 QVXYZ

A

(B)
different sets of checkerboards. An interesting solution, however, is that wherein the same message has been cryptographed by two texts but different cryptographic texts are available for examination. They are superimposed for this system, since no special conditions are prerequisite to the procedure set forth.

No. 1. G
No. 2. W
No. 3. L
No. 4. T

Suppose, for instance, that in this system two cryptograms of identical lengths and plain texts but different cryptographic texts are available for examination. They are superimposed for this system, since no special conditions are prerequisite to the procedure set forth.

48. Special solution.—a. The preceding example of solution constitutes the general solution for this system, since no special conditions are requisite to the procedure set forth. An interesting solution, however, is that wherein the same message has been cryptographed by two different sets of checkerboards.

b. Suppose, for instance, that in this system two cryptograms of identical lengths and plain texts but different cryptographic texts are available for examination. They are superimposed and appear as follows:

No. 2. W I L T G S I H W W A W K I N D C U W W A X L X


No. 2. L Q D U H K I R X Y I R N B A Y I I G T E W A Y I I L W N K I R

# FIRST COMPONENTS OF CRYPTOGRAM No. 1

| A | B | C | D | E | F | G | H | I | K | L | M | N | O | P | Q | R | S | T | U | V | W | X | Y | Z |
| A | T₁,₁ | T₂,₁ | | | | | | N₄,₁ | | | | | | | | | | | | | | | | | | |
| B | | | L₄,₁ | G₄,₁ | D₄,₁ | | | | | | | | | | | | | | | | | | | | | | |
| C | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| D | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| E | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| F | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| G | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| H | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| I | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| K | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| L | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| M | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| N | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| O | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| P | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Q | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| R | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| S | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| T | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| U | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| V | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| W | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| X | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Y | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Z | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

*SECOND COMPONENTS OF CRYPTOGRAM No. 1*

| A | B | C | D | E | F | G | H | I | K | L | M | N | O | P | Q | R | S | T | U | V | W | X | Y | Z |
| A | T₃,₁ | T₄,₁ | | | | | | N₅,₁ | | | | | | | | | | | | | | | | | | |
| B | | | L₅,₁ | G₅,₁ | D₅,₁ | | | | | | | | | | | | | | | | | | | | | | |
| C | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| D | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| E | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| F | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| G | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| H | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| I | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| K | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| L | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| M | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| N | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| O | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| P | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Q | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| R | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| S | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| T | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| U | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| V | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| W | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| X | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Y | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Z | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
c. Now consider the first few superimposed letters in these two cryptograms:

```
No. 1............. G C O D M G C E G
No. 2............. W I L T G S I H W
```

Take the pair of superimposed letters GW. The G is the cipher resultant of the recombination of two bipartite numerical components that apply to the recomposition checkerboard. The actual identities of these numerical components are not known, but, whatever they be, the first of them determines the first half of G, the second determines the second half of G. Therefore, for cryptanalytic purposes, the actual, but unknown, numerical components may be represented by the symbols G₁ and G₂, the former referring to the row coordinate of the recomposition checkerboard, the latter to the column coordinate. What has been said of the letter G applies also to the letter W, the equivalent of G in another checkerboard. It will be found that this manner of designating bipartite components by means of subscripts to the letters themselves is a very useful method of handling the letters.

d. Let the first few letters of the two cryptograms be replaced by these same cipher letters with their subscripts to indicate components. Thus:

```
No. 1............. C G C O D M G C E
Components...... C₁C₂ C₂C₂ O₁O₂ D₁D₂ M₂M₂ G₁G₂ C₁C₂ E₁E₂
No. 2............. I W I L T G S I H
Components...... I₁I₂ W₁W₁ I₁I₂ L₁L₂ T₁T₂ G₁G₂ S₁S₂ I₁I₂ H₁H₂
```

Now from the method of encipherment it is clear that C₂Q₁ and I₂W₁ represent the same plain-text letter, since both messages are assumed to contain identical plain texts. That is, C₂G₁ of cryptogram No. 1 = I₂W₁ of cryptogram No. 2. Likewise G₂C₁ = W₂I₁; C₂O₂ = I₂L₁; G₂D₁ = L₂T₁; and so on.

e. Let all the component pairs of the cryptograms be equated in this manner and let these pairs be distributed in a table, such as that shown in figure 86. It will be seen in figure 86 that, for example, A₂C₁ of cryptogram No. 1 = T₂I₁ and T₂N₁ of cryptogram No. 2. This means that I₁ and N₁ must represent the same row coordinate of the recomposition checkerboard for cryptogram No. 2; in other words I and N must be in the same row in that checkerboard. Again, in figure 86, it is seen that C₂G₁ = G₂W₁ and D₂W₁, which means that G and D must be in the same column in that checkerboard. Again, A₂S₁ = W₂A₁ = S₂C₁; this means that A and C are in the same row, W and S, in the same column, in the recomposition checkerboard for cryptogram No. 2. All these data in figure 86 are studied with the following results:

<table>
<thead>
<tr>
<th>In same row:</th>
<th>In same column:</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) I, N, L, W</td>
<td>(1) T, H, Y, Q, U</td>
</tr>
<tr>
<td>(2) A, C, S, D</td>
<td>(2) W, C, K, B, I</td>
</tr>
<tr>
<td>(3) Q, V, X, Y</td>
<td>(3) S, L, E, D, G</td>
</tr>
<tr>
<td>(4) B, G</td>
<td>(4) R, F, H, V</td>
</tr>
</tbody>
</table>
An attempt is now made to bring together these results to reconstruct the recomposition checkerboard for message No. 2. This yields the following:

\[
\begin{array}{ccc}
\text{U} & \text{K} & \text{R} \\
\text{T} & \text{L} & \text{M} \\
\text{N} & \text{W} & \text{K} \\
\text{D} & \text{C} & \text{S} \\
\text{G} & \text{H} & \text{B} \\
\text{V} & \text{Y} & \text{X} \\
\end{array}
\]  

Compare this with the recomposition checkerboard shown in figure 85 (B). Enough has been shown to illustrate the procedure. If there were just a little more text, probably all 25 letters of the checkerboard could be definitely placed.

By making a reciprocal table for equivalencies between component pairs in cryptogram No. 2, the data obtained would permit of reconstructing the recomposition checkerboard for cryptogram No. 1. Having these checkerboards completely or at least partially reconstructed, the reconstruction of the decomposition checkerboards is a relatively easy matter and follows the procedure described in paragraph 47f.

The complete solution of the two cryptograms, including the decomposition and recomposition matrices, is as follows:

\[
\text{No. 1}
\]

\[
\begin{array}{lcccc}
1 & 2 & 3 & 4 & 5 \\
R & E & F & L & C \\
T & I & N & G & P \\
O & A & B & D & H \\
K & I & M & Q & S \\
V & W & X & Y & Z \\
\end{array}
\]

\[
\text{No. 2}
\]

\[
\begin{array}{lcccc}
1 & 2 & 3 & 4 & 5 \\
W & I & L & N & O \\
T & U & E & R & M \\
A & M & B & C & S \\
R & W & G & D & H \\
Q & V & X & Z \\
\end{array}
\]
General principles underlying the solution.—a. It will be noted that the periods in the foregoing example contain an odd number of letters. The result of adopting odd-length periods is to impart a much greater degree of cryptographic security to the system than is the case when even-length periods are involved. This point is worth elaborating upon to make its cryptanalytic significance perfectly clear. Note what happens when an even period is employed:

<table>
<thead>
<tr>
<th>Plain Text</th>
<th>Cipher Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>ONEPLANE</td>
<td>LREHA</td>
</tr>
<tr>
<td>LREHA</td>
<td>LREHA</td>
</tr>
<tr>
<td>LREHA</td>
<td>LREHA</td>
</tr>
<tr>
<td>LREHA</td>
<td>LREHA</td>
</tr>
<tr>
<td>LREHA</td>
<td>LREHA</td>
</tr>
</tbody>
</table>

b. A different checkerboard may, of course, be employed for the recomposition process. Also, periods of any convenient length may be adopted; or, in a complicated case, periods of varying lengths may be employed in the same cryptogram, according to some prearranged key.

50. General principles underlying the solution.—a. It will be noted that the periods in the foregoing example contain an odd number of letters. The result of adopting odd-length periods is to impart a much greater degree of cryptographic security to the system than is the case when even-length periods are involved. This point is worth elaborating upon to make its cryptanalytic significance perfectly clear. Note what happens when an even period is employed:

<table>
<thead>
<tr>
<th>Plain Text</th>
<th>Cipher Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>ONEPLANE</td>
<td>LREHA</td>
</tr>
<tr>
<td>LREHA</td>
<td>LREHA</td>
</tr>
<tr>
<td>LREHA</td>
<td>LREHA</td>
</tr>
<tr>
<td>LREHA</td>
<td>LREHA</td>
</tr>
<tr>
<td>LREHA</td>
<td>LREHA</td>
</tr>
</tbody>
</table>

Now if each 6-letter cipher group is split into three parts and the letters are taken alternately from each section (Ex. L H L R E A = L R H E L A) the results are exactly the same as would be obtained in case a simple digraphic encipherment were to be employed with the 2-square checkerboard shown in figure 88. For example, ONp = LRe; EPp = HBL, and so on. Encipherment of this sort brings about a fixed relationship between the plain-text digraphs and their cipher equivalents, so that the solution of a message of this type falls under the category of the cryptanalysis of a case of simple digraphic substitution, once the length of the period has been established. The latter step can readily be accomplished, as will be seen presently. In brief, then, it may be said that in this system when encipherment is based upon even periods the cipher text is purely and simply digraphic in character, each plain-text digraph having one and only one cipher-text digraph as its equivalent.

b. But the latter statement is no longer true in the case of odd periods. Note, in the example under paragraph 49a, that the cipher equivalent of the first plain-text digraph of the first group, ON, is composed of the initial and final component of the letter L, the final component of the letter O, and the initial component of the letter E. That is, three different plain-text letters, L, O, and E, are involved in the composition of the cipher equivalent of one plain-text digraph, ON. Observe now, in the following examples, that variants may be produced for the digraph ONp.

<table>
<thead>
<tr>
<th>Plain Text</th>
<th>Cipher Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>ONEPLANE</td>
<td>LREHA</td>
</tr>
<tr>
<td>LREHA</td>
<td>LREHA</td>
</tr>
<tr>
<td>LREHA</td>
<td>LREHA</td>
</tr>
<tr>
<td>LREHA</td>
<td>LREHA</td>
</tr>
<tr>
<td>LREHA</td>
<td>LREHA</td>
</tr>
</tbody>
</table>

c. The foregoing examples fall into two classes. In the first, where the O of ON, falls in an odd position in the period, the first letter of the trigraphic cipher equivalent must be an L, the second must be one of the 5 letters in the second column of the substitution checkerboard, the third must be one of the 5 letters in the third row of the checkerboard.

51. Periodic fractionating systems.—a. Another type of combined substitution-transposition system involving fractionation is that in which the processes involved are applied to groupings of fixed length, so that the system gives external evidence of periodicity. One such system, commonly attributed to the French cryptographer Delastelle, is exemplified below. Let the bipartite alphabet be based upon the 25-cell substitution checkerboard shown in figure 80. Let the message to be enciphered be ONE PLANE REPORTED LOST AT SEA. Let it also be assumed that by agreement between correspondents, periods of 5 letters will constitute the units of encipherment. The bipartite equivalents of the plain-text letters are set down vertically below the letters. Thus:

<table>
<thead>
<tr>
<th>Letter</th>
<th>Cipher</th>
</tr>
</thead>
<tbody>
<tr>
<td>O</td>
<td>A</td>
</tr>
<tr>
<td>N</td>
<td>E</td>
</tr>
<tr>
<td>E</td>
<td>P</td>
</tr>
<tr>
<td>L</td>
<td>A</td>
</tr>
<tr>
<td>A</td>
<td>N</td>
</tr>
</tbody>
</table>

Recombinations are effected horizontally within the periods, by joining components in pairs, the first period yielding the pairs 41, 34, 32, 33, 31. These pairs are then replaced by letters from the original checkerboard, yielding the following:

<table>
<thead>
<tr>
<th>Plain Text</th>
<th>Cipher Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>ONEPLANE</td>
<td>LREHA</td>
</tr>
<tr>
<td>LREHA</td>
<td>LREHA</td>
</tr>
<tr>
<td>LREHA</td>
<td>LREHA</td>
</tr>
<tr>
<td>LREHA</td>
<td>LREHA</td>
</tr>
<tr>
<td>LREHA</td>
<td>LREHA</td>
</tr>
</tbody>
</table>

b. A different checkerboard may, of course, be employed for the recomposition process. Also, periods of any convenient length may be employed; or, in a complicated case, periods of varying lengths may be employed in the same cryptogram, according to some prearranged key.

52. General principles underlying the solution.—a. It will be noted that the periods in the foregoing example contain an odd number of letters. The result of adopting odd-length periods is to impart a much greater degree of cryptographic security to the system than is the case when even-length periods are involved. This point is worth elaborating upon to make its cryptanalytic significance perfectly clear. Note what happens when an even period is employed:

<table>
<thead>
<tr>
<th>Plain Text</th>
<th>Cipher Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>ONEPLANE</td>
<td>LREHA</td>
</tr>
<tr>
<td>LREHA</td>
<td>LREHA</td>
</tr>
<tr>
<td>LREHA</td>
<td>LREHA</td>
</tr>
<tr>
<td>LREHA</td>
<td>LREHA</td>
</tr>
<tr>
<td>LREHA</td>
<td>LREHA</td>
</tr>
</tbody>
</table>

Now if each 6-letter cipher group is split into three parts and the letters are taken alternately from each section (Ex. L H L R E A = L R H E L A) the results are exactly the same as would be obtained in case a simple digraphic encipherment were to be employed with the 2-square checkerboard shown in figure 88. For example, ONp = LRe; EPp = HBL, and so on. Encipherment of this sort brings about a fixed relationship between the plain-text digraphs and their cipher equivalents, so that the solution of a message of this type falls under the category of the cryptanalysis of a case of simple digraphic substitution, once the length of the period has been established. The latter step can readily be accomplished, as will be seen presently. In brief, then, it may be said that in this system when encipherment is based upon even periods the cipher text is purely and simply digraphic in character, each plain-text digraph having one and only one cipher-text digraph as its equivalent.

b. But the latter statement is no longer true in the case of odd periods. Note, in the example under paragraph 49a, that the cipher equivalent of the first plain-text digraph of the first group, ON, is composed of the initial and final component of the letter L, the final component of the letter O, and the initial component of the letter E. That is, three different plain-text letters, L, O, and E, are involved in the composition of the cipher equivalent of one plain-text digraph, ON. Observe now, in the following examples, that variants may be produced for the digraph ONp.

<table>
<thead>
<tr>
<th>Plain Text</th>
<th>Cipher Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>ONEPLANE</td>
<td>LREHA</td>
</tr>
<tr>
<td>LREHA</td>
<td>LREHA</td>
</tr>
<tr>
<td>LREHA</td>
<td>LREHA</td>
</tr>
<tr>
<td>LREHA</td>
<td>LREHA</td>
</tr>
<tr>
<td>LREHA</td>
<td>LREHA</td>
</tr>
</tbody>
</table>

c. The foregoing examples fall into two classes. In the first, where the O of ON, falls in an odd position in the period, the first letter of the trigraphic cipher equivalent must be an L, the second must be one of the 5 letters in the second column of the substitution checkerboard, the third must be one of the 5 letters in the third row of the checkerboard. Therefore, L may combine with 5X5 or 25 pairs of letters to form the first and second letters of the 3-letter equivalent of ON. In the other class, where the O of ON, falls in an even position in the period, the first letter of the equivalent must be one of the 5 letters in the fourth column of the checkerboard, the second must be one of the 5 letters in the first row, and the third letter must be E. Therefore, E may combine with 5X5 or 25 pairs of letters to form the first and second letters of the 3-letter equivalent of ON, in this position in the period. Hence, ON, may be represented by 50 trigraphic combinations; the same is true of all other plain-text digraphs. Now if the system based upon even periods is considered as a simple digraphic substitution, the foregoing remarks lead to characterizing the system based upon odd periods as a special type of digraphic substitution with variants, in which 3 letters represent 2 plain-text letters.

An example of the solution of a cryptogram of this type was given in Military Cryptanalysis, Part I, sec. IX.
d. However, further study of the odd-period system may show that there is no necessity for trying to handle it as a digraphic system with variants, which would be a rather complex affair. Perhaps the matter can be simplified. Referring again to the example of encipherment in paragraph 49c:

\[
\begin{array}{cccccc}
\text{ONEPL} & \text{ANERE} & \text{PORTE} & \text{DLOST} & \text{ATSEA} \\
\text{LHOEB} & \text{MDDEE} & \text{QTERR} & \text{HQTAW} & \text{APAGD}
\end{array}
\]

Now suppose that only the cipher letters are at hand, and that the period is known. The first cipher letter is L, and it is composed of two numerical components represented by the symbols Ll and L2, the former referring to the row coordinate of the substitution checkerboard, the latter to the column coordinate. The same thing may be done with the components of the second cipher letter, the third, fourth, and fifth, the respective components being placed into their proper positions in the period. Thus:

\[
\begin{array}{c}
\text{Cipher} \\
\text{Components}
\end{array}
\]

Now let the actual plain-text letters be set into position, as shown at the right in the two diagrams below:

\[
\begin{array}{c}
\text{Plain text} \\
\text{Components}
\end{array}
\]

By comparing the two diagrams it becomes obvious that Ll, H2, and 03 all represent the coordinate 4; H4, E5, and B5 all represent the coordinate 5, and so on. If this equivalence were known for all the 50 combinations of the 25 letters with subscript 1 or 2 there would be no problem, for the text of a cryptogram could be reduced to 25 pairs of digits representing monalphabetic encipherment. But this equivalence is not known in the case of a cryptogram that is to be solved; basically the problem is to establish the equivalence.

e. It is obvious that the vertical pair of components 02 represents \(O_2\), the vertical pair \(L_3\) represents \(N_5\), and so on. The complete example therefore becomes:

\[
\begin{array}{c}
\text{Plain text} \\
\text{Components}
\end{array}
\]

Thus, any given plain-text letter is of two classes: \(O_2\) and \(O_5\), and the two must be kept separate in cryptanalysis.

f. Now consider the equivalent of \(O_3\) in the first period. It is composed of \(L_4\). This is only one of a number of equivalents for \(O_3\) in an odd position in the period. The row of the substitution checkerboard indicated by \(L_4\) may be represented by 4 other components, since that row contains 5 letters. Therefore the upper component of the \(O_3\) equivalent of \(O_3\) may be any one of 5 letters. The same is true of the lower component. Hence, \(O_3\) in an odd position in the period may be represented by any one of \(5 \times 5 = 25\) combinations of vertical components in the sequence \(O_2\equiv O_5\). \(O_3\) in an even position in the period may be represented by any one of a similar number of combinations of vertical components in the reverse sequence, \(O_5\equiv O_2\). Thus, disregarding the position in the period, this system may be described as a mononalphabetic substitution with variants, in which every plain-text letter may be represented by any one of 50 different component-pairs. But in studying an actual cryptogram in this system, since the position (odd or even) is determined by a cipher letter in the period, this observation may be made. Because of the mechanics of encipherment in this system the first cipher letter and the first plain-text letter in each period must come from the same row in the substitution checkerboard. Since there are only 5 letters in a row in the checkerboard the probability that the two letters referred to will be identical is \(1/5\). (The identity will occur every time that the coordinate of the row in which the second plain-text letter stands in the checkerboard is the same as the coordinate of the column in which the first plain-text letter stands.) The same general remark applies to the second cipher letter and the third plain-text letter; as well as to the third cipher letter and the fifth plain-text letter; and in these two the letters must come from the same row in the checkerboard and the probability that they will be identical is likewise \(1/5\). (The identity in the former case will occur every time that the coordinate of the row in which the fourth plain-text letter stands in the checkerboard is the same as that of the column in which the third plain-text letter stands; in the latter case the identity will occur every time that the coordinate of the column in which the fifth plain-text letter stands is the same as that of the column in which the fifth plain-text letter stands.) The last of the foregoing sources of identity is exemplified in only 4 of the 9 cases mentioned in subparagraph e above. These
involve the fifth plain-text letter in the third, fourth, and fifth periods, and the first letter in the fifth period, wherein it will be noted that the 2nd component standing directly under the plain-text letter is identical with the latter in each case.

b. How are the other 5 cases of identity brought about? Analysis along the same lines as indicated above will be omitted. It will be sufficient to observe that in each of those cases it is the 2nd component which is identical with the plain-text letter involved, and again the probability of the occurrence of the phenomenon in question is 1/5.

c. Since the probability of the occurrence of the event in question is 1/5 for the 0th components and 1/5 for the 2nd components, the total probability from either source of identity is 2/5. This probability applies only to the letters occupying odd positions in the period, and it may be said that in 40 percent of all cases of letters in odd positions in the periods the one or the other of the two cipher components will be identical with the plain-text letter.

d. As regards the plain-text letters in even positions, analysis will show why only in very few cases will either of the cipher components coincide with the plain-text letter to which they apply, for the method of finding equivalents in the substitution checkerboard is to take the first component as the row coordinate indicator and the second component as the column indicator; a reversal of order will give wholly different letters in the cipher alphabet. The number of repeated occurrences will be a function of the length of the period; the interval between the letters constituting the repetition will be a function of the length of the period and the position of the repeated polygraph in two periods in which the two polygraphs occur.

Finally, what of the peculiar phenomenon to be observed in the case of the first column of the fifth period of the example in subparagraph 7? Here is a case wherein the plain-text value of a pair of superimposed components is unmistakably indicated directly by the cipher components themselves. Studying the cipher group concerned it is noted that it contains 2 A's separated by one letter, that is, the A's are 2 intervals apart. This situation is as though the plain-text letter were entirely self-enciphered in this case. Now it is obvious that this phenomenon will occur in the case of periods of 5 letters every time that within a period a letter occurs 3 times in an interval of 2, for this will bring about the superimposition of a 0th and 2nd component standing directly under the plain-text letter and therefore the plain-text letter is indicated directly. This question may be pertinent: How many times may this be expected to happen? Analysis along the lines already indicated will soon bring the answer that the phenomenon in question may be expected to happen 4 times out of 100 in the case of letters in odd positions and only 8 times out of 1,000 in the case of letters in even positions. In the latter cases the letters involved are those falling in the diagonal sloping from left to right in the substitution matrix.
letters separated by an interval $x$, it is probable that these repetitions represent repetitions of plain-text trigraphs which occupy homologous positions in the period. The interval $x$ (between the letters constituting the repetition in the cipher text) then gives a good clue to the length of the period: $p = 2x \pm 1$.

d. A special kind of index is prepared to facilitate the search for repetitions of the nature indicated. If tabulating machinery is available, an alphabetically arranged index showing say 10 succeeding letters after each $A_0, B_0, C_0, \ldots, Z_0$ is prepared for the cryptogram. Then this index is studied to see how many coincidences occur at various intervals under each letter. For example, under $A_0$ one looks to see if there are 2 or more cases in which the same letter appears 2, 3, 4, \ldots intervals to the right of $A$, a record being kept of the number of such cases under each interval. The same thing is done with reference to $B_0, C_0,$ and so on. The tallies representing coincidences may be amalgamated for all the letters $A, B, C, \ldots, Z$, only the intervals being kept segregated. When tabulating machinery is not available, the search for repetitions may be made by transcribing the cryptogram on two long strips of cross-section paper, juxtaposing the strips at $A, B, C, \ldots, Z$, and noting the coincidences occurring 1, 2, 3, \ldots up to say 10 letters beyond the juxtaposed letters. For example, beginning with $A_0$, the two strips are juxtaposed with the first $A$ on one against the first $A$ on the other. Note is made of any coincidences found within 10 letters beyond the $A$'s, and a record is kept of such coincidences according to intervals. Keeping one strip in position the other is slid along to the second $A$, and again coincidences are sought. All the $A$'s are treated in this way, then the $B$'s, $C$'s \ldots $Z$'s. The record made of the coincidences may consist merely of a tally stroke written under the intervals 1, 2, 3, \ldots 10. That interval which occurs more frequently than all the others is probably the correct one. This interval times 2 plus or minus 1 is the length of the period. There are, therefore, only two alternatives. A choice between the two alternatives may then be made by transcribing the text or a portion of it according to each hypothesis. That transcription which will most often throw the two members constituting a repetition into one and the same period is most likely to be correct.

e. Finally, for ascertaining the period there is one method which is perhaps the most laborious but surest. It has been pointed out that this system reduces to one that may be described as monoalphabetic substitution with variants. If the cipher text is transcribed into $\theta_1$ and $\theta_2$ components according to various assumed periods, and then a frequency distribution is made of the pairs of vertical components for each hypothesis, that period which gives the best approximation to the sort of distribution to be expected for a system of monoalphabetic substitution with 25 variants for each letter may be taken to be correct. For in the case of an incorrect period the resultant vertical bipartite components are not the equivalents of the actual plain-text letters; hence such repetitions as occur are purely accidental and the number of such cases would be rather small. But in the case of the correct period the resultant vertical pairs of components are the equivalents of the actual plain-text letters; hence repetitions are causal and fairly frequent. Were it not for variants, of course, the distribution would be perfectly monoalphabetic.
|       | A | B | C | D | E | F | G | H | I | K | L | M | N | O | P | Q | R | S | T | U | V | W | X | Y | Z |
| **A** |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| **B** |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| **C** |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| **D** |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| **E** |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| **F** |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| **G** |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| **H** |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| **I** |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| **K** |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| **L** |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| **M** |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| **N** |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| **O** |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| **P** |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| **Q** |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| **R** |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| **S** |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| **T** |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| **U** |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| **V** |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| **W** |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| **X** |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| **Y** |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| **Z** |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |

**Figure 54.**

**Second (56) Components**
**First (9b) Components**

|   | A | B | C | D | E | F | G | H | I | K | L | M | N | O | P | Q | R | S | T | U | V | W | X | Y | Z |
| 0 | A | B | C | D | E | F | G | H | I | K | L | M | N | O | P | Q | R | S | T | U | V | W | X | Y | Z |
| 1 |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 2 |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 3 |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 4 |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 5 |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 6 |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 7 |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 8 |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 9 |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 10|    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 11|    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 12|    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 13|    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 14|    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 15|    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 16|    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 17|    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 18|    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 19|    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 20|    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |

**Second (9b) Components**

|   | A | B | C | D | E | F | G | H | I | K | L | M | N | O | P | Q | R | S | T | U | V | W | X | Y | Z |
| 0 |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 1 |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 2 |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 3 |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 4 |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 5 |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 6 |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 7 |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 8 |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 9 |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 10|    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 11|    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 12|    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 13|    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 14|    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 15|    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 16|    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 17|    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 18|    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 19|    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 20|    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
52. Illustration of solution.—a. With the foregoing principles in mind, the following cryptogram will be studied:

```
1  K Z F B E I L  Y Y M O C B R B  L Z D O T G  B L P K Y W C U C E P Q L
A M E Y L Z Q X W H L R W Q Y  D R W B M T  I Z E B E L A Y E S O B R Y
Q V B B L Y X A  X N A B Q B D O  Y M Q D L W L  N A C O X C R R G A S W Q B
C A U G G X O A R T X X T S D A Y X H K  O L S X A B R K R P U Z W H O
M T D H T S G M L S L Q P O U N H C I C K K A Q B D O F L E K A P R G
S X U P O W A L M A V Q H L M L A X K P W S T M C X K Q V H S I X S L
L W X L X R S G Z D F K L N Y B X M R B N A D K T T B A E O B H W V L
Y S X M B O W P G X K O R Z I U C E A D Y I D B L Z M I T A N H C A I
D N C I D D O Y I B C N O L Y U U M C E P O T D M G B F U N A H L B D
W X N X K K C S C T O X T S D A Y X H K C N L D K R R F A Y A P M H C
A N M B V G R E Z Q A T C Y I M N D L R L G M T W E T R C V V K T E D
U F D E L X H Y Q V C B L Y U D U G Y A F H N Q L K F R U C N V D L H
L Z D R E L K F A X I Y D A K Z L X X O R R P E R R R R N C I E
```

b. The long repetitions noted in the text (intervals = 210 and 35) indicate a period of either 5 or 7. By transcribing several lines of text into their 01 and 02 components according to both of these alternatives and distributing the vertically superimposed pairs, it is soon found that a period of 7 produces many more repetitions than does a period of 5. The entire text is then transcribed into its θ1 and θ2 components according to a period of 7 (see fig. 89) and complete distributions of θ1θ2 and θ2θ1 vertical pairs are made, the distributions being, of course, kept separate. They are shown in figures 90 and 91. The individual distributions show many repetitions and the distributions as a whole are very favorable for a period of 7.
c. The text now being transcribed into periods of 7, with the 01 and 02 components indicated by the cipher letters in each period, the vertical pairs of components are examined to locate cases in which the basic letters of the 01 and 02超级 components are identical, since, upon the plain-text letters indicated are at once inserted into position. In this example 10 such cases are found, one each in periods 14, 22, 26, 35, 36, 52, 59, 68, and two in period 74. All of these, of course, involve letters in odd positions in the periods. The plain-text letters thus inserted may serve as clues for assuming probable words.

d. Now if only a few equivalences can be established between a few of the 01 components, or between a few of the 02 components, or between a few 01 and 02 components, a long step forward may be taken in the solution. Perhaps some information can be found by studying figures 90 and 91. A consideration of figure 90 will soon lead to the idea that each row of frequencies can indicate only 5 different plain-text letters, one of which coincides with the indicating letter at the left of the row. Moreover, in this same figure, while there are 26 rows in all, there are really only 5 different categories of rows, each category corresponding to a row in the substitution checkerboard.

e. To explain quite clearly what is meant and how the principle can be employed in this case, assume that figure 90, instead of applying to an unknown checkerboard, applied to a known one, say that shown in figure 87. The bipartite coordinates and the letters which would occupy the cells are as seen in figure 92:

<table>
<thead>
<tr>
<th>Period</th>
<th>Letters</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A M M A N F P U U F W M N A U N F A U M A N U F</td>
</tr>
<tr>
<td>2</td>
<td>B D B D E K K H K H B B E D E H K D H B D E H K</td>
</tr>
<tr>
<td>3</td>
<td>C T C C T R G T I G C C R T R I C T R I C T R I</td>
</tr>
<tr>
<td>4</td>
<td>D B B D E K K H K B B E D E H K D H B D E H K</td>
</tr>
<tr>
<td>5</td>
<td>E B D B D E K K H K B B E D E H K D H B D E H K</td>
</tr>
<tr>
<td>6</td>
<td>F A M M A N F P U U F W M N A U N F A U M A N U F</td>
</tr>
<tr>
<td>7</td>
<td>G T C C T R G T I G C C R T R I C T R I C T R I</td>
</tr>
<tr>
<td>8</td>
<td>H D B D E K K H K B B E D E H K D H B D E H K</td>
</tr>
<tr>
<td>9</td>
<td>I D B D E K K H K B B E D E H K D H B D E H K</td>
</tr>
<tr>
<td>10</td>
<td>J C C C R T R I C T R I C T R I C T R I</td>
</tr>
<tr>
<td>11</td>
<td>K D B D E K K H K B B E D E H K D H B D E H K</td>
</tr>
<tr>
<td>12</td>
<td>L A M M A N F P U U F W M N A U N F A U M A N U F</td>
</tr>
<tr>
<td>13</td>
<td>M T C C T R G T I G C C R T R I C T R I C T R I</td>
</tr>
<tr>
<td>14</td>
<td>N D B D E K K H K B B E D E H K D H B D E H K</td>
</tr>
<tr>
<td>15</td>
<td>O T C C T R G T I G C C R T R I C T R I C T R I</td>
</tr>
<tr>
<td>16</td>
<td>P E B D E K K H K B B E D E H K D H B D E H K</td>
</tr>
<tr>
<td>17</td>
<td>Q C C R T R I C T R I C T R I C T R I</td>
</tr>
<tr>
<td>18</td>
<td>R C C R T R I C T R I C T R I C T R I</td>
</tr>
<tr>
<td>19</td>
<td>S A M M A N F P U U F W M N A U N F A U M A N U F</td>
</tr>
<tr>
<td>20</td>
<td>T T C C T R G T I G C C R T R I C T R I C T R I</td>
</tr>
<tr>
<td>21</td>
<td>U H D B D E K K H K B B E D E H K D H B D E H K</td>
</tr>
<tr>
<td>22</td>
<td>V O T C C T R G T I G C C R T R I C T R I C T R I</td>
</tr>
<tr>
<td>23</td>
<td>W E B D E K K H K B B E D E H K D H B D E H K</td>
</tr>
<tr>
<td>24</td>
<td>X E C C R T R I C T R I C T R I C T R I</td>
</tr>
<tr>
<td>25</td>
<td>Y C C R T R I C T R I C T R I C T R I</td>
</tr>
<tr>
<td>26</td>
<td>Z C C R T R I C T R I C T R I C T R I</td>
</tr>
</tbody>
</table>
Now consider the A row and the F row. The 25 letters in both rows of cells are, from the very nature of the system, identical in their sequence and there are only 5 different letters involved, each appearing 5 times. Therefore it would seem that frequency distributions corresponding to these rows should show definite characteristics by means of which they could be compared statistically. Furthermore, the Ωi coordinates applying to these two rows, A and F, indicate that A and F are in the same row in the checkerboard. What has been said of the A and F rows also applies to the M, N, and U rows, for the letters A, F, M, N, and U are all in the same row in the checkerboard (fig. 87). Perhaps a statistical test can be applied to ascertain which rows of distributions in figure 90 are similar and this in turn may give clues to the letters which fall in the same row in the checkerboard applicable to the problem in hand.

j. Again, consider the columns in figure 90. What has been said of the rows applies equally to the columns, and therefore the same sort of test may also be applied to the columns of figure 90 for clues as to the composition of the columns of the checkerboard applicable to the problem under consideration. If there were sufficient text much of the labor of solving such cases would be reduced to a matter of statistical analysis. But what sort of statistical test should be used? Obviously it should be one based upon “matching” the distributions of figure 90, but specifically what should it be? Note the distributions in rows D and M; they appear to be similar. Is it correct to apply the usual χ-test for matching two frequency distributions? Consider the composition of the rows of figure 90, and specifically consider the A and F rows, composed as follows:

A__________________ A M M A
M__________________ M A
F__________________ F M M A
M__________________ M A
U__________________ U

Here the letters in opposite cells are identical and there are only 5 different letters involved: A, M, N, F, and U. Of these only 3 are high-frequency letters in normal plain text; 2 are of medium to low frequency. But the high-frequency letters in the A row match those in the F row, the low-frequency letters in the A row also match the low-frequency letters in the F row. Hence, if frequency distributions corresponding to these rows are tested statistically, they should yield a fairly high index of coincidence. But should the constant .0667 (probability of monographic coincidence in normal English text) be used in the test? Obviously not, for this constant is derived from statistics based upon the normal frequencies of all 26 letters of the alphabet, whereas here only 5 letters are involved and the exact 5 involved in any example is determined by the composition of the checkerboard. Again, consider the A and C rows of figure 90, composed as follows:

A__________________ A M M A
C__________________ C M M A
M__________________ M A
N__________________ N A
U__________________ U

Here is a case where, by chance, high-frequency letters stand opposite high-frequency letters (A and T, N and R); medium-frequency letters stand opposite medium-frequency letters (M and G). The only case of fairly marked difference is in that of the pairing of U and I. Hence, a statistical matching of frequency distributions applying to these two rows would be apt to yield a high index of coincidence. Yet, these two rows do not belong together and to assume that the letters A and C belong in the same row in the checkerboard would block or at least retard solution. In spite of the foregoing reasoning, there nevertheless remains the feeling that a statistical matching of the rows should be possible or should at least offer some clues as to the composition of the checkerboard.

p. In applying the usual χ-test for matching two distributions use is made of the constant .0667, the probability of monographic coincidence for normal English text. This constant may be modified to meet the special conditions of the present problem. If it be assumed that the mixing of the letters in the checkerboard is fairly good, in normal cases it may be assumed that there will be 1 high-frequency letter, 3 medium-frequency letters, and 1 low-frequency letter in each row and in each column of the checkerboard. Suppose the letters in each category be as follows:

\[
\begin{array}{cccccc}
\text{High frequency} & A & E & I & N & O & R & S \\
\text{Medium frequency} & B & C & D & F & G & H & L & M & P & U \\
\text{Low frequency} & K & Q & W & V & X & Z \\
\end{array}
\]

Adding the squares of the probabilities for separate occurrence of the letters in each category:

- For high frequency: 
  \[ \frac{0.667^2}{1} = 0.4445 \]
- For medium frequency: 
  \[ \frac{0.0667^2}{3} = 0.0113 \]
- For low frequency: 
  \[ \frac{0.0067^2}{5} = 0.0001 \]

Therefore, the constant that should be applied in the χ-test for the problem under consideration is .0667, which is the sum of the squares of the average probabilities of the letters occurring in each row of figure 90 as follows:

\[ \chi^2 = 0.4445 + 0.0113 + 0.0001 = 0.4559 \]

Since each row of figure 90 contains 25 letters, composed of 5 different letters each appearing 5 times, and it is assumed that each row of the checkerboard contains 2 high-frequency letters, 2 medium-frequency letters, and 1 low-frequency letter, the rows in figure 90 will be composed of 10 high-frequency letters, 10 medium-frequency letters, and 5 low-frequency letters. Therefore, the sum of the squares of the average probabilities of the letters occurring in each row of figure 90 is as follows:

\[ \chi^2 = 0.4559 \times 10 = 4.559 \]

This, then, is the constant that should be applied in the χ-test for the problem under consideration. Suppose, for convenience, the approximation .06 is used. This is considerably less than the normal constant .0667 and means that in the case of this problem two distributions can be considered to “match” even if the number of coincidences (value of χ) is considerably less than what would be expected in the case of the normal type of frequency distribution. However, it must be remembered that even if two distributions give an observed value for χ that is close to or even greater than the expected, one can still not be certain that the two distributions are identical in their sequence. In applying the usual χ-test to the composition of the checkerboard, it may happen that the composition of the checkerboard is such that two rows have letters of about the same frequency values, as pointed out above.

\[ A \]

With this reservation in mind, let figure 90 be examined. Take rows D and M, which on casual examination look a good deals alike, as seen in figures 92 and 93.

\[ \text{As given in the table on p. 114 of Military Cryptanalysis, Part 1, Appendix 2, par. 26 (1).} \]
Applying the $x$-test, the observed value of $x = 34$, the expected value is $0.05 (23 \times 16) = 18$. An excellent match is obtained, and the hypothesis that $D$ and $M$ are in the same row in the checkerboard seems promising. Can any confirmation be found in the cryptogram itself?

It has already been pointed out that this system reduces to monoalphabetic substitution with variants. This being the case it should be possible to find manifestations of equivalence between some of the variant forms of $O_1$, vertical pairs in the cryptogram. If the student will think over the matter he will quickly see that this manifestation of equivalence is but a reflection of the principle elucidated in paragraph 46, expressed in a little different way. In other words, establishing equivalence between two $O_1$ components means that the two base letters involved belong in the same row of the checkerboard; establishing equivalence between two $O_2$ components means that the two base letters involved belong in the same column of the checkerboard.

Note the following instances of apparent equivalence between $D_1$ and $M_1$:

<table>
<thead>
<tr>
<th>Period</th>
<th>$D_1$</th>
<th>$D_2$</th>
<th>$D_3$</th>
<th>$D_4$</th>
<th>$D_5$</th>
<th>$D_6$</th>
<th>$D_7$</th>
<th>$D_8$</th>
<th>$D_9$</th>
<th>$D_{10}$</th>
<th>$D_{11}$</th>
<th>$D_{12}$</th>
<th>$D_{13}$</th>
<th>$D_{14}$</th>
<th>$D_{15}$</th>
<th>$D_{16}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

It may be assumed $D_1 = M_1$, and the two distributions in figure 93 may be amalgamated.

$$D_1 + M_1$$

The only other row in figure 90 which gives indications of being similar to this distribution is the $A$ row. Applying the $x$-test individually to the $D_1$ and $M_1$ distributions, and then to the combined $D_1 + M_1$ distribution:

<table>
<thead>
<tr>
<th>Period</th>
<th>$A_1$</th>
<th>$A_2$</th>
<th>$A_3$</th>
<th>$A_4$</th>
<th>$A_5$</th>
<th>$A_6$</th>
<th>$A_7$</th>
<th>$A_8$</th>
<th>$A_9$</th>
<th>$A_{10}$</th>
<th>$A_{11}$</th>
<th>$A_{12}$</th>
<th>$A_{13}$</th>
<th>$A_{14}$</th>
<th>$A_{15}$</th>
<th>$A_{16}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

It certainly seems as though $A_1 = D_1 = M_1$, and that these letters are in the same row in the checkerboard. This tentatively will be assumed to be correct.
Among the most frequent combinations is the pair Y, E, appearing in the following sequences:

<table>
<thead>
<tr>
<th>Period</th>
<th>Components</th>
<th>Cipher-text components</th>
</tr>
</thead>
<tbody>
<tr>
<td>35</td>
<td>L, W, X, L</td>
<td>L, W, X, L</td>
</tr>
<tr>
<td>36</td>
<td>L, W, X, L</td>
<td>L, W, X, L</td>
</tr>
</tbody>
</table>

Note how M, A, E, D, and S all appear to be interchangeable. Are these the letters which belong in the same row? The probable equivalence among A, D, E, and M has been established by noting cases of equivalency in the text. A further search will be made to see if E and S also show equivalencies with A, D, and M.

Note the following:

<table>
<thead>
<tr>
<th>Period</th>
<th>Components</th>
<th>Cipher-text components</th>
</tr>
</thead>
<tbody>
<tr>
<td>21</td>
<td>C, A, G, G</td>
<td>V, V, H, S, I</td>
</tr>
<tr>
<td>22</td>
<td>D, H, H, A</td>
<td>L, W, X, L</td>
</tr>
<tr>
<td>23</td>
<td>D, H, H, A</td>
<td>L, W, X, L</td>
</tr>
</tbody>
</table>

Here are indications that E belongs to the same series, but not enough cases where S is interchangeable with A, D, E, or M can be found to be convincing. But perhaps it is best not to go too far in these early stages. Let it be assumed for the present that A, S, E, and M are in the same row of the substitution checkerboard. In period 16 there is the pair of vertical components D, E. Since D = R, this pair may be written E, E, whereupon the plain-text letter E is immediately indicated. All cases of this sort are sought in the text and the plain-text letters are inserted in their proper places, there being 7 such instances in all, but these yield the important letters, A, D, and E.

In a similar manner, by an intensive search for cases in which components appear to be equivalent because they occur in repetitions which are identical save for one or two components, it is established that C, O, M, and W are in the same column in the checkerboard. Note the bracketing of these letters occurring as components in the 4th column of the first list of sequences in subparagraph j. Likewise, B, H, and N are established as being in the same row. Again the text is examined for cases in which plain-text letters C, O, M, W, B, H, and N may be inserted. By carrying out this process to the full extent possible, the skeletons of words will soon begin to appear.

Enough has been demonstrated to show this line of attack. Of course, if there is a large volume of text at hand, the simplest procedure would be to construct frequency distributions of the types shown in figures 60 and 91, and use the statistical method to match the individual distributions. For this method to be reliable it would be necessary to have several hundred letters of text, but this in actual practice would not be too much to expect.

There is, however, another line of attack, based upon the probable-word method. It has been pointed out that, in the case of letters in odd positions in the periods, 40 percent of the time the plain-text letter is indicated by either its O or E component. This property affords a fair basis for assuming a probable word. For example, the cryptogram here studied shows the following two periods:

<table>
<thead>
<tr>
<th>Period</th>
<th>Components</th>
<th>Cipher-text components</th>
</tr>
</thead>
<tbody>
<tr>
<td>35</td>
<td>L, W, X, L</td>
<td>L, W, X, L</td>
</tr>
<tr>
<td>36</td>
<td>L, W, X, L</td>
<td>L, W, X, L</td>
</tr>
</tbody>
</table>

The word HOSTILE is suggested by the letters H, S, I, L. This word will be assumed to be correct and it will be written out with its components under the cipher components. Thus:

<table>
<thead>
<tr>
<th>Period</th>
<th>Components</th>
<th>Cipher-text components</th>
</tr>
</thead>
<tbody>
<tr>
<td>21</td>
<td>C, A, G, G</td>
<td>V, V, H, S, I</td>
</tr>
<tr>
<td>22</td>
<td>D, H, H, A</td>
<td>L, W, X, L</td>
</tr>
<tr>
<td>23</td>
<td>D, H, H, A</td>
<td>L, W, X, L</td>
</tr>
</tbody>
</table>

This word, if correct, yields the following equivalences: H = X = O; S = O; T = S; L = S; I = O; R = R. Again the text is examined for cases in which the plain-text letters may now be directly inserted; but only one case is found, in period 44, where I = L = S = L = L = R = R = R. This is unfortunate, so that additional words will have to be assumed. The 14th period shows a C, and the components after it suggest that the word CROSSROADS may be present. Thus:

<table>
<thead>
<tr>
<th>Period</th>
<th>Components</th>
<th>Cipher-text components</th>
</tr>
</thead>
</table>

Take the first letter R, represented by C, R.

Since R = C, R = R.

Therefore, R = C = R.

Hence R = C and R = R.

Therefore, R = R = C.

Again, in the case of the first O,

O = O = R.

But O = O = O = R.

Therefore, R = R = O = C.
The various equivalencies yielded are as follows:

\[ C_0 = R_0 = S_0 = R_1 = G_1 = W_1 = Q_1 \]
\[ B_1 = D_1 \]
\[ S_1 = W_1 = B_1 = T_1 \]
\[ H_1 = X_0 = G_1 \]
\[ Q_0 = Q_1 \]
\[ Q_0 = A_0 = D_1 \]

Let all the equivalencies found thus far from subparagraphs \( e, f \), and \( A \) be collected in two tables, as shown in figures 94, one for \( 0_1 \) \( 0_2 \) combinations, the other for \( 0_3 \) \( 0_4 \) combinations.


A study of the equivalencies indicates that:
1. \( A, D, E, W \) belong in the same row.
2. \( B, H, N \) belong in the same row.
3. \( G, R, S, Q \) belong in the same row.
4. \( R, C, O, M, W \) belong in the same column.
5. \( J, L, A, Q \) belong in the same column.
6. \( X, H, G \) belong in the same column.
7. \( T \) coordinates of \( R \) and \( A \) are identical and hence this letter occupies a cell along a diagonal sloping from left to right in the matrix.

Since a row or a column can contain only 5 letters, it is obvious that \( A, D, E, M; B, H, N; \) and \( G, R, S, W \) fall in 3 different rows; \( O, G, M, W, R \) and \( J, L, A, Q \) fall in different columns. A start may be made by an arbitrary placement of \( R \) in the position 1-1, and since \( R_1 = Q_1 \) =

\[ M_0 = W_0 \] this means that \( R, O, C, M, \) and \( W \) form one column in the substitution checkerboard, as shown in figure 95-A. The data also indicate that \( R, G, Q, \) and \( S \) must be in row 1, \( A, D, \) and \( E \) must be in row 2, \( 0_1 \) \( 0_2 \) and \( 0_3 \) \( 0_4 \) must be in row 3, \( 0_3 \) \( 0_4 \) must be in row 4, and that \( 0_3 \) for \( H \) and \( X \) must be 3. And since \( M_0 = X_0 = 0_4 \) for \( I \) and \( L \) must be 4. Substituting in the text the coordinates for the known values, additional plain-text words soon become evident. For example, taking the periods with the word HOSTILE, it becomes possible to insert

Period.............. 35
Plain text.......... HOSTILE
Components........... A T X S L 1
Cipher text......... V H S I X S L

the letters \( R_1 \) and \( 0_3 \) as the second and fourth letters after \( 0_3 \), suggesting that the word after HOSTILE is TROOP. This gives \( W_0 = T_0 \), which permits of placing \( T \) in position 5-3. Since \( T \) in HOSTILE is \( S_0 \), therefore \( S_0 = S_3 = S_4 \). Since \( S_0 \) is in row 1, and \( S_3 = S_4 \), \( S_0 \) must go in position 1-5. Since \( L_0 = 4 \) and \( L_3 = 3 \), \( L \) must go in position 3-4. Since \( 0_4 \) (the 1st 0 in TROOP) = \( x_1 \) \( x_0 \) and it is known that \( 0_3 = 3-1 \), therefore \( X \) must be in position 3-3. The checkerboard is now as shown in figure 96-B. From figure 94, \( X_0 = \bar{E} \). Now \( X_4 = 5 \), and since the \( E \) must be in row 4, \( \Theta_4 = 0_5 \)

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>R</td>
<td>S</td>
<td>G</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>C</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>O</td>
<td>X</td>
<td>L</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>M</td>
<td>A</td>
<td>D</td>
<td>E</td>
</tr>
<tr>
<td>5</td>
<td>W</td>
<td>T</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[ \Theta_4 = 0_5 \]

\[ \Theta_4 = 0_5 \]
it is evident that E must occupy cell 4-3, as seen in figure 95-C. There are now only 2 possible rows for H, either 1 or 2. It is deemed unnecessary to give further details of the process. Suffice it to say that in a few minutes the entire checkerboard is found to be shown in figure 95-D. It will decipher the entire cryptogram as it stands, but speculating upon the presence of WUTVZ in the last row, and assuming a keyword mixed sequence has brought this about, a rearrangement of the columns of the checkerboard is made to give TUVWZ, as shown in figure 95-E. The arrangement of the rows now becomes quite evident and the original checkerboard is found to be as shown in figure 95-F. It seems to be based upon the key phrase XYLOPHONIC BEDLAM.

\[ \begin{array}{cc|ccccc} 
\hline 
& \emptyset_2 & \emptyset_3 & \emptyset_4 & \emptyset_5 & \emptyset_1 & \emptyset_0 \\
1 & R & K & Q & S & 1 & X & T & L & O & P \\
2 & C & N & H & I & 2 & H & N & I & C & B \\
3 & O & Y & X & L & P & 3 & E & D & A & M & F \\
4 & M & D & E & A & F & 4 & G & K & Q & R & S \\
5 & W & U & T & V & Z & 5 & T & U & V & W & Z \\
\hline 
\end{array} \]

The completely deciphered cryptogram is as follows:

\[
\begin{align*}
&\text{p. The completely deciphered cryptogram is as follows:} \\
&\text{SITUATION} & \text{ONONFRO} & \text{NTOPTWE} & \text{NTYFOUR} \\
&\text{KZFBEIL} & \text{YMMOCBR} & \text{BLZDOTG} & \text{BLPKYWC} \\
&\text{THBRIA} & \text{DEASOFL} & \text{LOWSCOL} & \text{ONFIRST} \\
&\text{UCCEPQL} & \text{AMEYLQZ} & \text{XWHLRWQ} & \text{YDRWBM} \\
&\text{BATTALI} & \text{ONFORTY} & \text{SEVENTH} & \text{INFANTR} \\
&\text{IZEBELA} & \text{YSEOBRY} & \text{QBBLYX} & \text{NABQDO} \\
&\text{YHASREA} & \text{CHEDCRO} & \text{SSROADS} & \text{EVENFIV} \\
&\text{YMQLDL} & \text{NACOXCR} & \text{RGASWQB} & \text{FDDEBA} \\
\end{align*}
\]
that there are only factoral 5 cyclical-equivalent checkerboards and not factorial 10, is that whatever permutation is applied to the row coordinates must be the same as that applied to the column coordinates in order that the referred relationship hold true. If two checkerboards have the same row coordinates but different column coordinates certain portions of the cryptographic text will decipher correctly, others incorrectly. For this reason, in working with cryptograms of this type the cryptanalyst may successfully use a checkerboard which is incorrect in part and correct in part as he progresses with the solution. It may also be added that the actual permutation of digits applied to the side and top of the checkerboard is of no consequence, so long as the permutations are identical. In other words, the permutation 5-2-1-3-4 will work just as well as 3-2-4-1-5, or 1-2-3-4-5, etc., so long as the same permutation is used for both row and column coordinates. It is the order of the rows and columns in the checkerboard which is the determining element in this system. Any arrangement of (the letters within the checkerboard) which retains the original order as regards the letters within rows and columns will work just as well as the original checkerboard.

r. A final remark may be worth adding. After all, the security of cryptograms enciphered by the bifid fractionating method rests upon the secrecy inherent in a 25-cell matrix containing a single mixed alphabet. In ordinary substitution, a single mixed alphabet hardly provides any security at all. Why does the bifid system, which also uses only a single mixed alphabet, yield so much higher a degree of security? Is it because of the transpositional features involved? Thinking about this point gives a negative answer, for after all, finding the length of the periods and replacing the cryptographic text by components based upon the cipher letters is a relatively easy matter. The transpositional features are really insignificant. No, the answer to the question lies in a different direction and may be summed up as follows. In solving a simple mixed-alphabet substitution cipher one can attack a few cipher letters (the ones of greatest frequency) and find their equivalents, yielding fragments of good plain text here and there in the cipher text. Once a few values have been established in this manner, say 6 values, the remaining 20 values are found almost from the context alone. And in establishing these 6 values, the letters involved are not so interrelated that they can be found simultaneously. The cryptanalyst may establish the values one at a time. But in the case of the bifid system the equivalents of the plain-text letters are so interrelated that the cryptanalyst is forced to assume or establish the positions of several letters in the checkerboard simultaneously, lest one or the other. In other words, to use an analogy which may be only partially justified, the solution of a simple monoalphabetic substitution cipher is somewhat like forcing one's way into an inner chamber which has a number of doors each having a single lock; the solution of a bifid fractionated cipher is somewhat like getting into a vault—there is only one door which is provided with a complex 5-combination lock and all the tumblers of the lock must be positioned correctly simultaneously before the releasing lever can drop into the slot and the door opened. Fundamentally, this principle is responsible for the very much greater degree of security which the bifid system affords as compared with that afforded by the simple monoalphabetic system. It is a principle well worth remembering and speculating upon.
c. Apparently these two cryptograms contain almost identical texts. In order to bring the identities into the form of superimposed components, it is necessary to transcribe the texts into periods of 7 and to superimpose the two messages as shown in figure 96.

d. The shifting of the second cryptogram 2 intervals to the right brings about the superimposition of the majority of \( \theta_1 \) and \( \theta_2 \) components and it may be assumed that for the most part the texts are identical. Allowing for slight differences at the beginnings and ends of the two messages, suppose a table of equivalencies is drawn up, beginning with the eighth superimposed pairs. Thus, \( N_1 = N_2 \), hence \( Q_1 = Q_2 \) and \( D_1 = D_2 \); \( N_1 = H_1 \); hence \( H_2 = H_1 \) and \( B_1 = D_2 \). Going through the text in this manner and terminating with the 42d superimposed pairs, the results are tabulated as shown in figure 97.

e. From these equivalencies it is possible to reconstruct, if not the complete substitution matrix, then at least a portion of the matrix. For example, the data show that \( N, H, B, \) and \( I \) belong in the same row; \( E \) and \( F \) belong in the same row; \( N, D, U, Y, \) and \( K \) belong in the same column, and so on. Experimentation to make all the data fit one checkerboard would sooner or later result in reconstructing the checkerboard shown in figure 95-F, and the two messages read as follows:

1. SEVENTH INFANTRY IN POSITION TO ATTACK AT FOUR AM PLAN FOUR.
2. TENTH INFANTRY IN POSITION TO ATTACK AT FOUR AM PLAN THREEX.

f. The foregoing gives a clue to what would happen in the case of an extensive traffic in which long phrases or entire sentences may be expected to occur repeatedly. By a proper indexing of all the material, identical sequences would be uncovered and these, attacked along the lines indicated, would soon result in reconstructing the checkerboard, whereupon all the messages may be read with ease.

54. Solution of trifid systems.—a. In the trifid fractionating system the cipher alphabet is tripartite in nature, that is, the plain-text letters are represented by permutations of 3 components taken in groups of 3's, thus forming a set of 27 equivalents, such as that shown below:

\[
\begin{array}{ccc}
A=111 & J=211 & S=311 \\
B=112 & K=212 & T=312 \\
C=113 & L=213 & U=313 \\
D=121 & M=221 & V=321 \\
E=122 & N=222 & W=322 \\
F=123 & O=223 & X=323 \\
G=131 & P=231 & Y=331 \\
H=132 & Q=232 & Z=332 \\
I=133 & R=233 & ?=333 \\
\end{array}
\]

b. The equivalents may, of course, be arranged in a mixed order, and it is possible to use one tripartite alphabet for decomposition and a wholly different one for recomposition. One disadvantage of such an alphabet is that it is a 27-element alphabet and therefore some subterfuge must be adopted as regards the 27th element, such as that illustrated in the footnote to paragraph 57 of Special Text No. 166, Advanced Military Cryptography, whereinZA stands for Z and ZB for the 27th character.
<table>
<thead>
<tr>
<th>No. 1</th>
<th>Q Y B L Y X</th>
<th>No. 2</th>
<th>V R N H Y X D</th>
</tr>
</thead>
<tbody>
<tr>
<td>1, 2, 3, 4, 5, 6, 7</td>
<td>8, 9, 10, 11, 12, 13, 14</td>
<td>15, 16, 17, 18, 19, 20, 21</td>
<td>22, 23, 24, 25, 26, 27, 28</td>
</tr>
</tbody>
</table>

### Figure 96

\[ O_1 O_2 \ldots A_1 B_1 C_1 D_1 E_1 F_1 G_1 H_1 I_1 K_1 L_1 M_1 N_1 O_1 P_1 Q_1 R_1 S_1 T_1 U_1 V_1 W_1 X_1 Y_1 Z_1 \]

\[ A_2 B_2 C_2 D_2 E_2 F_2 G_2 H_2 I_2 K_2 L_2 M_2 N_2 O_2 P_2 Q_2 R_2 S_2 T_2 U_2 V_2 W_2 X_2 Y_2 Z_2 \]

### Figure 97

\[ O_1 O_2 \ldots A_1 B_1 C_1 D_1 E_1 F_1 G_1 H_1 I_1 K_1 L_1 M_1 N_1 O_1 P_1 Q_1 R_1 S_1 T_1 U_1 V_1 W_1 X_1 Y_1 Z_1 \]

\[ A_2 B_2 C_2 D_2 E_2 F_2 G_2 H_2 I_2 K_2 L_2 M_2 N_2 O_2 P_2 Q_2 R_2 S_2 T_2 U_2 V_2 W_2 X_2 Y_2 Z_2 \]

---

324155-41 (Face p. 182)
c. The various types of fractionation possible in bifid systems are also adaptable in trifid systems. For example, using the alphabet shown above for recomposition as well as decomposition the encipherment of a message in periods of 7 is as follows:

Plain text: RELIEF FO YOUR REGIMENT TOMORROW
Components: 2 1 2 1 1 2 1 3 2 3 2 2 1 1 1 2 1 2 3 3 2 2 2 2 2 2
Cipher text: KA QHORR HHFWLXIZ BF? NATN NNNWROIZ

Cryptogram
KA QHORR HHFWLXIZ BFZBN ATNNN WROIZ

d. The solution of a single cryptogram of this nature would be a quite difficult matter, especially if there were nothing upon which to make assumptions for probable words. But a whole series of cryptograms could be solved, following in general the procedure outlined in the case of the bifid system, although the solution is, admittedly, much more complicated. The first step is to ascertain the length of the period, and when this has been done, transcribe the cipher text into components, which in their vertical combinations then represent monoalphabetic equivalents, with, of course, many variants for each letter of the plain text. Then a study is made to establish component equivalents, just as in the bifid system. If the text is replete with repetitions, or if a long word or a short phrase may be assumed to be present, a start may be made and once this sort of entering wedge has been forced into the structure, its further disintegration and ultimate complete demolition is only a matter of time and patience.

55. Concluding remarks on fractionating systems.—a. It goes without saying that the basic principles of fractionation in the bifid and trifid systems are susceptible to a great deal of variation and complication. For example, instead of having periods of fixed length through the message it is possible to vary the length of the periods according to some simple or complex key suitable for this purpose. Or the bifid and trifid systems may be combined into a single scheme, enciphering a text by the bifid method and then reenciphering the cipher text by the trifid method and so on. Systems of this sort may become so complex as to defy analysis, especially if the keys are constantly and frequently varied so that no great amount of traffic accumulates in any single key. Fortunately for the cryptanalyst, however, such complex systems as these, if introduced into actual usage, are attended by so many difficulties in practice that the enemy cryptographic service would certainly break down and it would not be long before requests for repetition, the transmission of the same cryptogram in different keys, and so on, would afford clues to solution. Could such systems be employed successfully in field service there is no doubt that from the standpoint of security, the cryptograms would be theoretically secure. But the danger of error and the slowness with which they could be operated by the usual cryptographic clerks are such that systems of this complexity can hardly be employed in the field, and therefore the cryptanalyst may not expect to encounter them.

b. However, the simple bifid system, the ADFGVX system, and the like, are indeed practicable for field use, have been used with success in the past, and may be expected to be in use in the future. It is therefore advisable that the student become thoroughly familiar with the basic principles of their solution and practice the application of these principles as frequently as possible. In this connection, the attention of the student is directed to the fact that there is theoretically no reason why the bipartite components of the ADFGVX system cannot be recombined by means of the same or a different checkerboard, thus reducing the cryptographic text to a form wherein it consists of 25 different letters, and at the same time cutting the length of the messages
in half. The matter is purely one of practicability: it adds one more step to the process. But it must not be overlooked that this additional step would add a good deal of strength to the system, for it would shorten, mask, distort, or entirely eliminate similar beginnings and similar endings—the two most fruitful sources of attack on this system.

56. Concluding remarks on transposition systems.—a. Simple transposition systems hardly afford any security at all; complex ones may in the case of individual or single messages afford a high degree of security. But just as soon as many cryptograms in the same key are transmitted, if they are at all irregular or provide for nulls and blank spaces. Such devices are also useful in the Civil War period.

d. Transposition designs, diagrams, or patterns are susceptible of yielding cryptograms of good security, if they are as all irregular or provide for nulls and blank spaces. Such devices are particularly difficult to solve if frequently changed.

e. Transpositions effected upon fixed-length sequences of plain text yield a low degree of security but when a transposition is applied to the cipher text resulting from a good substitution system or to the code text of cryptograms first encoded by means of an extensive code book the increase in the cryptographic security of such cryptograms is quite notable. In fact, transposition methods and designs are frequently used to "supercipher" substitution text or code and play a very important role in this field. Their great disadvantage is that inherent in all transposition methods: The addition or deletion of a single letter or two often makes the entire cryptogram unreadable even with the correct key.

f. The clues afforded by messages with similar beginnings, endings, or internal portions, and by repetitions of incorrectly enciphered messages without paraphrasing the original text are often sufficient to make a solution possible or to facilitate a solution. For this reason the cryptanalyst should note all cases wherein clues of this sort may be applicable and be prepared to take full advantage of them.